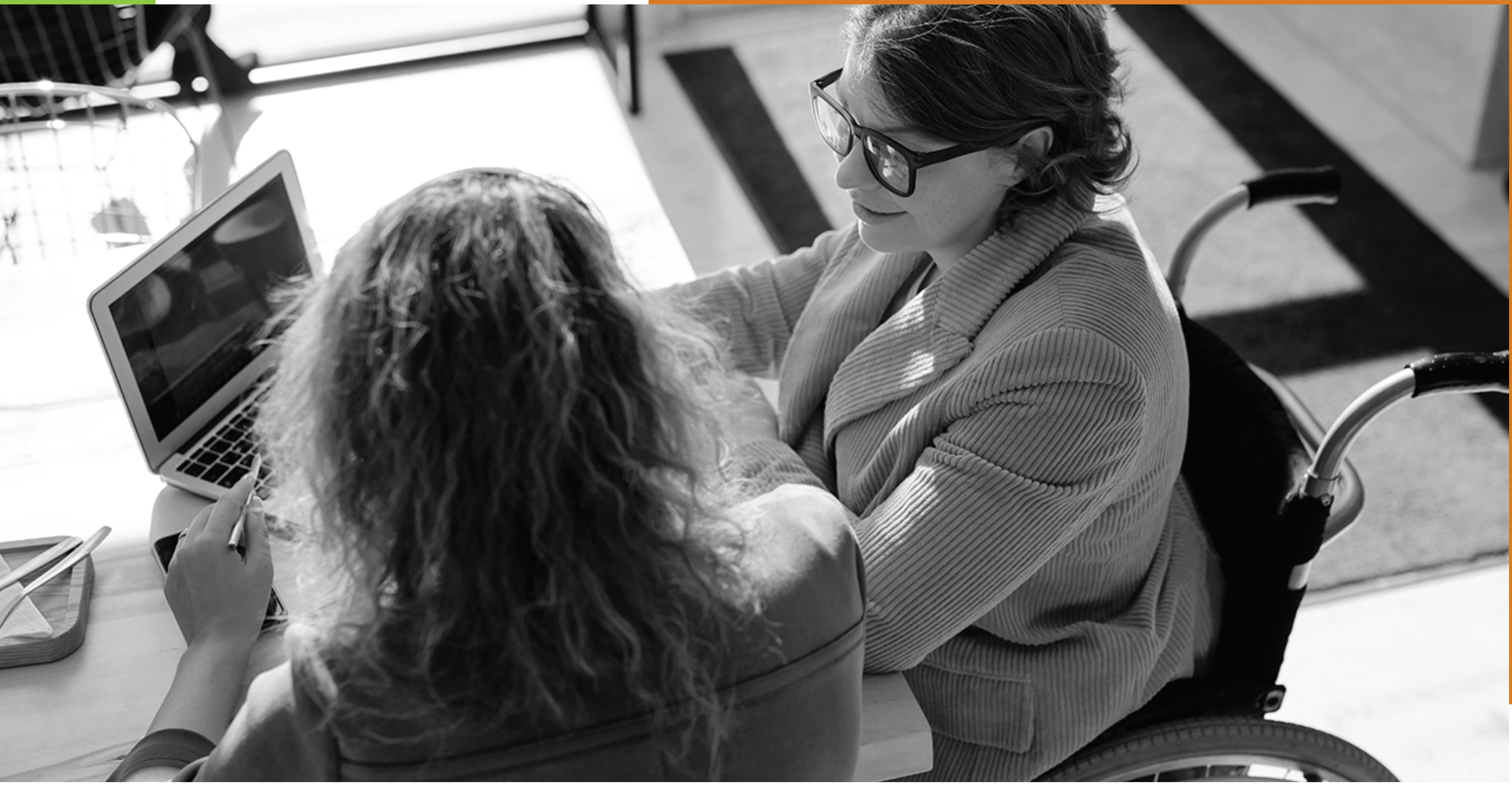
How to Forecast Skills in Demand:

A Primer



Future Skills Centre

LMIC LABOUR MARKET INFORMATION COUNCIL

Future Skills Centre

fsc-ccf.ca
@fsc_ccf_en
info@fsc-ccf.ca
Future Skills Centre

The Future Skills Centre (FSC) is a forward-thinking centre for research and collaboration dedicated to preparing Canadians for employment success. We believe Canadians should feel confident about the skills they have to succeed in a changing workforce. As a pan-Canadian community, we are collaborating to rigorously identify, test, measure, and share innovative approaches to assessing and developing the skills Canadians need to thrive in the days and years ahead. The Future Skills Centre was founded by a consortium whose members are Ryerson University, Blueprint ADE, and The Conference Board of Canada, and is funded by the Government of Canada's Future Skills program.

LMIC LABOUR MARKET INFORMATION COUNCIL

lmic-cimt.ca
@LMIC_CIMT
info@lmic-cimt.ca
LMIC / CIMT

The Labour Market Information Council (LMIC) is a non-profit research institute dedicated to ensuring Canadians have the necessary information and insights to navigate the changing world of work. Our mission is to empower Canadians with timely and reliable labour market information in an engaging way that supports their decision-making process.

This paper is funded by the Government of Canada under the Future Skills program.

Funded by the Government of Canada's Future Skills Program

Canada

Authors

Tony Bonen

Director of Research, Data and Analytics

Tony Bonen   @TonyBonen   tony.bonen@lmic-cimt.ca

Tony Bonen leads LMIC's team of labour economists. As part of the leadership team, he collaborates with colleagues across governments, academia and in the private sector to deliver high quality labour market information. He brings expertise in quantitative and data analyses, econometrics and research design.

Tony previously led the development and integration of housing price and macroeconomic stress test models for Canada Mortgage and Housing Corporation. He worked for the Schwartz Center for Economic Policy Analysis, the NATO Parliamentary Assembly, and CSR Europe.

Jacob Loree

Economist at Finance Canada / Finances Canada

The views expressed are not necessarily those of the Department of Finance

Jacob Loree

Jacob Loree is an economist with the Economic Studies and Policy Analysis Division at the Department of Finance. He works on labour market issues and provides statistical analysis to inform labour market policy. Previously, he studied at Ryerson University where he focused on occupation skills, mismatch, and search frictions.

Contents

Executive Summary

The demand for future skill needs is intensifying

Driving forces in the world of work—including population aging, climate change, globalization and technology, among others—have been disrupting employment for decades. The quantity and quality of jobs have evolved rapidly as workplaces adapt to new ways of working. In the past year, all of these changes have been exacerbated by the onset of the COVID-19 pandemic. Canadians are concerned about the future and anxious about how the world of work is changing.

For training and education systems and workforce development strategies to succeed, educators, employers, students (etc.) must understand the future of the labour market. How many jobs will be needed? Where will the jobs be, and what education and training levels will be required? To add to this complexity, while education remains a fundamental pillar of workforce development, data and information on skills are now considered essential to ensuring the country's future success. Accordingly, anticipating Canada's future skills needs is a key consideration in this regard.

There are three broad approaches to forecasting in-demand skills, each with strengths and limitations

Unfortunately, accurately forecasting future labour market demand, in general, and future skills, in particular, is an immense challenge. Already the process of forecasting jobs—also known as occupation outlooks—is complex and necessitates making considerable assumptions, not the least of which is how people transition from education into the world of work. Forecasting the skills that will be in demand presents unique challenges because unlike employment levels in certain occupations or industries, skills are not directly observed or measured. In fact, a wide variety of skills data and taxonomies exist that can introduce additional complexities and confusion. However, there are several approaches to forecasting in-demand skills that, with refinements and more testing, could lead to reasonably accurate and useful skills insights.

There are three major approaches to forecasting in-demand skills, each with unique strengths and weaknesses. These approaches include

- forecasting employment in occupations and mapping these data to skills
- forecasting skills based on those listed in online job postings
- modelling expert opinions about the future trajectories of skills to make broad predictions about those that will be in demand

Becoming conversant with these general approaches to skills forecasting and understanding the benefits and limitations of each will enable individuals and organizations to make better investment decisions in training and education today.

Approach 1: Forecast occupations in demand, then map to skills

The work of forecasting employment in occupations and mapping those occupations to skills relies on well-established forecasting methods. A wide variety of organizations, including provincial governments and industry-focused workforce development groups, use this approach to skills forecasting. Its main benefit is that the key variable—employment level by occupation—has been tracked for decades by high-quality sources like the Labour Force Survey and the census. Secondly, these forecast models can leverage other economic variables—such as GDP—whose relationships with employment have been studied for many years and are broadly accepted and understood.

However, applications of this approach assume a pre-established and fixed set of skills that are associated with each occupation. The most common skills framework used is the US O*NET system, which includes 35 skills rated in terms of their importance and complexity for more than 923 standardized (American) occupations. To apply O*NET skills to the Canadian context, one must take the additional step of translating US occupations into our National Occupational Classification (NOC) system. Once the link between occupations and skills has been established, the forecast employment growth can be converted to skills growth by calculating the average growth across occupations requiring, for example, numeracy or oral communication.

As a result, any predicted growth or decline in future skills demand depends entirely on the employment trajectories of the underlying occupations. For example, recent forecasts by the BC provincial government suggest that employment among registered nurses will increase over the next decade, meaning demand for "active listening"—their top O*NET skill—will also grow. Similarly, if all occupations that are associated today with "digital skills" are predicted to grow by, say, 10%, then "digital skills" will be predicted to grow at that same rate. As such, these forecasts do not reflect the fact that the skills required within occupations (their "skills composition") change over time. A second drawback (albeit one common to any standard economic forecast) is that this approach is unlikely to accurately forecast dramatic shifts in economic activity, whether a typical recession or structural change in the composition of the economy across sectors.

Approach 2: Directly forecast skills by leveraging data from online job postings

This method draws on an important new source of data: skills and other work requirements identified in online job postings. Extracting skills information from online job postings is achieved by using natural language processing (NLP) algorithms that categorize written text into a predefined taxonomy of work requirements (of which skills are one type). In this way, skills are directly observed at the level of each job posting. As such, they can be measured across or within occupations, industries, regions or any other feature linked to the job postings. Importantly, because there is a large volume of information available (given that there are more than two million online job postings per year in Canada), the skills data can be highly localized. In addition, with new job ads being posted each day, the data are available in near-real time.

In exploring the possibilities for skills forecasting using online job-posting data, it is crucial to understand the caveats and limitations that can limit the accuracy of forecasts. First and foremost, some work requirements are assumed to be obvious. As a result, they are often not explicitly included in the job posting. For example, only 19% of job postings for economists in 2019 listed "Microsoft Excel," likely because this is a basic requirement for economists. Second, the data are volatile and may be skewed toward certain types of jobs and regions, depending on the level of detail being explored. Finally, the algorithms used to scan, clean and categorize online job postings are typically proprietary, representing a "black box" when it comes to understanding how the skills information is created.

Despite these important caveats, direct observations of the skills in demand allow for greater flexibility in structuring forecast models. For example, given a sufficiently long history of job-posting data, one could forecast whether demand for data analysis is expected to grow among customer service-related occupations. This particular method would likely suffer from the same drawbacks noted above: namely, that shifts in economic activity would be difficult to predict accurately. In addition, as a recently developed data source, online job postings often lack a sufficient number of historical data points for standard economic forecasting. (Ideally, such models should be built upon observations from at least one full business cycle.) An alternative implementation would be to draw upon machine learning predictive models to leverage the large number of observations and data features. In either case, while the ability to forecast skills demand based on online job-posting data offers considerable promise, it remains in its infancy. A great deal of further development and testing are merited and should be explored to determine the validity of these and other ways to use these data for skills forecasting.

Approach 3: Draw on qualitative insights and expert opinions

Experts' predictions about the future trajectory of employment in selected occupations have been used as a key input for modelling broader labour market and skills trends. Typically, in this approach, experts make predictions about whether a handful of representative occupations will grow or contract in the coming months, years or even decades. In this way, the "forecasts" are qualitative assessments about the future prospects for a subset of occupations. These predictions are then projected onto all other occupations using machine learning models that link occupations based on the similarity of their skills composition. The core idea here is that expert predictions are highly responsive, forward-looking and bring together a range of historical and contextual knowledge. The hope is that these qualitative predictions can better account for trends that are important (but difficult to quantify) as well as future economic disruptions and structural shifts.

In principle, expert judgment about specific skill trajectories (rather than occupational ones) could be used as the key input. However, to our knowledge, this has not been done. Instead, applications of this approach build on the framework popularized by Frey and Osborne (2013) in which experts assess the likelihood that certain occupations will be rendered obsolete by automation. To extrapolate predictions about the subset of occupations to all sectors of the economy, they need to map occupations to a predefined, fixed set of skills. As in the first approach discussed above, this has typically been done using O*NET. The skills and other work requirements linked to a subset of occupations and assessed by experts are then used as the input data to train a machine learning algorithm that estimates the likelihood that similar occupations will grow or decline. Again, as in the first approach, skills-specific forecasts require one to map the occupational forecasts back to the set of underlying skills using a taxonomy like O*NET or Employment and Social Development Canada's (ESDC's) Skills and Competencies Taxonomy.

Projecting expert opinions to the broader economy is an innovative approach that the Brookfield Institute's Employment 2030 project has recently applied to the Canadian context. Although this approach relies on a fixed occupation–skills framework, it leverages nuanced insights that are otherwise difficult to capture when forecasting. Moreover, there are possibilities for allowing skills requirements to vary within occupations, as done by Arntz et al. (2017), who combine other skills data with the O*NET framework. It is important to note that while expert opinions about the future are nuanced, they are also limited. This was emphasized during the COVID-19 pandemic, which sent unprecedented shockwaves through labour markets in 2020.

Knowledge of skills in demand is not the same as a shortfall in skills

The focus here is on skills demand. While the current and future supply of skills is also important, the sources of information about supply are quite separate from those about demand. This difference is particularly relevant when tackling the issue of current and future skills shortages. While there are challenges in estimating the future demand for skills, as explored in here, forecasting the supply of skills also comes with unique challenges.

We can derive skills supply information by directly measuring individual skills through testing. The most well-known source of test-derived skills information is the Programme for the International Assessment of Adult Competencies (PIAAC), which rigorously tests individuals for a small number of general skills (e.g., literacy, numeracy and problem-solving) in a technology-rich environment. Given that these tests are done infrequently and that test specifications change over time, it is not clear that PIAAC and similar supply information can be used as inputs in forecasting models. Separately, there is some possibility of extracting skills information from resumés posted online, but this raises ethical issues about accessing and using personal data. Ultimately, skills supply forecasts represent an important source of information that needs to be developed with respect to the specific types of supply-side data available.

The way forward: Finding balance

There are advantages and disadvantages to each of the three approaches to forecasting skills. There is no “silver bullet” approach. Given the unique lens used by each, there is value in drawing insights from all three and using them in a complementary fashion. Further research is needed to explore the possibility of combining the strengths of the various approaches. One example might be using data from online job postings to map expert opinions to the prospects for changing skills demand. Another might be mapping occupation outlooks to skills using the work requirements identified in online job postings. As a final example, one could use expert opinions to develop scenario-based occupation outlooks in which the scenarios reflect qualitative assessments of future shocks by experts.

In the end, the usefulness of any approach depends on two considerations. First, as with any predictive model, is the forecast more accurate than what can be produced by using alternative approaches? Answering this question requires rigorous out-of-sample testing before selecting a particular approach and revisiting the predictions when data become available to learn, adjust and improve the forecast. Second, and specific in the case of skills, are the skills and other work requirements defined in such a way that actionable insights can be drawn from the results? This latter consideration applies not only to forecasting, but to any empirical research related to skills. An accurate forecast showing that “verbal communication skills” will be in greater demand is of little use to anyone (e.g., policy-makers, educators, job seekers) without further context to define what it means to have and apply this skill in a job.

To support the development of robust skills-demand forecasts, the Future Skills Centre and the Labour Market Information Council have launched a project that aims to equip front-line career service organizations with tools and insights to help Canadians navigate their career choices. Part of that work will entail testing and piloting new methods, including leveraging online job-posting data and other approaches, to better understand future skill requirements. Ultimately, drawing on the strengths of different and complementary approaches is the best way forward.

Introduction

As the economy continues to recover from the COVID-19 crisis, we need accurate predictions about the future work requirements of jobs more urgently than ever. In the spring of 2020, three million Canadians lost their jobs during an unprecedented public health crisis. Despite a strong initial rebound in employment, the recovery has been volatile as the second and third waves of COVID-19 infections forced the reintroduction of public health measures to contain the virus. The latest data available show employment remains nearly 300,000 (-1.5%) below its pre-crisis level, but these figures predate new lockdowns introduced in early April. Thankfully, vaccines are being rolled out across the country, which will be essential for a sustained and stable recovery.

Undoubtedly, many individuals will end up looking for work in new fields that demand new skills. At the same time, among other challenges, employers are likely to face skills shortages as they ramp up production. Accurate predictions about future work requirements can help Canadians make more informed choices about retraining and where to find jobs and make it easier for employers to find workers with the right skills. As well, predictions that focus on the future of the skills required for specific job types can help policy-makers design programs to correct potential skills shortages.

There are many challenges involved in economic forecasting, both in general and specifically regarding skills, including acquiring robust information about skills. A single incorrect prediction or unanticipated shock can drastically affect an entire series of forecasts. The recent economic disruptions caused by the COVID-19 pandemic are a stark example of an unanticipated shock. But even in relatively stable conditions, accurate forecasting is difficult because economies are highly complex, dynamic systems. A high-level, aggregated forecast (e.g. the employment level in Canada) stands a good chance of being accurate because it averages across the ups and downs in different sectors and regions, but it is unlikely to be informative enough for someone trying to plan a career pathway. Conversely, detailed forecasts of in-demand skills can deliver useful insights about the future, but the more granular level of analysis required presents unique challenges.

One of the reasons that it is far more challenging to forecast and predict in-demand skills than other economic indicators is because it is not possible to directly observe their demand and supply. Skills also differ from other unobserved variables in that we lack a shared taxonomy to categorize them. Without a shared language to describe skills, there will inevitably be contrasting and confusing statements about what skills are in demand now and will be in future. Skills taxonomies range from the conceptually sound but abstract (e.g. the U.S. O*NET system) to the messy but all-encompassing lists that can typically be obtained from online job postings. In the latter case, a wide variety of work requirements (of which skills are one type) ranging from "Microsoft Word" to "Leadership", and "Fast-pace setting" to "Communications skills" can be observed. Conversely, more abstract skills taxonomies include limited sets of skills. For example, Employment and Social Development Canada's (ESDC's) Skills and Competencies Taxonomy includes 47 skills among 300 work requirements. Since these taxonomies define the categories to be forecast, we must carefully consider which taxonomy (or mix of taxonomies) to use when identifying and categorizing skills information.

The different possible ways to categorize skills are closely connected with the data sources used to identify skills in demand. As a result, the potential approaches to forecasting skills depend largely on the type of skills-related observations used in the predictive model. Using this lens, three general approaches emerge regarding forecasting skills and other work requirements:

- forecasting employment in occupations and mapping these data to skills through a fixed taxonomy that links skills to occupations (e.g. O*NET)
- forecasting skills based on those listed in online job postings
- modelling expert opinions about the future trajectories of skills to make broad predictions about the skills that will be in demand

The first two approaches are entwined with the two canonical types of skills taxonomies mentioned above. The third approach—extrapolating expert opinions—is more agnostic (in principle, any skills data or taxonomy could be used). However, in practice, O*NET has primarily been used in such applications.

Although these approaches are fundamentally different from one another, they all draw on modern skills taxonomies that provide a shared benefit and drawback. Each approach enables the forecasting of "multi-dimensional skills" (see Box 1) that define a broad range of tasks applied to specific problems (e.g., equipment maintenance, time management, persuasion, data analysis etc.). This is in contrast to earlier economic analysis of skills that treated them as a single dimension: high or low (see Box 1). Analyses of multi-dimensional skills, including forecasts, can more accurately pinpoint what a worker does or what a job requires, and help identify potential skills shortages.[1] Conversely, any given taxonomy of skills posits

1 For more, see LMIC's reports. The report on skill and labour shortages is particularly relevant.

a fixed set of skill types. Although some may grow or contract, it is essentially impossible to identify and forecast *new* skills if they are not already among the categories of skills being forecast. For this reason, forecasts of a wider range of detailed work requirements (such as those based on online job postings) offer some advantages versus more narrowly defined taxonomies. However, the same general limitation still applies.

Before discussing the details of each skills-forecasting approach, the next section provides an overview of the three main sources of skills-related data. Two of these sources are used to forecast skills in demand (occupation- and job posting-derived data). The third—test-derived data—is not well suited for forecasting even though it is a crucial source of detailed and robust information about skills supply.

Box 1: Multi-Dimensional Skills

Throughout the 20th century, economists treated skills as a singular quantity, often with only two levels: that is, a job or a person was endowed with either “high” or “low” skill (first introduced by Roy, 1951). With certain jobs considered “high skill” and others designated as “low skill,” economic models could be simplified to characterize workers as possessing either a “high” or “low” level of skill. This one-dimensional approach to skills does not provide usable insights to policy-makers or individuals trying to figure out what training and education to invest in. For example, dentistry and programming would both be considered “high-skilled” occupations, but they have significantly different educational and training requirements and pathways. Moreover, the uni-dimensional treatment of skills reinforces an outdated and unfair value hierarchy that typically places university-educated, white-collar employees at the top. By identifying occupations with sets of skills (and potentially rating each required skill for its importance and/or complexity), we can better align economic theory with new and expanding datasets that contain granular information about skills.

Economists are beginning to develop modelling approaches that consider skills as multi-dimensional. For example, Loree (2020) models’ workers who possess different levels of several different skill types (e.g., cognitive, interpersonal and motor skills) to show that on-the-job training can lead to the development of new skills and/or the obsolescence of existing ones. Such approaches describe the work requirements of different occupations more accurately, so they are more applicable to informing labour market and skill development policies. Although adding dimensions to economic models increases their complexity, emerging dimension reduction techniques can help alleviate this issue (e.g., Neumann, 2006; Loree and Stacey, 2019).

Skills Data

There are two types of data related to the demand for skills: Data derived from occupation-specific skill requirements and data derived from skills listed in online job postings. The first two data sources feed, respectively, into the first two forecasting approaches listed above, whereas the expert-opinion approach can use either data type (though typically, it only uses occupation-specific data). Separately, skills supply data can be derived by directly measuring individual skills through testing. A brief discussion of the third data source—direct measurement—is provided below. However, given that direct skills testing lends itself to skills supply analyses (and the focus here is on forecasting skills demand), a more thorough discussion is left for later work.

Occupation-derived skills refer to datasets containing information on the skills linked to an occupational category as determined by occupational analysts. The O*NET database is the leading North American example.[2] O*NET contains characteristics (including skills) associated with 923 different U.S.-based occupations. Each job is rated by a set of occupational experts on two scales for each of the 35 skills in the dataset: one, "On a scale of 0–5, how important is the skill to the job?" and two, "On a scale of 1–7, what level of sophistication of a skill do you need for the job?" These two measures can be used to identify the skills workers are expected to use in specific occupations. Crucially, this type of skills information assumes that every job within that occupation demands the same skills at the same level of complexity and importance. It follows that any analysis—including forecasting—must be done at the occupations level, which in Canada means linking skills to the National Occupational Classification (NOC) system.[3] Doing so is the focus of a joint project between ESDC, Statistics Canada and the Labour Market Information Council (LMIC).

2 ESDC's Skills and Competencies Taxonomy contains information about Canadian-based occupations in a similar manner. However this taxonomy is not yet linked to Canada's National Occupation Classification (NOC) system.

3 For more information on how to use the O*NET system in a Canadian context, see LMIC's report on this topic.

Skills information derived from job postings has become increasingly available in recent years. The collection, cleaning and structuring of online job posting data are typically done by data analytics firms (e.g. Vicinity Jobs, Burning Glass Technologies and TalentNeuron) that scan multiple Canadian corporate websites and aggregators. The raw text information from job postings is then organized into a taxonomy of work requirements, including skills, work activities, knowledge domains, tools and technologies. However, there are important limitations that should be considered. The raw text in job postings reflects only what employers explicitly note, but employers often use inconsistent language. Some also assume that many work requirements will be obvious to job seekers, so leave them out. Despite these and other important limitations, online job postings offer a direct view of skills in demand. Handled properly, they can generate useful insights about employers' needs today and tomorrow.

Finally, there is test-derived skills information. In this case, data are generated from individual test results that focus on a small number of targeted skills areas. This means the data are related to the supply of skills. The most well-known source of test-derived skills information is the Programme for the International Assessment of Adult Competencies (PIAAC). PIAAC uses a standardized test across countries to measure information-processing skills like literacy, numeracy and problem-solving. PIAAC's rigorous methodology makes this skills information an excellent source for unbiased measurement of individuals' skill levels. Although it is theoretically possible to aggregate skills-tested individuals into, say, occupational groups and make inferences about occupational work requirements, the underlying source of the skills information would remain that which is supplied by individuals to those occupations, not what is demanded or needed.

A skills supply forecast would be an insightful complement to any skills-demand forecast. However, given the very different nature of skills supply data, exploring this issue is beyond the scope of the present discussion and is left for future work. In the remainder of this paper, we focus on forecasting approaches that draw on occupation-derived or job posting–derived skills data.

Summary of Approaches

Forecasting emerging skills demand is a relatively new and challenging area of applied economics. In this report, we will discuss three distinct approaches to forecasting skill demand. Table 1 summarizes these three approaches and the types of skills information required. Each method has distinct benefits and drawbacks.

Table 1: Three approaches to forecasting skills demand

	Forecasting skills through occupation outlooks	**Forecasting skills through job postings**	**Modelling expert opinion about the future**
Approach	Forecast employment change by occupation, then map to skills	Aggregate skills to level of analysis desired, then forecast	Collect expert opinion on the future trajectory of key skills or occupations, then predict trajectory of broader range of skills or occupations
Data required	Occupation-derived data (e.g. O*NET)	Online job posting–derived data	Occupation- or online job posting–derived data
Key benefits	Easy to use; enables use of standard labour market data (e.g. the Labour Force Survey)	Leverages skills information directly, not through proxy; new data updated regularly	Can account for broad trends, foreseeable disruptions and difficult-to-quantify expectations
Key drawbacks	In practice, skills are statically tied to each occupation; difficulty in predicting structural breaks in the economy	Important data limitations; methodologies untested	Forecast only as good as experts' opinion; typical methods focus on a subset of skill drivers (e.g. automation); typically assumes skills are statically tied to each occupation
Example	British Columbia's Occupation Outlook	No examples found	Brookfield Institute Employment 2030

Note that the approach that relies on modelling expert opinions about the future is, strictly speaking, not a forecast but the generalization of a small set of qualitative assessments to a larger set of data points. For this reason, the approach is not tied to any single source of skills data. However, occupation-derived skills data are the most common source used in practice.

Forecasting Skills Through Occupations

Forecasting future employment levels for various occupations is a common practice of federal, provincial and territorial governments as well as human resource councils that focus on employment in particular industries. These occupational forecasts can be linked to skills information through the expert-determined, skills-by-occupation data contained in sources such as O*NET. Therefore, in this approach, the forecast is of employment by occupation which is then mapped to skill levels.

Forecasts of employment by occupation—often called occupation outlooks—are typically made using standard econometric techniques, such as time series or macro-structural models (see Box 2). In either case, the models are time-dependent in the sense that last year's employment level helps predict this year's, and so on. Once the model has been set up (i.e. its structure and parameters are defined), future values are forecast by initializing the model with the most recent set of observations available and then "rolling forward" the predicted employment levels (i.e., this year's employment level predicts next year's, and next year's prediction is used to predict the employment level in two years' time). Evidently, any forecast error is compounded over time. Both intuitively and as a mathematical rule, the uncertainty of any forecast increases as the horizon of the prediction extends further into the future.

One of the most well-known of these occupational outlooks is ESDC's Canadian Occupational Projection System (COPS) which forecasts employment demand for more than 293 occupation groups in Canada over a 10-year period. These forecasts can be linked to skills demand by mapping from the occupational categories to expert-driven analyses of the skills needed for different occupations. For example, British Columbia has begun linking its occupational outlooks to O*NET skills thanks to a concordance between the Canadian NOC and US occupational system. The Future Skills Centre is undertaking similar work, led by the Conference Board of Canada's forthcoming Model of Occupations, Skills and Technology (MOST).

Box 2: Standard Econometric Forecasting Methods

Broadly, there are two approaches to standard econometric forecasting. The time series approach—which includes Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models—focuses on the statistical relationship of the key variable with itself over time (e.g., the temporal pattern of how GDP grows or contracts). The starting point for this approach is the assumption that the variable of interest will return to some long-run, stable trend observed in the historic data. Note that this "trend" can be defined as the level, growth rate, volatility, or any other statistical feature. Time series econometrics is a highly developed area of research, and many advanced, complex extensions to the basic ARIMA framework exist, including those that use exogenous regressors (so-called "ARMAX" models), multivariate models (e.g., vector autoregression [VAR] models), smooth-transition autoregression models (i.e. non-fixed, state-dependent model parameters) and dynamic panel data models.

Macroeconomic structural models also rely on historic, time-dependent relationships to generate forecasts. However, they do so through a series of interrelated variables or, formally, a system of equations. This system of two or more dynamic variables enables more complex (e.g. nonlinear) dynamics to be captured; unlike multivariate VAR models, each constituent equation in a structural model can be independently defined and parameterized. As such, structural models offer a great deal of flexibility in defining the relationships between the past and future values of variables. For example, yesterday's employment may not directly affect today's employment, but it might influence today's GDP, which in turn affects current employment. As the number of variables and equations increase, structural models can become quite complex and include features like state-dependent parameters. However, no matter how complex the structural model is, it must include at least one time series–type equation, whether an autoregressive or distributed lag equation.

The intuition behind this process is straightforward: if an occupation is growing in terms of the number of people likely to be employed in it, then demand for the skills associated with it must be increasing. As an example, if numeracy skills are important to the occupation of "programmer," and the occupation outlook predicts more programmers in the future, then the quantity of numeracy skills in demand is predicted to grow. By aggregating across occupations by their associated – and fixed[4] – skill levels, we can observe that, for example, employment among occupations requiring a high level of numeracy is expected to grow by X% in five years.

There are many benefits to this forecasting approach. Given that a large literature exists that forecasts employment, many different methodologies can be used.[5] These methodologies tend to be sound, as they have been vetted through years of applied research. Official statistics (such as those available from the Labour Force Survey and census) offer long histories of employment by occupations, which can make predictions more robust. As well, this exercise is useful even when not used to predict in-demand skills: forecasting employment levels by occupation was and remains an important source of labour market information independent of skills-related insights.

Adding the layer of skills information to existing forecasts provides insights that can allow workers to bridge gaps between different occupations more efficiently. For example, imagine two similarly skilled occupations, one growing and the other contracting: workers could plan to transition out of the shrinking occupation and into the growing one given that the retraining needed would be smaller than in a situation where they were moving to an occupation that required vastly different skills. In short, this type of forecasting is beneficial in determining both the skills that will be demanded in the future and the actual occupations that will grow.

Nonetheless, this forecasting approach has several significant shortcomings. Most importantly, forecasting employment and applying occupation skills without further adjustments assumes that the relationship between an occupation and the skills it requires is *constant*.[6] Since we are applying the occupation characteristics of a job in 2020 to the projected growth of that occupation, we are assuming that occupations will hold the same characteristics in the future. But this is unrealistic, because new technologies can change the com-

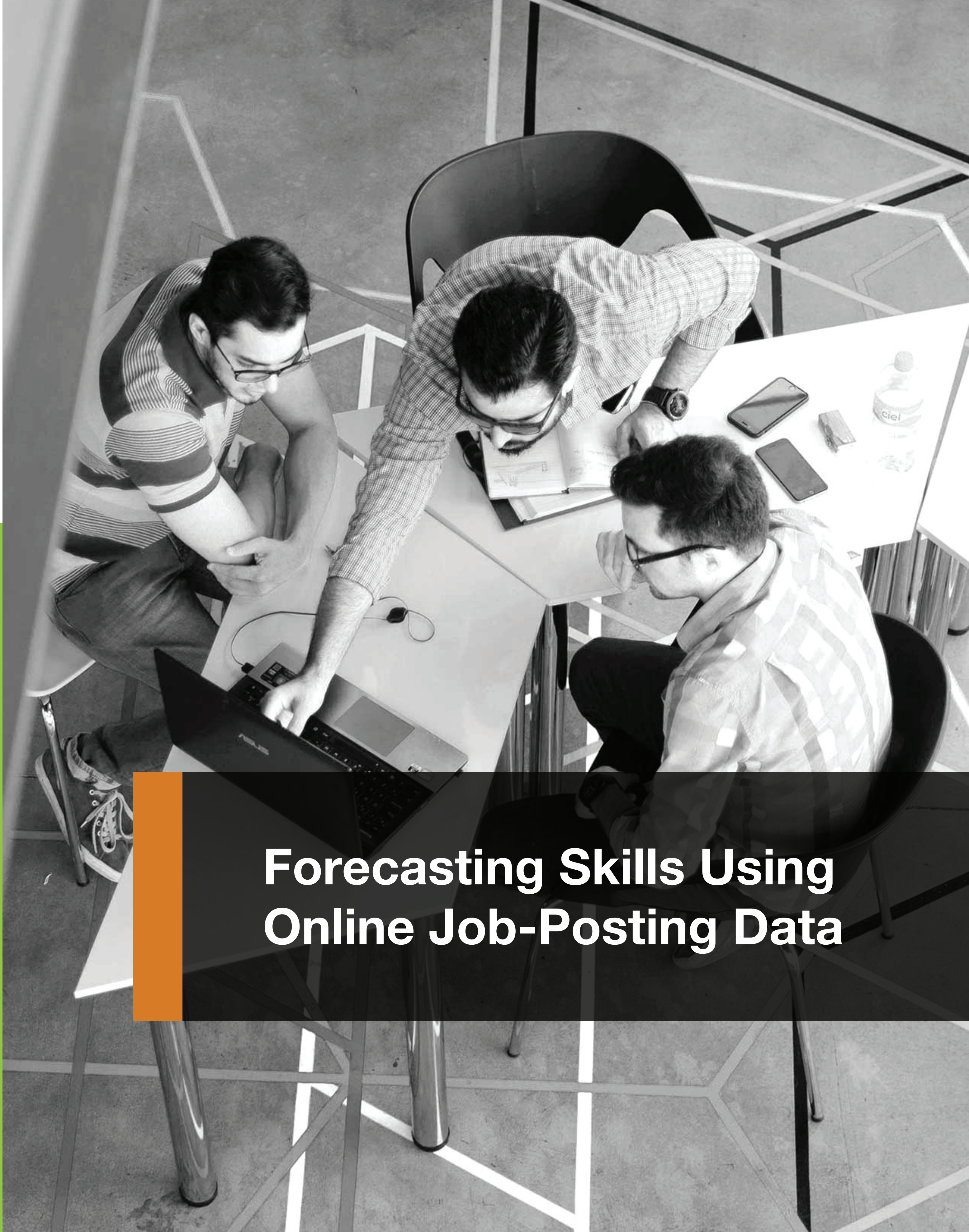
4 In theory, one could develop a model in which the skills composition within occupations changes over time, either structurally (through a separate model or simulation) or stochastically. To our knowledge, no models have been published that introduce this type of variability.

5 Lamb (2016) and (2017), the Advisory Council on Economic Growth (2017), and Brown et al. (2018) all do this as examples.

6 As mentioned in footnote 4, it is theoretically possible to allow the skill composition of occupations to vary. This is an important and valuable area for future research.

position of skills required in an occupation. While there can be many potential changes, the most common disruptive technology studied involves automation. Automation has already drastically changed the skills required by production workers, who once required mainly motor skills but increasingly require cognitive skills to oversee automated machines (Cutean et al., 2017). For another example, the skill set of an administrative assistant in the 1980s would have been focused on answering and directing phone calls. While today's administrative assistant may still do these tasks, they must now also use word processing or spreadsheet software to prepare memos, budgets and other documents. Over time, occupations' skill compositions have changed to include technological innovations.

Generally, the standard forecasting techniques discussed are reasonable predictors of future trends if the economy remains steady, but they weaken in the face of a sudden structural break, such as a deep recession (much like the COVID-19 crisis) or the emergence of new types of technology. By assuming that changes in the future will be smooth, these forecasts may underestimate the impact of emerging trends, such as automation.



Forecasting Skills Using Online Job-Posting Data

There are very few data sources that observe skills independently from occupational categories, but online job postings are one of them. In these postings, employers explicitly state the skills and other work requirements for positions they are looking to fill (see, for example, LMIC's Canadian Online Job Posting Dashboard). These data enable direct skills observations at the level of specific postings. As such, they afford a great deal of flexibility in applying modelling and forecasting techniques. However, because online job postings are an emerging data source, there is much work to be done in leveraging them for the purposes of forecasting.[7]

As mentioned in Section I, extracting skills information from online job postings is achieved through the use of natural language processing algorithms that categorize written text into a predefined taxonomy of work requirements (of which skills are one type). In this way, skills are directly observed at the level of each job posting and can be measured across or within occupations, industries, regions or any other feature linked to the postings. Importantly, because there is a large volume of information available (given that there are more than two million job postings per year in Canada), the skills data can be highly localized. In addition, with new job ads being posted every day, the data are available in near-real time.

Indeed, it is the granularity of online job postings that provide a degree of flexibility not available in occupation-derived skills information. The frequency with which any skill or other work requirements appear in online job postings can be organized by any other feature included in the data set, including, for instance, location, industry and (of course) occupation. This type of granularity is a huge benefit in providing analysis tailored to specific research questions. Similarly, the time dimension can be monthly, annual or even daily.

More generally, the freedom to explore the pattern of work requirements demanded across multiple dimensions is a key benefit of the data that can and should be exploited across an array of forecasting methodologies. Traditional econometric forecasts can be applied directly to observed skills (see Box 2). This means one could, for example, project whether the demand for "Data Analysis" in customer service occupations will grow. Alternatively, machine learning techniques could be applied to predict future skills compositions. Here, one could produce (for example) scenario-based forecasts in the form of "If demand for Oral Communication were to grow by 10%, then we expect demand for Written Communication to grow by 5%." However, in spite of these promising uses, online job posting data have yet to be leveraged for robust skills forecasting. This is largely due to three key data limitations.

7 The UK innovation think-tank Nesta has some had some success in this area by mapping job posting data to occupations with the purpose of forecasting (Djumalieva and Sleeman, 2018).

First, skills can only be identified by what is explicitly stated in postings. For example, a job posting for an accountant may not mention the need for numeracy skills because applicants are widely expected to understand that these are required. As such, these data may under-represent certain skills that the economy requires. We must also assume that employers themselves understand the skill requirements of the jobs they post online.

Second, data quality remains an issue with online job posting information. Recent research (LMIC, 2020) has shown that postings tend to over-represent certain occupations (e.g., white-collar, higher-paid service sector jobs). Any forecast based on these data must account for the fact that such biases or risks may misrepresent current and future skills demand.

Third, most data providers use their own skill taxonomies to classify the work requirements found in online job postings. The result is that it is difficult to draw comparisons across job-posting data because each database may characterize the same job posting differently. Similarly, the specific algorithms for scanning, cleaning and categorizing the data are proprietary. As such, they represent a “black box” vis-à-vis publicly available information about where skills information comes from.

The first limitation is a fundamental feature of the data that is unlikely to be overcome. However, we believe the second and third limitations can be treated through analytical and data cleaning techniques. Proprietary taxonomies can be mapped to public taxonomies or otherwise made open and accessible. The issue of data skewness can, in theory, be accounted for by estimating the magnitude of bias, adjusting aggregated skills information accordingly and using additional data sources. Determining the best approach to do so will require further research, but in theory this would involve using techniques similar to applying weights to survey data to correct for sampling bias. Adjusting for known biases in econometric models has been done in other areas, such as the corrections made for Nickell bias in dynamic panels (Bun & Carree, 2005). Such corrections will not resolve every limitation, but given the potential benefits afforded by the breadth and flexibility of online job-posting data, these methods merit further research.

Modelling Skills Through Expert Opinion

The modelling of expert opinion does not use—and thereby avoids the key limitations of—traditional forecasting approaches. Instead, the experts themselves do the forecasting and a predictive algorithm identifies similar occupations or skills.

Whereas standard forecasting uses a series of historic observations to roll predictions forward through time, this unique approach maps a cross-section of predictions about the future (i.e. experts' opinions) to similar observations that experts do not explicitly evaluate. Typically, this is done at the level of occupations, such that experts are asked to gauge the likelihood that an occupation will grow or contract in the future (or be lost due to automation) for a small set of occupations. By linking these occupations to skills (e.g. via O*NET), we can apply a machine learning algorithm to predict the likelihood that other occupations will grow or contract based on the similarity or dissimilarity of their skills.

The first major application of this approach was by Frey and Osbourne (2013), who use experts' opinions about which occupations (from a selection of 70) are likely to be automated in the future. The "eyeballing" of these occupations is done using the O*NET task and job description data. After this process is complete, they apply a machine learning algorithm to generalize the results across the entirety of occupations. This is done with the end goal of determining which occupations are least likely to be performed by computers in the future. Frey and Osbourne find that occupations that involve creativity and social intelligence are least likely to be automated.

Subsequent research sought to extend and improve on the Frey and Osborne framework by considering differences in the tasks performed within an occupation. Notably, Arntz et al. (2017) allow for tasks to differ for the same occupation across workplaces and countries by using PIACC skill proficiency scores. Similarly, using the Canadian portion of the PIAAC, Frenette and Frank (2020) apply the same methodology and find that nearly 11% of Canadian workers are at high risk of being automated out of their jobs. These studies, as with Frey and Osborne (2013), focus on whether or not an occupation is likely to be automated. If the model predicts automation for the occupation or for tasks within occupations, then employment will fall, and demand for the skills associated with those occupations or tasks will drop. In a Canadian study in a similar vein, the Brookfield Institute's Employment 2030 (2020) uses expert opinion on the growth and decline of both nationally and regionally important occupations to identify which skills will be important in the future. Notably, this approach seeks expert opinion on trends in employment growth or decline for any reason rather than limiting the question to the occupation's risk of automation.

In theory, a key advantage of the modelling of expert opinion is that it can account for structural changes that are notoriously difficult to model. Whether or not this approach actually generates more accurate predictions in general, or of structural changes in particular, remains an open question. A key limitation, of course, is that the accuracy of these forecasts depends enormously on the quality of experts' opinion. For example, in 2019, experts did not foresee a global pandemic in 2020, but the COVID-19 pandemic led to the greatest disruption of labour markets since World War II.

Finally, as noted above, this is the only approach that is agnostic about the data input. The only requirement is expert opinion. How those qualitative data points are mapped to other skills or occupations is an open methodological question. Yet this mapping is a complex process that leads most researchers to rely on well-known skills sources, such as O*NET. Therefore, in practice, the modelling of expert opinions draws on the same skills data used to forecast employment by occupation, with the same benefits and limitations of occupation-derived data (namely, that the skill composition of each occupation is fixed over time). That said, future research could implement this still-evolving methodology by drawing on job posting–derived data.

Conclusion: The Way Forward

Reliably forecasting skills is an important and challenging endeavour that needs to be further explored and implemented. Usable forecasts of skills in demand can help Canadians make more informed decisions about training and education today. This is more important than ever in the context of the COVID-19 crisis.

To that end, we have explored three approaches to forecasting skills. First, we considered skills linked to occupations. This approach relies on occupation-derived skills data and forecasts employment by occupation, which can then be mapped back to skills. The key limitation in this approach is that the skills composition of every occupation is fixed through time. Second, we explained how job skills are forecast using online job posting data, an approach that has only recently become somewhat feasible. While this emerging data source offers many new possibilities, there are ongoing concerns related to data quality and the reliability of third-party, proprietary data. Finally, the modelling of expert predictions was evaluated as a data source–agnostic approach that could be extended to online job posting data or other data sources. Here, the key limitations are the accuracy of the experts' predictions and the ability to map these predictions to specific skills in demand.

Each approach discussed has benefits and drawbacks. Importantly, the validity of all approaches depends enormously on the specific assumptions used in applied models, discussion of which is beyond the scope of this summary paper. That said, the usefulness of any approach and/or specific model implementation depends entirely on two factors: whether the skills forecast is accurate, and whether it helps people (e.g., policy-makers, educators, job seekers) make more informed decisions. The former requires us to assess the out-of-sample forecast accuracy of any proposed model versus its alternatives. The latter requires us to carefully identify, define and contextualize skills in a way that is relevant for real-world decision-making. As new methods and data sources are developed and tested, forecasts of in-demand skills should be evaluated continually against the fundamental criteria.

Acknowledgements

We thank Félix Simoneau (l'Université du Québec à Montréal), Naomi Pope (British Columbia Ministry of Advanced Education, Skill and Training; Labour Market & Immigration Analysis), Kelly Gallagher-MacKay, (Future Skills Centre), Michael Burt (Conference Board of Canada), Steven Tobin (LMIC) and Diana Rivera (Brookfield Institute for Innovation + Entrepreneurship) for their feedback.

References

Advisory Council on Economic Growth. (2017). *Learning nation: Equipping Canada's workforce with skills for the future*. Government of Canada.

Arntz, M., Gregory, T., & Zierahn, U. (2017). Revisiting the risk of automation. *Economic Letters*, 159 (July), 157–160.

Borwein, S., & Refling, E. (2014). *The Great Skills Divide*. Higher Education Quality Council of Canada.

Brandes, P., & Wattenhofer, R. (2016). *Opening the Frey/Osborne black box: Which tasks of a job are susceptible to computerisation?*. ETH Zurich. Zurich, Switzerland.

Brown, J., Gosling, T., Sethi, B., Sheppard, B., Stubbings, C., Sviokla, J.,Williams, J., & Zarubina, D. (2018). *Workforce of the future: The competing forces shaping 2030*. PricewaterhouseCoopers.

Bun, M., & Carree, M. (2005). Bias-corrected estimation in dynamic panel data models. *Journal of Business & Economic Statistics, 23*(2), 200–210.

Calero, J., Murillo-Huertes, I., & Raymond-Bara, J. (2016). *Education, age, and skills: An analysis using the PIAAC survey*. Barcelona Institute of Economics.

Cörvers, F., & Heijke H. (2004). *Forecasting the labour market by occupation and education: Some key issues*. University of Maastricht.

Cutean, A. (2017). *Autonomous vehicles and the future of work in Canada*. Information and Communications Technology Council.

Djumalieva, J., & Sleeman, C. (2018). The first publicly available data-driven skills taxonomy for the UK. Economic Statistics Centre of Excellence.

Frenette, M., & Frank, K. (2020). *Automation and job risk in Canada: Who's at risk?* (Analytical Studies Branch Research Paper Series, no. 448).

Frey, C., & Osborne, M. (2013). The future of employment: How susceptible are jobs to computerisation? University of Oxford.

Ingramm, B., & Neumann, G. (2006). The returns to skill. *Labour Economics,* 13, 235–59.

Lamb, C. (2016). *The talented Mr. Robot: The impact of automation on Canada's workforce*. Brookfield Institute for Innovation + Entrepreneurship.

Lamb, C. and Doyle, S. 2017. *Future-proof: Preparing young Canadians for the future of work*. Brookfield Institute for Innovation + Entrepreneurship.

Loree, J., & Stacey, D. (2019). *Occupational skill portfolios and wages*. (Working paper).

Loree, J. (2020). *Multidimensional skill accumulation and mismatch over the lifecycle*. (Working paper).

Nedelkoska, L., & Quintini, G. (2018). *Automation, skills use, and training*. (OECD Social, Employment, and Migration Working Papers No. 202).

Richardson, S., & Tan, Y. (2007). *Forecasting future demands: What we can and cannot know*. National Centre for Vocational Education Research.

Roy, A. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2), 135–146.

The Expert Panel on STEM Skills for the Future. (2015). *Some assembly required: STEM skills and Canada's economic productivity*. Council of Canadian Academies.

Urban, M., & Johal, S. (2020). *Understanding the future of skills: Trend and global policy responses*. Skillsnext.

Future Skills Centre

LMiC LABOUR MARKET INFORMATION COUNCIL