TAYLOR FRY

Data-led approach to identifying skills shortages

New Zealand Productivity Commission

11 April 2022

Document classification: Client use

ACN 087 047 809
ABN 29 087 047 809
www.taylorfry.com.au

Taylor Fry Pty Ltd

11 April 2022

Ben Temple
Principal Advisor
New Zealand Productivity Commission
Ben.Temple@productivity.govt.nz

Dear Ben,

New Zealand skills shortage model feasibility

We are happy to provide this report scoping the feasibility of a data-based model of skill and occupational shortages in New Zealand.

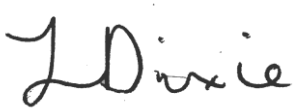
The report explores the broader literature available on the subject before considering the data already available in New Zealand to build a model.

We also discuss our views on the extent to which a model could be dis-aggregated by occupational level and/or geographical unit, before discussing some practical elements of model development and design.

Finally, we draw conclusions on the feasibility of building a model in New Zealand.

We look forward to discussing this report with you further.

Yours sincerely,


Hugh Miller
Principal

Laura Dixie
Director

Dan Stoner
Director

Table of contents

1 Executive summary

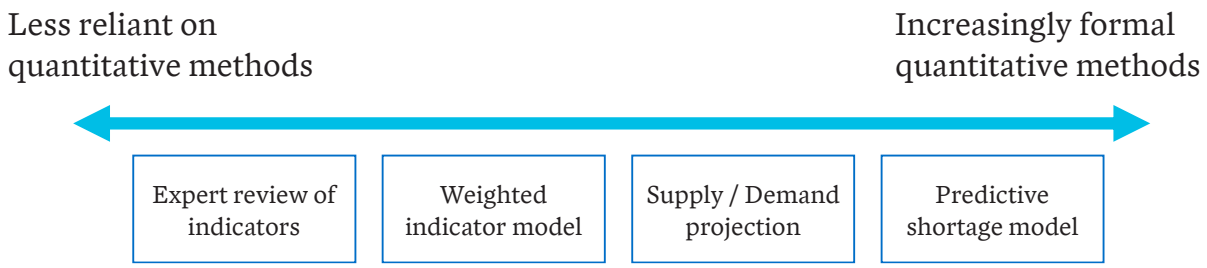
Taylor Fry has been asked to explore the feasibility of an occupational skills shortage model, based on objective data. If feasible, this would be an important input into both migration and education policy. We have approached this topic by:

- Performing a literature review, with a focus on understanding international practice
- Performing a data survey of New Zealand data assets that could support a skills shortage model
- Assessing feasibility and related considerations.

Literature review

Countries take a range of approaches to estimating skills shortages. While some are less quantitative, many are formal, bringing in a wide range of data sources to come to an estimate of shortages. We have categorised different countries according to this range, depicted below.

Figure 1 – Common approaches for developing a skills shortage list



We have found that:

- Many countries have decided to formally create a skills list based on quantitative indicators. However, there is significant variation in approaches taken. A concrete list is judged a valuable tool, particularly for migration applications. Many sources also point to the direct implications for education and training policy too. The Australian list is explicitly designed to serve both policy areas.
- While there is some variation in 'skills' and 'occupations' across sources, most forecasts and applications are done using occupational groupings.
- There is some consistency across the types of labour market indicators that are useful for identifying shortages. Unemployment, employment growth, wage growth and vacancy information are generally included. Beyond this, there are more differences in the exact form of the indicators. Indicators can be driven by other factors beyond pure labour shortages, so overreliance on single factors is mostly avoided.
- Most processes, even if they primarily rely on quantitative measures, leave some room for expert judgement and stakeholder consultation. This is particularly relevant if rapid changes are occurring, since labour market indicators can lag behind the changes. Often there is a dedicated independent body to perform this function.
- Limitations identified for quantitative approaches to estimate skills shortages are reasonably consistent across different countries:
  - There are inherent limitations attached to defining shortages itself. For example, the data is often imprecise at the occupation level as occupation groups are not necessarily homogeneous.

- Many indicators rely on survey data, which make conclusions for small regions and smaller occupations more uncertain.
- There are limits to the level of accuracy that can be expected from a skill shortage model. Having some form of validation or uncertainty estimate is useful for understanding the extent of this.

In addition to noting the limitations of skills shortage models the review also discusses alternative approaches such as setting migration policy purely based on income thresholds.

Data survey

We split our discussion into primary measures of a skill shortage and other secondary indicators that are useful for predicting shortages.

A primary measure directly attempts to answer the question of a shortage. Based on the literature review, we believe the preferred primary measure should be whether **employers are able to fill specific occupational vacancies.** While there is nuance to this (there are various reasons why vacancies might be hard to fill), it aligns well to the potential policy levers (migration and training) that might be used to remedy a shortage.

The two most common ways to directly measure employers' ability to fill specific skills vacancies are:

- **Vacancy-based surveys –** Surveying people recently posting vacancies and asking if they were able to fill the position. This is the approach taken for the Australian SERA survey, where the sample frame is based on classifying job advertisements and then calling up the contact to see if the position was able to be filled.
- **Employer based surveys –** Asking employers if they advertised for specific occupations and then whether they were able to fill the position.

New Zealand does not have a suitable survey that could be used as a primary measure of shortage, which makes it difficult to validate and refine any proposed skills model. This is consistent with many other countries. Perhaps the most practical way to create a measure is to add occupational detail to the New Zealand Business Operations Survey, which is an annual survey with significant scale and already has some questions related to the ease of recruitment.

New Zealand has a very good range of data sources that cover the secondary indicators most commonly used overseas (as found in our literature review). This is summarised in the table below. Most of these have proven validity to provide insight into skills shortage, even if they are somewhat indirect. Across the literature, a selection of these which are available (rather than all) are used in each context.

Table 1 – Summary of potential labour market indicators, all ideally at an occupational level

Indicator	Suitable data?	Source
Number employed	✓	Household Labour Force Survey (HLFS) Census
Employment growth	✓	HLFS
Working hours and change	✓	HLFS
Overtime hours	Partly	HLFS, partly
Education level of those employed	✓	HLFS
Study diversity	✓	Census
New hires fraction	✗	Not directly available, but proxies in the HLFS
Unemployment rate	✓	HLFS
Unemployment duration	✓	HLFS

Indicator	Suitable data?	Source
Wage premium (how much an occupation is paid relative to expected, given training and skills)	×	HLFS Labour Cost Index (LCI)
Wage growth	✓	HLFS income supplement Potentially HLFS + linkage to income tax data LCI
Annual change in vacancy numbers	✓	MBIE Jobs Online Index
Vacancy fill rate	×	
Vacancy required education, or other job ad characteristics	Likely	Job boards data (e.g. Burning Glass)
Migrant numbers	✓	Immigration NZ visa data HLFS
Fraction of foreign workers	Maybe	Immigration NZ visa data HLFS

Our review of the potential data sources suggest that they are high-quality and available in significant detail (e.g. at a unit record level through the Integrated Data Infrastructure).

The table above is not intended to be exhaustive, and there is the potential for other data sources to be added (and formally tested if a primary measure is developed).

The size of datasets, particularly the sample size of the Household Labour Force Survey (HLFS), is enough to give reasonably good, but not total, resolution of occupations nationally and grouped occupations regionally.

Feasibility

Shortages in specific skills and occupations have the potential to cause disruption and hinder economic growth. The Government has the potential to smooth shortages through the use of skilled migration policy, apprenticeship incentives and training funding.

While a skills model will have limitations, we believe that there is significant value to be gained. A skills model would enable:

- Better coordination between migration and training policies
- Better management of lower-wage but essential occupations, compared to policies based on income thresholds
- Detailed labour market monitoring.

We regard these benefits as substantial, potentially reducing some of the friction in labour markets.

Overall, we think that **a weighted indicator model of skills shortage approach is feasible and is likely to add insight.** In coming to this conclusion, we note:

- New Zealand lacks a good primary indicator of skills shortage to build a formal predictive model, but this could be developed
- There is a very good range of secondary indicators to build a weighted indicator model
- International sources are a good place for starting to build a weighted indicator model
- Model accuracy is likely to be reasonable

In terms of model granularity, the resolution that can be delivered will depend on a combination of data-specific issues and a user's tolerance for uncertainty. However, we believe it is feasible to develop a

weighted indicator model for occupations based on Australian and New Zealand Standard Classification of Occupations (ANZSCO) classifications that disaggregates with reasonable credibility to:

- ANZSCO-6 (full occupation list, 1,000 categories) at a national level
- ANZSCO-4 (unit groups, 360 categories) with a geographical breakdown for key regions. Some smaller regions may need to be aggregated.

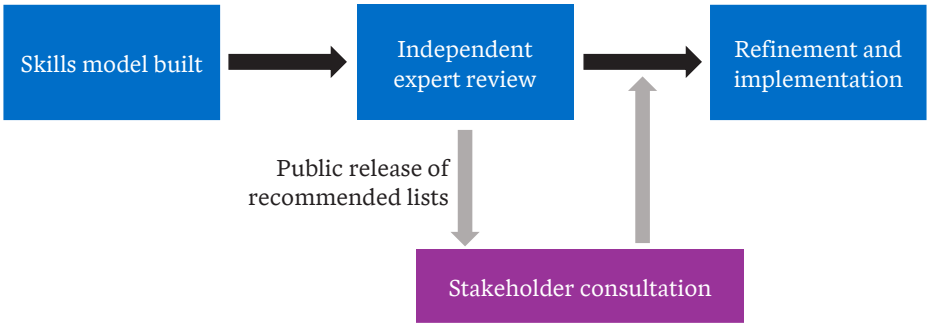
This is detailed enough to be useful for both national and regional planning.

Considerations

There are a range of conceptual and practical issues that should be undertaken when building a skills model. For example, should a model be predicting today's shortages or whether they will exist in the future? Or what occupational or skills taxonomy should be used to construct a list? How accurate is the model likely to be? We work through such questions in Section 5.3 of the report.

One particularly important consideration is the overall structure of how a skills model is translated into policy, which includes the institutional framework. While we do not recommend a specific structure, we recognise the value of transparency and some form of independent review process to reduce the risk of regulatory capture in any stakeholder engagement and refinement process. A conceptual structure is shown below.

Figure 2 – Potential process for moving from a skills model to implementation



Specific issues, such as whether a new separate independent body is required, or what governance is required around the stakeholder consultation and refinement stage, would have to be carefully worked through.

Potential next steps

If the decision was made to build a weighted indicator model, the high-level next steps would be around:

- Establishing the specific purpose of the model, what it is intended to be used for and who 'owns' the model and its maintenance.
- Detailed exploration and testing of data sources.
- Establishing the basis for determining which indicators to include and what weightings to apply, including consultation with subject matter experts.
- Developing the model.
- Validating the model.
- Implementing the model, including documentation and potential publication of information.

These steps relate specifically to building a model. Broader process and consultation will no doubt be required to address some of the conceptual and practical issues identified in Section 5.3.

2 Introduction

2.1 Background

The Productivity Commission ('you') has asked Taylor Fry ('us') to provide advice on whether a data-led approach to identifying skills shortages in New Zealand is feasible and if so, how this could operate as a means of informing migration and education policy. In particular, you are interested in:

- The value proposition of any approach, including limitations
- What other comparable countries do to identify skills shortages
- How existing New Zealand data sources could be used for such an approach and/or adjusted in the future, and
- Practical implementation considerations.

Currently, skills shortages are not explicitly measured and reported on in New Zealand. But occupational shortage lists are maintained for immigration purposes. The current three skills shortage lists are:

- The long-term skill shortage list
- The regional skill shortage list
- The construction and infrastructure shortage list.

These lists are maintained by the Ministry of Business, Innovation and Employment (MBIE) and based on qualitative judgements e.g. using information provided by affected industries, employers, unions and other interested parties. While used for immigration policy, they do not directly influence education and training policy.

Relatedly, MBIE use a high paid/low paid distinction in policy setting, based on median wage linked thresholds. Specifically, labour market test requirements for some visas, including the Essential Skills Work Visa and the new Accredited Employer Work Visa, depend on how wage relates to median wage linked thresholds (among other factors).

In theory, a data-led approach to estimating shortages could supplement or replace elements of the current approach. It can provide new evidence of occupations that are in shortage that were not previously identified, and also provide evidence when a particular shortage has subsided. The primary purpose of our advice is to identify what is feasible.

2.2 How we have structured this report

We have structured the rest of this report following a bottom-up approach. We start with a literature review of the construction of skill and occupation lists, before considering potential predictive indicators of skills shortages in New Zealand. We then draw conclusions about what is feasible in New Zealand and, given this, the value proposition of a data-led approach. More specifically, the following sections cover:

- **Section 3 – Literature review** – a rapid search of relevant literature on the construction of skill and occupation shortage lists, detailing approaches in other countries, thematically linked by style of approach
- **Section 4 – Data** – this section explores:
  - The existence (and options) for a primary measure of skills shortage. This can be used to measure directly or guide the construction of a predictive skills shortage model.
  - The range of other labour market indicators that could contribute to a skills shortage model.
- **Section 5 – Value proposition and feasibility of a data-led approach in NZ** – this section covers:
  - The value proposition of a data-led approach given

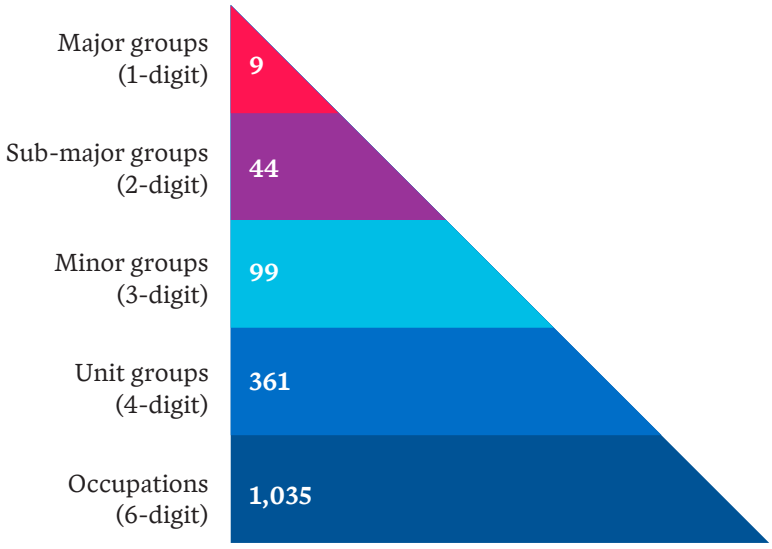
  - The feasibility of a New Zealand skills shortage model
  - Conceptual and practical issues to consider for model development
  - Potential next steps for building a model.
- **Section 6 – References** – a list of literature referenced in this report
- **Appendix A – Detailed indicator lists from literature review** – specifically from examples of models used in the United Kingdom, Malaysia, Ireland and Australia.

2.3 Terminology

Some key terms and ideas used throughout the report are:

- **Skills versus occupations –** There is an obvious conceptual different between an occupational shortage (e.g. cannot find a computer programmer to hire) and a skills shortage (e.g. cannot find a programmer who can code in Python). However, skills tend to be tied to occupations so the distinction can be blurry. We discuss the distinction further in Section 5.3.2, but we have generally used the terms without distinction throughout our report.
- **Occupation lists and hierarchies –** We have generally used the Australian and New Zealand Standard Classification of Occupations (ANZSCO) when talking about occupational groups. ANZSCO is a skill-based classification used to classify all occupations and jobs in the Australiana and New Zealand labour markets. Occupations are organised into progressively larger groups on the basis of their similarities in terms of both skill level and skill specialisation. ANZSCO codes are arranged in a 5-layer hierarchy (e.g. Technicians and Trades Workers → Automotive and Engineering Trades Workers → Automotive Electricians and Mechanics → Motor Mechanics → Diesel Motor Mechanic), as described in Figure 3. All hierarchies, including ANZSCO, have some limitations (see Section 5.3.5). Current data collection and encoding practices mean that ANZSCO is the most practical choice for a New Zealand model.

Figure 3 – ANZSCO levels and number of categories within each



2.4 Status of this report

This is the finalised version of the report.

3 Literature review

3.1 Approach to literature review

We have performed a rapid search of relevant literature on the construction of skill and occupation lists. This has been a combination of:

- Relevant literature already known to the authors
- Other literature identified by the Productivity Commission and other discussants
- Search of academic and grey literature through web search (search engine and google scholar)
- Some snowballing of relevant sources referred to in other work.

Our focus has been identifying actual approaches adopted by countries, and the methodology underpinning this.

We start by reviewing New Zealand sources. Then we explore other countries' approaches, thematically linked by style of approach. Alternatives and criticism of skills lists are then discussed, and institutional arrangements considered, followed by a summary. A full reference list is provided in Section 6, and useful weblinks are also included as footnotes.

3.2 Relevant New Zealand literature

As early as 2006, the (then) Department of Labour was considering what indicators of skills shortage might look like and how they could be used to inform immigration policy (Department of Labour, 2006)[1]. Their report, *Indicators of Skill Shortage*, identified a series of principles for building a shortage list:

- Identifiers of a skills shortage – This included vacancy fill rates (using a now discontinued survey), vacancy volumes and wage growth.
- Judgment on whether immigration policy is an appropriate lever for an occupation – This included a view on the degree of labour market competition, whether demand is sustained and the potential of education to fill a shortage.
- The future risk of shortage, with regard to the degree of specialisation.

Mason et al. (2012), *A Good Worker is Hard to Find*[2] is a technical report published by the Ministry of Economic Development. It describes a firm level analysis (rather than occupation-level) of what can predict a firm struggling to fill a vacancy.

They relied on the Business Operations Survey (BOS), including a specialist module developed for the study, and Inland Revenue (IR) data. They were able to build a model predicting shortages based on variables from the:

- BOS – market focus, training provided to staff, ownership of overseas businesses, foreign ownership, competition, investment in expansion, Research & Development, occupational break down and innovation undertaken.
- IR data – employees/size, self-employment, employee churn and employee growth, wages (including growth and premium relative to the industry), business activity.

---

[1] Indicators of Skill Shortage, 2006. Infometrics Ltd. Available from: https://thehub.swa.govt.nz/assets/documents/Indicators%20of%20skill%20shortage.pdf

[2] A Good Worker is Hard to Find, 2002, Mason et al. Available from: https://www.researchgate.net/publication/254448872_A_Good_Worker_is_Hard_to_Find_Skills_Shortages_in_New_Zealand_Firms

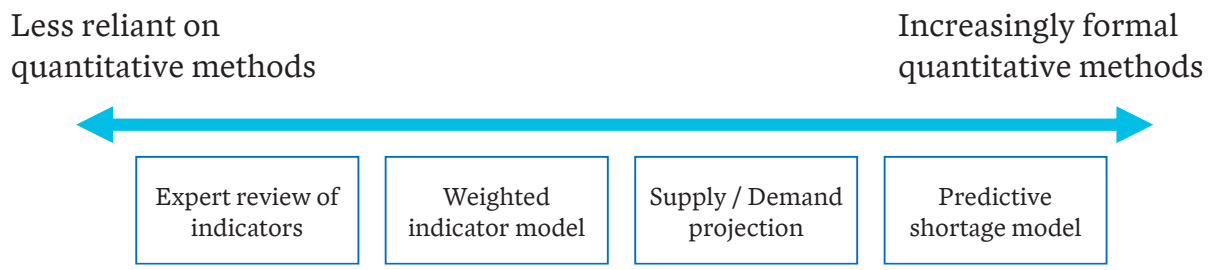
The paper includes modelling all vacancies (which are not all hard to fill) and also distinguishes between skills-related shortages (vacancies that are hard to fill because applicants lack specific training or experience) and other types of shortages (vacancies that are hard to fill for other reasons, such as too-low wages or a tight labour market). It found good evidence that the indicators used were predictive of shortages.

Ball et al. (2020) of the Reserve Bank recently published *Using Job Transitions Data As A Labour Market Indicator*[3]. While not directly related to skills shortages, it is an interesting demonstration of the ability of large administrative datasets (here from IR) and modern modelling approaches to provide models which match or outperform conventional econometric models. The authors used job transitions observable in IR data to provide a nowcast of the unemployment rate.

3.3 Broad categories of approach

There is a spectrum of approaches used to create skills shortage lists, which we have classified by their relative reliance on quantitative methods. These are summarised in Figure 4.

Figure 4 – Common approaches for developing a skills shortage list



Briefly, the types of classifications are:

- **Expert review of indicators –** This recognises that labour markets are dynamic and multi-faceted, so aims to provide an expert review of a range of indicators that can be qualitatively combined to determine absolute or relative shortages.
- **Weighted indicator model –** This attempts to collate the range of relevant factors into a single list by giving each indicator some weight (often equal) and combining.
- **Supply and Demand projection –** This approach models and forecasts the flows of people into and out of an occupation. Entries incorporate people being trained, migration and other people transferring into an occupation. Exits are due to retirements from the workforce or transfers out of an occupation. Combining components allows imbalances to be identified.
- **Predictive shortage model –** This approach is similar to a weighted indicator score approach, but there is an explicit target variable ('primary measure for skills shortage') that signifies shortage, with indicators calibrated into a model that predicts the target variable. Effectively, the target variable helps determine the 'best' choice of indicators and weights.

There are also potential approaches to managing skilled migration or training without reference to a specific skills or occupation list. Section 3.6 sets out some of the arguments against such lists and potential alternative approaches.

The approaches above are also rarely stand-alone; in practice, any quantitative approach can still be layered with subject matter expertise and other contextual factors.

[3] Using Job Transitions Data As A Labour Market Indicator, Ball et al. 2020. Available from: https://www.rbnz.govt.nz/-/media/ReserveBank/Files/Publications/Analytical%20notes/2020/AN2020-02.pdf?revision=766dd357-89e6-48c5-a787-29d82d500870

3.3.1 Expert review of indicators

With an expert review of indicators approach, quantitative measures are eschewed in favour of subject matter expertise.

Ireland

Ireland's migration policy relies on their Critical Skills and Ineligible Occupation Lists. These are reviewed twice a year and feed into reduced salary requirements for work permits. Defining the lists involves an extensive consultation process with an online submission form.

Each year the Skills and Labour Market Research Unit (SLMRU) publishes a National Skills Bulletin (McNaboe et al, 2021), which is a key source used by the Department of Jobs, Enterprise and Innovation to identify occupations to be included on either list. The National Skills Bulletin examines a range of labour market indicators across 95 occupational groups. Indicators include employment numbers, employment growth, other employment profiles (female employment rates, full-time rates, proportion aged over 55, tertiary qualification rates, proportion non-Irish nationals), difficulty in filling vacancies (based on an employer survey), training supply and job vacancy numbers.

While there are a range of indicators considered in the National Skills Bulletin, we are not aware of a formal mechanism for combining them. We assume that skill lists are derived through consideration of these, combined with the consultation process, ultimately using subjective judgement to combine information.

Other research

Mavromaras et al. (2013)[4] propose a suite of labour market indicators that can be used to identify labour market disequilibrium. These are divided into aggregate labour market indicators (e.g. employment rates), recruitment indicators (e.g. vacancy fill rate), student indicators (e.g. number of course completions) and labour market entry indicators (e.g. graduate employment outcomes). The paper itself also has a literature review of related work.

3.3.2 Weighted indicator model

The principle of a weighted indicator model is that there is no single source of truth.

The main drawback is that there is some arbitrariness to the way measures are combined; it may be that certain factors are a better gauge of labour market conditions, but there is limited ability to determine this and choose appropriate weights.

United Kingdom

In the United Kingdom (UK), the Migration Advisory Committee (MAC) advises UK government on migration issues including input into the UK Shortage Occupation List (Migration Advisory Committee, 2019, 2020). Filling occupations on the shortage list is an eligibility criterion for a skilled worker visa. MAC require three criteria for inclusion of an occupation to all be met:

- It is **skilled** at a level required by the Government. Skills are typically defined in terms of the length of training required, mapped to their occupation list.
- There is a **shortage** as indicated by their indicators.
- It is **sensible** to fill the shortage with migrant workers (qualitative assessment).

[4] Accessible at https://webarchive.libraries.tas.gov.au/20140617045312/http://www.skills.tas.gov.au/employersindustry/industryresources/workforce-planning-and-development-resources/A_system_for_monitoring_skills_imbalance_and_surpluses_in_the_market_for_skills.pdf

Nine indicators are used for a set of 369 occupations. Indicators span wages (change over time, occupations wage premium), vacancy rates, employment growth, hours worked, new hire rates and unemployment rates. These indicators are selected after being judged both relevant to skills shortage and of sufficient coverage and quality. Each indicator is converted to a ranking by occupation and then average rank used as an overall indicator.

Some noted limitations to the methodology are:

- Recognition that rankings should be considered alongside the key indicators they aggregate over for vacancies, wages and employment
- Constraints on the granularity of data (by occupation code)
- Timeliness of data
- Volatility attached to small sample sizes
- Indicators being affected by conditions other than skills shortages.

Malaysia

Since 2015, the Malaysian Critical Skills Monitoring Committee has developed a Critical Occupations List (COL) on an annual basis. The COL defines critical occupations according to three main criteria:[5]

- **Skilled –** primarily focused on growing the knowledge-based economy, assessment is based on a list with skill ratings, analogous to ANZSCO
- **Sought-after** – demand outstrips supply, assessment of qualitative indicators (top-down) and stakeholder consultation (bottom-up) combined
- **Strategic** – filling the shortage must be consistent with strategic economic development objectives.

The COL is created through analysis of quantitative and qualitative evidence collected through both top-down and bottom-up approaches. These are then combined over two rounds of dovetailing.

The top-down approach uses a range of indicators and data sources, spanning labour force surveys (employment numbers, hours, education), wage data, and online job postings. Each indicator has a threshold set and scores are based on the number of indicators above a threshold by occupation. This is combined with bottom-up survey work with employers and focus groups as well as consultations with employer and industry groups.

The World Bank (Moroz et al., 2019) reviewed the approach. It found many strengths, including regular updating and improvement, rigorous evidence base, transparency and private-public collaboration. Noted challenges included regional disaggregation and the standardisation of occupational data. The World Bank Report noted future applications included informing development and funding of training programs, career advice, employment services and immigration policy.

3.3.3 Supply / Demand projection

Canada

In Canada, the Economic Policy Directorate (EPD) of Employment and Social Development Canada (ESDC) produces the Canadian Occupational Projection System (COPS).[6] This is a 10-year labour market forecast at the national level, produced every two years. The overall **labour supply and demand model**

[5] Critical occupations List 2020/2021 Technical Report, TalentCorp. Available from: https://www.talentcorp.com.my/clients/TalentCorp_2016_7A6571AE-D9D0-4175-B35D-99EC514F2D24/contentms/img/Documents/COL%202020-2021_Technical%20Report_FINAL.pdf

[6] Canadian Occupational Projection System (COPS). See: http://occupations.esdc.gc.ca/sppc-cops/l.3bd.2t.1ils@-eng.jsp

makes formal estimates of the supply and demand of people within occupations, as a way of identifying shortages and over supply. The methodology involves:

1. A qualitative review of a range of indicators
2. Projecting numbers of job openings and job seekers
3. Combining the two results into a projected outlook.

The **qualitative review of indicators** considers a range of factors, including:

- Changes in unemployment rate
- Changes in job vacancies
- Wage growth
- Employment growth
- The proportion of employment insurance recipients
- Paid workers working overtime hours.

Indicators are compared to occupational historical trends and overall indicator levels to determine imbalances. The indicators are mainly sourced from the Labour Force Survey, a monthly cross-sectional survey of 54,000 households using a rotating panel design.

Estimates are made across 293 occupation groups, with some smaller groups combined (from an initial list of 500) for tractability.

There may be new more occupation-specific models to come. Recently, the Future Skills Centre Canada and the Labour Market Information Council launched a new project focussed on helping Canadians navigate their career choices[7]. This is planned to involve piloting new methods, including leveraging online job-posting data and other approaches, to better understand future skills requirement.

Some review work has been done on the COPS projects. Foot and Meltz (1992) found an acceptable level of accuracy for major occupation groups, and further value in other aspects of the model. Ignaczak (2017) found that the model produces reliable and credible outputs at an occupational level, but also identified challenges. For example, credible scenarios for the labour market under full-employment assumptions are difficult. And it is difficult to validate assumptions for how education maps to occupational skills.

Other research

Mavromaras et al. (2013), in an Australian context, argue that accurately measuring imbalances through separate demand and supply estimates is 'fraught with identification difficulties'. They suggest that a search for labour market disequilibrium, using labour market indicators, is more efficient and feasible. See Section 3.3.1 for further information.

Thomas (2015) notes that caution should be used when analysing the information obtained from any occupational forecast and that errors can creep into occupational projections quite easily due to the complex estimation methodology.

The Australian Department responsible for training policy released a discussion paper[8] on supply and demand models as a potential option for apprenticeship planning (Department of Employment, Skills, Small and Family Business, 2019). It judged the approach as feasible for select occupations at an ANZSCO-4 level, despite some data challenges.

---

[7] https://fsc-ccf.ca/engage/future-skills-centre-and-labour-market-information-council-announce-3-million-partnership-for-creation-of-front-line-career-guidance-tools/

[8] https://www.dese.gov.au/review-australian-apprenticeships-national-skills-needs-list/resources/review-australian-apprenticeships-national-skills-needs-list-methodology-discussion-paper

3.3.4 Predictive shortage models

Australia

In Australia, the National Skills Commission (NSC) has responsibility for the Skills Priority List. This list provides:

- Both a current labour market rating and a future demand rating for occupations nationally
- Current labour market ratings for occupations at a State and Territory level.

Developed in 2021, this list provides the backbone piece of labour market analysis on occupations that will be a key input to a range of Australian Government policy initiatives. This includes targeting of skilled migration as an input into the Priority Migration Skilled Occupation List. The scope is limited to occupations classed as Skill Level 4 and above. This excludes Skill Level 5 occupations which generally do not require significant post-school education and training.

The list is constructed by augmenting the **Skills Priority List (SPL) Indicator Model**[9] with an estimate of future growth (demand) plus the results of consultations with industry groups and qualitative employer survey results.

The SPL Indicator Model is notable in that it explicitly targets the 'vacancy fill rate' as the working definition of occupational shortage. The fill rate is measured from the Survey of Employers who have Recently Advertised (SERA). This survey, as the name suggests, contacts employers and recruiters who have recently advertised for certain occupations. The survey generates quantitative information (the fraction of vacancies that were filled and statistics on the number of applications), as well as qualitative information about the hiring attempts. It does not capture wage information. This provides a direct measure on whether there is adequate labour supply to meet demand.

However, the SERA is only undertaken for a minority of occupations, so estimating shortages more broadly relies on the SPL Indicator Model, which can generalise the learnings from SERA to the full set of occupations.

The advantage of having a primary measure for an occupational shortage measure is that a larger set of labour market variables can be tested and refined, leading to a better choice of indicators and weights. A total of 190 variables (many variants of underlying information) were tested for predictiveness. On historical data the model provided good discrimination, with about a third of occupations having a predicted fill rate below 55% and a fifth with a predicted fill rate greater than 70%.

The methodology also included credibility components so that different information could be blended:

- Regional labour market variables could be layered, balancing the additional signal with the additional volatility in these indicators
- SERA results (including from previous years) could be incorporated into a combined estimate.

The credibility approach relies on explicit estimates of uncertainty from various sources so they can be sensibly combined.

One relatively unique feature built into the model is the 'study diversity' indicator. This attempts to capture the fact that occupations with many training paths tend to have more resilience to shortage, since they can 'pull' more students in from different places. The score is higher when those currently employed in the occupation come from a range of study fields, and low when there is a single pathway.

Dawson et al. (2020) adopted a similar philosophy in predicting skills shortage. Using historical SERA results as the primary target, they used a range of indicators to predict shortages more generally, finding good predictive performance. The strong autocorrelation of shortages over time in specific occupations is one notable observation in the paper.

---

[9] Methodology at https://www.nationalskillscommission.gov.au/skills-priority-list-methodology

3.4 Indicator lists

Table 2 summarises the types of indicators used across the countries covered.

Table 2 – Summary of indicators used in various skills shortage approaches

Occupational-level variable	Malaysia	United Kingdom	Ireland	Canada (qualitative)	Australia
Number employed			✓		✓
Employment growth	✓	✓	✓	✓	
Working hours growth	✓	✓			
Overtime hours				✓	
Education levels of those employed (level or change)	✓		✓		
Study diversity					✓
New hires fraction		✓			
Unemployment rate (or variant)		✓	✓	✓	✓
Unemployment duration					✓
Wage premium (modelled)	✓	✓			
Wage growth		✓		✓	✓
Vacancy numbers (level or change)	✓		✓	✓	
Vacancy rate (# vacancies ÷ employment) Level and/or change	✓	✓			✓
Vacancy fill rate			✓		✓ (target)
Vacancy required education, change	✓				✓
Migrant numbers			✓		✓
Fraction of foreign workers			✓		

3.5 Institutional arrangements

For a data-led approach to identifying skills shortages to deliver on its value proposition, it needs appropriate institutional arrangements to facilitate and empower its use. Institutional arrangements may be embedded within core government agencies, or through dedicated entities focussing on specific related topics. The degree of independence, decision-making vs advisory capacity, governance and remit are key considerations in establishing arrangements. Overseas examples of dedicated entities include:

- Australia – National Skills Commission (NSC) – The NSC was established in 2020 with the broad purpose of providing expert advice and national leadership on Australia's labour market and current, emerging and future workforce skills needs. The creation was one of the key recommendations of

Steven Joyce's expert review of Australia's Vocational Education and Training System[10]. The NSC plays an advisory role and does not have policy setting powers directly. The commission is responsible for the skills priority list, which in turn is used as a key input in the skilled migration occupation list.

- United Kingdom – Migration Advisory Committee (MAC) – The MAC is an independent public body that advises the UK government on migration issues. The core committee is a group of academics specialising in the labour market and migration. It explicitly reviews and recommends updates to the Shortage Occupation List. It does not have policy setting powers.

Whatever arrangements are in place, the application of judgement typically continues to play a role in final decisions. This makes it difficult to tell exactly how much the data and model results have influenced final settings. We discuss institutional arrangements further in Section 5.4.

3.6 Alternatives to a managed skills list

Our literature review focuses on ways to list out occupations in shortage. However, there are alternatives that have been advocated.

In **Australia,** the authors of the Grattan Institute report (Coates et al., 2021) *Rethinking permanent skilled migration after the pandemic*[11] argue that Australian migration policy should prioritise skilled workers, with requirements being based on age, English capability and salary – rather than prioritising any particular occupations – as these migrants bring higher economic value to the country. The key objections raised regarding occupational shortages and lists are:

- Skill shortages can be hard to define – arguing that:
  - Relying on wages is not enough, although we note the SPL Indicator Model, for example, does not rely only on wages
  - ANZSCO classifications are not appropriate for skills shortage modelling as:
    - Timely wage datasets are not available at the six-digit level (The Australian Tax Office publishes wages by 4-digit ANZSCO)
    - Workers have multiple tasks and cannot be classified into black-and-white categories
    - It is slow to update to new developments and emerging occupations.
  - There are cases of similar occupations with similar skill and income levels being classified differently by priority occupation lists.
- Most skill shortages are temporary, whereas migrants stay long-term.
- While allowing migrants to fill a shortage in occupations with lower incomes can enable government sectors to control costs, it incurs a shadow tax in terms of the lost tax revenue from a higher income migrant.
- Incorporating consultation is problematic as stakeholders have different goals and are potentially incentivised to argue for a shortage in given occupations.

A migration policy based purely on salary bypasses many of these issues.

[10] https://www.pmc.gov.au/resource-centre/domestic-policy/vet-review/strengthening-skills-expert-review-australias-vocational-education-and-training-system

[11] https://grattan.edu.au/report/rethinking-permanent-skilled-migration-after-the-pandemic/

In the **United Kingdom**, Sumption (2022)[12] argues that shortages are difficult to define and measure as there are no direct measures in the UK. Therefore, visas should be based on skill and salary thresholds and be independent of occupation. Noted issues with measuring skills shortages include:

- The data is often imprecise at the occupation level as occupation groups are not necessarily homogeneous and the data can be out of date by publication date
- The measures are indirect and can have multiple interpretations. For example, increasing employment numbers could be high demand, but demand may be being met.
- Qualitative evidence is problematic because it is not objective and there are perverse incentives.
- There are no objective thresholds to distinguish between no shortage, a shortage or a severe shortage.
- It is challenging for businesses to plan if occupations can move on and off priority lists.

We include discussion of these alternatives in Section 5.1 - Why is a skills shortage data-led approach useful?.

3.7 Themes from the literature review

There are a few themes that emerge:

- Many countries have decided to formally create a skills list based on quantitative indicators. However, there is significant variation in approaches taken. A concrete list is judged a valuable tool, particularly for migration applications. Many sources also point to the direct implications for education and training policy too. The Australian list is designed to serve both policy areas.
- While there is some variation in 'skills' and 'occupations' across sources, most forecasts and applications are done using occupational groupings.
- There is some consistency across the types of labour market indicators that are useful for identifying shortages. Unemployment, employment growth, wage growth and vacancy information are generally included. Beyond this there is more variation in the exact form of the indicators. Indicators can be driven by other factors beyond pure labour shortages, so overreliance on single factors is mostly avoided.
- Most processes, even if they primarily rely on quantitative measures, leave some room for expert judgement and stakeholder consultation. This is particularly relevant if rapid changes are occurring, since labour market indicators can lag the changes. There can be a dedicated independent body to perform this function.
- Limitations identified are reasonably consistent:
  - There are inherent limitations attached to defining shortages itself. For example, data is often imprecise at the occupation level as occupation groups are not necessarily homogeneous.
  - Many indicators rely on survey data, which make conclusions for small regions and smaller occupations more uncertain.
  - There are limits to the level of accuracy that can be expected from a skills model. Having some form of validation or uncertainty estimate is useful for understanding the extent of this.

[12] https://academic.oup.com/oxrep/article/38/1/97/6514754

4 Data availability

4.1 Introduction

A data-based view of skills shortage is not possible if there is not sufficient data to support it. This section explores:

- The existence (and options) for a primary measure of skills shortage in New Zealand. This can be used to measure skills shortages directly or guide the construction of a predictive skills shortage model.
- The range of other labour market indicators that could contribute to a skills shortage model in New Zealand.

A key focus is exploring the **granularity** and **quality** of data. An indicator might be good in principle, but if it is only available at a high level (broad occupations, or national only), it will not be able to offer some of the desired insight. Additionally, measures based on smaller surveys may carry significant volatility that reduces the reliance that can be put on them.

We draw conclusions about the feasibility of building a skills shortage model using available data in Section 5.

4.2 Primary measures for skills shortage

The range of indicators covered in the literature review reflects that there is no universal indicator for skills shortage. However, our consolidated view is that **measures related to employers being able to fill specific skills vacancies is the most direct representation of skills shortage**. This section looks at avenues for sourcing this information in New Zealand. This includes existing private surveys, which are potentially still relevant as primary measures.

The two most common ways to directly measure employers being able to fill specific skills vacancies are:

- **Vacancy-based surveys –** Surveying people recently posting vacancies and asking if they were able to fill the position. This is the approach taken for the Australian SERA survey, where the sample frame is based on classifying job advertisements and then calling up the contact to see if the position was able to be filled.
- **Employer based surveys –** Asking employers if they advertised for specific occupations and then whether they were able to fill the position.

Any direct measure has limitations. Both of the above suffer from the dynamics of how individual occupations relate to job advertisements. For instance, the rate of self-employment is high in many trades, so a shortage may not be as visible through vacancies. Some occupations may see vacancies filled through direct recruitment, rather than through general job boards and websites.

Further, low vacancy fill rates also might reflect different types of shortage, which will not be obvious from the raw statistics (but can be explored through secondary questions on a survey). A low fill rate could be a genuine shortage of people in that occupation, or that the wages on offer are too low relative to alternatives, or that the specific skills or experience requirements within that occupation are an obstacle, or perhaps that the overall labour market is tight.

There are some practical difficulties with the vacancy-based surveys. There may be some ambiguity as to specific occupations (e.g. if a job advertisement is general), so some selectivity is needed in classification and allocation to the survey. Some job ads are also given to recruiters, which means that they are not publicly advertised, or can even be advertised multiple times if it goes to multiple recruiters. And in occupations with a small number of larger employers the survey design may need to be cognisant of repeat sampling rules – if a survey rule is not to contact the same employer for 6 months to avoid hassling, this can impact on the ability to draw a large enough sample of vacancies. Finally, if an employer believes that there is a shortage, they might choose not to post a vacancy at all.

Employer-based surveys tend to be less precise, in that they typically ask about general hiring experience rather than specific positions and locations. Employers also must make the judgement of what occupation their experience corresponds to. General perceptions of hiring has greater risk of respondent bias, where an employer may have a vested interest in suggesting there is (or is not) a particular shortage. While vacancy-based surveys might carry some respondent bias, focusing on factual questions around specific vacancies reduces the risk.

We are not aware of any ongoing vacancy-based surveys in New Zealand – The Department of Labour used to maintain a Survey of Employers who have Recently Advertised, but this was discontinued around 2007. However, there are several employer surveys relevant to the discussion.

New Zealand Business Operations Survey

The New Zealand Business Operations Survey (BOS)[13] is an employer survey run by Statistics New Zealand. The stratification is by industry (being business-focused rather than occupation), covering most private industries (government departments are excluded). It is annual, surveying around 7,000 businesses from a population of 40,000-50,000 business with at least six employees. There is currently one question around difficulty in recruitment for four high-level occupational categories:

Over the last financial year, to what extent did this business experience difficulty in recruiting new staff for any of the following occupational groups?

- *Managers and professionals*
- *Technicians and associate professionals*
- *Tradespersons and related workers (including apprenticeships)*
- *All other occupations*

This is not detailed enough to build an occupations shortage list, but modifications could be used to produce detailed numbers or enough information for validating other indicators.

National Survey of Employers

The National Survey of Employers[14] is a telephone survey conducted annually by MBIE, covering about 2,000 businesses stratified by industry and size. Recent surveys have a few questions related to skills shortages:

- If they had job vacancies in the last 12 months, were there difficulties finding staff?
- Whether they sought a migrant for a hard to fill vacancy?
- Whether they had any recent migrants on staff?

The survey is dynamic, with questions reviewed each year. To be useful for shortage monitoring, questions related to shortages would need to ask about specific occupations and they would need to be maintained from year to year.

NZIER Quarterly Survey of Business Opinion

The New Zealand Institute of Economic Research[15] conducts a quarterly survey of about 1,000 businesses about business conditions. It does not cover all industries; the survey samples manufacturers, builders, architects, wholesalers and retailers, and service sector firms. Information from these industries provides

[13] https://catalogue.data.govt.nz/dataset/business-operations

[14] https://www.mbie.govt.nz/business-and-employment/employment-and-skills/labour-market-reports-data-and-analysis/national-survey-of-employers/

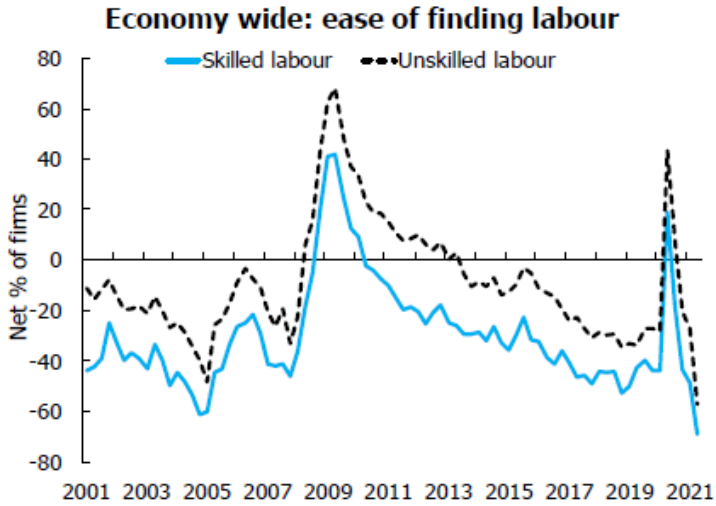
[15] https://nzier.org.nz/ABout%20QSBO/

useful indicators of future investment patterns, and the likely direction and composition of economic growth in coming quarters.

One specific question of interest asks businesses 'in general, do you find that getting labour you want is easier, the same, or harder than it was three months ago?' Results are split into unskilled and skilled, covering the employer rather than specific occupations. These are converted to an index – see Figure 5.

The survey has a long history dating back to the 1950s. Reporting and data is accessible through an annual subscription model ($2,450 per annum). While the survey provides useful insight into businesses' experience recruiting labour, the lack of detailed occupational data limits its use for building a skills shortage model. The industry categories are useful in general, but not sufficient for the purposes of building an occupational level model of skills shortages. Some occupations will tend to be confined to certain industries (e.g. mine manager) while others are diffuse (e.g. accountant, receptionist).

Figure 5 – NZIER index of ease of finding labour



Source: NZIER, https://nzier.org.nz/media/nziers-qsbo-shows-a-strong-pick-up-in-demand-and-confidence-quarterly-survey-of-business-opinion-july-2021

Auckland Business Chamber New Kiwis Employer Survey

This is a relatively new survey with one iteration to date. Businesses were included in this survey who had either:

- Used the www.newkiwis.co.nz website
- Been members of the Auckland and National Regional Chambers of Commerce
- Responded to marketing about the survey.

There were 865 responses. Questions focused on understanding how employers recruited, and the experience with employment of migrants, such as barriers and retention.

The sampling process means that the survey is less useful as a primary measure but may be of interest for validation purposes.

4.3 Labour market indicators related to skills shortage

4.3.1 Coverage of main skill shortage indicators in NZ data

Table 2 in Section 3.4 shows the range of indicators typically used as indirect measures of skills shortage, which are often in turn aggregated to provide a combined measure. We have used this as a starting point for considering the indicators available in New Zealand.

When considering availability of regional information, we have assumed that 10 regions (as used by MBIE in their Jobs Online collection) is appropriate. This aligns with New Zealand regions, but combines some smaller regions, particularly when labour markets are likely to be shared.

Overall, there is a good set of detailed indicators that could support some form of indicator model (whether a weighted indicator model or predictive model). The Household Labour Force Survey (HLFS) encapsulates a large fraction of the common candidate indicators, and a range of other survey and administrative data captures much of the rest.

Table 3 summarises our view on the indicators with further comments on specific data sources following.

Table 3 – Summary of labour market indicators

Indicator	Possible?	Source	ANZSCO level	Region level	Comments
Number employed	✓	HLFS Census	6 6	Y Y	Census figures might be useful in contexts where the relative size of occupations is important, where up-to-date figures are less important, avoiding the volatility of the HLFS.
Employment growth	✓	HLFS	6	Y	Appropriate smoothing or credibility is needed since the volatility at ANZSCO-6 and regional level would be very high.
Working hours and change	✓	HLFS	6	Y	
Overtime hours	Partly	HLFS, partly	6	Y	Some information could potentially be inferred from the difference between hours actually worked and hours usually worked.
Education level of those employed	✓	HLFS	6	Y	HLFS has both highest school qualification and post-school qualifications.
Study diversity	✓	Census	6	Y	Census collects highest educational attainment and field of study. These can be compared to occupation (particularly for younger workers) to develop study diversity measures. These are fairly static over time so recency of data less an issue.
New hires fraction	✗	Not directly available, but proxies in the HLFS	6	Y	E.g. Question: has your job changed in past 3 months?
Unemployment rate	✓	HLFS			The HLFS also has an underutilisation measure (people looking for additional hours), which might be useful for testing a broader measure than unemployment.
Unemployment duration	✓	HLFS	6	Y	HLFS collects when a person started looking for work, as well as both the occupation sought and previous occupation.

Indicator	Possible?	Source	ANZSCO level	Region level	Comments
Wage premium	✗	HLFS Labour Cost Index (LCI)			A wage premium model involves estimating which occupations have above average pay after adjusting for factors such as skill and training. It is unclear whether data volumes in the HLFS would support a model. The LCI could be an alternative, but regional information is not available and occupational dis-aggregation is only to ANZSCO-2 level.
Wage growth	✓ although best option unclear	HLFS income supplement Potentially HLFS + IR linkage in the IDI LCI	6	Y	Inland Revenue (IR) tax data does not have occupation fields but could be linked to HLFS.
Annual change in vacancy numbers	✓	MBIE Jobs Online Index	4	Y	Note: Regional breakdown only available at ANZSCO-1 level.
Vacancy fill rate	✗				Would require a dedicated survey, or expansion of employer surveys.
Vacancy - education level required	Likely	Job boards data (e.g. Burning Glass)	6	Y	Relies on metadata being collected from job vacancy data.
Migrant numbers	✓	Immigration NZ visa data HLFS	6	Y	Migration data includes occupation at entry for some visa types e.g. Essential Skills and Work to Residence; large numbers can be indicative of a shortage. HLFS also asks the number of years a person has lived in New Zealand.
Fraction of foreign workers	Maybe	Immigration NZ visa data HLFS	6	Y	Occupation on application is captured in the visa data. HLFS also asks the number of years a person has lived in New Zealand. If reliant on the link to HLFS, likely to be very volatile.

Household Labour Force Survey (HLFS)

The HLFS[16] is one of the largest regular surveys, collecting detailed information on labour force status. Approximately fifteen thousand (15,000) households take part in this survey every three months. A house is selected using a fair statistical method to ensure the sample is an accurate representation of New Zealand. Every person aged 15 years or over living in a selected house needs to take part.

Underlying data is collected at an ANZSCO-6 and Meshblock of residence (the smallest Stats NZ geographic unit) level; so, the granularity is powerful, limited only by the size of the survey itself.

Income data by occupation

Income is not part of the quarterly HLFS, but there are options for building wage-based indicators:

- The HLFS includes an income component once per year
- Via linkage, the HLFS could be linked to IR taxation income data
- The Labour Cost Survey which is used to form the Labour Cost Index (LCI) has occupation-level data (unknown level ANZSCO) but not regional information.

The Household Economic Survey (HES) separately measures detailed income information but does not have occupation recorded.

At this stage we do not have a strong view on which avenue is the most useful. Data volumes are likely to be important, and the HLFS with income component is the most complete self-contained source. The value of wages (both premium and growth measures) as an indicator for shortage has been popular elsewhere, so we would expect some form of measure of wages should be included in a New Zealand model.

Census data

Census data suffers from frequency issues, but still potentially plays an important part in constructing a model. 'Structural' labour market features that tend to move slowly can rely on census data, where coverage is good. For instance:

- Combining and splitting occupational groups using the ANZSCO (or alternative) hierarchy requires different 'weights' for various occupation sizes. Low volatility in weights is more important than timeliness, so census data for weighting makes sense.
- Mapping fields of study to later occupations requires large datasets, and likely evolves slowly over time. The census has the fields required to do this attribution.
- Estimating the number of people in each occupation at a regional level will tend to triangulate data from different sources. Census could be the starting point, with trends in HLFS and other sources providing the evolution over a few years.

While there are known issues with Census 2018, it still provides the most comprehensive basis to set some of these structural assumptions.

Detailed job advertisement data

Various databases exist that collate and process job ad data. One prominent example is Burning Glass, which maintains datasets across different countries. Burning Glass Technologies[17] is a private company that provides a database of structured job ads data (Burning Glass data) collected from online job postings.

---

[16] https://www.stats.govt.nz/help-with-surveys/list-of-stats-nz-surveys/about-the-household-labour-force-survey/

[17] https://www.burning-glass.com/au/. Note that Burning Glass Technologies merged with EMSI in 2021.

These are classified and coded in systematic ways including required skills, required education, required experience, as well as wage information. This data was used in the Australian Priority Skills List Indicator Model, the Malaysian Core Occupation List and the UK MAC model.

We have not confirmed indicative pricing, but we have confirmed that Burning Glass does maintain unit-record on New Zealand job posting data back to 2012. There are likely some systematic data issues in the collection – Cammeraat and Squicciarini (2021)[18] identify some year-to-year volatility in the New Zealand data compared to other countries. This means absolute volumes are likely to be more reliable from MBIE Jobs Online, but some of the detailed variables (e.g. education requirements) will still be of potential value. Therefore, Burning Glass data is likely to complement, rather than replace, MBIE Jobs Online data.

Migration data

Immigration New Zealand migration data is available in the Stats NZ Integrated Data Infrastructure (IDI), covering migrants and international visitors applying for a visa to enter New Zealand. This includes all resident visa applications and those applying for a temporary stay (work, study, and visitor). This data forms part of the basis for MBIE's Migrant Employment Data (MED) reporting.

Visa applications capture people's occupation on application which, in theory, could be used as a proxy for occupation when working in New Zealand. The accuracy of the proxy could be judged if there is a suitably large pool of new migrants captured in the HLFS.

4.3.2 Other indicators

We note the list in Table 3 is not intended to be exhaustive. If a primary measure is established, other indicators can be tested and potentially validated. For example:

- Census data could be used to estimate retirement rates from professions, and a source of future demand.
- Similarly, numbers studying in different fields (as a fraction of that occupation size) can be used as a proxy supply indicator.
- It is unclear whether welfare data might be useful. However, trends in numbers, education level and region will have some implications for the labour market.
- The MBIE survey of employment intentions[19] is potentially another demand indicator, although some mapping between industry and occupation may be needed to provide trends at an occupational level.
- IR tax or HLFS data could be used to derive employment churn (rates of movement) at an occupational level, which will have implications for specific occupations

Note that when developing the Skills Priority List Indicator Model for the NSC in Australia, nearly 200 potential indicators were tested. Many of these were alternative versions of the same underlying data. The indicators retained were proven to have the most predictive power.

4.3.3 Implications of survey sample sizes

One challenge in selecting data for a model is that there will be natural limits to detail arising from the size of datasets. While the desired granularity (e.g. ANZSCO-6 at a regional level) may be available, the volume of contributing data may be too low to infer reliable evidence at that level of disaggregation.

[18] https://www.oecd.org/sti/burning-glass-technologies-data-use-in-policy-relevant-analysis-cd75c3e7-en.htm

[19] https://www.mbie.govt.nz/business-and-employment/employment-and-skills/labour-market-reports-data-and-analysis/national-survey-of-employment-intentions/

One of the main sources identified is the HLFS. This covers about 15,000 households (and 30,000 individuals) each quarter. Given 1,000 occupations and 16 regions, there are 16,000 potential 'cells' of data desired, more than the number of data points.

In practice dataset size limitations will mean that:

- Data is used hierarchically to provide stabilisation, with national numbers contributing to regional estimates and higher-level occupation groups contributing to the lower-level occupational group estimates.
- Uncertainty ranges should be tracked, so that estimates with unreasonably large confidence intervals can be deprecated.

We note that the HLFS compares reasonably well with other surveys internationally. For example, the Australian Labour Force Survey covers 26,000 households and 50,000 people. The survey component for Queensland (roughly the same size as New Zealand in population) is 4,300 households. The Australian model was able to provide credible estimates at a sub-state level, even for some smaller states. This means that National-level results in New Zealand will be significantly more accurate than state-level in Australia, and that incorporating some regional results should be feasible.

5 Value proposition and feasibility of a data-led approach in NZ

5.1 Why is a skills shortage data-led approach useful?

Shortages in specific skills and occupations have the potential to cause disruption and hinder economic growth. The Government has the potential to smooth shortages through the use of skilled migration policy, apprenticeship incentives and training funding.

The ability of Government to enact such targeted support is dependent on being able to identify which occupations or skills need support. This is the role of a skills shortage model.

At its core, a skills shortage model provides an objective basis for identifying which occupations, or underlying skills, are in shortage. This has direct policy implications:

- **Migration policy –** An explicit list of occupations that receive priority in the migration system (e.g. needed to meet a skilled migrant category) is common internationally and in current operation in New Zealand.
- **Training and education policy –** The government has the ability to offer training incentives in different fields to encourage a pipeline into a shortage area. This reflects that there can be market inefficiency in how students select courses relative to later employment prospects. Incentives could range from course funding, apprenticeship or wage subsidies. For instance, the Australian Labour Party has made an election promise that "anyone pursuing a TAFE course which is on the National Skills Commission's priority list will do it for free", as a direct link between such a list and training policy[20].

A model also has potential non-policy uses, such as labour market signalling and monitoring, and general information provision.

A common criticism of skills lists is that they represent an attempt to 'pick winners' and that it is too difficult to usefully maintain a list in a complex labour market. An alternative in prioritising migrant applications would be to replace a skilled migration list with an income threshold (e.g. position secured for a skilled migrant applicant must be for an income of at least $75,000). We have some sympathy for the argument and understand the logic of placing reliance on market forces to balance competing intentions. We also recognise inherent inaccuracies in skills modelling. However, we note:

- **A skills list enables coordination between migration and training policies.** Migration may not always be the best solution if the aim is to ensure good training and opportunities for young New Zealanders. While migrants taking up high-paying jobs contribute to GDP growth and government tax revenue, there may also be a perception that this reduces opportunities for other residents.
- **A skills model is useful for understanding essential occupations with lower wages.** Occupations can be in shortage even if the typical wages on offer lie below a migration threshold (e.g. disability and aged care roles). A skills list grants greater flexibility to address acute shortages.
- **A skills model also enables labour market monitoring.** Even if a skills list was not used to tune specific policies, there would still be significant value in understanding what shortages were emerging over time. For instance, there may be particular regional or ongoing shortages despite material migration. An objective basis for monitoring, using existing data sources, is an attractive way to do this. This also provides a framework of indicators of which some could be monitored by demographic splits (such as gender or Indigeneity) to provide further insight into the labour market.
- **Hybrid models are also possible.** For instance, Ireland uses income thresholds for work permits. The threshold is lowered in cases where the application relates to an occupation assessed as in shortage

[20] https://www.smh.com.au/politics/federal/free-tafe-more-uni-places-in-1-1-billion-labor-promise-for-post-pandemic-skills-training-20211202-p59ec9.html

using a weighted index model. We also see a place for data-led approaches to supplement consultation-based approaches to adding and removing occupations from skills lists.

A second criticism is that labour market conditions 'at the coal face' are not well captured by government statistics, and that a list based on expert review and stakeholder submissions should be preferred. We agree there will always be a role for this type of consultation. However, without objective data for comparison there are significant risks of inefficiencies in this process. Peak bodies may be better at advocating for a particular occupation to be added or removed from a list, whereas other occupations may lack the same level of coordination. It can also be difficult to remove an occupation from a migration list without objective information showing that there is not a shortage.

Conditional on an indicator model having proven validity (as has been done overseas), we believe that there is genuinely useful information that can be gained from some form of model. This would allow government to be more tactical with policies and potentially reduce friction in the labour market.

5.2 Overall feasibility of a New Zealand Skills model

In Section 4, we explored a range of data sources for a potential primary measure of skills shortage and secondary potential predictors of skills shortage. From this we conclude that:

- **New Zealand lacks a good primary indicator of skills shortage, but this could be developed.** This means that there is no natural target variable for a predictive shortage model, or something that could be used for ongoing validation. Amending the BOS is probably the best option if such a measure is desired:
  - The best approach to amending the BOS is likely to be replacing or supplementing the existing question on recruiting difficulty (See Section 4.2) with, say, 3-5 fields where respondents can list specific occupations they have advertised for over the last 12 months, the number of positions sought in these occupations, and their assessment of difficulty filling the vacancies.
  - The specific design of the question(s) would need to be carefully considered to minimise ambiguity and ensure appropriate capture of information.
  - Given the survey is not online, conversion of occupational responses to ANZSCO classifications would be required.
  - For large organisations who recruit in more occupations than is reasonable to collect information on in the survey, some guidance on which occupations to choose could be given.

  Survey results could be used immediately for validation – for example, testing how well a proposed weighted indicator model separates some high- and low-shortage occupations. However, it would take several years to build up a longitudinal dataset sufficient to form the basis of a predictive shortage model.
- **There is a very good range of secondary indicators to build a weighted indicator model** or a predictive shortage model (if a primary measure of skills shortage existed). In particular:
  - The HLFS is a rich source of information with the required occupation and regional information.
  - The IDI provides a useful way to work with granular data safely, allowing modellers to move across different levels of ANZSCO and geography.
- **International sources are a good place for starting to build a weighted indicator model** – this would be the most practical way to build a model quickly. For example, the Australian Skills Priority List Indicator Model could act as a guide as to what indicators to include and what weights to assign each.
- **Model accuracy is likely to be reasonable.** We discuss further in Section 5.3.7, but the Australian experience suggests we can expect somewhere between half to three quarters of variability (in vacancy

fill rates) to be explained using an indicator model. While uncertainties definitely remain, this is enough to substantially improve labour market understanding.

Based on this, we think that **a weighted indicator model of skills shortage approach is feasible and is likely to add insight** to related migration, and training and education policy considerations. This could be done relatively quickly and cheaply, given the availability of secondary indicators and typical constructions used internationally. If an occupational measure is added to an employer survey, this can be used to validate and improve the model over time.

Ideally a model would provide estimates at a very detailed occupational level (6-digit ANZSCO level), as well as a detailed regional level (e.g. regional council). This would allow fine-grained examination of labour markets, with the ability to identify specific acute shortages. There are inherent limitations to full granularity:

- Some indicators will not be available at full granularity.
- Many measures carry natural statistical noise and variation, accentuated on more granular results. Even if an indicator has data at a granular level, it may be very volatile – reporting outputs using the most detailed level of data may detract from the model. This is particularly true for survey-based measures.

What can be delivered is a combination of data-specific issues and a user's tolerance for uncertainty. With the data available we believe it is feasible to develop a weighted indicator model that disaggregates to:

- ANZSCO-6 at a national level
- ANZSCO-4 with a geographical breakdown for key regions. Some smaller regions may need to be aggregated.

Estimation of confidence intervals would allow 'pockets of certainty' to be identified. For instance, some specific larger occupations might be known accurately at a regional level, even if other occupations are uncertain.

5.3 Conceptual and practical issues to consider for model development

There are a range of conceptual and practical issues associated with developing a data-based skills model, some of which were touched on in the literature review.

5.3.1 Predicting current versus future skills shortages

Ideally a model would be able to predict future shortages, since this gives government time to act through migration, training or other levers. This is particularly true for training, since there are significant delays between a person starting training and being work-ready. However, predicting future shortages is more difficult, and carries greater uncertainty, than identifying current shortages since it requires extrapolation of trends. Given the significant amount of evidence that shortages tend to persist over time, a 'now-casting' approach appears more feasible; using secondary labour market indicators to predict current shortages. We have not seen compelling evidence that a future-focused measure produces materially different and improved estimates compared to a 'now-casting' approach. That's not to say that forecasting skills shortages isn't possible. The Canadian Occupational Projection System (COPS) is a good example and has been in use for several decades. But it is relatively dependent on assumptions that are set using the current state of the labour market; according to their summary, over 90% of occupations receive the same shortage classification as their current market assessment.

5.3.2 Occupations versus skills

In this report we have used the terms skill shortage and occupation shortage loosely. However, the distinction is important. Skills shortages talks to the inherent employee capabilities which may be important for a job and are not uniquely tied to occupations. Two different occupations might share skills

(for example, finance manager and accountant might both need skills in finance software), enabling easier transitions between the two jobs. Alternatively, the same occupation might require different skills (for example, a computer programming vacancy might require skills in Python programming, which not all programmers have). And the skills required of an occupation might evolve over time (e.g. farming might need increased skills in computerised monitoring tools).

There is a natural attraction to skills shortage, since this provides a deeper labour market understanding and directly ties to training content. However, the challenges of monitoring skills and identifying gaps is far harder, particularly as most labour-market data collection is currently geared towards occupations. Migration lists and courses are often geared towards occupations, often on the assumption that the underlying skills will evolve with the needs of the labour market.

That said, there are examples of relatively detailed mappings of competencies and skills to occupations, such as the Australian Skills Classification[21]. It identifies specialist tasks, technology tools and core competencies for every occupation. The ASC is partly based on the US Department of Labor sponsored O*NET OnLine[22] classification system.

While further development of skills classification systems may yield data-based opportunities in the future, particularly in the context of education and training, we believe that occupation-level shortages are the most useful and practical to consider, consistent with much of the existing literature. Any data-based approached focussing on occupation-level shortages could, in theory, be supplemented with skills classification information to accommodate potential transferability of skills across occupations.

5.3.3 Relative versus absolute shortages

A skills model can be used to talk about relative shortage (e.g. the score for plumbers is higher than librarian) or for absolute measures (e.g. both occupations sit above a given threshold for defining a shortage). With a suitable range of indicators both views will be visible. Some consideration should be given to how the list is used relative to the aggregate labour market. When a labour market is tight (low unemployment, high wage growth), it may appear that a much larger fraction of occupations are in shortage than usual. In such a case, judgement is required to decide if this translates into a longer migration or targeted training list, or broader macroeconomic policy considerations.

5.3.4 Skilled versus unskilled labour

Many skills shortage lists focus on the more skilled occupations (e.g. skill levels 1, 2 and 3 in the ANZSCO classification) on the basis that labour supply for unskilled labour is more fluid and does not have the same implications for education and training policy. A skills model could still be applied to lower skill levels, since indicators will still exist. This may be useful when specific essential occupations with low skill classifications are considered (e.g. Aged and disability care workers, classified as skill level 4).

5.3.5 Choice of occupation taxonomy

Assuming an occupation-based approach to skills modelling (see section 5.3.2), an occupation taxonomy is needed. While the labour market and specific occupations evolve over time, such a standardised list allows the labour market to be grouped and trends tracked over time. ANZSCO is the natural choice, as this is used for much of the existing data collection in Australian and New Zealand. While there was a recent small update to ANZSCO, typically updates over the past 15 years have been limited which means that some newer job titles (e.g. data scientist) are not represented. It is also relatively fine-grained, with the consequence that many jobs (e.g. nurse) are advertised at a less granular level, potentially making it

[21] https://www.nationalskillscommission.gov.au/australian-skills-classification

[22] https://www.onetonline.org/

unwieldy. There are also ambiguities around the assignation of skill levels to different occupations (e.g. Whether a skill level 2 ICT technician can be relabelled as a skill level 1 ICT manager).

One key international alternative is the 2008 International Standard Classification of Occupations (ISCO-08). This 4-digit hierarchy can be extended down to the unit group level (rather than occupation level), with fewer fine-grained occupational distinctions (e.g. Unit group 2212 for specialist medical practitioners groups different specialties such as anaesthetists, cardiologists and obstetricians). This classification has been used in other international work such as the International Assessment of Adult Competencies (PIAAC) survey. There are some advantages to a finer-grained occupation list like ANZSCO, since there may be occupation-level differences to recognise.

Other national categorisations exist, such as the Standard Occupational Classifications (SOC) used in the United States. Such taxonomies can be useful in broader research, such as the O*NET database, which among other things maps skills onto occupations.

Most practically, ANZSCO is used for most current statistical work around occupations. So, a departure from ANZSCO would involve substantive effort in mapping to a new coding basis, as well as ambiguities when mappings are not perfect. For this reason particularly, it remains the most realistic option for skills shortage modelling in New Zealand.

Further, we note:

- Version 1.2 of ANZSCO, with some adjustments, is currently used by Immigration New Zealand. Version 1.3 is typically used in Australia. The differences between versions 1.2 and 1.3 mainly relate to the skill levels assigned to occupations. For the purpose of developing a skills shortage model it may be necessary to make some adjustments to ANZSCO to reflect the recruiting practice of specific occupations. For example, for occupations that are consistently advertised at an aggregated ANZSCO level, it may make sense to not model below that level for that occupational grouping.
- The Australian Government has announced a review of ANZSCO as part of the 2022-23 Budget. This will prove a timely update to recognise the evolving labour market and newer occupations.

5.3.6 Data access and timeliness

The modelling envisaged requires more detail from the underlying data than is routinely published (e.g. the HLFS does not publish below ANZSCO level 1 currently). The Stats NZ Integrated Dataset Infrastructure (IDI) is the obvious place for accessing most of the required information and provides a good environment for much of the work. Some data is not in the IDI (e.g. Jobs Online data, or private job ad data). This means that either this data is imported into the IDI to do the modelling in a single workspace, or semi-aggregated indicators are exported from the IDI and the model finalised outside that environment.

Unless data is real-time, there is a lag between the time point or period data relates to, and the time point the data is accessed and incorporated into a model. Large data lags may impact the usefulness of data and any model that relies on it. The IDI operates on a three-times a year data refresh process. This means that data lags can be significant. For example, the most up to date HLFS data available in the IDI at the date of this report is for the June 2021 quarter HLFS. This compares to the output from the December 2021 quarter HLFS which has been publicly available on Stats NZ's website since early February 2022.

We also note that our assessment of available data in Section 4 is a current state assessment. Surveys and other data sources can change over time or be withdrawn. For a model to have longevity, data sources need to be protected and maintained. We note that a key data source identified in Section 4, the HLFS, is relatively stable, given that it is a core component of labour market reporting. Other data sources maybe more prone to change. Upon building a model, discussions should be had with the custodians of key data sources, to ensure continuity.

5.3.7 Model accuracy

While hard to judge from the international literature, accuracy is a key question; if a model is built that can only explain 10% of the variation in shortage across occupations, it is not a useful planning tool. Conversely if it explains 90% then it can be given significant weight and trust in decision making. Results in Dawson et al. (2020) are consistent with the experience of this report's authors; it might be reasonable to expect somewhere between half to three quarters of variability can be explained using an indicator model. This still leaves room for error and potentially expert review, but represents a relatively strong basis for decision making. A model that also quantifies the uncertainty of a particular estimate will aid interpretation.

Further, we also believe that finding ways to validate a model is important**;** for instance, even if a regular primary measure is not used to build a predictive model, ad hoc collections can still be used to establish estimates of accuracy and the ongoing credibility of the model.

5.3.8 Role of stakeholder and expert judgment

Given a skills model will be subject to some inaccuracies and potentially data lags, we agree with the view that there remains a role for expert judgement and stakeholder feedback in developing a shortage list. Having an objective starting point from a skills shortage model will significantly improve the robustness around such a process. This is consistent with most overseas processes.

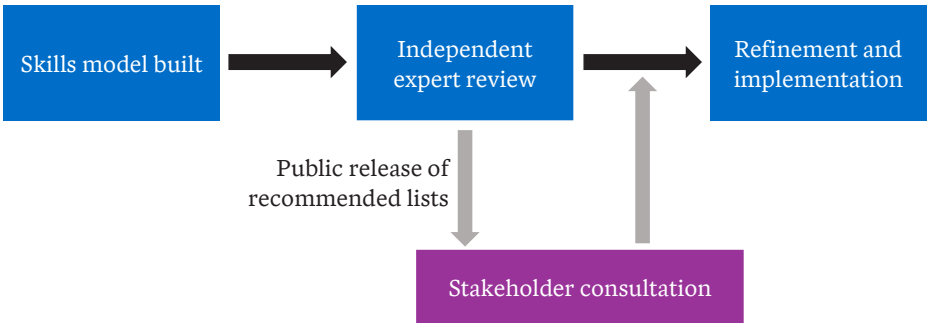
5.4 Overall structure for developing a skills list

Much of our attention in the report is focused on the 'skills model'. However, this would need to fit into a broader institutional framework and process for deriving specific policy outputs. We do not argue for a specific institutional framework here, but we do believe there is significant value in an expert review stage that is independent of the final policy users and the stakeholder consultation. For example:

- In Australia, the National Skills Commission (NSC) defines the priority list before they are then passed to and used by Education and Immigration.
- In the UK, the Migration Advisory Committee (MAC), comprised mainly of academics, provides and publishes independent views.

A broad representation of the process is shown diagrammatically in Figure 6. The independent expert review stage provides an important steppingstone that can improve the credibility of the process. Publicly releasing results at this stage provides transparency that, we believe, reduces the risk of regulatory capture and skill lists unduly influenced by advocacy.

Figure 6 – Potential process for moving from a skills model to implementation



Specific issues, such as whether a new separate independent body is required, or what governance is required around the stakeholder consultation and refinement stage, would have to be carefully worked through.

5.5 Potential next steps for building a skills shortage model

If the decision was made to build a weighted indicator model, the high-level next steps would be similar to the following:

- Establish the specific purpose of the model, what it is intended to be used for and who 'owns' the model and its maintenance
- Detailed exploration and testing of data sources
- Establish the basis for determining which indicators to include and what weightings to apply, including consultation with subject matter experts
- Develop the model
- Validate the model
- Implement the model, including documentation and potential publication of information.

These steps relate specifically to building a model. Broader process and consultation will no doubt be required to address some of the conceptual and practical issues identified in Section 5.3.

6 References

- Ball, C., Richardson, A., & van Florenstein Mulder, T. (2020). Using job transitions data as a labour market indicator (No. AN2020/02). Reserve Bank of New Zealand.
- Cammeraat, E., & Squicciarini, M. (2021). Burning Glass Technologies' data use in policy-relevant analysis: An occupation-level assessment. OECD Working paper 2021/05 accessed February 2022 at https://www.oecd.org/sti/burning-glass-technologies-data-use-in-policy-relevant-analysis-cd75c3e7-en.htm
- Coates, Brendan, Henry Sherrell, and Will Mackey. Rethinking permanent skilled migration after the pandemic. Grattan Institute, Carlton, Victoria, 2021.
- Critical Skills Monitoring Committee (2021). *Critical occupations List 2020/2021 Technical Report.* Accessed February 2022 at https://www.talentcorp.com.my/clients/TalentCorp_2016_7A6571AE-D9D0-4175-B35D-99EC514F2D24/contentms/img/Documents/COL%202020-2021_Technical%20Report_FINAL.pdf
- Dawson, N., Rizoiu, M. A., Johnston, B., & Williams, M. A. (2020). Predicting Labor Shortages from Labor Demand and Labor Supply Data: A Machine Learning Approach. *Preprint https://arxiv.org/abs/2004.01311*
- Department of Employment, Skills, Small and Family Business, 2019. Review of the Australian Apprenticeships National Skills Needs List – Methodology Discussion Paper. Accessed February 2022 at https://www.dese.gov.au/review-australian-apprenticeships-national-skills-needs-list/resources/review-australian-apprenticeships-national-skills-needs-list-methodology-discussion-paper
- Department of Labour NZ, Infometrics (Firm). (2006). Indicators of skill shortage. Accessed February 2006 at https://thehub.swa.govt.nz/assets/documents/Indicators%20of%20skill%20shortage.pdf
- Department of Prime Minister and Cabinet, 2019. Expert Review of Australia's Vocational Education and Training System. Accessed April 2019 at https://www.pmc.gov.au/resource-centre/domestic-policy/vet-review/strengthening-skills-expert-review-australias-vocational-education-and-training-system
- Foot, D., & Meltz, N. (1992). An ex-post evaluation of Canadian occupational projections, 1961-1981. Relations Industrielles/Industrial Relations, 47(2), 268-278.
- Ignaczak, L. (2017). Projecting labour market needs in Canada: The COPS national model after 30 years. *Human Resources and Skills Development Canada, Ottawa, viewed*, *27*.
- Mavromaras, K., Healy, J., Richardson, S., Sloane, P., Wei, Z., & Zhu, R. (2013). A system for monitoring shortages and surpluses in the market for skills. *Report prepared by NILS, Flinders University for the Australian Workforce and Productivity Agency*, 1-65.
- Migration Advisory Committee. (2019). Full review of the shortage occupation list. Migration Advisory Committee. Accessed February at https://www.gov.uk/government/publications/full-review-of-the-shortage-occupation-list-may-2019
- Migration Advisory Committee. (2020). Review of the shortage occupation list. Migration Advisory Committee. Accessed February at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/927352/SOL_2020_Report_Final.pdf
- McNaboe, J., Burke, N., Condon, N., Hogan, A. M., Shally, C., & Walls, D., (2021). National Skills Bulletin 2021. *Further Education and Training Authority (SOLAS)*. Accessed February 2022 at https://www.solas.ie/f/70398/x/fcee571661/solas_nsb_report.pdf

- Mok, P., Mason, G., Stevens, P., & Timmins, J. (2012). A Good Worker is Hard to Find: Skills shortages in New Zealand firms. *Occasional Papers, 12*(5).
- Moroz, H. E., Schmillen, A. D., Abdur, R., & Amanina, B. (2019). Monitoring occupational shortages: Lessons from Malaysia's Critical Occupations List. The Malaysia Development Experience Series. Washington, DC: World Bank Group. Accessed February 2002 at https://documents.worldbank.org/en/publication/documents-reports/documentdetail/877271568138256963/monitoring-occupational-shortages-lessons-from-malaysias-critical-occupations-list
- National Skills Commission (2021). *Skills Priority List Methodology*. Accessed February 2022 at https://www.nationalskillscommission.gov.au/skills-priority-list-methodology.
- Sumption, M. (2022). Shortages, high-demand occupations, and the post-Brexit UK immigration system. *Oxford Review of Economic Policy*, *38*(1), 97-111.
- Thomas, J. (2015). Review of best practices in labour market forecasting with an application to the Canadian Aboriginal population. Centre for the study of living standards, Ontario, Canada.

Appendix A Detailed indicator lists from literature review

Table 4 – Indicators in the MAC model, United Kingdom

Indicator	Data source
Percentage change of median real pay over: ▪ 1 year ▪ 3 years	Annual Survey of Hours and Earnings
Return to occupation, the predicted hourly wage for a set of reference characteristics relative to the average predicted wage for the same characteristics over all occupation codes. Separately calculated for 3 skill levels.	Annual Population Survey
Vacancies / Employment	Employer Skills Survey Annual Population Survey
Vacancy postings / Employment	Burning Glass Annual Population Survey
Percentage change of employment level over 1 year	Annual Population Survey
Percentage change of median hours worked over 3 years	Annual Survey of Hours and Earnings
Change in proportion of workforce which are new hires over 1 year	Annual Population Survey
Weighted stock of unemployment and inactive / Employment, unemployment and inactive	Annual Population Survey

Table 5 – Indicators and data sources used in the COL, Malaysia

Indicator	Data source
Employment growth (1 year and 3 year)	Labour Force Survey
Working hours growth (1 year and 3 year)	Labour Force Survey
Median education level decrease (1 year and 3 year)	Labour Force Survey
Wage premium (using regression at individual level) (1 year and 3 year)	Salaries and Wages Survey
Number of vacancies	Online Job Posting Data
Vacancy rate	Online Job Posting Data
Required education increase over 1 year	Online Job Posting Data

Table 6 – Indicators considered in the National Skills Bulletin 2021, Ireland

Indicator	Data source
An annual average of employment figures over the four quarters in 2020	Labour Force Survey
The annualised rate of employment growth for the period 2015-2020	Labour Force Survey
The percentage of females employed in an occupation	Labour Force Survey
The percentage of persons who work full-time in an occupation	Labour Force Survey
Proportion aged 55 years and over	Labour Force Survey
Proportion of non-Irish nationals	Labour Force Survey
The percentage of persons who have attained a third level qualification	Labour Force Survey
Employment permits issued to non-EEA nationals in 2020	Department of Enterprise, Trade and Employment
Mentions of difficult-to-fill vacancies in the Recruitment Agency Survey	SLMRU Recruitment Agency Survey
Job vacancies and growth	Online jobs board
Supply of skills from education and training	
Unemployment rates	Labour Force Survey

Table 7 – Predictors and data sources used in the SPL Indicator Model

Predictor	Data source
Change in the number of average weeks spent looking for work for unemployed person, relative to the average over the last three years	ABS Labour Force survey
Number of job vacancies reported online standardised by size of occupation according to number employed	Internet Vacancy Index
Employment numbers: ▪ Number of persons employed ▪ Change in the number of persons employed, relative to the average over the last three years	ABS Labour Force survey
Change in the unemployment rate from the previous year	ABS Labour Force survey
Change in mean annual salaries from previous year	Australian Tax Office
Proportion of job listings that specify at least a bachelor degree as a minimum education requirement	Burning Glass
Temporary skilled visas granted relative to occupation size	Home Affairs
Study diversity of education for occupation	Census

TAYLOR FRY

www.taylorfry.com.au