

# **Disaggregated Sectoral Employment Dynamics in Australia**

A thesis submitted for the degree of  
Bachelor of Commerce (Honours)

by

**Zhixiang Yang**

30306396



Department of Econometrics and Business Statistics  
Monash University  
Australia

October 2022

# Contents

<b>Acknowledgement</b>	<b>1</b>
<b>Declaration</b>	<b>3</b>
<b>Abstract</b>	<b>5</b>
<b>1 The Australian COVID-19 Pandemic Background</b>	<b>7</b>
1.1 Research Aim and questions . . . . .	7
1.2 Thesis Structure . . . . .	8
<b>2 Review of literature</b>	<b>9</b>
2.1 Sectoral Impact of COVID-19. . . . .	9
2.2 Baysian VAR . . . . .	10
2.3 Setting Minnesota Prior with Shrinkage . . . . .	10
<b>3 Data collection and exploratory analysis</b>	<b>11</b>
3.1 Data Introduction and wrangling . . . . .	11
3.2 Preliminary Exploratory Data Analysis . . . . .	13
<b>4 Methdology</b>	<b>15</b>
4.1 Proposed Model . . . . .	15
4.2 Prior and Shrinkage . . . . .	16
4.3 Selecting the hyperparameter of Minnesota Prior . . . . .	17
<b>5 Sectoral Employment Analysis</b>	<b>21</b>
5.1 Long-run Multiplier Analysis . . . . .	21
5.2 Evaluations after COVID-19 . . . . .	23
<b>6 Discussions and conclusion</b>	<b>29</b>
6.1 Limitation . . . . .	30
6.2 Future Extension: An Application of Machine Learning in Hierarchical Forecasting	30
<b>A An Example of Bayesian VAR and Minnesota Prior</b>	<b>33</b>
<b>B Graphs</b>	<b>37</b>
<b>C Supplementary Materials (Data &amp; Codes)</b>	<b>41</b>
<b>Bibliography</b>	<b>43</b>



# Acknowledgement

I would like to express gratitude to my supervisor Professor Farshid Vahid and my coordinator Professor Heather M Anderson for their selfless support and devoted care along the way. I would also like to thank my family and Xiefei Li for their continuous supports and encouragement.

Matlab version. 2022a codes are used to generate the estimated coefficients for the proposed Bayesian VAR model.

R 4.1.2 "Bird Hippie" (R Core Team, [2021a](#)) and Rstudio 2022.07.1+554 "Spotted Wakerobin" (RStudio Team, [2020](#)) are used for data analysis in this thesis.

R package used and their versions in this research include `matrixStats` 0.61.0 (Bengtsson, [2021](#)), `ggplot2` 3.3.5 (Wickham, [2016](#)), `zoo` 1.8.9 (Zeileis and Grothendieck, [2005](#)), `tidyverse` 1.3.1 (Wickham et al., [2019](#)), `dplyr` 1.0.8 (Wickham et al., [2022](#)), `stats` 4.1.2 (R Core Team, [2021b](#)), `readr` 2.1.2 (Wickham, Hester, and Bryan, [2022](#)), `pracma` 2.3.8 (Borchers, [2022](#)), `lubridate` 1.8.0 (Grolemund and Wickham, [2011](#)), `fpp3` 0.4.0 (Hyndman, [2021](#)), `ggpubr` 0.4.0 (Kassambara, [2020](#)), `ggrepel` 0.9.1 (Slowikowski, [2021](#)), `directlabels` 2021.1.13 (Hocking, [2021](#)), `gghighlight` 0.3.3 (Yutani, [2022](#)). (*Note: In no particular order above*)

This thesis is produced using the R package `bookdown` 0.25 (Xie, [2016](#)) with the Rmarkdown template package `monash` 0.0.0.9000 for the honours thesis in the Department of Econometrics and Business Statistics, Monash University (Tanaka and Hyndman, [2022](#))



# Declaration

I declare that this thesis contains no material which has been submitted in any form for the award of any other degree or diploma in any university or equivalent institution, and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

– Signature: Elvis Zhixiang Yang



# Abstract

I investigate the performance of Australian employment at a disaggregated level and compare the post-COVID situation with the no-COVID situation. A multivariate time series model is utilised to determine the long run employment spillovers to total employment at the two-digit subsectoral level. My results suggest that in the two-digit subsectoral level, “Other Store-Based Retailing”, followed by “Administrative Services”, “Furniture and Other Manufacturing” will generate the strongest positive spillovers to the total economy. One of the main contributions of this paper is that the government can stimulate these high and positive spillovers to recover total employment effectively and efficiently. Moreover, I evaluate the impacts of COVID-19 on the Australian labour market with counterfactual analysis. The outcomes demonstrate that the pandemic has a long-lasting effect on the economy. In addition, I point out the main reason behind the lowest unemployment rate record and compare the no-COVID scenario for unemployment rate. The spillover analysis together with the evaluation, thus, provide another aspect for policymakers to improve the Australian economy at the post-pandemic stage.





## **Chapter 1**

# **The Australian COVID-19 Pandemic Background**

The COVID-19 pandemic has had a massive impact on global economies. Millions of workers have been furloughed or even fired across the globe as businesses struggled to survive (Lewis and Hsu, 2020). The same situation occurred in Australia, due to increased restrictions, many businesses closed their doors, while employees worked few hours or were laid off by businesses. As a result of the continued “lockdown” periods in 2020, the Australian Bureau of Statistics (ABS, 2021) estimated that 72% of businesses generated less revenue and the underemployment rate reached a historical high of 13.8% by the end of April, 2020, only one month after the outbreak of COVID-19.

Our research is motivated by the lack of quantitative research on the employment of two-digit disaggregated industry subsectors in Australia, despite the fact that many studies have focused on the aggregated employment rate. The loss of hierarchical information is a common issue with aggregated studies, which can lead to biased conclusions or the illusion of prosperity in the labour market. In this manner, a quantitative analysis of sectoral employment will alleviate this problem, allowing us to better assess the effects of COVID-19 in Australia.

### **1.1 Research Aim and questions**

This research will expand on Anderson et al. (2020) by using data from 87 two-digit industry sectors rather than the 19 sectors that they used previously. I will develop a model for the

two-digit subsectors to assess the long run effect and the COVID-19 post-impacts. In addition, I will present a counterfactual analysis based on the optimistic assumption that no COVID-19 or major events occurred. The two-digit sectoral data will provide us with more valuable information, allowing us to gain a deeper understanding of employment dynamics in Australia at a disaggregated level.

The overall goal of the research is to provide estimates of two-digit sectoral employment based on historical data. My specific objectives are as follows:

1. Create a time series model of employment in 84 two-digit subsectors of the Australian economy.
2. To conduct a counterfactual analysis using this model.
3. To use this model to determine which two-digit subsectors have the highest long-term impact (or positive spillover) on employment growth in the long run.

### **1.2 Thesis Structure**

This thesis examines Australian Employment at a disaggregated level before estimating the long run effects of COVID-19 on the sectoral employment rate in Australia. The rest of the thesis is organised as follows. First, in chapter 2, I begin by reviewing the existing literature in the relevant fields. Second, in chapter 3, I will provide data sources as well as an exploratory data analysis. Then, in chapter 4, I will present our selected model. After we have chosen our model, I will conduct a counterfactual analysis to assess the impact of COVID-19 and provide useful insights on key beneficial industries after COVID-19 in chapter 5. Finally, in chapter 6, I will provide a brief conclusion and discuss the limitations and potential future extensions.

## Chapter 2

# Review of literature

We focused our review of the existing literature on two major topics:

1. The COVID-19 sectoral impacts and economic modelling
2. Bayesian VAR Modelling of large numbers of time series

### 2.1 Sectoral Impact of COVID-19.

Most existing studies have concentrated on assessing the effects of COVID-19 on broad sectors of large economies such as the US and Europe. To predict the impacts of COVID-19, Ludvigson, Ma, and Ng (2021) developed a disaster series to translate the macroeconomic impact of costly and deadly disasters in recent US history and model them as sectoral shocks. They concluded that the shock would result in a cumulative loss of 20% in industrial production, 39% in public services and a 12.75 percent reduction in US GDP by the end of 2020. Gregory, Menzio, and Wiczer (2020) conducted simulations under various scenarios via a search theoretic model using US data and discovered that the recovery in the U.S. would be L-shaped, with employment remaining lower than pre-covid for a long period. They also expanded the research to a disaggregated level of 20 sectors, predicting that the “arts and entertainment” and “accommodation and food services” sectors would experience the hardest hit during the pandemic.

In Australia, Anderson et al. (2020) developed a multivariate time series model for 19 main sectors (as a typical small open economy) using a Bayesian VARX model. According to their findings, “Manufacturing” and “Construction” have the greatest positive spillovers for the

overall economy. Meanwhile, they also applied a “conditional forecasting” method proposed by Waggoner and Zha (1999) to forecast different scenarios of total employment in Australia after the pandemic. However, their research did not use a finely disaggregated level in Australia (two-digit subsectors of main sectors), which may be less informative in macroeconomic analysis.

## 2.2 Bayesian VAR

The Bayesian Vector Autoregression model (BVAR) is commonly used in the literature for high-dimensional multivariate modelling (e.g. Anderson et al., 2020; Litterman, 1986; Bańbura, Giannone, and Reichlin, 2010). The BVAR model is appealing because it allows us to estimate a large number of parameters when the sample size is small, in a statistically coherent way. (Litterman, 1986; Woźniak, 2016).

To utilize the Bayesian VAR estimators, Litterman et al. (1979) proposed the Minnesota Prior, which decreases the weight of the lagged variables with the lag length. The prior mean on the first own lag is set to unity and the rest are set to zero so that (a) the most recent lag should provide more information than distant lags; and (b) own lags should explain more than the lags of other variables.

## 2.3 Setting Minnesota Prior with Shrinkage

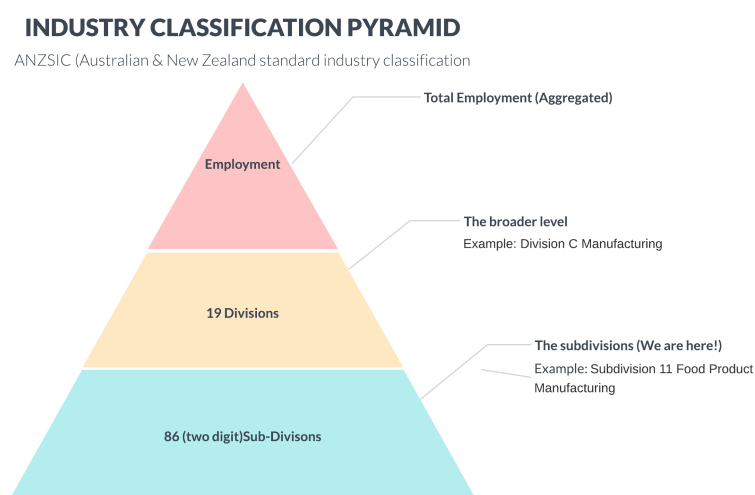
The literature suggests that a significant improvement in the predicting performance of large BVAR dynamic models can be made by more careful choice of the prior assumptions (Bańbura, Giannone, and Reichlin, 2010; Litterman, 1986). Additionally, Robertson and Tallman (1999) and Kadiyala and Karlsson (1997) proposed a Normal-inverse-Wishart prior to put the Minnesota prior in our estimated model, by preserving its main characteristics. Particularly, Bańbura, Giannone, and Reichlin (2010) suggested incorporating dummy observations into the BVAR system as an easier way to apply the Normal-inverse-Wishart prior (see Appendix A for details).

## Chapter 3

# Data collection and exploratory analysis

### 3.1 Data Introduction and wrangling

The data I use for this thesis comes from the Australia Bureau of Statistics (ABS), which covers 86 industry sub-divisions of main jobs. The ABS records the employment (measured in thousands of people ('000)) from 1984 : Q4 to 2021 : Q4 with a structure shown in Figure 3.1.



**Figure 3.1:** *Australian Industry Pyramid plot by (ANZSIC)*

Although seasonally adjusted data is available in (ABS, 2022a), however, the seasonal adjustment of post-COVID data is problematic as shocks caused by lockdowns would affect the post-covid seasonality. In addition, the seasonally adjusted subsectoral data do not add up to

total employment, which prevents the model from maintaining internal consistency (such as forecasting coherence) in the model. Therefore, rather than using the seasonally adjusted data provided by the Australian Bureau of Statistics, I use the original quarterly employment data with transformations (see description below) to capture any potential changes in seasonal patterns.

In this thesis, I only downloaded the data from the fourth quarter of 1984 to the second quarter of 2022. I also use Microsoft Excel to extract both the employment for two-digit disaggregated levels as well as total employment. It should be noted that some of the data are zero, which makes a log transformation impractical. As a result, I will make the changes below to get rid of them while preserving the coherence of the data structure. The cleaned data can be seen in the *ABSEmp.xlsx* file at ([https://github.com/elvisssyang/Disaggregated\\_Employment](https://github.com/elvisssyang/Disaggregated_Employment)).

- Merge the two-digit subsector 57- *Internet Publishing and Broadcasting* and 54- *Publishing (except internet)* to a new combined subsector called 54 *Publishing and broadcasting*.
- Combine the 96 *Private Households Employing Staff and Undifferentiated Goods and Service Producing Activities of Households for Own Use* and 95 *Personal and Other Services* as 95 *Personal and other services (include activities for own use)*.

It is obvious that the ABS employment data contains few non-classified series (nfd). Because ABS offers no additional information, I will not discuss these series in this thesis. If this is incorporated into the model, it will have an impact on the sectoral dynamics. To make the forecasts coherent (sum to total employment) and analysis universal, I will not take them into account at this stage. As a result, my total employment data differs from the total employment data that has been published. The largest discrepancy, however, concerns only a small portion of the real total employment. Accordingly, this will not have a big impact on our analysis.

When there are no zeros, a log transformation can be used to explain the percentage change in employment. To generate accurate forecasts from the estimated VAR model, we need the data to be stationary. Thus, I will further apply a fourth difference (seasonal difference) to eliminate the seasonality and nonstationarity for the original quarterly data.

Finally, in *Chapter 6*, I also combine the following data together with the number of employment by subsectors (ABS, 2022a) to further support our counterfactual analysis:

- Total Labour Force: Downloaded from ABS website (see ABS (2022a)) <sup>1</sup>.
- Unemployment Rate: Downloaded from ABS website (see ABS (2022b))

## 3.2 Preliminary Exploratory Data Analysis

Figure B.1 illustrates the changes in the raw data for 19 key sectors in Australia between 2010 and 2022. As a result of business closings and travel bans in 2020:Q2, we can observe that the number of total employment dropped substantially (from around 13,200,000 to 12,200,000 in 2020 : Q2). Figure B.1 illustrates how most industries behaved similarly and underwent significant changes during the pandemic. The “Accommodation & Food”, “Media & telecom” and “Arts” industries have experienced a severe loss of employment when compared to the previous data of these industries, and have not yet fully recovered to the pre-covid level. Some industries, such as “Financial” and “Healthcare”, on the other hand, essentially remained the same as pre-COVID period and displayed an ongoing upward trend.

When compared to the sectoral level, the employment of subsectors at the two-digit level illustrates a different pattern (see Figure B.2). As shown in Figure B.2, the subsectors “Food and Beverage Services”, “Construction Services”, “Professional, Scientific and Technical Services (Except Computer System Design and Related Services)” and “Other Store-Based retailing” are relatively large subsectors with various losses. For example, the “Food and Beverage Services” subsector increased steadily from 2010 Q1 to 2020 Q1, but was significantly impacted by the first lockdown in 2020 Q2, which resulted in a loss of about 210,000 employed people. On the other hand, “Construction Services”, which belongs to the construction sector, did not experience a sizable decline during the pandemic. Similar to the “Professional, Scientific and Technical Services” subsector, it recovered the loss in just two quarters and is expected to increase in the future.

However, only taking into account the 19 broad sectors has a disadvantage in that the two-digit subsectoral dynamics of these sectors might not be consistent with their aggregated sectoral changes. For instance, one might assume that the corresponding subsectors would exhibit the same pattern based on the aggregated performance of the “Education” and “Healthcare” sectors at the 19 sectoral level (see Figure B.1). However, the reality is that while there is an overall upward trend in the “Healthcare” sector or in the “Education” sector, some of their two-digit

---

<sup>1</sup>This includes the number of employment by subsectors, which is the original data we use in this thesis.



subsectors perform differently (e.g. “Residential Care Services” and “Tertiary Education” ) (see Figure B.2). This indicates that not all two-digit subsectors behave in the same way as the sectoral level.

The top five and bottom five two-digit subsectors in terms of their year-on-year growth rate of “2022 Q2” are shown in Table 3.1. From Table 3.1, we can see that “Other Transport”, “Non-Metallic Mineral Mining and Quarrying” and “Sports and Recreation Activities” were the industries that affected most by the lockdown in “2020 Q2”. However, not all subsectors suffered a lot in “2020 Q2”. During this period, both “Water Transport” and “Gas Supply” had a remarkable increase, followed by “Broadcasting (except Internet)” and “Exploration and Other Mining Support Services”.

Date	Sector	YoY growth rate
2020: Q2	48 Water Transport	81.01%
2020: Q2	27 Gas Supply	55.14%
2020: Q2	56 Broadcasting (except Internet)	34.49%
2020: Q2	10 Exploration and Other Mining Support Services	32.87%
2020: Q2	63 Insurance and Superannuation Funds	27.78%
Date	Sectors	YoY growth rate
2020: Q2	50 Other Transport	-78.98%
2020: Q2	09 Non-Metallic Mineral Mining and Quarrying	-66.92%
2020: Q2	91 Sports and Recreation Activities	-60.48%
2020: Q2	55 Motion Picture and Sound Recording Activities	-55.19%
2020: Q2	03 Forestry and Logging	-54.40%

**Table 3.1:** *The highest and lowest five two-digit subsectors’ employment percentage change for 2020:Q1 to 2020:Q2*

## Chapter 4

# Methdology

### 4.1 Proposed Model

I plan to use a Bayesian VARX model based on the method used in Anderson et al. (2020). The VARX model is particularly useful in modelling dynamic behaviours of the inter-variable relationships (Warsono et al., 2019). In the model, each sector is affected by the lags of sectoral annual growth and a lag of the total employment growth. In addition to serving as an economy-wide factor, the inclusion of the lag of total employment growth ensures the model's self-consistency (e.g., forecasting coherence).

In many time series models, the number of lags is selected based on the patterns of the time series (i.e. seasonality, cycle, or trend). Four lags are typically used for quarterly data (see Anderson et al. (2020) and Stock and Watson (2001)). However, because of the high-dimensionality and relatively small sample size in my case, I will use one lag of 84 sectors and one lag of the total employment.

Therefore, the suggested BVAR model is:

$$\mathbf{y}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{y}_{t-1} + \Gamma x_{t-1} + \mathbf{u}_t$$

where  $\mathbf{y}_t$  is an  $84 \times 1$  vector of two-digit subsectoral employment growth rate at time  $t$  and  $\mathbf{x}_{t-1}$  is a  $1 \times 1$  vector stands for one lag on the growth rate of the total employment (this vector of variables are predetermined at time  $t$ ),  $\mathbf{c}$  is a vector of constants,  $\mathbf{A}_1$  is an  $84 \times 84$  parameter

matrix.  $\Gamma$  is an  $84 \times 1$  matrix and  $\mathbf{u}_t$  is a vector of reduced form errors with the mean equal to zero and independent variance  $\mathbf{u}_t \sim (\mathbf{0}, \Sigma)$ . (See Appendix A for details)

In the proposed model, the use of seasonality unadjusted data allows the estimates to be coherent (i.e. the sum of subsectoral employment equals total employment). Moreover, the share of each subsector also changes endogenously as the employment changes over time. Therefore, even if we have one lag of the growth rate of total employment, there is no multicollinearity because the shares of subsectors change over time.

## 4.2 Prior and Shrinkage

By imposing prior beliefs on the parameters, a Bayesian VAR aids in overcoming the curse of high dimensionality (Bańbura, Giannone, and Reichlin, 2010). I will estimate the employment dynamics using a the Bayesian VAR model by specifying a Minnesota type prior (e.g. Anderson et al., 2020; Litterman, 1986; Robertson and Tallman, 1999), which is defined as follows:

$$E[a_i^{jk}] = E[\gamma_i^j] = 0$$

$$Var[a_i^{jk}] = \begin{cases} \frac{\lambda^2}{i^2}, & j = k \\ \frac{\lambda^2}{i^2} \frac{\sigma_j^2}{\sigma_k^2}, & otherwise \end{cases}$$

$$Var[\gamma_i^j] = \frac{\lambda^2}{i^2} \frac{\sigma_j^2}{\sigma_e^2}$$

where in the proposed model (see Chapter 4.1), the number of lag is  $i = 1$ . Therefore, the  $a_1^{jk}$  and  $\gamma_1^{jk}$  are  $j, k^{th}$  of  $A_1$  and  $\Gamma_1$  matrices. The degree of shrinkage is governed by  $\lambda$ , where  $\frac{1}{i^2}$  down-weights the distant lags in general notation and the  $\frac{\sigma_j^2}{\sigma_k^2}$  adjusts for different scale of the data.  $\sigma_e^2$  is the variance after fitting an AR model on total employment growth.

Bańbura, Giannone, and Reichlin (2010) also suggests a natural conjugate Normal-Inverse-Wishart prior, which retains the principle of Minnesota prior. This will greatly simplify the steps of adding Minnesota prior to the Bayesian VAR model. Its posterior moments can be calculated either analytically or by adding the dummy observations. I will use dummy observations to

estimate the BVAR (Bańbura, Giannone, and Reichlin, 2010). More details are provided in the Appendix.

### 4.3 Selecting the hyperparameter of Minnesota Prior

Specifically, the Minnesota prior has the following beliefs about the variances in our estimated one lag Bayesian VARX model:

$$Var[a_1^{jk}] = \begin{cases} \lambda^2, & j = k \\ \frac{\lambda^2 \sigma_j^2}{\sigma_k^2}, & otherwise \end{cases} \dots (4.3.1)$$

$$Var[\gamma_1^j] = \frac{\lambda^2 \sigma_j^2}{\sigma_e^2} \dots (4.3.2)$$

where  $\lambda$  is a hyperparameter specified based on how far we will shrink the estimates and  $\frac{\sigma_j^2}{\sigma_k^2}$  adjusts for the different scale of the data. To effectively scale the estimator  $\gamma_1^j$  and  $a_1^{jk}$ , I obtain  $\sigma_n^2$  by fitting an AR(4) model on the  $n$ -th variable using least squares, which is commonly used in many literatures (Anderson et al., 2020; Bańbura, Giannone, and Reichlin, 2010; Koop, 2013).

As the Minnesota prior is defined from the previous equations (see 4.3.1 and 4.3.2), the hyperparameter  $\lambda$  controls the overall tightness (variance) of the prior distribution (Bańbura, Giannone, and Reichlin, 2010). If  $\lambda \rightarrow 0$ , we can see that the prior assumption is influential, which means that the posterior is approaching to the prior. That is, the data do not affect the estimation. On the other hand, if  $\lambda \rightarrow \infty$ , the posterior expectations will approach the ordinary least squares (OLS) estimates. In many macroeconomic VAR forecasting scenarios, the data has a large dimension. As the dimension grows, we want to shrink more in order to prevent over-fitting (De Mol, Giannone, and Reichlin, 2008).

Undoubtedly, by regulating the degree of shrinkage, the hyperparameter  $\lambda$  is crucial in increasing forecast accuracy. For example, Bańbura, Giannone, and Reichlin (2010) points out that a gain in efficiency could be made by applying Bayesian shrinkage in estimating large multivariate VAR models. Additionally, they also conclude that large Bayesian vector autoregressions (BVARs) with shrinkage are helpful for constructing structural analyses.

Based on applied experience, Litterman came to the conclusion that the shrinkage estimate  $\lambda = 0.2$  is adequate to handle a large number of empirical cases (Litterman, 1986). More importantly, the data size needs to be considered as well when determining the degree of shrinkage (Bańbura, Giannone, and Reichlin, 2010). The two-digit level (84 subsectors) of employment in Australia is more complex with 84 subsectors than the one-digit level (19 sectors) studied by Anderson et al. (2020). Therefore,  $\lambda = 0.2$  might not be appropriate in this multivariate case. Here, I will give a thorough breakdown of the approach I employed to select the optimal  $\lambda$ .

Due to the fact that disaggregated subsectors have different scales, the commonly used scale-dependent error measurement (e.g. MAE, MSE) may not work when comparing forecast accuracy between subsectors. Despite being unit-free, mean absolute percentage error (MAPE) is inaccurate in subsectors with relatively small shares (e.g. a small change in a relative small subsector will significantly increase the MAPE for that subsector). When  $y_t$  is close to zero, MAPE will likely have extreme values or become undefined. Accordingly, I will sum the employment across all sectors to total employment and minimise the total employment forecast error to select the optimal  $\lambda$ .

The error measurement I will use is the root mean squared forecast error RMSFE. It is calculated via an out-of-sample forecasting experiment, which is a similar approach used in many empirical cases (Bańbura, Giannone, and Reichlin, 2010; Koop, 2013). Here, I denote  $H$  as the longest forecast horizon to be evaluated, both  $T_b$  and  $T_e$  as the length of the training set and testing set, respectively. Give the forecast horizon  $h$ , hyperparameter  $\lambda$  and model  $m$ , for each given period between  $T_b$  and  $T_e$  ( $T = T_b, \dots, T_e - h$ ), I compute  $h$ -step-ahead forecasts  $y_{i,T+h|T}^{(\lambda,m)}$  using only the information up to time  $T$ . I then compute the forecast error  $y_{i,T+h}$  by subtracting the actual data  $y_{i,T+h}$ .

Then, out-of-sample forecast accuracy is measured in terms of the root mean squared forecast error (RMSFE) as:

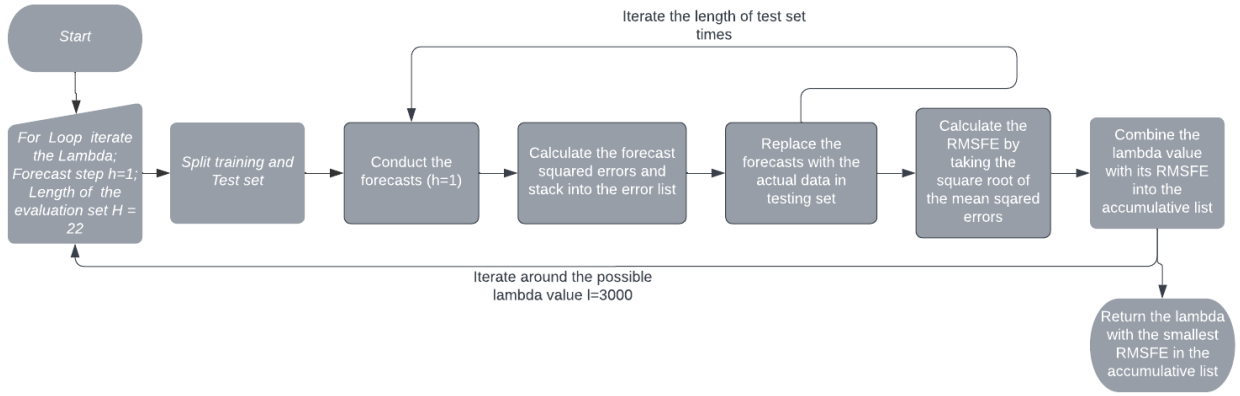
$$RMSFE_h^\lambda = \sqrt{\frac{1}{T_e - h - T_b + 1} \sum_{T=T_b}^{T_e-h} (y_{T+h|T}^\lambda - y_{T+h})^2}$$

where  $y_{T+h|T}^\lambda$  is defined as the  $h$ -th steps ahead forecast (total employment in this case) given the information up to time  $T$  and  $y_{T+h}$  is the actual data for the  $h$ -th steps ahead forecast (total

employment in this case). Here,  $\lambda$  stands for the evaluated RMSFE, conditioned on a specific model and the hyperparameter  $\lambda$ .

In this section, I will set up an effective searching algorithm to search for the optimal shrinkage parameter  $\lambda$ . For our purposes, I want to provide accurate forecasts of total employment based on the scenario in which there was no pandemic happened to conduct the counterfactual analysis. Therefore, the pre-covid total employment data (before 2020 Quarter 2) of length  $T_e = 142$  is divided into training and test portions, with a training set length ( $n = 120 = T_b$ ) and a test set length ( $H = 22 = T_e - h - T_b + 1$ ). In addition, I use a one-step forecasting experiment with the  $h = 1$ .

Here is a brief description of the proposed algorithm:



**Figure 4.1:** Proposed algorithm for selecting the optimal  $\lambda$

To mitigate the adverse impact of high-dimensionality, I start the algorithm from 0.0001, use steps of 0.0001 and stop at 0.3. There are 3000 different lambdas considered, and the algorithm will automatically return the lambda with minimum RMSFE in forecasting total number of employment (see Matlab code for this algorithm in my github file-Links in the Appendix ).

From the return value of our searching algorithm (see Figure 4.1). The estimated hyperparameter from the algorithm is  $\lambda = 0.0808$ , which certainly has the lowest mean scaled forecast error (RMSFE) as designed. Based on this, I choose  $\lambda = 0.0808$  for subsequent analysis.



## Chapter 5

# Sectoral Employment Analysis

### 5.1 Long-run Multiplier Analysis

Because industries are interdependent, changes in one subsector can affect both other subsectors as well as total employment. In this chapter, the sectoral employment dynamics for each subsector will be captured using the dynamic structure of the multivariate Bayesian VAR model.

The following analysis is based on the estimated BVAR model and 84 two-digit disaggregated data. I assume that the structure of Australian economy will not change after COVID-19.

At each time point, the growth rate of total employment and the growth rates of sectoral employment satisfy the following identity:

$$GR_T = \sum_{i=1}^{84} w_i \times GR_i$$

where  $w_i$  is the share of subsector  $i$ ,  $GR_T$  is the growth rate of total employment and  $GR_i$  is the growth rate in employment of subsector  $i$ .

In particular, if there is a one percent increase in employment in subsector  $i$ , total employment will increase by the corresponding share( $w_i$ ) simultaneously. Additionally, given an increase in total employment, it may also have indirect effects on other sectors in consecutive periods, especially those with close economic ties. In accordance with Anderson et al. (2020)'s definition, I define the employment long-run employment multiplier as the effect of an initial increase in sector  $i$  on total employment in the long run. If the subsector has a greater long-run effect on



total employment than its immediate effect, then stimulating this sector will have a positive spillover effect on total employment.

I simulate the long-term employment multiplier for each sector over time horizons of one year, two years and ten years using the estimated Bayesian VARX model. Subsequently, the differences between the simulated ten-year multipliers and the initial shares are the spillovers of the disaggregated subsectors. I list the 10 subsectors with strongest positive spillovers alongside the 10 subsectors with strongest negative spillovers in Table 5.1. A comprehensive list can be found in Appendix B (see Table B.2).

**Table 5.1:** *Disaggregated Subsectoral Long-run Employment Multipliers*

Sector/ Sub-sector	M10-M0	Sector/ Sub-sector	M10-M0
75 Admin/ Public Administration	-0.0356327	42 Retail/ Other Store-Based Retailing	0.01773361
80 Educ/ Preschool and School Education	-0.0320646	72 Admin/ Administrative Services	0.01339419
81 Educ/ Tertiary Education	-0.0215685	25Manu/ Furniture and Other Manufacturing	0.01141447
84 Health/ Hospitals	-0.016484	39 Retail/ Motor Vehicle and Motor Vehicle Parts Retailing	0.00966318
85 Health/ Medical and Other Health Care Services	-0.0112633	56 Info/ Broadcasting (except Internet)	0.00736334
82 Educ/ Adult, Community and Other Education	-0.0109159	18 Manu/ Basic Chemical and Chemical Product Manufacturing	0.00674883
46 Trans/ Road Transport	-0.0108955	33 Wholesale/ Basic Material Wholesaling	0.00646865
58 Info/ Telecommunications Services	-0.0084901	13 Manu/ Textile, Leather, Clothing and Footwear Manufacturing	0.00619223
11 Manu/ Food Product Manufacturing	-0.0082163	52 Trans/ Transport Support Services	0.00617277
01 Agri/ Agriculture	-0.0076574	94 Other/ Repair and Maintenance	0.00583089

Note: M10 is the 10-year long-run total employment multiplier and M0 is the shares of each sector. ;The M10-M0 is the spillover effect.; The M10/M0 is the spillover effect relative to the size of sector.

Comparing the long-term multipliers with the shares, we discover that “Other Store-Based Retailing”<sup>1</sup> will generate the greatest positive spillover effects on total employment, followed by “Administrative Services”<sup>2</sup>, and “Furniture and Other Manufacturing”. At the broadest level, these subsectors fall under the “Retailing”, “Administrative and Support Services” and “Manufacturing” sectors respectively. These results suggest that total employment will increase over and above the initial increase in these sectors if there is an exogenous increase in these sectors.

<sup>1</sup>This subsector contains the following groups: 421. Furniture, Floor Coverings, Houseware and Textile Goods Retailing; 422. Electrical and Electronic Goods Retailing; 423. Hardware, Building and Garden Supplies Retailing; 424. Recreational Goods Retailing; 425. Clothing, Footwear and Personal Accessory Retailing; 426. Department Stores; 427. Pharmaceutical and Other Store-Based Retailing

<sup>2</sup>This subsector contains the following groups: 721. Employment Services; 722. Travel Agency and Tour Arrangement Services; 729. Other Administrative Services.

There are also some interesting points to be noticed here. First, both the sign and the magnitude of spillover effect are not solely dependent the size of subsectors. For instance, “Fishing, Hunting and Trapping” is the smallest sector (see Table B.2). Nevertheless, it generates positive spillovers. In addition, “Furniture and Other Manufacturing” is a relative small subsector but generate stronger spillover than “Food and Beverage Services”, a relative large subsector<sup>3</sup>. Contrarily, both the “Professional, Scientific and Technical Services” and “Construction Services”, which are large subsectors but generate negative spillovers in the long run.

Second, I discover that subsectors in the “Construction” sector<sup>4</sup> may not bring strong positive spillovers in the two-digit level, which is in contrast to the result concluded by Anderson et al. (2020). The “Construction Service” subsector, for example, will not have positive spillovers even though the overall spillover effect of the “Construction” sector being positive<sup>5</sup>. This further demonstrates the necessity of expanding the research to a finer partition (more disaggregated level)

Third, both “Tertiary Education” and “Adult, Community and Other Education” generate negative spillovers, which implies that stimulating the education industry will probably reduce total employment in the long run. This is mainly because if one decides to pursue a postgraduate degree or a certificate, then the focus will shift away from working/finding jobs. The results are also in line with the broadest level (19 sectors) analysis that undertaken by Anderson et al. (2020).

## 5.2 Evaluations after COVID-19

### 5.2.1 Losses of Total Employment

Based on the fact that COVID-19 can spread quickly and there is no wonder drug to prevent outbreaks of new variants so far, we will never be able to completely halt the virus’ spread or prevent infection unless there is no human contact. As a result, it is anticipated that COVID-19 has caused massive losses to the Australian labour market and these negative effects will be persistent in the long run. To prove my idea and raise awareness about COVID-19, I will conduct

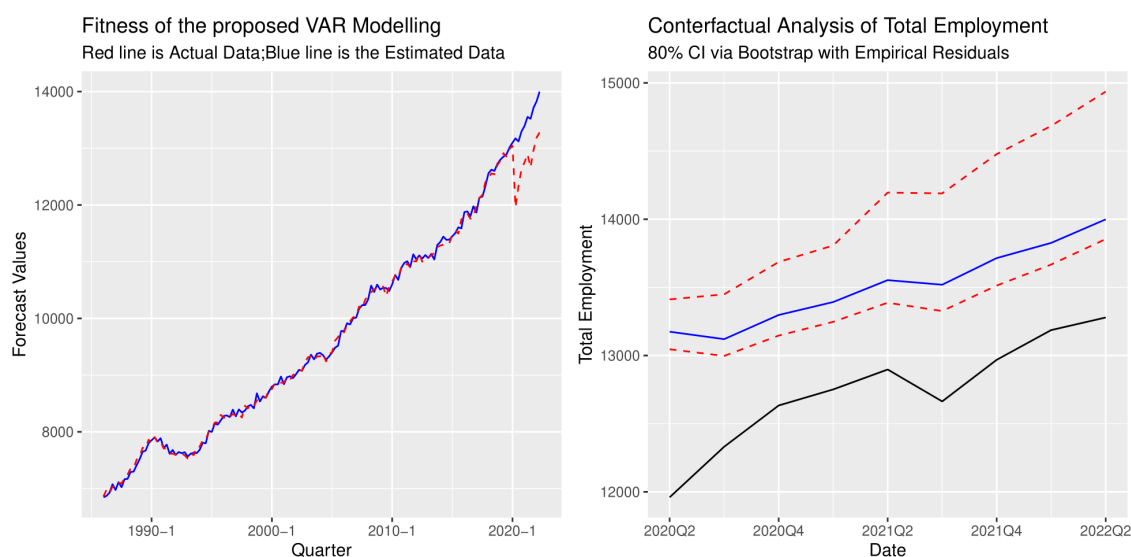
---

<sup>3</sup>The share of “Furniture and Other Manufacturing” is 0.00496526, whereas the share of “Food and Beverage Services” is 0.06348661

<sup>4</sup>Construction Sector contains three subsectors: 30. Building Construction; 31. Heavy and Civil Engineering Construction; 32. Construction Services

<sup>5</sup>The spillover for the “Construction” sector is equal to the sum of all its subsectors  $(-0.0004619 + 0.00271218 + 0.00066987 = 0.00292015)$ . In this case, it is still positive but no longer strong.

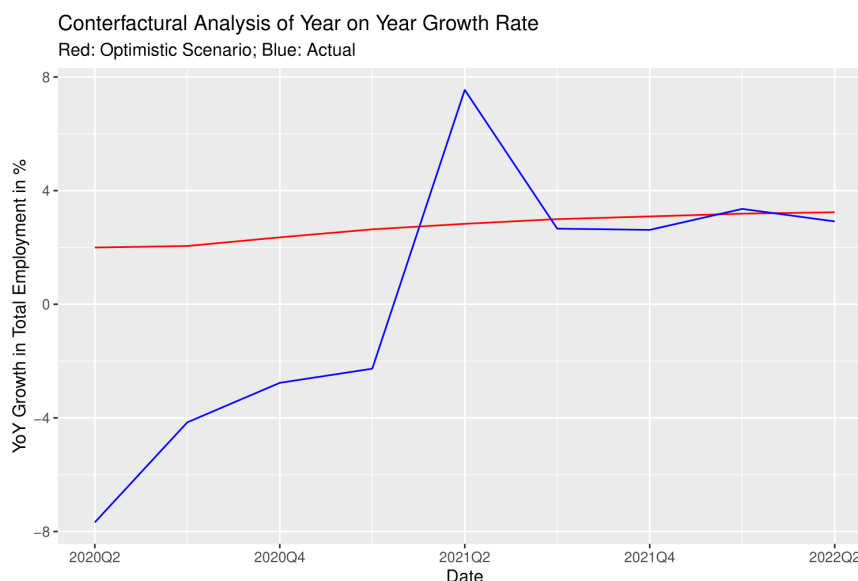
a counterfactual analysis by evaluating the difference of total employment with and without the pandemic. In this section, I use the estimated model to conduct a counterfactual analysis for total employment after the pandemic. Similar to Anderson et al. (2020), I will consider a “no-COVID” scenario. In light of the fact that the pandemic has already occurred, other scenarios are not taken into account. Recent data will be used to compare and evaluate the effects of COVID-19 on the labour market.



**Figure 5.1:** Counterfactual analysis of total employment(in thousands) with confidence interval generated via bootstrap (Blue is counterfactual scenario (no-COVID);Black is the actual data)

Figure 5.1 displays COVID-19 has caused a continuous structural shock of total employment in Australia since the outbreak of COVID-19. Based on the point forecasts together with an 80% confidence interval (via empirical bootstrap<sup>6</sup>), our model indicates that employment losses remained at approximately 750,000 persons below where it would have been without the pandemic (see the differences in Figure 5.1). By comparing the trend of the “no-pandemic” scenario and that of the actual data, the essentially parallel trend revealed that we may not expect total employment to attain the forecasts under the no-COVID case in the anticipated future. That is, COVID-19 has had a long-lasting effect on the economy even after two years of the primary lockdown in 2020 “Quarter 2”.

<sup>6</sup>As there are 7224 parameters to estimate, we do not use the Bayesian prediction intervals concerning the time complexity required when taking high-dimensional integrals.



**Figure 5.2:** Counterfactual analysis of Year-on-Year growth rate (forecasts generated via the estimated BVAR model)

Additionally, I have also considered the Year-on-Year growth rate for quarterly total employment data from 2020 Q2 to 2022 Q2 (up-to-date at the time of collecting). According to Figure 5.2, the actual employment growth was far below from the expected growth, particularly in 2020 Q2, when Australia was experiencing the first lockdown. After that, the year-on-year growth rate gradually recovered back to “no-COVID” forecasts, but total employment is still lower than it would be under the “no-COVID” situation. Again, it has been emphasized the finding above and we may not expect the influences of COVID-19 to disappear unless there is a higher year-on-year growth rate and persists for a while in the future. Therefore, it is reasonable to assume that COVID-19 has indeed a significant impact on the labour market both in the short term as well as in the long term based on the counterfactual analysis of employment and the year-on-year growth rate.

### 5.2.2 An explanation for the historically low unemployment rate

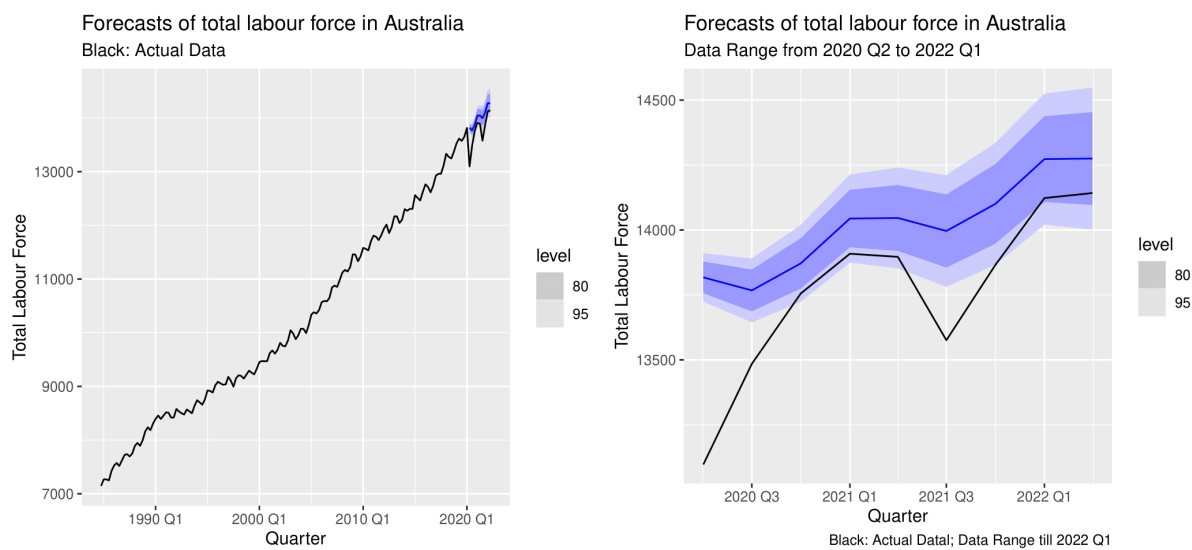
In June 2022, Australia fell to its the lowest unemployment rate since August 1974 (ABS (2022b)). Then one might be interested in learning what the underlying reason for this incredibly low unemployment rate. Are the stimulus policies during COVID-19 that have contributed the most to the low unemployment rate? In answering this question, I will use a counterfactual analysis of the unemployment rate to exploit why and what the unemployment rate would have been without the pandemic case.

To give an accurate interpretation of the low unemployment rate, the answer should refer to the definition, which is the percentage of people who are in the labour force but are unemployed. Mathematically, it is:

$$\text{Unemployment Rate} = \frac{\text{Total Labour Force} - \text{Number of Employed People}}{\text{Total Labour Force}}$$

It is obvious that the unemployment rate depends on both the total labour force and the number of employed people. According to the estimated BVAR model suggests that employment is less than it would have been in the absence of COVID (see *Chapter 5*). Therefore, given the low unemployment rate and fewer employed people than in the no-COVID scenario, one explanation might be a significant drop in the total labour force after the pandemic.

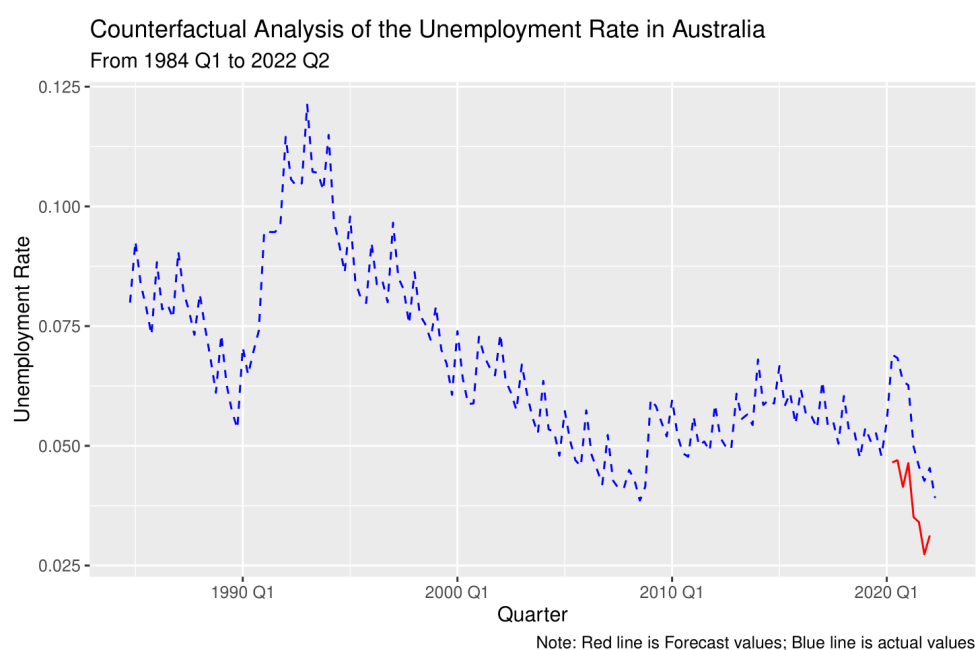
To further support this hypothesis, I use quarterly labour force data from ABS from “1984 Q4” to “2022 Q2” (ABS, 2022a). In addition, a stepwise ARIMA model (Hyndman, 2021) is used to fit the “no-COVID” data between “1984 Q4” and “2020 Q1” to forecast the total labour force under the “no-COVID” scenario (see Figure 5.3). Compared with the actual data, it is evident that the real total labour force is lower than the “no-COVID” forecasts at the time when the unemployment rate is at its lowest historically. Therefore, rather than the effects of stimulus policies during the pandemic period, the main reason for the low unemployment rate is the decline in the total labour force.



**Figure 5.3:** Counterfactual analysis of the total labour force (Blue: the forecasts under the no-COVID scenario, Black is the Actual Data)

After looking into the underlying reason for the low unemployment rate, I also studied how the unemployment rate would perform in the absence of COVID-19. The forecasts of the no-COVID unemployment rate is calculated as the difference between the total labour force and employment rate over the total labour force under the no-COVID scenario. It is noteworthy that the anticipated unemployment rate (under no-COVID scenario) can still be the lowest on record, even with a larger labour force and total employment.

Overall, the findings suggest that we would have experienced the historically low unemployment rate even in the absence of COVID-19. This further assures us that the policies directed at stimulating employment during the pandemic (e.g. Jobkeeper program) are unlikely to be responsible for the current historically low unemployment rate.



**Figure 5.4:** *Counterfactual analysis of the unemployment rate in Australia*



## Chapter 6

# Discussions and conclusion

In conclusion, I have developed a dynamic Bayesian VAR (BVAR) system to analyse the two-digit employment dynamics in Australia. This allows me to view the Australian employment market from a new perspective. First, an out-of-sample forecasting algorithm was proposed to select the most optimal hyperparameter in Minnesota Prior and also to improve the forecast accuracy. Second, I conducted the spillover analysis using the estimated model to discover how total employment reacts in the long run when the subsectors are stimulated. Based on the result, “Other Store-Based Retailing”, followed by “Administrative Services”, “Furniture and Other Manufacturing” will generate strong positive spillovers in employment. In particular, they will make greater contributions to total employment than their original sizes (i.e. the shares of subsectors). It is helpful for policymakers to consider industries that could operate efficiently, and act as the backbone of the economy (i.e. high-spillover industries) when formulating policies to stimulate the economy.

Furthermore, I evaluate the shocks of COVID-19 to the Australian labour market with the estimated BVAR model. This counterfactual analysis ends up with a conclusion that these structural shocks in fact have a long-term negative impact. As a result, it will be challenging to completely mitigate the influences of the pandemic in the foreseeable future. To shed light on the reason for the low unemployment rate in Australia, I conduct an empirical analysis and compare the estimated unemployment rate in the absence of the pandemic with actual values. Results suggest that labour force is less than it would have been without the pandemic, and we would have experienced a low unemployment rate even in the absence of COVID-19. Given



these points, I identify that the low unemployment rate is mainly driven by the loss of the total labour force rather than the stimulus policies during the pandemic.

## 6.1 Limitation

There are few non-classified data (nfd data) under each sector in the ABS employment by subdivision dataset (ABS, [2022a](#)). Data that are not classified have not been considered in this study. This is mainly because the data is not evenly distributed, which will be problematic if we simply distribute it into our system by the share of each subsector. I haven't yet thought of a good way to distribute the data. In the future, it would be great to come up with some effective ways (e.g. Contact ABS about detailed information) to trace the source of them. Consequently, the total employment data used in this thesis is not the same as the published total employment data as we discussed in *Chapter 3.1*. Nevertheless, we only consider the ANZSIC subdivision in our case (two-digit sectors). In the future, it would be beneficial to consider using both the Division level (the broadest level) and the group and classes (the finest level) using hierarchical forecasting. The predictions we examined are based on the assumption that the parameter of the model has been constant thought the time before COVID-19 and structure will not change after COVID-19, which might be restrictive. An appropriate time-varying model could be implemented to have more flexibility in modelling dynamics.

## 6.2 Future Extension: An Application of Machine Learning in Hierarchical Forecasting

In this research, the Bayesian VAR modelling method has provided a useful analysis of employment dynamics in Australia. However, VAR modelling may be inefficient, when there is more data and hierarchies are considered. Therefore, a more efficient machine learning technique could be considered as an extension of the hierarchical forecasting.

To improve the forecast accuracy, we could use machine learning to give a new way that accounts for both accuracy and interpretability. First, pick up the common features or data types from the data and cluster them into groups based on them. Second, conduct group-based forecasting for each new cluster. Then, we can reconcile them to be coherent, which will benefit the decision and policy implementation processes.

Here, I will give an example of the machine learning algorithm.

1. Cluster-bottom level data based on common features (e.g. domain-specific features, time series characteristics etc.) via possible machine learning algorithms (e.g. manifold learning, k-means).
2. Create a model to forecast each time series within the same cluster. Although it seems restrictive, there are two points should be helpful in proving both accuracy and flexibility. First, the choice of model is flexible and can be complex. There are many types of models to choose from, depending on the domain types and time series patterns. Moreover, the types of models are not limited, which can be either a single model (e.g. univariate, multivariate, and ML) or a combined model (e.g. Weighted average of various forecasting methods). Second, forecasting methods can be different for different clusters of the data, due to some unique patterns in time series.
3. Reconcile the forecasts produced by the machine learning algorithm to make them coherent. To get more interpretable results, reconciliation is required to fit all forecasts produced in our clusters into the original structure of the data (i.e. original groups).



## Appendix A

# An Example of Bayesian VAR and Minnesota Prior

The VARX model is:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{c} + \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{\Gamma}_1 \mathbf{x}_{t-1} + \mathbf{u}_t \\ &= \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} a_1^{11} & \cdots & a_1^{1n} & \gamma_1^1 \\ \vdots & \ddots & \vdots & \vdots \\ a_1^{n1} & \cdots & a_1^{nn} & \gamma_1^n \end{bmatrix} \begin{bmatrix} \mathbf{y}_{t-1} \\ x_{t-1} \end{bmatrix} \\ &\quad + \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{n,t} \end{bmatrix} \end{aligned}$$

where  $\mathbf{y}_t = \ln(\mathbf{z}_t) - \ln(\mathbf{z}_{t-4})$  and  $\mathbf{z}_t$  is the number of employment in subsectors;  $\mathbb{E}(\mathbf{u}_t \mathbf{u}_t') = \mathbf{\Sigma}$  and  $\mathbb{E}(\mathbf{u}_t \mathbf{u}_{t-1}') = 0$ . Here the  $n$  represent the number of sectors (in our case this will be 84) and  $\mathbf{c}$  represents the vector of constants. There is one lag ( $p=1$ ) included of the total employment growth for each equation ( $x_{t-1}$ ) as a predetermined variable at time  $t$ .

We estimate the VARX using Bayesian method via a natural-conjugate-Normal-Wishart prior which preserve the properties of the Minnesota prior. We apply the shrinkage to the VAR slope coefficients using a Minnesota Prior specification as follows:

$$E[a_i^{jk}] = E[\gamma_i^j] = 0$$

$$Var[a_i^{jk}] = \begin{cases} \frac{\lambda^2}{i^2}, & j = k \\ \frac{\lambda^2}{i^2} \frac{\sigma_j^2}{\sigma_k^2}, & otherwise \end{cases}$$

$$Var[\gamma_i^j] = \frac{\lambda^2}{i^2} \frac{\sigma_j^2}{\sigma_\epsilon^2}$$

The  $\sigma_\epsilon^2$  is estimated by taking the residual variances after fitting an AR(4) on the  $l^{th}$  variable using least squares, which is common practice (see Anderson et al. (2020); Bańbura, Giannone, and Reichlin (2010)). The degree of shrinkage is governed by  $\lambda$  and the  $i$  stands for number of lags.  $\frac{\sigma_j^2}{\sigma_\epsilon^2}$  adjust different scales of the data. The  $\lambda$  we apply in this thesis is selected in *Chapter 4.3*.

The nature conjugate Normal-Inverse-Wishart prior implies the posterior moments can be calculated either analytically or by using the dummy observations.

Then we implement our VAR by defining  $(np + n + 1)$  dummy observations:

To estimate the BVAR using dummy observations, we rewrite the estimated model as

$$Y = X\beta + u$$

where we set the prior parameters as  $Y = [y_1, \dots, y_T]$ ,  $X = [X_1, \dots, X_T]$  where  $X_t = [y_{t-1}, x_{t-1}]$  and  $u = [u_1, \dots, u_T]$ .

The Normal-Wishart prior distribution then take the form:

$$vec(\beta) | \Sigma \sim N(vec(\beta_0), \Sigma \otimes \Omega_0), and$$

$$\Sigma \sim IW(S_0, a_0)$$

where we set the prior parameters  $\beta_0$ ,  $\Omega_0$ ,  $S_0$  and  $a_0$  such that they are consistent with the Minnesota Prior setting.

The expectation of  $\Sigma$  being  $diag(\sigma_1^2, \dots, \sigma_n^2)$ . We follow Anderson et al. (2020) to set our prior by defining dummy observations:

$$\begin{aligned}
 Y_d &= \begin{pmatrix} \mathbf{0}_{np+p,n} \\ \text{diag}(\sigma_1, \dots, \sigma_n) \\ \mathbf{0}_{1 \times n} \end{pmatrix}, \\
 X_d &= \begin{pmatrix} J_p \otimes \text{diag}(\frac{\sigma_1}{\lambda} \dots \frac{\sigma_n}{\lambda}, \frac{\sigma_e}{\lambda}) & \mathbf{0}_{(np+p) \times 1} \\ \mathbf{0}_{n,np+p} & \mathbf{0}_{n \times 1} \\ \mathbf{0}_{1,np+p} & \epsilon \end{pmatrix},
 \end{aligned}$$

where  $Y_d$  and  $X_d$  are the dummy observations chosen according to the Minnesota Prior assumption (consistent with the mean and variance setups above).

$$\begin{aligned}
 J_p &= \text{diag}(1, \dots, p), \\
 S_0 &= (Y_d - X_d B_0)'(Y_d - X_d B_0), \\
 B_0 &= (X_d' X_d)^{-1} X_d' Y_d, \quad \Omega_0 = (X_d' X_d)^{-1} \text{ and} \\
 a_0 &= T_d - np - p - 1,
 \end{aligned}$$

where  $T_d$  is the number of rows for both  $Y_d$  and  $X_d$ .  $\epsilon$  is a very small number to impose an uninformative and diffused prior on the constants. The first block of the dummy observations imposes the prior belief on the VAR slope coefficients and the second block contains the prior for the covariance matrix and third block imposes the prior belief on the constants.

Then we augment the original BVAR model with the estimated dummy observations. We can get:

$$\begin{aligned}
 Y^* &= X^* \beta + \mu^* \quad \text{where :} \\
 Y^* &= [Y', Y_d']'; \quad X^* = [X', X_d']'; \quad \mu^* = [\mu', \mu_d']'
 \end{aligned}$$

Then we can estimating the BVAR by conducting least squares regression of  $Y^*$  on  $X^*$ . The posterior distribution then has the form of

$$\begin{aligned}
 \text{vec}(\beta) | \Sigma, Y &\sim N(\text{vec}(\tilde{\beta}), \Sigma \otimes (X^{*'} X^*)^{-1}) \text{ and} \\
 \Sigma | Y &\sim IW(\tilde{\Sigma}, T_d + T - np + 2)
 \end{aligned}$$

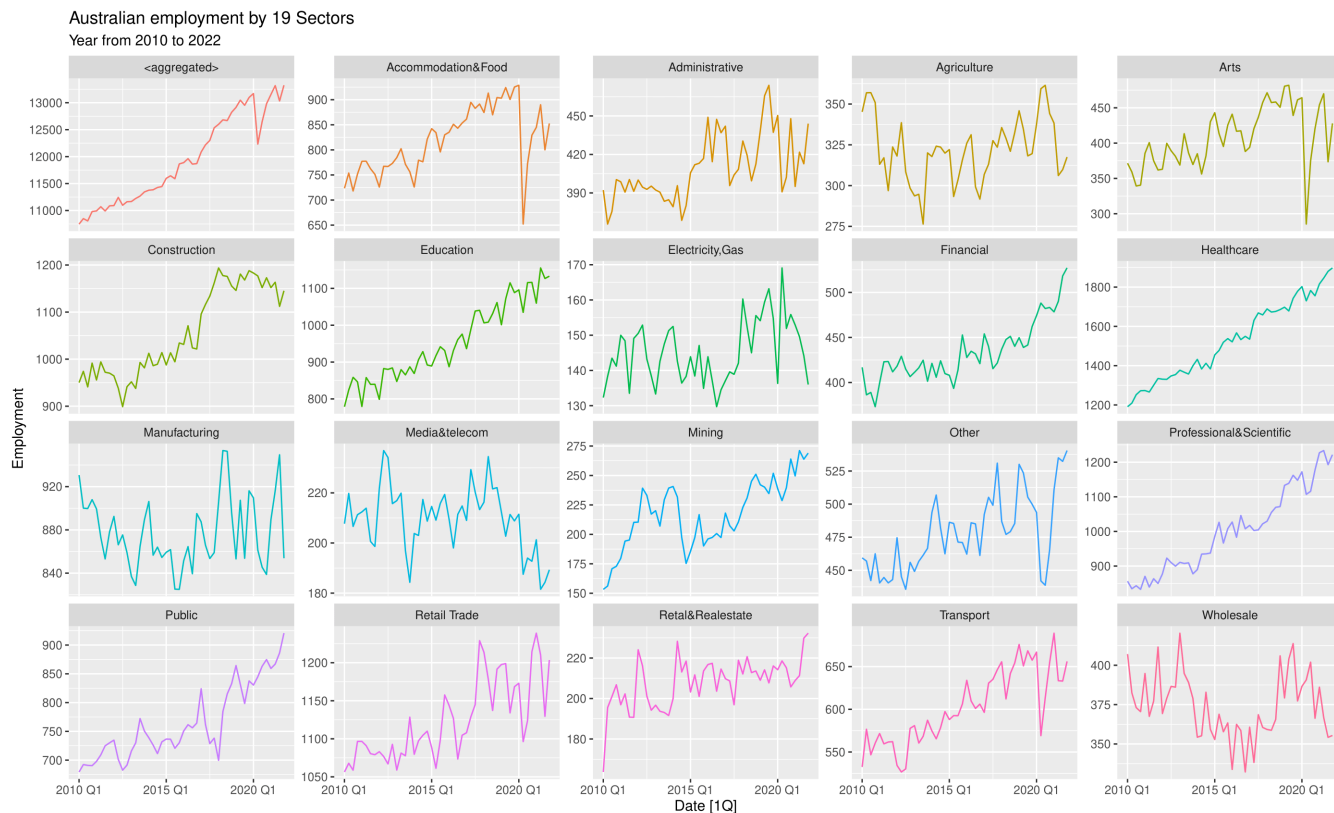
where  $\tilde{\beta} = (X^{*'} X^*)^{-1} X^{*'} Y^*$  and  $\tilde{\Sigma} = (Y^* - X^* \tilde{\beta})'(Y^* - X^* \tilde{\beta})$



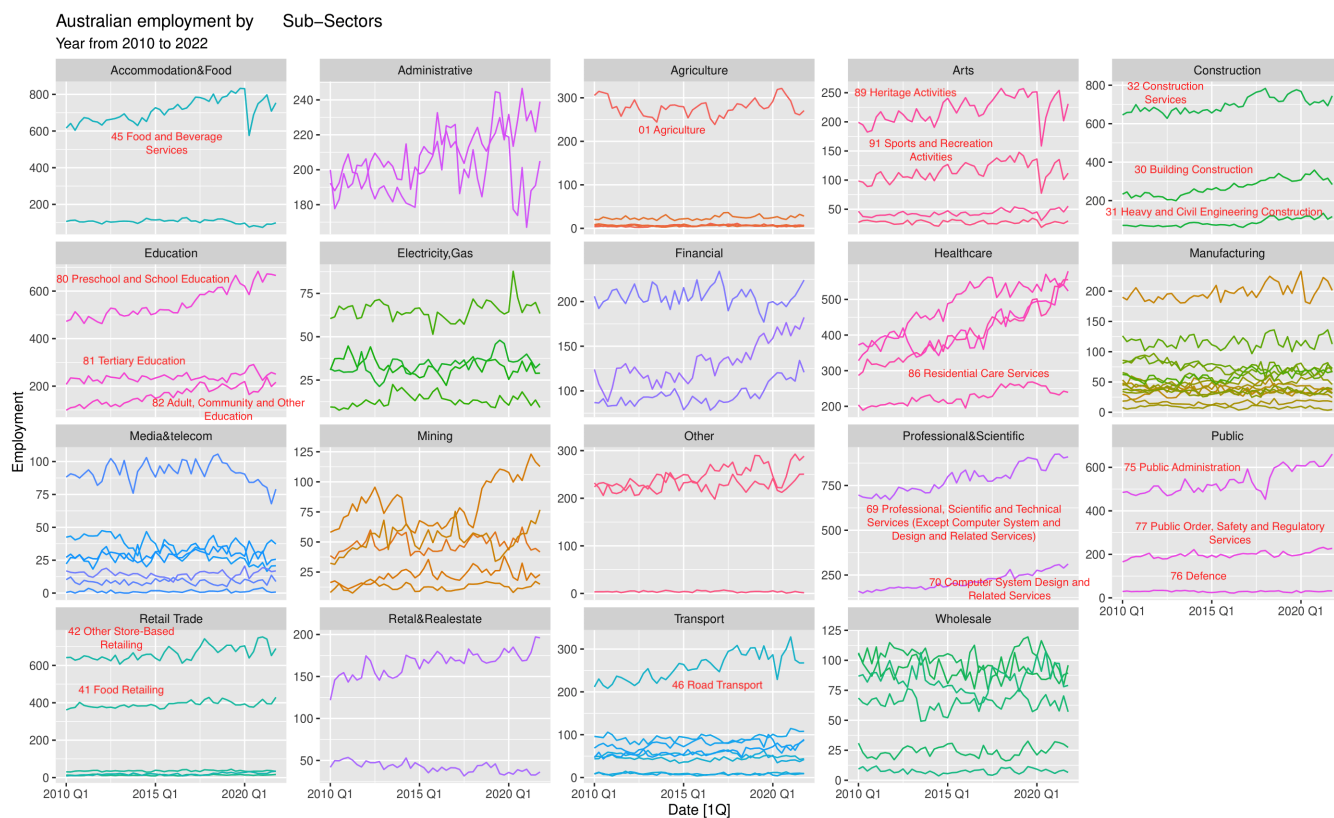
## **Appendix B**

### **Graphs**





**Figure B.1:** Employment('000) of 19 sectors in Australia from 2010 to 2022



**Figure B.2:** Employment('000) of two-digit subsectors in Australia from 2010 to 2022

**Table B.1: The long run Employment Multipliers Analysis (84 Disaggregated Sectors) Sorted by Shares of Subsectors**

Sub-Sector	Shares	1 Year	2 Years	5 Years	10 Years	M10/M0	M10-M0
69 Professional, Scientific and Technical Services (Except Computer System) Design and Related Services)	0.06931674	0.0485337	0.062318	0.06666942	0.06671967	0.96253315	-0.0025971
45 Food and Beverage Services	0.06348661	0.04693738	0.06311697	0.0666367	0.06668314	1.05034955	0.00319652
32 Construction Services	0.05913525	0.04299346	0.05568259	0.05863971	0.05867338	0.99218952	-0.0004619
42 Other Store-Based Retailing	0.05378582	0.04963343	0.06655111	0.07145667	0.07151944	1.32970786	0.01773361
80 Preschool and School Education	0.0492665	0.02091953	0.01855257	0.01721796	0.01720194	0.34916098	-0.0320646
75 Public Administration	0.04641708	0.01553789	0.01250828	0.01080403	0.01078435	0.23233592	-0.0356327
85 Medical and Other Health Care Services	0.04116996	0.02431334	0.02916687	0.02989798	0.02990662	0.72641828	-0.0112633
84 Hospitals	0.0369132	0.01785159	0.0199097	0.02042347	0.02042915	0.55343755	-0.016484
87 Social Assistance Services	0.03680509	0.02207106	0.02929537	0.0311832	0.03120476	0.84783798	-0.00056003
41 Food Retailing	0.03040546	0.01923095	0.02775936	0.03134743	0.0313965	1.03259348	0.00099102
30 Building Construction	0.02365835	0.01966493	0.02362512	0.02432029	0.02432822	1.02831441	0.00066987
46 Road Transport	0.02212746	0.01012106	0.01115604	0.01122802	0.011232	0.50760455	-0.0108955
01 Agriculture	0.02177417	0.00998719	0.01235378	0.01409829	0.01411674	0.64832471	-0.0076574
95 Personal and other services (include activities for own use)	0.02127031	0.0127474	0.01505848	0.01597865	0.01599074	0.75178666	-0.0052796
86 Residential Care Services	0.02028961	0.0108231	0.01516439	0.01583014	0.0158379	0.78059125	-0.0044517
70 Computer System Design and Related Services	0.01989965	0.01064273	0.01220485	0.0126649	0.01267126	0.63675763	-0.0072284
81 Tertiary Education	0.01945371	0.00166075	-0.0015457	-0.0021075	-0.0021148	-0.1087109	-0.0215685
72 Administrative Services	0.01790544	0.02068182	0.02911507	0.03127308	0.03129963	1.74805115	0.01339419
94 Repair and Maintenance	0.01777003	0.01434611	0.02128838	0.02357418	0.0236012	1.32812526	0.00583089
73 Building Cleaning, Pest Control and Other Support Services	0.01736876	0.01180556	0.01573853	0.01716652	0.01718243	0.98927169	-0.0001863
11Food Product Manufacturing	0.01649038	0.00702254	0.00846642	0.00827582	0.00827409	0.50175246	-0.0082163
82 Adult, Community and Other Education	0.01566605	0.00554203	0.00471415	0.00475267	0.00475016	0.30321348	-0.0109159
77 Public Order, Safety and Regulatory Services	0.01521818	0.00871417	0.01309719	0.01440524	0.01442184	0.94767203	-0.0007963
62 Finance	0.01472397	0.00857652	0.01062002	0.0110417	0.01104891	0.75040274	-0.0036751
67 Property Operators and Real Estate Services	0.01358111	0.00573538	0.00632144	0.00644085	0.00644122	0.47427795	-0.0071399
64 Auxiliary Finance and Insurance Services	0.01229925	0.00710233	0.00875331	0.00932324	0.00933078	0.75864648	-0.0029685
91 Sports and Recreation Activities	0.01027608	0.00887835	0.01209347	0.01274308	0.01275193	1.24093321	0.00247585
24Machinery and Equipment Manufacturing	0.0087394	0.00384771	0.00486625	0.00522805	0.00523156	0.59861717	-0.0035078
34 Machinery and Equipment Wholesaling	0.00863129	0.00463038	0.00550646	0.00586781	0.00587311	0.68044432	-0.0027582
39Motor Vehicle and Motor Vehicle Parts Retailing	0.00844596	0.01088201	0.01646315	0.01809008	0.01810914	2.14411792	0.00966318
08 Metal Ore Mining	0.00838419	0.00574686	0.00556897	0.0053132	0.00530956	0.6332827	-0.0030746
31 Heavy and Civil Engineering Construction	0.00832434	0.00731709	0.01027104	0.01102691	0.01103652	1.32581315	0.00272118
63 Insurance and Superannuation Funds	0.00805214	0.00443011	0.00497832	0.00501181	0.00501281	0.62254327	-0.0030393
51 Postal and Courier Pick-up and Delivery Services	0.00757723	0.00121361	0.00131748	0.00140896	0.00141185	0.18632811	-0.0061654
44 Accommodation	0.00753476	0.0032441	0.00545	0.00634005	0.00635098	0.84289032	-0.0011838
37 Other Goods Wholesaling	0.0071892	0.00459914	0.00631193	0.00661623	0.00662058	0.92090624	-0.0005686
58 Telecommunications Services	0.00707916	0.00011311	-0.001374	-0.0014087	-0.001411	-0.1993137	-0.0084901
33 Basic Material Wholesaling	0.00697878	0.00828557	0.01222171	0.01343315	0.01344743	1.92690232	0.00646865
52 Transport Support Services	0.00675677	0.00753772	0.01162295	0.01291333	0.01292955	1.91356851	0.00617277
21Primary Metal and Metal Product Manufacturing	0.0058726	0.00462923	0.00499628	0.00494162	0.00494054	0.84128732	-0.0009321
22Fabricated Metal Product Manufacturing	0.00538804	0.0066765	0.00944416	0.01058564	0.01059958	1.96724154	0.00521154
53 Warehousing and Storage Services	0.00527028	0.00263016	0.00272967	0.00287253	0.00287296	0.54512411	-0.0023973
23Transport Equipment Manufacturing	0.00514673	0.00299308	0.00469464	0.00556819	0.0055781	1.08381385	0.00043137
26Electricity Supply	0.00502125	0.00491081	0.00586741	0.00563882	0.005636	1.12243022	0.00061475
25Furniture and Other Manufacturing	0.00496526	0.00907029	0.01434533	0.01635773	0.01637974	3.29886574	0.01141447
36 Grocery, Liquor and Tobacco Product Wholesaling	0.00491893	0.00061785	-0.0006336	-0.0008216	-0.0008238	-0.1674704	-0.0057427
49 Air and Space Transport	0.00419885	0.00350454	0.00513191	0.00542985	0.00543509	1.29442326	0.00123624
18 Basic Chemical and Chemical Product Manufacturing	0.00395754	0.00606204	0.00950461	0.01069071	0.01070637	2.70531007	0.00674883
06 Coal Mining	0.00383591	0.00383353	0.00407393	0.00408138	0.00408139	1.06399293	0.00024547
47 Rail Transport	0.00381661	0.00126166	0.00105419	0.00112838	0.00112925	0.29587899	-0.0026874
90 Creative and Performing Arts Activities	0.00360425	0.00468914	0.00613651	0.00651375	0.00651912	1.80878281	0.00291486
14Wood Product Manufacturing	0.00347877	0.00302624	0.00302602	0.00331986	0.00332296	0.9552116	-0.0001558
29Waste Collection, Treatment and Disposal Services	0.00339383	0.00565936	0.00756816	0.0078664	0.0078681	2.31835492	0.00447427
89 Heritage Activities	0.00301159	0.00369792	0.0051636	0.00528106	0.00528232	1.7539992	0.00227074
10 Exploration and Other Mining Support Services	0.00299228	-0.0006873	-0.0017035	-0.0019049	-0.0019072	-0.6373821	-0.0048995
55 Motion Picture and Sound Recording Activities	0.00294209	0.00032963	-0.0005498	-0.0005605	-0.0005605	-0.1905091	-0.0035026
40 Fuel Retailing	0.00291506	-0.0012046	-0.0024966	-0.0027078	-0.0027087	-0.9292227	-0.0056238
66 Rental and Hiring Services (except Real Estate)	0.00287066	0.003369	0.00341671	0.00336895	0.00336725	1.17298563	0.00049658
19Polymer Product and Rubber Product Manufacturing	0.00283012	0.00233992	0.00331413	0.00359534	0.00359859	1.27153097	0.00076847
20Non-Metallic Mineral Product Manufacturing	0.00277221	0.00357344	0.00319618	0.0030411	0.0030398	1.0965273	0.00026759
16Printing (including the Reproduction of Recorded Media)	0.00262163	-0.000264	0.00122375	0.00182629	0.00183417	0.69963121	-0.0007875
12Beverage and Tobacco Product Manufacturing	0.00252896	0.00277597	0.00231169	0.00223403	0.00223473	0.88365682	-0.0002942
28Water Supply, Sewerage and Drainage Services	0.00249614	0.00282008	0.00239868	0.00201716	0.0020111	0.80568129	-0.000485
13Textile, Leather, Clothing and Footwear Manufacturing	0.00244016	0.00487731	0.00773749	0.00862047	0.00863239	3.53763202	0.00619223
92 Gambling Activities	0.00243437	0.00430809	0.00667533	0.00741282	0.00742126	3.04853477	0.00498689
07 Oil and Gas Extraction	0.00232626	-0.0017933	-0.0034336	-0.0039433	-0.0039495	-1.6977721	-0.0062757
56 Broadcasting (except Internet)	0.00225097	0.00562035	0.00883845	0.00960654	0.00961432	4.27118764	0.00736334
35 Motor Vehicle and Motor Vehicle Parts Wholesaling	0.00208109	0.00130996	0.0019761	0.0022859	0.00228915	1.09997643	0.00020806
76 Defence	0.00203668	0.00108328	0.00081797	0.00081458	0.00081483	0.40007524	-0.0012219
05 Agriculture, Forestry and Fishing Support Services	0.00201545	0.00410222	0.00612542	0.00684863	0.00685748	3.40246046	0.00484203
54 Publishing and Broadcasting	0.00199035	0.00146161	0.00254755	0.00299741	0.00300331	1.50893341	0.00101296
43 Non-Store Retailing and Retail Commission Based Buying and/or Selling	0.00198842	0.00008765	-0.0006368	-0.0010072	-0.0010108	-0.5083429	-0.0029992
15Pulp, Paper and Converted Paper Product Manufacturing	0.00134556	0.00088098	0.00127857	0.00145108	0.00145231	1.07933418	0.00010675
60 Library and Other Information Services	0.00113321	0.00262681	0.00408402	0.00438669	0.00438973	3.8737221	0.00325652
09 Non-Metallic Mineral Mining and Quarrying	0.00109846	0.00301024	0.00035283	0.00047586	0.00047618	0.43350017	-0.0006223
27Gas Supply	0.00093436	0.00119372	0.00200704	0.00226239	0.00226542	2.42456066	0.00133106
38 Commission-Based Wholesaling	0.00073359	0.00117105	0.00148834	0.00157279	0.00157404	2.14566446	0.00084045
59 Internet Service Providers, Web Search Portals and Data Processing Services	0.00071236	-0.0002143	-0.0006387	-0.0006332	-0.0006331	-0.8887063	-0.0013454
48 Water Transport	0.00067182	-0.0003845	-0.0003261	-0.0002884	-0.0002886	-0.4295098	-0.0009604
17Petroleum and Coal Product Manufacturing	0.00066795	0.00026082	-0.0002759	-0.0005609	-0.0005649	-0.8457507	-0.0012329
50 Other Transport	0.00062162	0.00128362	0.00155246	0.00160919	0.00160982	2.58970933	0.0009882
03 Forestry and Logging	0.00053668	-0.0012108	-0.0017314	-0.0018066	-0.0018077	-3.3683565	-0.0023444
02 Aquaculture	0.00050193	-0.0001496	-0.0003808	-0.0004863	-0.0004873	-0.9709079	-0.0009893
04 Fishing, Hunting and Trapping	0.00046139	0.00209291	0.00329398	0.00352863	0.00353112	7.65320735	0.00306973

*Note:* The long run employment multiplier for subsector  $i$  is the effect of a 1% increase in employment of subsector on total employment in the long run.

M10 is the long term multiplier; M0 is the share of subsector  $i$  in total employment. M10-M0 is the spillover of the subsector  $i$ .

M10/M0 is the spillover of the subsector  $i$  relative to the share of it.

**Table B.2: Disaggregated Sectoral Long-Run Employment Multipliers: Full list of 84 sectors (Sorted by Shares)**



## **Appendix C**

# **Supplementary Materials (Data & Codes)**

Supplementary materials include figures, codes and documentations can be found in the Github repo of this project: [Link: [https://github.com/elvisssyang/Disaggregated\\_Employment/tree/main/Matlab](https://github.com/elvisssyang/Disaggregated_Employment/tree/main/Matlab)]. Please contact me via Github or email me at `zyan0056@student.monash.edu` if something is unclear.



# Bibliography

- ABS (2021). *One year of COVID-19: Aussie jobs, business and the economy*. <https://www.abs.gov.au/articles/one-year-covid-19-aussie-jobs-business-and-economy> (visited on 04/29/2022).
- ABS (2022a). *Labour Force, Australia, Detailed, May 2022*. <https://www.abs.gov.au/statistics/labour/employment-and-unemployment/labour-force-australia-detailed/latest-release> (visited on 06/27/2022).
- ABS (2022b). *Unemployment rate falls to 3.5 percent*. <https://www.abs.gov.au/media-centre/media-releases/unemployment-rate-falls-35> (visited on 10/01/2022).
- Anderson, HM, G Caggiano, F Vahid, and B Wong (2020). Sectoral Employment Dynamics in Australia and the COVID-19 Pandemic. *Australian Economic Review* **53**(3), 402–414.
- Bañbura, M, D Giannone, and L Reichlin (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics* **25**(1), 71–92.
- Bengtsson, H (2021). *matrixStats: Functions that Apply to Rows and Columns of Matrices (and to Vectors)*. R package version 0.61.0. <https://CRAN.R-project.org/package=matrixStats>.
- Borchers, HW (2022). *pracma: Practical Numerical Math Functions*. R package version 2.3.8. <https://CRAN.R-project.org/package=pracma>.
- De Mol, C, D Giannone, and L Reichlin (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics* **146**(2), 318–328.
- Gregory, V, G Menzio, and D Wiczer (2020). Pandemic recession: L-shaped or V-shaped. *Covid Economics* **88**.
- Grolemund, G and H Wickham (2011). Dates and Times Made Easy with lubridate. *Journal of Statistical Software* **40**(3), 1–25.

- Hocking, TD (2021). *directlabels: Direct Labels for Multicolor Plots*. R package version 2021.1.13. <https://CRAN.R-project.org/package=directlabels>.
- Hyndman, R (2021). *fpp3: Data for "Forecasting: Principles and Practice" (3rd Edition)*. R package version 0.4.0. <https://CRAN.R-project.org/package=fpp3>.
- Kadiyala, KR and S Karlsson (1997). Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics* **12**(2), 99–132.
- Kassambara, A (2020). *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.4.0. <https://CRAN.R-project.org/package=ggpubr>.
- Koop, GM (2013). Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics* **28**(2), 177–203.
- Lewis, B and T Hsu (2020). *The Collateral Damage of Coronavirus*. <https://www.nytimes.com/2020/05/09/business/economy/coronavirus-unemployment.html> (visited on 04/29/2022).
- Litterman, RB et al. (1979). *Techniques of forecasting using vector autoregressions*. Tech. rep.
- Litterman, RB (1986). Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics* **4**(1), 25–38.
- Ludvigson, SC, S Ma, and S Ng (2021). COVID-19 and the Costs of Deadly Disasters. In: *AEA Papers and Proceedings*. Vol. 111, pp.366–70.
- R Core Team (2021a). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>.
- R Core Team (2021b). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>.
- Robertson, JC and EW Tallman (1999). Vector autoregressions: forecasting and reality. *Economic Review-Federal Reserve Bank of Atlanta* **84**(1), 4.
- RStudio Team (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC. Boston, MA. <http://www.rstudio.com/>.
- Slowikowski, K (2021). *ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'*. R package version 0.9.1. <https://CRAN.R-project.org/package=ggrepel>.
- Stock, JH and MW Watson (2001). Vector autoregressions. *Journal of Economic Perspectives* **15**(4), 101–115.
- Tanaka, E and R Hyndman (2022). *monash: Consolidated tools and templates for staff at Monash University*. R package version 0.0.0.9000.

- Waggoner, DF and T Zha (1999). Conditional forecasts in dynamic multivariate models. *Review of Economics and Statistics* **81**(4), 639–651.
- Warsono, W, E Russel, W Wamiliana, W Widiarti, and M Usman (2019). Vector autoregressive with exogenous variable model and its application in modeling and forecasting energy data: Case study of PTBA and HRUM energy. *International Journal of Energy Economics and Policy* **9**(2), 390–398.
- Wickham, H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, H, M Averick, J Bryan, W Chang, LD McGowan, R François, G Grolemund, A Hayes, L Henry, J Hester, M Kuhn, TL Pedersen, E Miller, SM Bache, K Müller, J Ooms, D Robinson, DP Seidel, V Spinu, K Takahashi, D Vaughan, C Wilke, K Woo, and H Yutani (2019). Welcome to the tidyverse. *Journal of Open Source Software* **4**(43), 1686.
- Wickham, H, R François, L Henry, and K Müller (2022). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.8. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, H, J Hester, and J Bryan (2022). *readr: Read Rectangular Text Data*. R package version 2.1.2. <https://CRAN.R-project.org/package=readr>.
- Woźniak, T (2016). Bayesian vector autoregressions. *Australian Economic Review* **49**(3), 365–380.
- Xie, Y (2016). *Bookdown: Authoring Books and Technical Documents with R Markdown*. ISBN 978-1138700109. Boca Raton, Florida: Chapman and Hall/CRC. <https://bookdown.org/yihui/bookdown>.
- Yutani, H (2022). *gghighlight: Highlight Lines and Points in 'ggplot2'*. R package version 0.3.3. <https://CRAN.R-project.org/package=gghighlight>.
- Zeileis, A and G Grothendieck (2005). zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software* **14**(6), 1–27.