CENTRE FOR SOCIAL DATA ANALYTICS

AUT

MINISTRY OF SOCIAL DEVELOPMENT
TE MANATŪ WHAKAHIATO ORA

Nowcasting the current rate of unemployment using administrative data

March 2021

Authors

David Rea and Tim Maloney

Acknowledgements

We are grateful for comments and feedback from Robert Templeton, Christopher Ball, Amapola Generosa, Richard Sullivan, Yifei Lyu, Andrew Neal and Ryan Greenaway-McGrevy.

Disclaimer

The views expressed in this paper are those of the authors and not the Ministry of Social Development or Auckland University of Technology.

Published

March 2021

Centre for Social Data Analytics
Auckland University of Technology
Private Bag 92006
Victoria Street West
Auckland 1142
www.csda.aut.ac.nz

ISBN

978-1-927184-83-7 (PDF)

Contents

Table of figures

Table of tables

Introduction

Like other countries, New Zealand is experiencing an economic downturn associated with the COVID-19 pandemic. The number of people who are jobless has increased and will likely remain high by recent historical standards.

The speed and potential scale of the economic downturn in 2020 highlighted the importance of timely information about the labour market during the onset of a recession. Over this coming year there will also be considerably uncertainty about the timing and extent of any improvement in labour market outcomes that will also demonstrate the value of timely information.

The official unemployment rate is an important indicator of labour market and wider economic conditions. It is based on data collected by the Household Labour Force Survey (HLFS) over a three-month period. Because it takes time to collect, analyse and publish the results of the survey there is a natural delay before this official statistic can be released.

This note describes our nowcasting model that aims to provides more timely information about measured unemployment. At monthly intervals during each quarter the UR-NOW model predicts the headline unemployment rate prior to the release of the official HLFS results. The model uses high frequency administrative data on benefit receipt and filled jobs.

Background on the official unemployment rate

Across OECD countries, the unemployment rate is an important indicator of basic economic conditions and social outcomes.

Measuring unemployment is important because joblessness has a major adverse impact on people. Joblessness involves not just a loss of earnings and lower living standards but is also associated with a decline in life satisfaction and mental distress (Winkelmann and Winkelmann, 1998; Krueger and Mueller, 2012).

Measuring unemployment is also important from a macroeconomic perspective as it provides information about the extent of underutilisation of labour in the economy. The unemployment rate provides a key indicator of overall economic performance and the ability of the economy to deliver job opportunities.

The approach to the measurement of surveyed unemployment is standardised across countries through International Labour Organisation resolutions of the International Conference of Labour Statisticians. The current measurement framework can be traced back to pioneering analysis of the Works Progress Administration and the US Census Bureau in the 1930s (Card, 2011).

Consistent with the international standards, to be categorised as unemployed in New Zealand, a person must:

- be aged 15 years or over and usually resident in New Zealand (and not institutionalised)
- not be in a paid job or working unpaid in a family business

- be available to start work if offered a job
- have been actively seeking work in the previous four weeks, or due to start a new job in the next four weeks.

Actively seeking work means a person is applying for jobs or contacting employers and is not merely looking at posted job vacancies on the internet or other locations.

The unemployment rate is the proportion of people in the labour force who are unemployed (rather than the total population which is sometimes called the unemployment-population ratio). The 'headline' unemployment rate is also seasonally adjusted.

There are other important measures of joblessness as well as surveyed unemployment. These include available potential jobseekers (people wanting and available for work, but not actively searching) unavailable jobseekers (people actively searching for work but not able to start immediately) and the underemployed (individuals who are working part-time, and who are available and would like to work more hours). A more comprehensive indicator is the underutilisation rate which uses these wider measures of joblessness with the extended labour force as the denominator.

Administrative data on benefit receipt provides an alternative measure of joblessness for people in low income families. More details on the nature and differences between these measures are set out in Annex 1.

The unemployment rate is derived from the quarterly HLFS which has a sample size of approximately 15,000 households and includes roughly 30,000 adults. The target group for the survey is the resident non-institutionalised population who are 15 years and older. The survey has a rotating panel structure over eight quarters, and interviews are conducted on a rolling basis throughout the three months of a given quarter. The unemployment rate and other labour market measures are designed to represent an average outcome over the three months of each quarter (Statistics New Zealand, 2017)

The survey estimates have a margin of error due to sampling variability. For the seasonally unadjusted unemployment rate the 95% confidence interval is at least +/- 0.3 percentage points.

Improving the timeliness of information about the unemployment rate

The unemployment rate is a factor in decision-making in a range of areas including:

- monetary and fiscal policy
- specific budgeting and expenditure decisions of government
- investment decisions by firms, community organisations and government agencies
- decisions by individuals related to education, training, migration and job search.

Decision-makers in these different areas have information about the unemployment rate that is always slightly out-of-date due to the official series being published five weeks after the end of each quarter.

Most of the time this slight information lag is inconsequential because economic conditions are slow to change.

However, there are times when delays in updating labour market information is problematic. As the current downturn has shown, when the environment is changing rapidly, and in ways that are difficult to forecast, being able to accurately monitor real-time labour market conditions is immensely beneficial.

Information about the state of the labour market is particularly critical at the onset of a recession when decisions about the timing and scale of potential monetary and fiscal stimulus measures are required (Boushey et al., 2019). This is particularly true given the well-known impact lags for these policy interventions. Similarly, real-time information about a potential recovery is also important when decisions need to be taken about scaling back fiscal or monetary stimulus measures.

Many other OECD countries including the UK, USA and Australia run labour force surveys that provide monthly estimates of unemployment.

An alternative approach, which we explore in this paper, is to use higher frequency administrative data to create a monthly 'nowcast' of the unemployment rate. Importantly, nowcasting supplements rather than replaces the existing survey-based method of measuring unemployment. With nowcasting there is an ongoing need for survey data as the prediction model requires continued estimation and validation against survey results.

Nowcasting aims to improve the timeliness of information without the costs of additional data collection. At the same time, it also provides an independent check on unusual survey results from the HLFS. The trade-off is that nowcast predictions have a wider forecast confidence interval than estimates from the survey.

Figure 1 provides an example of the timeline for the collection and release of data for the June and September 2020 quarters of the HLFS. As can be seen, the unemployment rate for the June quarter (which spans the months of April through June) was published on the 5 August. This was 17 weeks after the start and 6 weeks after the end of this quarter.

Figure 1: Timeline for collection and publishing HLFS data

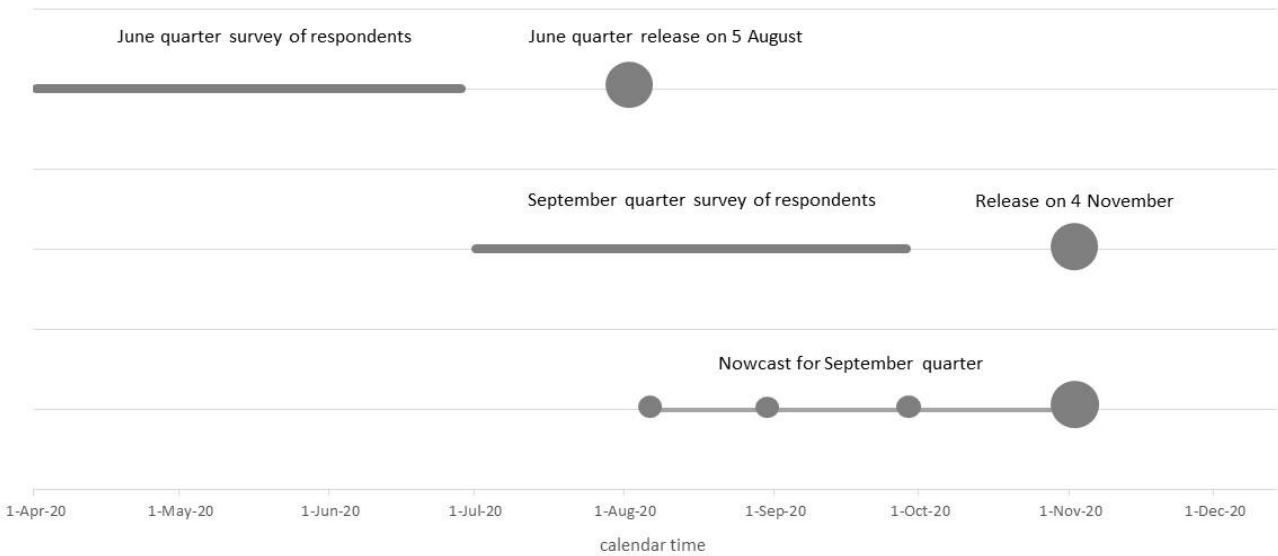


Figure 1 also shows how a monthly 'nowcast' of the unemployment rate for the September 2020 quarter could provide earlier information. Our nowcast model is configured to estimate the unemployment rate just after the end of each month of the quarter. It may be feasible, however, to eventually produce nowcast estimates of the unemployment rate on a weekly or even daily basis throughout a quarter.

Our paper builds on an extensive international literature on this topic (eg see Choi and Varian, 2009; Askitas and Zimmermann, 2009), as well as two important New Zealand papers that have investigated nowcasting the unemployment rate.

Karagedikli and Özbilgin (2019) demonstrate the ability of a mixed-frequency data sampling approach (MIDAS) to nowcast the unemployment rate and other labour market indicators. The paper develops a prototypical example that combines sub-models using monthly data on dwelling permits, motor vehicle registrations, international migration and the ANZ business confidence survey. Individual and combined models are shown to improve forecast accuracy relative to a simple first-order autoregressive (AR1) benchmark.

Ball et al. (2020) investigate if Inland Revenue data in the Integrated Data Infrastructure (IDI) could be used to nowcast the unemployment rate. Individual transitions between different sources of income each quarter are constructed from the data, and machine learning models are used for prediction. Out-of-sample validation shows that combining the output of the nowcast models is more accurate than a simple AR1 benchmark. The paper creates a strong case for improving the timeliness of the Inland Revenue tax records in the IDI to enable nowcasting of the unemployment rate.

Data and methods

Like previous research, our nowcasting uses higher frequency administrative data to estimate the overall rate of unemployment. The UR-NOW model differs from the existing research in that we use administrative data on both the number of current income-tested main benefits, as well as filled jobs from the Monthly Employment Index (MEI). The model is also operational with a regular nowcast occurring each month.

Our model development has been through numerous iterations, but broadly we have used a three-stage process. Stage one involved assessing the accuracy of different types of models and covariates for predicting the unemployment rate. Stage two involved fine-tuning our preferred models with different variables and lag structures using standard in-sample selection procedures. For stage three, we compared the performance of different models using out-of-sample forecast errors generated from a roll-forward approach.

An important requirement for properly assessing the out-of-sample accuracy of any forecast is to ensure that the model only uses information available at the time of the forecast. Often data revisions and seasonal adjustments mean that the historical or vintage data will differ from the time series data currently available. For our modelling we do not have access to vintage datasets, and this represents a limitation of our methodology. We do, however, ensure that data used as covariates in the models are only seasonally adjusted or smoothed using the historical data available prior to the forecast.

Our preferred approach turned out to be an ensemble method that averaged the predictions of two separate sub-models. A key reason for model averaging is the existence of model specification error. Combining multiple models rather than estimating a large single model is often a better strategy where there is both a high level of multicollinearity between predictors, and the possibility of including variables that are only spuriously correlated. We used a simple average to combine the predictions, noting that other weighting schemes could also have been used (Timmermann, 2006; Hansen, 2008).

Our preferred approach combines the predictions from the two different sub-models described in Table 1. The dependent variable for the models is the quarterly change in the headline unemployment rate, and covariates are the percentage change in benefit receipt or filled jobs. We discuss this further below, but our basic approach with the covariates is to use the 'new information' that becomes available each month.

Sub-model 1 uses the monthly average number of primary recipients of income tested main benefits for people aged 18 to 64 years. The series is seasonally adjusted within the model.

The number of people receiving any type of income tested main benefit is used rather than just those on the Jobseeker Support – Work Ready benefit because this former series is more highly correlated with unemployment in the HLFS. Benefit data also includes an adjusted number of people receiving the temporary COVID Income Relief Payment that began on 8 June 2020. Including the adjusted number of people in receipt of this payment assumes that many individuals would have been eligible and received another income tested main benefit if the COVID Income Relief Payment had not been implemented.

The sub-model uses the percentage change in benefit receipt compared to the average in the previous quarter. There is no time lag in the benefit data, so the percentage change relates to the month of each forecast. This means that after the end of the first month of the quarter the variable is the increase in the first month relative to the average over the previous quarter. This approach is also repeated after the end of the second and third months of the quarter.

The model uses the percentage change in benefit receipt, but of interest there is also a statistical relationship between the percentage of the population who are primary recipients of income tested main benefits and the unemployment rate. This relationship is more fully described in Annex 2, and it provides a readily understandable 'rough and ready' relationship between the two indicators.

Sub-model 2 uses filled jobs from the Monthly Employment Index (MEI).[1] The unadjusted series is converted by the model into a trend series. The model uses the percentage change in trend monthly employment observed in the previous month relative to the average over the previous quarter. The one-month lag occurs because of the time delay in the availability of the index.

The unadjusted MEI monthly filled job series is the average count of employees with wages and salaries that are taxed at source. It is based on administrative data from the Employer Monthly Schedule (April 1999 to April 2019) and payday filing (from May 2019). The filled jobs series is a count of both full-time and part-time employees who incur more than $50,000 PAYE and employer superannuation.

---

[1] The MEI methodology is described at https://www.stats.govt.nz/methods/about-new-employment-indicator-series

Table 1: Sub-models used in nowcasting the unemployment rate

Description of sub model	Description of administrative data variable
Benefit sub-model $\Delta UR_q = \alpha + \gamma\%\Delta B_m + u_q$	Benefit receipt data sourced from MSD. The monthly series which is derived from daily data commences August 1996. Seasonal adjustment of the series is undertaken within the model. There is no time lag in data availability at the forecast point. We use the percentage change in average monthly benefit numbers compared to the average over the previous quarter.
Employment sub-model $\Delta UR_q = \alpha + \gamma\%\Delta E_{m-1} + u_q$	MEI 'actual' filled jobs from Statistics New Zealand. The series commences in April 1999 and is transformed into a trend series within the model. Information is available for the month prior to the forecast point. The model uses the percentage change in filled jobs over the previous month compared to the average over the previous quarter.
Note: ΔUR_q is the change in unemployment from the previous quarter. B_m refers to benefit receipt observed in the current month. E_{m-1} refers to employment in the previous month	

We also investigated using the monthly All Vacancy Index data from the Ministry of Business, Innovation and Employment (MBIE).[2] The All Vacancy Index is constructed from unique vacancies posted on Seek, Trademe, Education Gazette and Kiwi Health Jobs, but is only available from May 2007. When used in a similar manner to the benefit and employment sub-models it was a marginally significant predictor of changes in the unemployment rate in the final month. However, when assessed in the out-of-sample validation of the models this predictor performed worse than the AR1 benchmark and was subsequently discarded.

An ongoing issue with nowcasting using administrative data is that derived indicators are not necessarily measuring the same concept over time. The information that is captured by administrative data can change across time because of new policies, differences in how services are delivered, or new IT or data collection systems. This means that changes such as the implementation of the COVID Income Relief Payment require some care to ensure that the data is used consistently through time. Measurement instability means that the nowcasting process requires continuous assessment to monitor the performance of the model.

Estimation and validation

Table 2 sets out estimates of the relationship between our selected predictors and the quarterly unemployment rate using data over the period 1996Q4 to 2020Q1. Each sub model is estimated separately for each month of the quarter. Rather than estimates of 'causal' relationships, the coefficients should be interpreted as statistically significant correlations that can then be used for prediction purposes.

[2] The Jobs Online data methodology is described at https://www.mbie.govt.nz/assets/82f9a170cc/jobs-online-methodology-2018.pdf

Table 2: Parameter estimates from sub-models nowcasting changes in the seasonally adjusted unemployment rate

	End of month 1*		End of month 2		End of month 3	
Benefit sub-model	Estimate	p value	Estimate	p value	Estimate	p value
Intercept	0.01	0.76	0.01	0.84	0.01	0.81
Percentage change in total benefit numbers	0.07	<0.01	0.05	<0.01	0.05	<0.01
Adjusted R^2	0.18		0.21		0.23	
Employment sub-model	Estimate	p value	Estimate	p value	Estimate	p value
Intercept	0.16	<0.01	0.18	<0.01	0.19	<0.01
Percentage change in employment (lagged)	-0.88	<0.01	-0.49	<0.01	-0.34	<0.01
Adjusted R^2	0.16		0.20		0.21	

Note: n = 95. The dependent variable is the quarterly change in the seasonally adjusted unemployment rate. *End of month 1 refers to the first Wednesday after the end of the first month to coincide with the public release of the previous quarter's HLFS results.

We assessed the predictive accuracy of the models using an out-of-sample 'roll-forward' strategy. For the assessment we generate nowcasts starting in 2002Q1 and finishing in 2020Q3. The roll-forward strategy involves iteratively estimating the models using only the data available prior to the nowcast quarter. At each iteration the estimated model is then used to predict the current quarter and the errors are recorded. Table 3 shows the root mean square errors of these predictions. We compare the two models as well as the ensemble forecast that was a simple average of both sub-models.

Table 3: Out-of-sample nowcast errors (RMSE)

	End of month 1*	End of month 2	End of month 3
AR1 benchmark	0.32	0.32	0.32
Benefit	0.31	0.31	0.27
Employment	0.31	0.30	0.27
Ensemble model	0.29	0.29	0.26
Statistical significance of ensemble model versus AR(1) benchmark#	0.39	0.25	0.03

*End of month 1 refers to the first Wednesday after the end of the first month to coincide with the public release of the HLFS results from the previous quarter. The model commences forecasts in 2002Q1 and finishes in 2020Q3. For assessment purposes we impute the 2020Q2 unemployment rate. #Modified Diebold-Mariano test (Diebold and Mariano, 1995; Harvey, Leybourne and Newbold, 1997).

The key result shown in Table 3 is that the ensemble model has forecast errors that are smaller than the AR1 benchmark and the individual sub-models. Forecast accuracy also improves over the course of the quarter. But based on the modified Diebold-Mariano test, the difference in the forecast errors of the ensemble model compared to the AR1 benchmark are only statistically significant at the end of the final month.

The forecast errors imply an average confidence interval in the final month of +/- 0.52 percentage points. This relatively wide interval needs to be put in the context of the margin of error of the official surveyed unemployment rate. As mentioned, for the non-seasonally adjusted series this varies but averaged just over +/- 0.30 percentage points during 2020.

Figure 2 shows the actual and out-of-sample predicted rate of unemployment over the period using the ensemble model. Crucially, the model appears to have been reasonably accurate when unemployment increased during the Global Financial Crisis in 2008.

Figure 2: Actual and out-of-sample predicted unemployment rate (1996Q3-2020Q3)

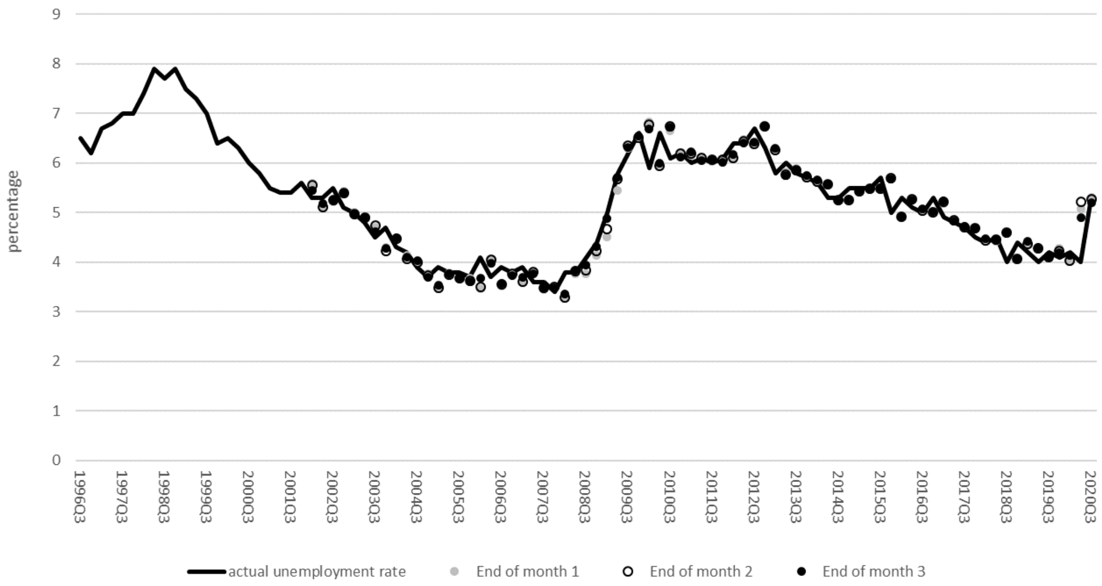
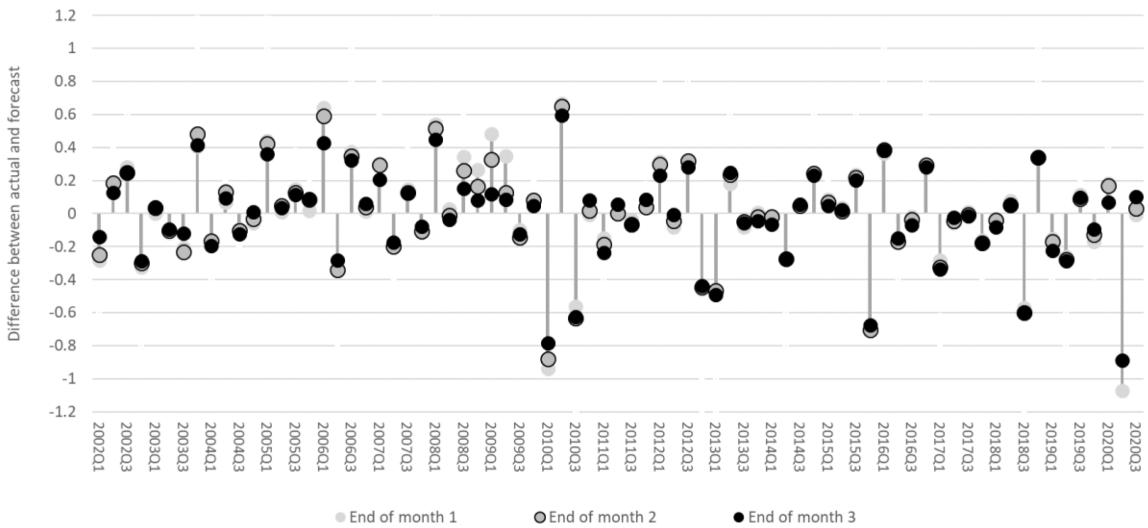


Figure 3 shows a more detailed view of the difference between the actual and forecasts rates of unemployment over the same period.

Figure 3: Difference between actual and out-of-sample predicted unemployment rate (1996Q3-2020Q3)



Nowcast results in 2020

It is useful to look in more detail at the performance of the model during the first three quarters of 2020 as the COVID-19 crisis began to adversely impact New Zealand.

The severity, scale and speed of the crisis demonstrated the need for nowcasting of labour market outcomes to achieve more timely information. However, assessing the accuracy of the model during this highly unusual period is not straightforward because the crisis and the public policy responses also affected the HLFS. The alert level restrictions created difficulties for the usual operation of the HLFS survey with a slight delay in surveying. It also necessitated some changes to sample selection (Statistics New Zealand, 2020a). In addition, the alert level restrictions affected the ability of jobless respondents to look and be available for work which are important elements of the measurement of unemployment (Statistics New Zealand, 2020b).

Table 4 sets out the model predictions and HLFS estimates of the unemployment rate for the first three quarters of 2020. As can be seen the model was highly accurate for the March 2020 quarter, being out by only 0.1 percentage point.

Superficially the model performance appears poor in the June quarter as the reported unemployment rate dropped to 4%. However, additional analysis by Statistics New Zealand suggests that the HLFS estimate of the headline unemployment rate was suppressed by the restrictions associated with Alert Level 4 and 3 during the first half of the quarter. Some respondents to the survey would have been classified as unemployed in normal circumstances, but because of the 'lock down' they did not meet the search and availability requirements of the definition of unemployment.

Table 4: Nowcast and actual results for the unemployment rate in 2020

	2020Q1*	2020Q2	2020Q3
End of month 1*	4.0%	5.1%	5.3%
End of month 2	4.0%	5.2%	5.3%
End of month 3	4.1%	4.9%	5.2%
Actual	4.2%	4.0%	5.3%
Forecast error at end of month 3	-0.1%	0.9%	-0.1%
Imputed seasonally adjusted unemployment rate using the under-utilisation rate	4.1%	4.9%	5.5%

*End of month 1 refers to the first Wednesday after the end of the first month to coincide with the public release of the HLFS results from the previous quarter.

To overcome this issue Statistics New Zealand also reported an alternative measure called the 'extended unemployment rate' for the June 2020 quarter (Statistics New Zealand, 2020c). This included people who indicated that COVID-19 or the lockdown were the reason for not meeting the job search and availability criteria for surveyed unemployment. On a seasonally adjusted basis, the extended unemployment rate was likely around 4.7%.

Table 4 also reports an alternative approach for estimating the headline unemployment rate without the COVID-19 alert level restrictions. This uses the underutilisation rate as a benchmark which may not have been affected as much by the lockdown. The historical correlation between the two indicators is high (0.9) and based on the previous relationship between these two measures we estimate that the June quarter unemployment rate should have been around 4.9 percent.

These alternative benchmarks suggest that the UR-NOW model predictions might have been relatively close in the June 2020 quarter, but given the unprecedented and uncertain circumstances it is difficult to be sure.

Summary and next steps

The unemployment rate is an important economic and social indicator with implications for decision-making in many areas of the economy.

An important issue for decision-making is that the official HLFS measure of surveyed unemployment is always slightly out-of-date due to the time needed for data collection, analysis and publication.

Our nowcasting of the unemployment rate aims to improve the timeliness of information by using high-frequency administrative data to predict the current rate of unemployment.

The UR-NOW model makes monthly predictions of the current unemployment rate. These predictions tend to be slightly more accurate than a simple AR(1) benchmark. However, it is only at the end of the final month of the quarter that the difference against the benchmark is statistically significant. After this final month the UR-NOW model predicts the unemployment rate with a 95% confidence interval of +/- 0.54%.

Overall, our view is that output from the model is a useful addition to the existing range of labour market indicators.

There is considerable scope for adding further sub-models to improve the performance of the overall model. These might use high-frequency data related to areas of economic activity such as consumer spending, road transport, and construction consents and permits. The advantage of the ensemble approach is that the results from these additional regressions can be simply added to those from the sub-models already in use to improve overall predictive accuracy.

The current model nowcasts the current quarter's unemployment rate monthly. There is no reason why this could not be done on a weekly or even daily basis. The only constraint here is the frequency of the updated administrative data used to nowcast the unemployment rate.

It would also be useful to widen the model to nowcast other labour market indicators including other measures of joblessness, the employment rate, and the labour force participation rate.

In addition, there is considerable scope to develop models that nowcast the unemployment rate and other indicators on both a sub-group and regional basis. This may be particularly beneficial because we know that the HLFS statistics for sub-units are prone to considerable measurement error. For example, a regional nowcasting approach using population-level administrative data may provide even more accurate and timely indicators of local labour market conditions.

For the future we aim to make monthly nowcasts of the unemployment rate publicly available through the Centre for Social Data Analytics at AUT https://csda.aut.ac.nz/. We are also aiming to make the underlying data publicly available and welcome any critique or suggestions for using additional variables or improving the modelling.

References

Askitas N. and Zimmermann K.F. (2009) Google econometrics and unemployment forecasting. *Applied Economic Quarterly* 55: 107–120.

Ball, C., Richardson, A., and van Florenstein Mulder, T. (2020) Using job transitions data as a labour market indicator. *Reserve Bank of New Zealand Analytical Notes* No. 02. Wellington.

Boushey, H., Nunn, R., & Shambaugh, J. (2019) Recession Ready: Fiscal Policies to Stabilize the American Economy. Retrieved from https://www.hamiltonproject.org/assets/files/AutomaticStabilizers_FullBook_web_20190508.pdf

Card, D. (2011) Origins of the Unemployment Rate: The Lasting Legacy of Measurement without Theory. *American Economic Review*, 101 (3): 552-57.

Choi, H. and Varian, R. (2009) Predicting initial claims for unemployment insurance using Google Trends. *Google Technical Report*.

Diebold F. X. and Mariano R. S. (1995) Comparing predictive Accuracy. *Journal of Business and Economic Statistics,* 13: 253–63.

Hansen, B. (2008) Least-squares forecast averaging. *Journal of Econometrics,* 146 pp342–350.

Harvey, D., Leybourne, S., and Newbold, P. (1997) Testing the equality of prediction mean squared errors. *International Journal of forecasting*, 13(2): 281-291.

Karagedikli, Ö., & Özbilgin, M. (2019) Mixed in New Zealand: Nowcasting Labour Markets with MIDAS, *Reserve Bank of New Zealand Analytical Notes* No. 04. Wellington.

Krueger, A. B., & Mueller, A. I. (2012) Time use, emotional well-being, and unemployment: Evidence from longitudinal data. *American Economic Review*, 102(3): 594-99.

Statistics New Zealand (2017a) Household Labour Force Survey sources and methods: 2016. Retrieved from www.stats/govt.nz

Statistics New Zealand (2017b) Guide to unemployment statistics, Third Edition, Statistic New Zealand https://www.stats.govt.nz/methods/guide-to-unemployment-statistics-third-edition.

Statistics New Zealand (2020a) COVID-19 and labour market statistics in the June 2020 quarter, Statistics New Zealand Information Release https://www.stats.govt.nz/methods/covid-19-and-labour-market-statistics-in-the-june-2020-quarter.

Statistics New Zealand (2020b) COVID-19 lockdown has widespread effects on labour market, Statistics New Zealand Information Release https://www.stats.govt.nz/news/covid-19-lockdown-has-widespread-effects-on-labour-market.

Statistics New Zealand (2020c) COVID-19 slows labour market activity, Statistics New Zealand Information Release https://www.stats.govt.nz/news/covid-19-slows-labour-market-activity.

Timmermann, A. (2006) Forecast Combination. *Handbook of Economic Forecasting*, Vol 1:135- 196.

Winkelmann, L., and Winkelmann, R. (1998) Why are the unemployed so unhappy? Evidence from panel data. *Economica*, 65(257): 1-15.

Annex 1: Unemployment statistics

There are a variety of different ways to measure unemployment using both survey and administrative data. This annex builds on the excellent Statistics New Zealand paper 'A Guide to Unemployment Statistics' (2017) and explores the nature of different indicators.

There are several important distinctions that need to be kept in mind when understanding indicators of 'unemployment'.

First, the concept can be defined in different ways. For example, it might be restricted to people who were not currently employed, or alternatively it might also include people who are working part-time but would like to work more hours.

Second, different datasets and methodologies can be used to construct the indicators. The differences here include the way data are collected (surveys versus administrative data), the target population, the time period over which the measurement occurs, and the unit of measurement (for example individuals versus families).

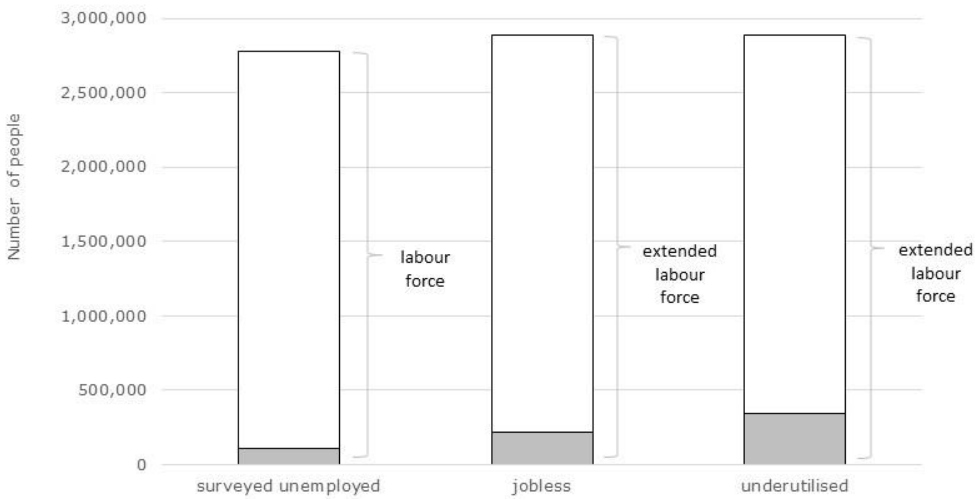
Third, indicators are often expressed as a percentage of those at-risk, but how the 'at risk' population is defined differs. As mentioned earlier, the official unemployment rate is expressed as a percentage of the labour force, while other measures are expressed as a proportion of the extended labour force or total population.

The HLFS has been designed to provide information about both narrow and wider concepts of unemployment. The key measures in this suite of indicators are:

- surveyed unemployed
- the jobless which as well as the surveyed unemployed also includes available potential jobseekers, and unavailable job seekers (the jobless rate is the number of people jobless as a percentage of the extended labour force)
- underutilised which as well as the jobless also includes the underemployed who are defined as individuals working fewer than 30 hours a week who would like and are available to work more hours (the underutilisation rate is expressed as a percentage of the extended labour force).

Figure A1 shows the size of each of these categories of 'unemployment' for the June 2020 quarter. It also shows the total labour force which is the denominator used for the calculation of rates for these different indicators.

Figure 4: Different measures of unemployment (June 2020)



The number of people receiving different types of income-tested main benefits are alternative measure of the number of people out of work.

However, it is important to recognise that what is being measured here is both conceptually and statistically different to surveyed unemployment. Some of the key differences between benefit receipt and survey measures of unemployment are that:

- generally, to be eligible for a benefit a person must not be in full-time employment, but they can be working part-time

- income-tested main benefits differ in terms of the requirement for a recipient to be available and seeking work. For *Job Seeker Support (Work Ready)*, a primary recipient must be looking and available for full-time work. For the Sole Parent Support Benefit there is a requirement to look for part-time or full-time work depending on the age of the youngest child. For *Job Seeker Support (Health Condition or Disability),* a person must be willing to undertake full-time employment, but be limited in their current capacity to seek or undertake employment because of a health condition, injury, or disability. For primary recipients of the *Supported Living Payment,* there is no requirement to be available or look for work

- individuals in receipt of working age main benefits must also meet an income test that takes account of a partner's income where applicable. Individuals might meet the HLFS definition of unemployment but not be eligible for a benefit if their partner is employed in a modestly paying job.

- there are also several other eligibility requirements that mean that someone unemployed is not eligible for a benefit. Individuals aged under 18 years of age are not eligible for some benefits. In most instances, individuals cannot receive an income-tested main benefit and be studying full-time.

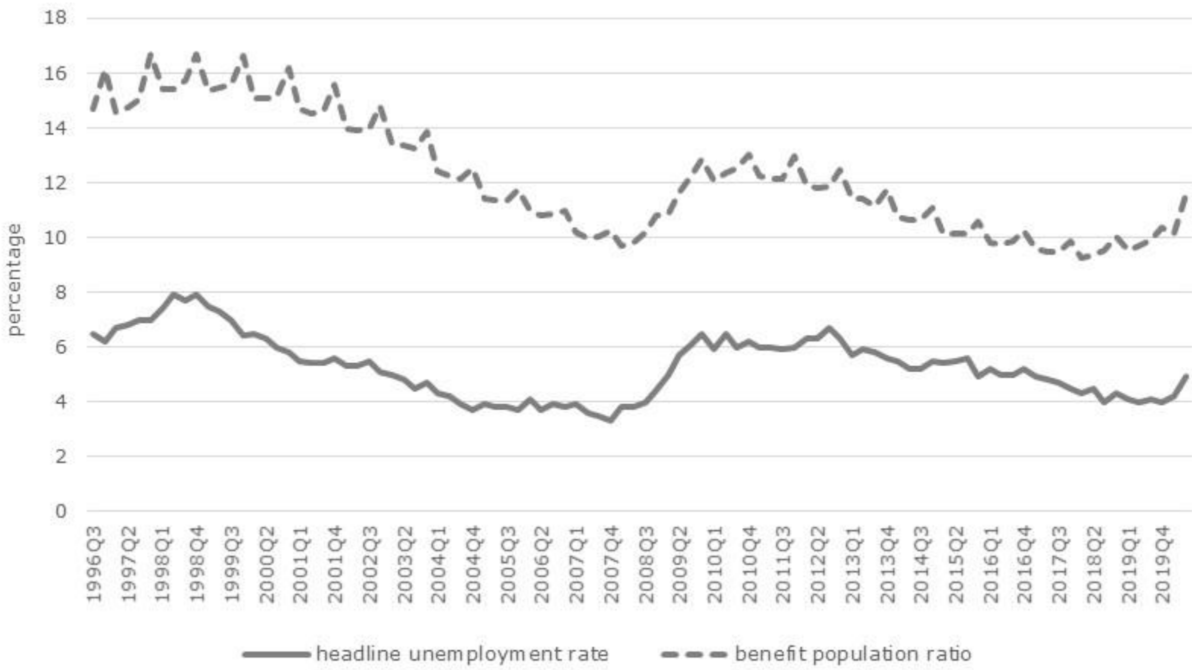
There are also some features of benefit data and the way that it is reported that are important to note. Typically reporting of benefit receipt for monitoring purposes is a count of families rather than individuals. The administrative data on benefit receipt is also typically reported as the count of primary recipients who are eligible to be paid 'on a day' (rather than an average over a quarter as occurs with the HLFS).

Annex 2: Indicative relationship between benefit population ratio and the unemployment rate

As part of investigating how best to nowcast the unemployment rate we looked at a range of different relationships between benefit receipt and unemployment.

One relationship we investigated was between the unemployment rate and the benefit population ratio (total benefit numbers as a percentage of the population 18-64 years). Figure 5 sets out trends in the two series since late 1996. As can be seen, the two indicators are reasonably closely correlated, but do not always move in the same direction.

Figure 5: Unemployment rate and benefit population ratio (18-64)



While the benefit population ratio did not perform as well as the change in benefit receipt variable we used in the final model, it is possible to use the relationship between the two indicators to provide a 'rough and ready' estimator of the current rate of unemployment.

To do this we estimated a simple regression to predict the headline unemployment rate. The covariates in this regression are the total benefit population ratio, the proportion of the population 15-65 who are aged under 25 years, quarter dummy variables, and a dummy variable for a break in the series that occurred after June 2008 (we are not exactly sure about the reason for this apparent structural break).

The results set out in table 5 suggest that holding other factors constant, for every percentage point change in the benefit population ratio the unemployment rate will change by 0.62 percentage points. However, as with any model, there will be forecast errors when using this approach to predict the current rate of unemployment.

Table 5: Parameter estimates from regression model exploring the relationship between the benefit population ratio and the unemployment rate

	Model 1		Model 2	
Benefit sub-model	Estimate	p value	Estimate	p value
Intercept	0.68	0.10	-14.05	<.0001
Benefit population ratio	0.39	<.0001	0.62	<.0001
June	0.00	0.99	0.00	1.00
September	-0.06	0.80	-0.16	0.22
December	-0.37	0.11	-0.59	<.0001
Proportion of the population 15-64 aged under 25	-	-	0.52	<.0001
Post June 2008 dummy	-	-	1.70	<.0001
Adjusted R^2	0.49		0.83	

Note: Dependent variable is the seasonally adjusted unemployment rate. n=94. Heteroscedasticity consistent standard errors.