

## Methodology

To estimate transition probabilities between three labour market states, we use the Labour Force Survey microdata for February and March of 2020 and 2019. March 2020 data was used as many important policy announcements have not yet been made, and thus likely capture the direct impact of COVID-19 on labour market decisions than subsequent months where financial support announcements were made for both the working population and current students.

Specifically, we are interested in estimating conditional probability of being in a particular labour state in March, conditional on being in a particular labour state in February. Given that we have three main labour market states of interest (Employment, Unemployment, and Not In Labour Force), we have nine probabilities of interest:

		Labour Market State in March (2019/2020)		
		Employed (E)	Unemployed (U)	Not in LF (L)
Labour Market State in February (2019/2020)	Employed (E)	$P(E_t E_{t-1})$	$P(U_t E_{t-1})$	$P(L_t E_{t-1})$
	Unemployed (U)	$P(E_t U_{t-1})$	$P(U_t U_{t-1})$	$P(L_t U_{t-1})$
	Not in LF (L)	$P(E_t L_{t-1})$	$P(U_t L_{t-1})$	$P(L_t L_{t-1})$

Within the Labour Force Survey, we focus on the following set of variables to identify some of these conditional probabilities:

Variable	Description	Answered by
LFSSTAT	Labour Force Status for the reference week	All respondents
EVERWORK	If respondent has ever held employment	Those unemployed or outside of labour force
DURUNEMP	Duration of unemployment (in weeks)	Unemployed respondents



FLOWUNEM	Previous labour state before unemployment	Unemployed respondents
DURJLESS	Duration since respondent last held employment (in months)	Those unemployed or outside of labour force

Specifically, the Labour Force Survey measures four different labour states: Employed At Work, Employed Not At Work, Unemployed, and Not in the labour force: for the purpose of this analysis, we group Employed at work and Employed not at work into a singular state of being employed.

Much of this analysis will rely on applications of [Bayes' rule](#), for example,  $P(U_t|E_{t-1})$  can be decomposed into:

$$P(U_t|E_{t-1}) = \frac{P(E_{t-1}|U_t)P(U_t)}{P(E_{t-1})}$$

The other conditional probability is relatively easier to estimate, as for all those unemployed in March, we have access to their previous labour state, as well as the duration with which they have been in this current labour state (using a combination of FLOWUNEM and DURUNEMP). Specifically for this conditional probability, we use the following reasons to identify workers who were previously in employment: “Job losers” (temporary or permanent layoff), “Job leavers”, and “Job leavers (reason unknown) who last worked more than 1 year ago”. In addition, we only consider those whose unemployment duration has been less than or equal to 4 weeks. We then use the estimated conditional probability of each labour state using March and February data to get our estimate. The combination of these two requirements mean that those who last worked more than 1 year ago will be excluded.

Similar calculation was used in estimating  $P(U_t|L_{t-1})$ . Of note is that LFS captures a special labour force state for being “unemployed” - those who are currently without a job but will begin working in the next four weeks. For the purpose of this analysis we classify this worker to have been outside of the labour force previously - they represent 4.4% of all those who are unemployed. We then exploit the following identity to estimate  $P(U_{t-1}|U_t)$  from which we can use to estimate  $P(U_t|U_{t-1})$ :

$$P(U_{t-1}|U_t) + P(E_{t-1}|U_t) + P(L_{t-1}|U_t) = 1$$

To estimate  $P(L_t|E_{t-1})$ , we focus on using a combination of DURJLESS and EVERWORK, where we count those who have been without a job for at most 1 month (the minimum value possible in



the dataset) for those who have had previous employment experience. This is an appropriate way to estimate this conditional probability, as those who are recorded to be outside of the labour force who have been without a job for one month and have previous experience working must have worked in the previous months given how the dataset is coded.

We then use  $P(L_t|E_{t-1})$ ,  $P(U_t|E_{t-1})$ , and the following identity to estimate  $P(E_t|E_{t-1})$ :

$$P(U_t|E_{t-1}) + P(E_t|E_{t-1}) + P(L_t|E_{t-1}) = 1$$

The previous 5 conditional probabilities are the only probabilities for which we can estimate directly from the data, as the previous labour market state for those currently employed or those outside of the labour market is not captured. To estimate this, we focus on the following system of equations (obtained by decomposing the number of people in a specific labour state into individual inflow and outflows):

$$E_t = E_{t-1} - E_{t-1} \times P(L_t|E_{t-1}) - E_{t-1} \times P(U_t|E_{t-1}) + U_{t-1}P(E_t|U_{t-1}) + L_{t-1}P(E_t|L_{t-1})$$

$$U_t = U_{t-1} - U_{t-1} \times P(L_t|U_{t-1}) - U_{t-1} \times P(E_t|U_{t-1}) + E_{t-1}P(U_t|E_{t-1}) + L_{t-1}P(U_t|L_{t-1})$$

Putting all terms that we can calculate or estimate we end up with the following equations:

$$U_{t-1}P(E_t|U_{t-1}) + L_{t-1}P(E_t|L_{t-1}) = E_t - E_{t-1} + E_{t-1} \times P(L_t|E_{t-1}) + E_{t-1} \times P(U_t|E_{t-1})$$

$$U_{t-1} \times P(L_t|U_{t-1}) - U_{t-1} \times P(E_t|U_{t-1}) = -U_t + U_{t-1} + E_{t-1}P(U_t|E_{t-1}) + L_{t-1}P(U_t|L_{t-1})$$

Unfortunately, this is a system of two equations with three unknowns:  $P(E_t|U_{t-1})$ ,  $P(E_t|L_{t-1})$ , and  $P(L_t|U_{t-1})$ . Though other equations exist using other known identities and decomposition, they are all linearly dependent on these two equations. To solve for the three unknowns, we follow a similar strategy to [Nakamura, Nakamura, Phong, and Steinsson \(2019\)](#), by focusing on the fact that direct transition from being out of the labour force to being in the labour force is a comparatively rare event, that is,  $P(E_t|L_{t-1})$  is close to zero. If for example, we set  $P(E_t|L_{t-1})=0$ , we assume that all in flow into employment were from previously unemployed people, and allows us to solve this system of equations. We use such estimates to present our baseline results, qualifying that when  $P(E_t|L_{t-1})=0$ ,  $P(E_t|U_{t-1})$  attain its maximum value, and  $P(L_t|U_{t-1})$  attains its minimum value.

To ensure the soundness of the approach, we also estimated the maximum possible value for  $P(E_t|L_{t-1})$  by setting  $P(E_t|U_{t-1})=0$  instead. Under this scenario,  $P(L_t|U_{t-1})$ 's (maximum) value only



## **WELCOME TO THE COVID-19 LABOUR MARKET: METHODOLOGY**

reaches 0.66% for young workers in 2020, and 0.41% for the general population in 2020. It does increase to 3.58% for young workers in 2019, though this is still lower than the variations observed for the other two probabilities of interest, likely reflecting the fact that young people, by construction, are entering the labour force for the first time.

