International Labour Organization

ILO brief

ILO Brief

May 2022

Alternative Data Sources for Labour Market Diagnostics and Policy Response[1]

Applications in the COVID-19 era and beyond

Key points

- Policy makers have struggled with capturing the labour market impacts of the pandemic in a relevant and timely manner. This challenge underlines the need for more investment in developing alternative data sources that help estimate essential labour market indicators, while preparing for future crises when labour force survey data may be unavailable.
- Non-conventional data sources include administrative data, leading economic indicators, dedicated surveys, and Big Data. One key underused source is employment and unemployment registry data, compiled by governments in most countries to track hires and unemployment claims. It provides policy makers with a high-frequency source of data with which to assess rapid changes in employment and unemployment trends by region and sector.
- Often leveraging the power of artificial intelligence and machine learning, Big Data analysis allows to identify signals or correlations within large datasets. One primary source of Big Data is mobile phone location data, collected either through phone signals or by GPS in phones and apps.
- Another useful Big Data application is through data from internet search, social media and online job portals. Data from online job portals allows researchers and policy makers to better understand job preferences, employer demands and worker skill profiles
- Rooted in analytical techniques used in weather forecasting – termed nowcasting – is being used increasingly by economic policy makers. A prime example of this is the ILO nowcasting approach used in the ILO Monitor series to estimate hours worked.
- Developing the capacity to effectively understand and employ these approaches will be challenging and require substantial efforts. Policy makers need to be aware of privacy and other ethical issues while creating partnerships with the technology companies.
- Beyond providing an expanded diagnostic capability, non-conventional data sources – particularly Big Data – have an important role in helping to improve the targeting, efficiency, and impact of policies and programmes directed at improving labour market outcomes (York and Bamberger, 2020). This is particularly the case in terms of job placement and worker upskilling programmes, where policy makers and social partners are increasingly using Big Data techniques. However, alternative data sources are not a replacement for labour force surveys and should be seen as complementary.

[1] This document was developed by Paul Dyer, Independent Consultant. Sher Verick and Drew Gardiner from the ILO provided overall guidance and feedback. Samantha Watson from the ILO reviewed the brief. This report has been funded by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Sector Project "Employment Promotion in Development Cooperation", on behalf of the Federal Ministry for Economic Cooperation and Development (BMZ). The analysis, results and recommendations in this paper represent the opinion of the author(s) and are not necessarily representative of the position of GIZ or BMZ.

I. Introduction

The COVID-19 pandemic and the set of policy responses with which policy makers sought to address it effected an unprecedented economic shock with stark consequences for workers across the globe. As documented by the ILO Monitor, the pandemic and the public health response – with policy makers resorting in many cases to long-term economic closures and periodic stay-at-home mandates – decreased employment, increased unemployment, and put pressure on wages (ILO, 2020a). Moreover, the crisis has compounded existing inequalities in terms of income, job security, and working hours. Many of the most vulnerable – often deemed essential workers under legal mandates related to the pandemic response – were forced to decide between working to meet basic economic needs and risking exposure to COVID-19. On the other hand, the pandemic response saw a shift for many to working at home, accelerating a trend that is poorly understood in how it impacts workers over time (ILO, 2021a).

While the broad impacts of the COVID-19 crisis on workers are generally understood, policy makers have struggled with capturing the labour market impacts of the pandemic in a way that allows them to rapidly develop and implement a policy response that is effective and targeted to the needs of the most vulnerable. Accurately identifying labour market impacts, an essential first step in developing national employment policies, has been difficult both because of the unprecedented nature of the shock and the difficulty of fielding traditional face-to-face household labour force surveys during the pandemic. In many countries, labour force surveys were delayed during government shutdowns or not completed for multiple periods during the pandemic, leaving gaps in long-term trends data. Nearly half of countries had to suspend labour force survey data collection at some point during 2020 (ILO, 2021b). In other cases, the fielding of these surveys failed to capture the direst economic impacts because they were fielded before or after economic conditions were most acutely affecting those workers most in need.

This information deficit underlines the need for national statistics agencies and policy makers more broadly to invest in developing supplementary data sources that can help them estimate essential labour market indicators when labour force surveys cannot be fielded or when they do not capture the effects of economic shocks in a timely manner. These non-conventional data sources include administrative data, existing economic indicators, dedicated surveys, and Big Data. Moreover, policy makers, employment specialists, and social partners can use these resources to expand their understanding of labour market dynamics and better inform labour market policy decisions, whether in times of crisis or not. This brief explores the opportunities and challenges posed by non-conventional sources, offering policy makers a resource in how to improve employment diagnostics and resultant policy response.

II. Understanding the Value of Conventional Labour Force Surveys

Labour force surveys remain the primary source for data used by policy makers for tracking labour market activity within the economy and understanding outcomes for workers. Traditionally deployed using face-to-face interviews in households selected at random from a national survey frame, these surveys provide policy makers with a dependable source of unbiased, representative data, including quantitative indicators like labour force participation, employment, and unemployment and indicators of the quality of work like wages, access to social protection, and underemployment. The structure and sample size of the labour force survey allows for employment diagnostics at the national, regional, and local level, while allowing analysts to assess outcomes for youth, women, and groups identified as being vulnerable to poor labour market outcomes. Importantly, the structure of questions within labour force surveys allows for the construction of indicators in line with international statistical standards[2], allowing for a consistent comparison of labour market activities and outcomes over time and geography.

Despite the importance of the labour force survey, the instrument itself poses several challenges in terms of the timeliness of the data it provides. First, the survey is generally fielded at set intervals (annually, quarterly, or monthly) or continuously, depending on the country's capacity. This periodicity has value for regular analyses; however, it means that fielded surveys often miss the onset of economic shocks or temporary distortions in the labour market (such as the COVID-19 pandemic and temporary economic shutdowns implemented in early Spring 2020). Second, the survey is limited in total sample size and the number of questions that can be asked of respondents. This restricts the depth and granularity of labour market issues that can be addressed by the survey. Finally, the face-to-face nature of how labour force surveys generally are fielded means that their implementation is expensive and logistically complex. This also means that these surveys are vulnerable during crises like COVID-19, when mobility restrictions and restricted work hours limited the ability of national statistics agencies to run face-to-face interviews and process collected data.

III. Using Employment-Related Data from Existing Surveys and Administrative Data

In the absence of access to timely labour force survey data, most governments have alternative sources of labour-related data to which they can refer. These include administrative data such as tax records and applications for government services, as well as government surveys of businesses aimed at calculating leading economic indicators and informing broader economic policy. At the same time, there may be numerous conventional data collection efforts delivered by private survey firms, chambers of commerce, and academia. Together, in the context of the COVID-19 crisis, these sources provided policy makers in many countries with an important set of resources to apply to analyses aimed at filling in gaps in labour market survey data and capturing the labour market impacts of the crisis. Going forward, they provide an important means for policy makers and national statistics agencies to capture alternative perspectives on labour market activity and outcomes, while preparing for future crises when labour force survey data may

[2] See the guidance provided by the International Conference of Labour Statisticians (ICLS) - ILOSTAT

be unavailable. Key examples are reviewed below.

Employment and Unemployment Registry Data

Administrative data collected while processing government permits, applications, and service claims offer an important resource for policy makers seeking to understand labour market activity, particularly when they require frequent data. Here, employment registries and unemployment insurance claims are particularly important. In most (high-income) countries, businesses are required to register new hires and often dismissals with revenue and regulatory agencies, ensuring worker access to social protection and providing revenue agencies with relevant income tax information. Similarly, unemployment insurance claims document initial and subsequent claims for unemployment support for workers who have been laid off. Both sources provide evidence of key labour market indicators (employment, unemployment), while also providing information about a worker's skill level, occupational category, experience, job title, and salary, as well as relevant demographic data.

Depending on how data from digital files are compiled, registry data can be available weekly, monthly, or in real-time, allowing policy makers with a high-frequency source of data with which to assess rapid changes in employment and unemployment trends by region and sector, particularly when authorities are able to link data from both sources (OECD, 2020). This proved important for many countries during the early stages of the pandemic (Baek et al., 2020; Chen et al., 2020).

When using employment registry data or unemployment insurance claims as a proximate indicator for total employment and unemployment, policy makers should understand that these data are only relevant for formal sector firms complying with labour market regulations. They likely do not cover those working in the informal sector, seasonal workers, or the self-employed. As such, they likely do not allow policy makers to diagnose outcomes for the most vulnerable. Importantly, in many developing countries, the informal sector often makes up a large share of total employment. Moreover, since women are more likely to work in the informal sector than men, there are important gender implications for any dependence on data reflecting formal sector outcomes.

Employer and Enterprise Surveys

Governments and private firms routinely survey employers to identify trends in hiring, overall employment, and skills demands in the labour market. The World Bank also maintains a series of enterprise surveys through which it seeks to collect comparable data across multiple countries about firm business activities, expenditures, and employment (World Bank, 2021). These surveys may be industry specific or focus on particular geographies rather than entire nations and their periodicity depends on local resources and demand for economic intelligence. Often, they are delivered less frequently than labour force surveys; however, they may provide timely data in the context of a crisis or a measure of labour market activity between labour force surveys.

Given their focus on informing government policy on tracking changes in labour market demand and emerging skills gaps to inform skills training policy, they provide policy makers with an alternative source of information for labour market demand by location, skills, and education, as well as overall growth of employment. Importantly, when one is interpreting results and comparing them with traditional indicators of employment, one must realize that these surveys are generally biased towards formal sector workers and non-farm employees. Like administrative data described above, these surveys often exclude the self-employed, seasonal, and informal workers.

Purchasing Managers Index

The purchasing managers index (PMI) is a widely used leading indicator of the economy. It is available for 44 countries across the globe, the majority of which are produced by the firm IHS Markit[3]. Built on a monthly survey of purchasing managers and supply chain managers in private firms, the index is constructed around a series of sub-indices measuring changes in firm production, new orders, supplier deliveries, inventory, and employment. Building the index on decisions that firms are making on the basis of costs, supply chain status, and expectations of growth ensures that it is usually an early signal of the overall health of the economy (Kuepper, 2022). Moreover, given that the survey is monthly, it provides policy makers with a relatively high-frequency indicator.

For policy makers interested in assessing labour market activity, the PMI tracks with employment growth. It has some predictive capacity, as employment growth tends to lag economic growth (Williamson, 2021). As such, the PMIs employment sub-index may be more useful to policy makers as an input for assessing employment and wage growth in the absence of the labour force survey. During COVID-19, the PMI provided policy makers in covered countries with an alternative means of estimating the impact of the crisis on employment and wages, as well as general economic growth, both at a national level and by international organizations like the ILO (see below). It should be emphasized that the PMI is a reflection of the formal market and, for many countries, limits its coverage to a specific sector (e.g., manufacturing, services, construction), which may limit its representativeness of overall economic trends depending on the structure of the economy as a whole.

IV. Big Data and Innovative Sources for Labour Market Information

In addition to administrative data and more traditional survey-based data sources and indicators, Big Data provides an alternative source of information that can supplement labour force survey-based indicators of labour activity for policy makers. The term Big Data refers broadly to a set of new data sources and analytical techniques made possible by digital technology with which data scientists analyse and extract information from volumes of data that are too large or complex for traditional data analysis software (Mezzanzanica and Mercorio, 2019). In addition to volume, Big Data generally is characterized by the use of data that are constantly being produced or changing in real-time.

Leveraging the power of artificial intelligence and machine learning, Big Data analysis allows

[3] IHS Markit produces PMIs for 44 countries. There is also a PMI for the United States produced by the Institute for Supply Management, while the Singapore Institute of Purchasing and Materials Management (SIPMM) produces a PMI for Singapore. The China Federation of Logistics and Purchasing (CFLP) produces China's official PMI.

analysts to identify signals or correlations within large datasets that would be interpreted as noise in traditional statistical analysis, in essence mining for hidden patterns and relationships that enable the predictive capabilities of these weak signals. Importantly, given the frequency and volume of data produced through these digital sources, Big Data analyses can provide policy makers with a 'real-time' analysis and response capability while allowing them to explore economic issues with a level of specificity and granularity not allowed by traditional survey data. It is a capability that greatly informed and improved the policy responses in the wake of the onset of the COVID-19 pandemic in countries that had invested in a Big Data capacity before the crisis (Nitschke et al., 2021).

For policy makers eager to work with Big Data, it is important to recognize that its application remains experimental, and its potential is largely untapped when it comes to broader issues of economic development. An exception may be central banks' regular use of Big Data as part of the analytical process in shaping monetary policy (Doerr et al., 2021). Its application to other realms of policy has been restricted, in part because of the newness of its approach (in contrast with traditional statistical analysis). In this regard, data drawn from labour force surveys provides policy makers with an accurate, representative assessment on which to base important policy decisions, whereas in most cases evidence drawn from Big Data analysis may capture only certain aspects of the labour market, often with uncertain bias. Moreover, most policy makers have become accustomed to working with key indicators derived from labour force surveys as a basis for labour market diagnostics and policy response and remain wary of effecting policy based on more experimental approaches.

It is also relevant for policy makers to understand that analysts working with Big Data generally employ techniques that mine available data for signals rather than building on economic theory to test empirical support for hypothesized causal relationships in the way that traditional statistical analyses do (Mezzanzanica and Mercorio, 2019)[4]. In short, the process is inductive rather than deductive. Given the exploratory nature of Big Data analytical techniques, there are few off-the-shelf indicators available to policy makers resulting from Big Data to use directly in labour market analysis, although a limited number of aggregated indicators have been developed (see below). Rather, the growing body of digital data available to policy makers offers new opportunities for exploratory data analysis that is likely to offer new insights and inform labour market policy decisions in a variety of ways. The following explores core areas in which data scientists have been using Big Data to support labour policy and programme development.

Mobile Phone Location Data

By design, mobile telephones generate large volumes of usage data, particularly on the location of the user. Mobile network operators (MNOs) collect data which reports which mobile

[4] To some degree, there is a potential overlap here in terms of the types of data being used. For example, high-frequency administrative data may be considered Big Data given the volume of data included and means by which it can be analysed, particularly in conjunction with other sources of data. We differentiate here between more conventional sources of data used by governments in macroeconomic analysis and non-conventional, new sources generated as part of the digital economy.

phone tower a user's telephone is connected to when users make calls or text messages, as well as signals the provider occasionally sends to the user's mobile phone to identify whether it is receiving a network signal or not. By triangulating signals received by various mobile phone towers, these data can be used to identify the user's location. The accuracy of this location data depends on the density of mobile phone towers and how the system is configured, with data from rural areas and smaller towns generally proving less accurate that data from urbanized areas, where data analysts can often identify whether a user is in a specific building or a specific part of a building. In addition to MNO-collected location data, mobile telephones often create data trails through the digital signals they send to and receive from global position system (GPS) satellites, Wi-Fi networks and Bluetooth devices. Anonymized data drawn from such resources are now widely available through various data brokers.

Google and Apple provide a similar resource for location data. Various applications on the Android and iOS operating systems track user location and mobility as a means of providing enhanced services to the user while collecting data that can be sold to advertisers to better target advertisements. During the COVID-19 pandemic, both Google and Apple have made anonymized, aggregated mobility data available to the public for most countries and regions within countries. The data sources were intended to inform COVID-19 policy responses, and they have helped policy makers assess the impact of public health measures as a measure of social distancing (Chen et al., 2020). At the same time, this aggregated data has allowed analysts, like those at the ILO, to assess labour market activity more accurately during COVID-19.

▶ Box 1: South Korea's widespread use of mobility data

The impact of using mobility data can be seen in a review of South Korea's use of mobile technology in contact tracing and quarantine efforts during the COVID-19 pandemic (Dyer, 2021). With the onset of the pandemic, South Korean public health authorities began working with mobile application creators to design a set of mobile phone applications that would allow them to identify and track individuals who had tested positive for COVID-19, to trace where they had been prior to their diagnosis, to identify individuals who had come in contact with them during their infectious period, and to help enforce quarantine. At the same time, data collected from the application allowed data scientists to track the mobility of users, including when they left for work and how long they were at work. In this regard, it provided a unique source of data on how the pandemic impacted employment and working hours in South Korea.

Big Data offers policy makers a unique resource that can help them understand patterns of behaviour among the population, including providing insights about their labour market activity (Putra and Arina, 2020). Depending on the quality of data available in any particular country context, these data may provide information about employment status, work location, working time and commuting, and broader issues of mobility. Sudden changes in patterns of behaviour like telephone usage or changes in routine movements may be tied to changes in employment (Toole et al., 2015). The data can also be used to analyse user engagement with government services as they track movement into government offices. For example, Morowoki (2020) uses GPS data from mobile telephones to estimate unemployment trends in Japan based on changes in foot traffic at unemployment claims offices. Those countries that are behind the curve in terms of developing digital citizen services might be able to use mobile signals (or localized Wi-Fi signals) to measure changes in the uptake of citizen services like unemployment insurance claims, job centres, or active labour market programmes (ALMPs). This, in turn, would help governments track rapid changes in employment, unemployment and other labour market information.

There are important caveats of which policy makers should be aware when it comes to the application of mobile location data to labour markets. First, while mobile telephone usage has expanded rapidly in recent years, it is not universal and those who do not use mobile telephones are likely to be different in some underlying characteristics than those who do. It is particularly likely that data depending on mobile telephone usage do not necessarily capture the economically vulnerable. In many countries, as well, women do not have the usage rates that men do, suggesting further bias in terms of gender. Finally, one should consider that families often share telephones or may have various telephones under a common subscription; depending on local regulations, this may affect the quality of collected data.

Internet Search and Social Media

For many, internet search engines and social media have become primary sources for information discovery and sharing their personal news and opinions with broad networks of family and friends. Together, in turn, search engines and social media like Facebook and Instagram provide an increasingly important source of data on a variety of labour market issues, as users search for labour market information, post about job changes, or provide information about their own experience as economic actors. Here, LinkedIn is unique source for labour market data because, as a social networking site, it focuses on professional networking and has increasingly positioned itself as a job portal.

Data scientists are exploring the volume of data provided within the collective body of our internet searches and social media posts, experimenting with means of accurately measuring labour market activity comparable with traditional indicators or identifying early signals of labour market activity within them.

▶ **Box 2: Research innovation examples working with internet search data**

Kern et al. (2019) are using social media to improve matching of people to ideal jobs. Baker and Fradkin (2017) have used Google search data to construct a job search index and assess whether access to unemployment insurance limits job search activity. Other researchers are using social media data to estimate the income of individual users based on the content and frequency of their posts (Matz et al., 2019). More broadly, researchers have developed programmes that crawl through social media to identify topics related to employment and unemployment to calculate trends in employment, unemployment, and job openings. (Antenucci et al., 2014; Ryu, 2018).

A readily available resource for policy makers in this regard is Google Trends, which offers data on the frequency of search queries and changes over time. Data are available at a national, regional, and municipal level for most countries. This enables policy makers to analyse a broad array of employment issues from interest in particular sectors to unemployment concerns to salary inquiries. An increase in searches for information about the labour market may be correlated with real changes in the labour market. For example, during the COVID-19 pandemic, Brave, Butters and Fogarty (2020) showed a strong correlation between Google Trends data on unemployment-related searches with initial unemployment insurance claims in the United States.

Online Job Portals

Online job portals provide an increasingly important means of matching employers with job seekers, given their ability to bring these two parties together at scale and to allow both to focus or refine their searches. In doing so, these portals collect digital job opening descriptions for millions of jobs and, in turn, collect digital resumes for millions of applicants. Working with machine learning techniques, analysts can train artificial intelligences to crawl through these digital documents to identify data points related to demographics, skills, education, job titles, job activities, work history, and salaries. In turn, analysts can use this data to understand skills demand in the labour market, the skills of job seekers, and evidence related to matching activities on both sides (Horton and Tambe, 2015). An increasing body of research is focused on using Big Data in this context to enhance the ability of ALMPs and jobs programmes to match employers and jobseekers and upskill workers (Nitschke et al., 2019).

These online job portals are now pervasive in developed countries, where they are commonly used as an intermediary in job matching for employers and jobseekers. Examples include, but are not limited to, Monster, Indeed, Ziprecruiter and LinkedIn. There are also a host of similar sites that are country- or -region specific, some run by private firms and others run either by government or development organizations. Across all of these platforms, there is a general bias in the types of employers and jobseekers that self-select into these online job portals. Both types of users tend towards technology, high-skilled jobs, and white-collar professions more generally. Even when coverage of wage employment is fairly representative, these platforms tend to underrepresent lower-skilled jobs, informal work, and seasonal work, which still depend largely on informal networks. In developing countries, this bias – and limitations on coverage of the labour market more generally – may be more pervasive. Another important bias is that usage of such portals requires a significant literacy level, meaning that usage in countries with low literacy rates is significantly biased. This bias may be more skewed when gender and age are taken into account.

There have been significant investments in creating online job portals for workers in developed countries that are beginning to shift this and, in turn, the ability of labour market specialists to use them to diagnose labour market issues therein.

In addition to these standard online job portals, there are also a growing number of platforms that manage crowd-sourced, short-term jobs and portals that support self-employed service

▶ Box 3: Analytic application of job board data in Asia

The Bong Pheak Jobs Portal in Cambodia targets low-skilled, informal workers, matching them with jobs using data science across a variety of sectors (Winrock International, 2019). Notably, the Asian Development Bank used job postings on local online job portals in Bangladesh and Sri Lanka to capture the impact of COVID-19 on local labour markets (Hayashi and Matsuda, 2020). The World Bank has been exploring the use of Big Data in aligning market skill demand with the skills of job seekers in Pakistan using an online job portal, work which they updated to understand labour outcomes in the wake of COVID-19 (Matsuda, Ahmed, and Nomura, 2019; Tas et al., 2021).

providers by matching them with short-term opportunities or gigs. The growing gig economy – from ride-sharing through applications like Uber to food delivery (Grubhub, Postmates, etc.) to applications that manage micro-jobs like Amazon Mechanical Turk – offers unique insights to policy makers, particularly as these crowd-sourcing applications put more emphasis on the data they create as part of their business model. To date, Uber and other ride-sharing companies have focused more on the provision of traffic and road condition data to policy makers. However, these applications collect a significant amount of relevant labour market data from participating workers, data that can be used to estimate trends in self-employment, part-time work and supplementary work, and wage expectations of workers (Abraham et al., 2018; Kässi and Lehdonvirta, 2018). While most research drawing on these types of applications have focused on income and job quality for self-employed gig workers, they may increasing provide data resources that meet the broader data needs of policy makers seeking to better understand wage trends and the balance between employment and self-employment among informal sector workers.

V. Practical Policy Applications for Non-Conventional Labour Market Data Sources

In seeking to identify key indicators that accurately reflect labour market trends and outcomes that are representative of the full labour force, policy makers have little alternative to conventional labour force surveys. Even as policy makers struggle with understanding labour market outcomes during a time of crisis, the labour force survey offers a source of quality data that cannot be replaced by existing data alternatives or innovative Big Data analysis. However, these non-conventional provide an increasingly important source of supplemental data analysis that can be used to estimate current labour market outcomes when labour market surveys are not available or to enhance our collective understanding of labour market issues more broadly.

Employment Diagnostics

The non-conventional sources of labour market data described above can provide data-driven insights into a range of labour market issues, from employment and unemployment to wages and benefits, as well as perspectives on worker well-being and early signals about a desire to change jobs. From this panoply of data resources, statisticians/analysts can construct diagnostic indicators that help policy makers fill in their understanding of the labour market impacts of economic shocks and inform broader labour market policy. These include indicators that closely track traditional labour market indicators derived from the labour force survey. While recognizing their inherent biases (i.e., formal sector vs. informal sector, urban vs. rural, gender, etc.), policy makers may find that such indicators provide useful, high-frequency measures of change within the labour market. Importantly, the validity of any indicator as a labour market signal depends on its being tested against the historical record provided by labour force survey data to see how well results track. This underlines the importance of developing alternative labour market measures before any economic shock hits. Given their biases, policy makers may find that these alternative data sources are more informative when deployed as one of a set of indicators included in more complex economic models.

Enhanced Modelling Capacity for Estimates and Projections

Given delays in the production of most traditional sources of macroeconomic data, and the periodic nature of releases of this data, economists have long sought the capacity to accurately estimate current economic conditions and make projections for the near future. This is particularly the case for key labour market indicators like employment growth and unemployment rates. Traditional approaches to economic modelling have been heuristic, depending on past observation and periodic data to calculate probable outcomes. However, the proliferation of non-conventional data, particularly high-frequency data in large volumes, coupled with advancements in analytical techniques that allow for concurrent analysis of low-frequency and high-frequency data, now allows analysts to construct models that generate rapid, highly accurate estimates and projections of current and near-term economic outcomes.

Improving Employment Policies and Interventions

Beyond providing an expanded diagnostic capability, non-conventional data sources – particularly Big Data – have an important role in helping to improve the targeting, efficiency, and impact of policies and programmes directed at improving labour market outcomes (York and Bamberger, 2020). This is particularly the case in terms of skills development, training and worker upskilling programmes, where policy makers and social partners are increasingly using Big Data drawn from job portals to improve linkages between labour market demands and the skills that government programmes are working to provide new entrants and established workers.

▶ **Box 4: The emergence of the nowcasting approach**

Rooted in analytical techniques used in weather forecasting, this approach to modelling – termed nowcasting – is being used increasingly by economic policy makers (Bańbura, Giannone, and Reichlin 2010; Bok et al. 2017). This is particularly the case for central banks, which have adopted nowcasting to inform macroeconomic and monetary policy. For example, the Bank of Finland has developed a nowcasting model that uses 50 variables with various frequencies of production to create real-time estimates of gross domestic product (Itkonen and Juvonen, 2017). New Zealand and Australia have applied nowcasting to labour market analysis to provide policy makers with more rapid assessments of labour market outcomes so as to inform and better target government programmes designed as part of the post-COVID recovery (Australian Government, 2020; Karagedikli and Özbilgin, 2019). Throughout the COVID-19 pandemic, the ILO has used nowcasting to provide more accurate estimates of the labour market impacts of the crisis on an international scale.

The ILO's nowcasting model builds on existing labour force survey data produced by national statistics agencies. Where data are available, ILO analysts build onto this backbone with higher frequency data including job registry and unemployment registry data, purchasing managers index data, national accounts data, and data from consumer and business confidence surveys (ILO, 2021c). To fine tune the model's ability to capture some of the more acute impacts on workers of the COVID-19 pandemic, particularly in countries for which the above data sources are lacking, the model takes into the Oxford Stringency Index, which captures the intensity of government-mandated closures and economic lockdowns as a response to the spread of COVID-19, and daily Google Mobility Reports data derived from mobile phone location data. The inclusion of this daily data and data on government-mandated closures together provides the ILO with an effective control for capturing restrictions on working hours and, in turn, adjustments for estimates of full-time equivalent employment.

Many national programmes cannot be evaluated through experimental approaches because it is infeasible to restrict access to policy and programmes. Applications of Big Data analytical approaches offer an alternative approach, providing evaluators with the capacity to evaluate programmes and policy at scale and using the volume of data available through Big Data analytics to control for comparability between participants and non-participants. To this end, for example, the EU and OECD are working to link administrative sources of data on country-level employment and unemployment to allow for broad-based evaluation of the impact of skills training and employment programmes on unemployment status over time (OECD, 2020).

▶ **Box 5: Using big data to anticipate skills demands: Examples from Australia and Singapore**

Several countries have begun using Big Data, particularly data from online job portals, to inform and improve government strategies and programmes like skills training. For example, Australia has begun using Big Data to supplement existing skills data in helping match workers with opportunities and inform young Australian about emerging opportunities (Australia National Skills Commission, 2022). Singapore is using Big Data drawn from LinkedIn to identify emerging skills in its efforts to implement an Artificial Intelligence Strategy, and it uses evidence drawn from Big Data to improve the career counselling provided to Singaporean youth (Singapore Government, 2019).

VI. Conclusion: Building an Enhanced Capacity for Employment Diagnostics

The COVID-19 pandemic proved the long-term need for strong labour market information systems, which can include enhanced labour market indicators and alternative sources of data.. There are numerous examples of how policy makers in countries and international organizations have taken advantage of early progress in adapting alternative economic indicators and Big Data analytical techniques, positioning themselves to rapidly evaluate and respond to resultant labour market outcomes. For others, it underlines the need to invest in this additional capacity, creating indicators and avenues for analysis that support established data sources and analytical practices.

Towards this end, it should be emphasized that building an enhanced capacity for labour market analysis using non-conventional data sources can take a significant investment, both in terms of financial resources and time. While there are areas in which existing data resources like purchasing managers indices or aggregated mobile phone mobility data can be leveraged to enhance existing labour market analysis and modelling efforts, efforts to build additional capacity will generally require significant time in designing alternative indicators, building the instruments for collecting and analysing data, and testing results against past labour market data. Big Data, in particular, may impose high upfront costs related to contracting data analysts, data storage, and the design of programmes used in the analysis (Nitschke et al., 2021).

Still, rather than waiting for a new economic shock or public health crisis, the time to begin putting in place the foundations for this enhanced capacity is now. Towards this end, policy makers should reach out for assistance. The ILO and other international organizations are keen to support countries in building their statistical and data analysis capabilities, building on their own experience in this area. At the same time, university-based researchers and graduate students provide an oft-overlooked resource for policy makers seeking to experiment with innovative data analytics.

Likewise, there are several private firms – from Google to data analytics firms like EBG and SAS – that can leverage years of expertise to provide efficient technical support for policy makers in this area.

Finally, it is important to emphasize the need for clear ethical guidelines related to privacy and the use of data collected on citizens, often without their realization or expressed consent. While traditional surveys are governed by a code of research ethics that ensures that participation is voluntary, that one's identity is protected, and that data will not be used to harm the participant, such approaches generally do not apply to data collected from digital technologies. This should be understood in a broader context wherein data collected from digital sources can directly or indirectly reveal information about people that can expose them to personal risks or embarrassment. Here, South Korea's experience with contact tracing during the pandemic is a lesson. Putting in place protections for privacy are an essential first step in developing a non-conventional data collection and analysis strategy.

References

Abraham, Katharine, John Haltiwanger, Kristin Sandusky, and James Spletzer. 2018. “Measuring the Gig Economy: Current Knowledge and Open Issues,” NBER Working Paper No. 24950. Cambridge, Massachusetts: National Bureau of Economic Research. August.

Australia Government. 2020. A Snapshot in Time: The Australian Labour Market and COVID-19. Canberra: Australian Government and National Skills Commission. July 1.

Australia National Skills Commission. 2022. “Jobs and Education Data Infrastructure.” Available online at: https://www.nationalskillscommission.gov.au/topics/jedi

Automatic Data Processing. 2022. “ADP National Employment Report” Roseland, New Jersey: Automatic Data Process Research Institute. Available online at: https://adpemploymentreport.com/

Baek, ChaeWon, Peter McCrory, Todd Messer, and Preston Mui. 2020. “Unemployment Effects of Stay-at-Home Orders: Evidence from High-Frequency Claims Data,” IRLE Working Paper No. 101-20. http://irle.berkeley.edu/files/2020/07/Unemployment-Effects-of-Stay-at-Home-Orders.pdf

Baker, Scott and Andrey Fradkin. 2017. “The Impact of Unemployment Insurance on Job Search: Evidence from Google Search Data,” The Review of Economics and Statistics, Vol. 99, No. 5.

Bok, Brandyn, Daniele Caratelli, Domenico Giannone, Argia Sbordone, and Andrea Tambalotti. 2017. “Macroeconomic Nowcasting and Forecasting with Big Data,” Federal Reserve Bank of New York Staff Report No. 830. New York: Federal Reserve Bank of New York. November.

Bańbura, Marta, Domenico Giannone, and Lucrezia Reichlin. 2010. “Nowcasting,” ECARES working paper 2010-021. Brussels: ECARES.

Brave, Scott, Andrew Butters, and Michael Fogarty. 2020. “A Closer Look at the Correlation between Google Trends and Initial Unemployment Insurance Claims,” Chicago Fed Insights. Chicago: Federal Reserve Bank of Chicago, July 17. Available online at: https://www.chicagofed.org/publications/blogs/chicago-fed-insights/2020/closer-look-google-trends-unemployment#:~:text=In%20doing%20so%2C%20we%20find,data%20seen%20in%20figure%201

Chen, Sophia, Deniz Igan, Nicola Pierri, and Andrea Presbitero. 2020. “Tracking the Employment Impact of COVID-19 and Mitigation Policies in Europe and the United States,” IMF Working Paper No. WP/20/125. Washington, DC: International Monetary Fund.

Dyer, Paul. 2021. “Policy and institutional responses to COVID-19: South Korea,” Middle East and North Africa (MENA) COVID-19 Response Project Report. Doha: Brookings Doha Center. June 15.

Hayashi, Ryotaro and Norihiko Matsuda. 2020. “COVID-19 Impact on Job Postings: Real-Time Assessment Using Bangladesh and Sri Lanka Online Job Portals,” ADB Brief No. 135. Manila: Asian Development Bank. May.

Horton, John and Prasanna Tambe. 2015. “Labor Economists Get Their Microscope: Big Data and Labor Market Analysis,” Big Data. Vol. 3, No. 3 (September).

International Labour Organisation. 2021a. "Teleworking arrangements during the COVID 19 crisis and beyond," Paper prepared for the 2nd Employment Working Group Meeting under the 2021 Italian Presidency of the G20. Geneva: ILO, April.

International Labour Organisation. 2021b. "Keeping labour data flowing during the COVID-19 pandemic," ILOSTAT brief. Geneva: ILO, September 30. Available online at: https://ilostat.ilo.org/keeping-labour-data-flowing-during-the-covid-19-pandemic/

International Labour Organisation. 2021c. "COVID-19 and the world of work," ILO Monitor, 8th edition. Geneva: ILO. October 27.

International Labour Organisation. 2020a. "COVID-19 and the world of work: Impact and policy responses," ILO Monitor, first edition. Geneva: ILO, March 18.

Itkonen, Juha and Petteri Juvonen. 2017. "Nowcasting the Finnish economy with a large Bayesian vector autoregressive model," Bank of Finland Economics Review. Available online at: https://helda.helsinki.fi/bof/handle/123456789/14979

Karagedikli, Özer and Murat Özbilgin. 2019. "Mixed in New Zealand: Nowcasting Labour Markets with MIDAS," Reserve Bank of New Zealand Analytical Note Series No. AN2019/04. Wellington: Reserve Bank of New Zealand.

Kässi, Otto and Vili Lehdonvirta. 2018. "Online Labour Index: Measuring the Online Gig Economy for Policy and Research," Technological Forecasting and Social Change, Vol. 137 (C).

Kern, Margaret, Paul McCarthy, Deepanjan Chakrabarty, and Marian-Andrei Rizoiu. 2019. "Social media-predicted personality traits and values can help match people to their ideal jobs," PNAS. December 11. Available online at: https://www.pnas.org/doi/10.1073/pnas.1917942116

Kuepper, Justin. 2022. "What Is the Purchasing Managers Index (PMI)?" The Balance, February 24. Available online at: https://www.thebalance.com/what-is-the-purchasing-managers-index-pmi-1978996

Kurmann, André, Etienne Lalé and Lien Ta. 2020. "The Impact of COVID-19 on Small Business Employment and Hours: Real-Time Estimates with Homebase Data," Working Paper No. 2020-09. Montreal: Department of Economic Sciences, University of Quebec. August 5.

Matsuda, Norihiko, Tutan Ahmed, and Shinsaku Nomura. 2019. "Labor Market Analysis Using Big Data : The Case of a Pakistani Online Job Portal," Policy Research Working Paper No. 9063. Washington, DC: World Bank.

Matz, Sandra, Jochen Menges, David Stillwell, and Andrew Schwartz. "Predicting individual-level income from Facebook profiles," Plos One, Vol. 14, No. 3.

Mezzanzanica, Mario and Fabio Mercorio. 2019. Big Data for Labour Market Intelligence: An Introductory Guide. Turin, Italy: European Training Foundation.

Nitschke, Julia, Layla O'Kane, Bledi Taska, Nyerere Hodge, Emmanuel San Andres, and Andre Wirjo. 2021. "Big Data for the Labor Market: Sources, Uses, and Opportunities," Issues Paper No. 13. APEC Policy Support Unit. Singapore: Asia-Pacific Economic Cooperation. December.

Organisation for Economic Co-operation and Development (OECD). 2020. Impact evaluation of labour market policies through the use of linked administrative data (VS/2019/0261- Joint OECD-EU analysis of labour market policies. Paris: OECD, December.

Putra, Randra and Silvia Arini. 2020. "Measuring the Economics of a Pandemic: How People Mobility depict Economics? An Evidence of People's Mobility Data towards Economic Activities," https://www.imf.org/-/media/Files/Conferences/2020/8th-stats-forum/paper-rendra-putra-and-silvia-arini.ashx

Ryu, Pum-Mo. 2018. "Predicting the Unemployment Rate Using Social Media Analysis," Journal of Information Processing Systems, Volume 14, No 4.

Singapore Government. 2019. National Artificial Intelligence Strategy. Singapore: Singapore Government. Available online at: https://www.smartnation.gov.sg/files/publications/national-ai-strategy.pdf

Sebastian Doerr, Leonardo Gambacorta and José María Serena Garralda. 2021. "Big data and machine learning in central banking," BIS Working Paper No. 930. Basel: Bank for International Settlements, March 4.

Tas, Emcet, Tanima Ahmed, Norihiko Matsuda, and Shinsaku Nomura. 2021. "Impacts of COVID-19 on Labor Markets and Household Well-Being in Pakistan: Evidence from an Online Job Platform." Washington, DC. World Bank.

Toole, Jameson, Yu-Ru Lin, Erich Muehlegger, Dnaiel Shoag, Marta Gonzalez, and David Lazer. 2015. "Tracking Unemployment Shocks Using Mobile Phone Data." Journal of the Royal Society. June 6. Available online at: https://royalsocietypublishing.org/doi/10.1098/rsif.2015.0185.

Williamson, Chris. 2021. "Understanding … headline PMI vs. subindices: how signals can be lost by focusing exclusively on the headline PMI," PMI Commentary. London: IHS Markit. September 16. Available online at: https://ihsmarkit.com/research-analysis/understanding--headline-pmi-vs.-subindices-how-signals-can-be-lost-by-focusing-exclusively-on-the-headline-pmi.html

Winrock International. 2019. "Bong Pheak Learning Action Brief," North Little Rock, Arkansas: Winrock International. Available online at: https://winrock.org/document/bongpheak-learning-action-brief/

World Bank. 2021. Enterprise Surveys. Microdata Library. Washington, DC: World Bank. Available online at: https://microdata.worldbank.org/index.php/catalog/enterprise_surveys/.21

Contact details

International Labour Organization
Route des Morillons 4
CH-1211 Geneva 22
Switzerland

T: +41 22 799 7239
E: emplab@ilo.org