Sectoral Employment Dynamics in Australia

(Elvis) Zhixiang Yang (ID:30306396)
Email :zyan0056@student.monash.edu

Farshid Vahid
Email :farshid.vahid@monash.edu

Report for
EBS Honours Research Plan, supervised by **Farshid Vahid**

9 May 2022

1 Introduction

The COVID-19 pandemic has had a massive effect on economies around the world. Across different countries, millions of workers were furloughed or even lost their jobs as businesses struggled to survive (Lewis & Hsu, 2020). Same situation happened in Australia, due to more restrictions, many businesses closed their doors, while employees were working with less hours or being dismissed by companies. As a result of the continuous "lockdown" periods in 2020, the estimates made by the Australian Bureau of Statistics (ABS, 2021) concluded that 72% of businesses generated less revenue and the underemployment rate hit a historical high with 13.8% by the end of April, 2020, only one month after the COVID-19 outbreak.

Our research is motivated by the lack of quantitative research on the employment of two-digit sectors in Australia, as many studies have focused on the aggregated employment rate. A general problem of aggregated researches is the loss of hierarchical information, which may result in a biased conclusion or "an illusion of employment prosperity". Thus, a quantitative analysis of the sectoral employment will overcome this problem, giving us a better scope to evaluate the impacts of COVID-19 in Australia.

2 Research aim and questions

This research will extend Anderson et al. (2020) by using data on 87 two-digit subsectors instead of 19 that they used. I will develop a model for the two-digit subsectors to evaluate the long run effect and the COVID-19 post-impacts. I will also provide a contrafactual analysis based on the assumption "if there is no pandemic". The two-digit subsectoral data will provide us with more information, which will assist in getting a better understanding of employment dynamics in Australia.

The overall research aim is to provide estimates of two-digit subsectoral employment based on historical data. Specifically, my goals are:

1. To construct a time series model of employment in 87 two-digit subsectors of the Australian economy.

2. To use this model to conduct a contrafactual analysis.

3. To use this model to determine which sectors have the highest impact (or positive spillover) on employment growth in the long run.

3 Review of literature

Our review of literature mainly focuses on two areas:

1. The COVID-19 sectoral impacts and modelling to Economies.

2. Modelling of large numbers of time series.

3.1 Sectoral Impact of COVID-19.

Most existing studies have focused on the evaluation of the impacts of COVID-19 on broad sectors of large economies such as the US and Europe. Ludvigson, Ma, and Ng (2020) developed a disaster series to translate the macroeconomic impact of costly and deadly disasters in recent US history and modelled them as sectoral shocks to predict COVID-19, concluding that the shock would lead to a cumulative loss of 20% in industrial production, 39% in public services and also reduce the US GDP by 12.75 per cent at the end of 2020. Gregory, Menzio, and Wiczer (2020) conducted simulations under different scenarios via a search theoretic model using US data and found the recovery in the US is L-shaped, with employment remaining lower than pre-covid for a long period. They also extended their studies at a disaggregated level of 20 sectors, showing that "arts and entertainment" and "accommodation and food services" sectors would have the biggest shock during the pandemic.

In Australia, Anderson et al. (2020) developed a multivariate time series for 19 main sectors in Australia (as a small open economy) using a Bayesian VARX model. Their research concluded that "Manufacturing" and "Construction" have the highest positive spillovers for the aggregate economy. Meanwhile, they also applied a "conditional forecasting" method proposed by Waggoner and Zha (1999) to simulate different scenarios for the pandemic in Australia. However, this research does not use a more disaggregated level in Australia (two-digit subsectors of main sectors), which can be extremely useful in macroeconomic analysis.

3.2 Modelling

3.2.1 Baysian VAR

The Bayesian Vector Autoregression model(BVAR) is commonly used in the literature for multivariate, large-size data (e.g. Anderson et al., 2020; Litterman, 1986; Bańbura, Giannone, & Reichlin, 2010). The BVAR model is attractive because it allows us to estimate a large number of parameters, when the sample size is not large, in a statistically coherent way.(Litterman, 1986; Woźniak, 2016).

In order to utilize the Bayesian VAR estimators and decrease the weight of the lagged variable with the lag length, Litterman et al. (1979) proposed the Minnesota Prior. 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. (b) own lags should explain more than the lags of other variables.

3.2.2 Improvement of BVAR

The literature suggests that a significant improvement can be made in the large BVAR dynamic model by imposing a stronger shrinkage parameter (Bańbura, Giannone, & Reichlin, 2010; Litterman, 1986). Robertson and Tallman (1999) and Kadiyala and Karlsson (1997) proposed a Normal-inverse-Wishart prior which retains the principal of Minnesota prior. Meanwhile, Bańbura, Giannone, and Reichlin (2010) suggested an easier way to apply the Minnesota prior via adding dummy observations in the BVAR system.

4 Data collection and exploratory analysis

4.1 Data Introduction

Our data source is the ABS Employment by industry subdivision of main job (ABS, 2022), which records employment (measured in *thousands*) by the ANZSIC industry sub-division of their main jobs from $1984 : Q4$ to $2021 : Q4$. Our data structure is provided via Figure 5.

Although seasonally adjusted data is available in (ABS, 2022), I will work with the original data to capture any possible changes in seasonal patterns. I will apply a seasonal difference the logarithm of the original series in order to make them stationary and eliminate seasonality, which will make it easier to conduct further steps.

4.2 Preliminary Exploratory Data Analysis

Date	Sector	Percentage Changes
2020:Q2	27 Gas Supply	50.48%
2020: Q2	57 Internet Publishing and Broadcasting	44.44%
2020: Q2	56 Broadcasting	41.43%
2020: Q2	26 Electricity Supply	36.34%
2020: Q2	20 Non-Metallic Mineral Product Manufacturing	32.37%

Date	Sectors	Percentage Changes
2020: Q2	03 Forestry and Logging	-55.88%
2020: Q2	96 Private Households Employing Staff	-49.09%
2020: Q2	60 Library and Other Information Services	-46.60%
2020: Q2	02 Aquaculture	-44.58%
2020: Q2	91 Sports and Recreation Activities	-43.71%

Table 1: *The highest and lowest five two-digit subsectors' employment percentage change for 2020:Q1 to 2020:Q2*

Figure 1 illustrates the changes of raw data for 19 main sectors in Australia during the pandemic. Due to the closedown of businesses and travel bans in 2020:Q2, we can observe that the total employment number dropped substantially (from around 13,200,000 to 12,200,000 on 2020 : Q2). Most industries behaved similarly with significant changes shown in Figure 1 . Comparing with the previous data of these industries," Accommodation & Food", "Media & telecom" and Administrative industries have experienced a severe loss of employment and have not fully recovered to the pre-covid level. However, some industries like "Financial" and "Electricity & Gas" showed a continuously increasing trend as before.

Nevertheless, there is a drawback of considering the 19 board sectors only; because the two-digit subsectoral dynamics of these sectors may not be homogeneous with their aggregated sectoral changes. For example, when observing the aggregated performance of the "Manufacturing" and "Mining" sectors from the 19 sectoral level (see Figure 3), we may believe that their corresponding subsetors should illustrate the same pattern. However, the reality is that while there is a decreasing trend in the "Manufacturing" sector or an increasing trend in the "Mining" sector, some of their two-digit subsectors are performed differently (see Figure 4). This means that not all two-digit subsectors follow the same pattern with the aggregated sectoral level.

Table 1 shows the top five and bottom five two-digit subsectors in terms of their employment growth during the pandemic. From 1 we can conclude that the "Forestry and Logging" experienced a severe shock after the lockdown happened on 2020:Q2, followed by "Private Households Employing Staff" and "Library and Other Information Services". Figure 2 demonstrates the performance of each industry at a more disaggregated two-digit subsectoral level, we can observe that many two-digit subsectors have also shown huge decreases in employment in 2020:Q2.

5   Methdology

5.1   Proposed Model

I plan to use a Bayesian VARX model based on a method proposed by Anderson et al. (2020). In the model, each sector is affected by the lags of sectoral growth and a lag of the total employment growth. The lag of aggregate employment growth acts as an economy-wide factor which will affect each sector.

Under the assumption that the structure of the Australian economy will not change during COVID-19, we suggest the BVAR model as

$$\mathbf{y}_t = \mathbf{c} + \mathbf{A_1}\mathbf{y}_{t-1} + \mathbf{A_2}\mathbf{y}_{t-2} + \mathbf{A_3}\mathbf{y}_{t-3} + \mathbf{A_4}\mathbf{y}_{t-4} + \mathbf{\Gamma}\mathbf{x_{t-1}} + \mathbf{u_t}$$

where $\mathbf{y_t}$ is an 87×1 vector of two-digit subsectoral employment growth rate at time t and $\mathbf{x_{t-1}}$ is the 4×1 vector of 4 lags of the growth rate of aggregate employment (this vector of variables are predetermined at time t), $\mathbf{c}$ is a vector of constants, $\mathbf{A_{1,2,3,4}}$ are 87×87 parameter matrices. $\mathbf{\Gamma}$ is a 87×4 matrix and $\mathbf{u_t}$ is a vector of reduced form errors with the mean equals to zero and independent variance $\mathbf{u_t} \sim (\mathbf{0}, \mathbf{\Sigma})$. (see Appendix)

5.2   Prior and shrinkage

Bayesian VAR helps to overcome the curse of high dimensionality via the imposition of prior beliefs on the parameters (Bańbura, Giannone, & Reichlin, 2010). I will estimate the VARX using Bayesian methods by specifying a Minnesota prior (e.g. Anderson et al., 2020; Litterman, 1986; Robertson & Tallman, 1999). In order to set up the Minnesota prior in our BVAR model, Bańbura, Giannone, and Reichlin (2010) suggests a Minnesota-type prior that applies shrinkage to the VAR slope coefficients as follows:

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

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

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

where the degree of shrinkage is governed by λ, $\frac{1}{i^2}$ to down-weight more distant lags and the $\frac{\sigma_j^2}{\sigma_k^2}$ adjusts for different scale of the data. σ_e^2 is the variance after fitting an AR model on total employment growth.

Bańbura, Giannone, and Reichlin (2010) also suggested that a natural conjugate Normal-Inverse-Wishart which retains the principle of Minnesota prior will help in adding Minnesota prior to the Bayesian VAR system. Its posterior moments can be calculated either analytically or through adding the dummy observations. I will use dummy observations to estimate the BVAR (e.g. Bańbura, Giannone, & Reichlin, 2010; Anderson et al., 2020; Woźniak, 2016). More details are provided in the Appendix.

5.3   Sectoral Employment Multiplier Analysis

I will also conduct a two-digit subsectoral employment multiplier analysis based on the estimated BVAR model and the 87 two-digit subsectoral data.

At each time period, we have :

$$GR_T = \sum_{j=1}^{87} w_j \times GR_j$$

where w_j is the share of employment of two-digit subsector j in the total employment, GR_T is the growth rate in total employment and GR_j is the growth rate in employment of two-digit subsector j.

Due to the interconnection of macroeconomic two-digit subsectors, when a two-digit subsector j has an increase in employment, in the long run, it may have spillover effect onto other two-digit subsectors, especially those that have high connections with the two-digit subsector j (Anderson et al. (2020)).

If this long-run effect is larger than the immediate effect, then this two-digit subsector will have positive spillover effect on total employment, and *vice versa*. I will use the interconnection of BVAR parameters to estimate two-digit subsectors which have strong positive spillovers for the total employment. Then, if the government makes policies to stimulate these two-digit subsectors, the total employment will recover more effectively.

Table 2: *Completed work*

Timeline	Tasks
Week 2	Background reading about Multivariate VAR model
Week 3	Collect two-digit subsectoral data from ABS website
	Prepare materials reltaed to two-digit subsectoral employment evaluations
Week 4	Prepare literature review related to BVAR model and setting priors
	Analysis about the MATLAB code done by Andereson et al. (2020)
Week 5	Read articles about Conditional Forecasting
Week 6	Prepare the conditional forecasting coding and hierarchical forecasting
Week 7	Exploratory data analysis on sectoral and two-digit subsectoral data
Week 8	Select suitable plots and test softwares to prepare the proposal
week 9	Write research proposal and prepare the first presentation

Table 3: *Research plan*

Timeline	Tasks
June - July	Modelling of two-digit subsectoral employment
August	Train the model using pre-covid data with BVAR model and compare the accuracy with other models
September	apply forecasts using BVAR , conduct both conterfactual analysis and sectoral multiplier analysis
October	Write my thesis and prepare for the second presentation

6   Timeline and expected outcome

Table 2 summarises the work done to date. Table 3 maps out the plan for completing the thesis research.

7   Acknowledgement

I would like to gratitude my supervisor Professor Farshid Vahid and my coordinator Professor Heather Anderson very much for their selfless support and devoted care along the way.

8   Appendix

The VARX model is:

$$\mathbf{y}_t = \mathbf{c} + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2 + \cdots + \mathbf{A}_4\mathbf{y}_{t-4} + \mathbf{\Gamma_1}\mathbf{x_{t-1}} + \mathbf{u_t} \tag{1}$$

$$= \begin{bmatrix} c_1 \\ \vdots \\ c_2 \end{bmatrix} + \begin{bmatrix} a_1^{11} & \cdots & a_1^{1n} & \cdots & a_4^{11} & \cdots & a_4^{1n} & \gamma_1^1 & \cdots & \gamma_4^1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_1^{n1} & \cdots & a_1^{nn} & \cdots & a_4^{n1} & \cdots & a_4^{nn} & \gamma_1^n & \cdots & \gamma_4^n \end{bmatrix} \begin{bmatrix} \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-4} \\ x_{t-1} \\ \vdots \\ x_{t-4} \end{bmatrix} \tag{2}$$

$$+ \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{n,t} \end{bmatrix} \tag{3}$$

$$\tag{4}$$

where $E(\mathbf{u_t u_t'}) = \Sigma$ and $E(\mathbf{u_t u'_{t-1}}) = \mathbf{0}$. Here the n represent the number of sectors (or in our case will be 87) and c represents the vector of constants. There are 4 lags included for the total employment $(x_{t-1}, \cdots, x_{t-4})$ as predetermined variable at time t.

Then we implement our VAR by defining dummy observations.

$$Y_d = \begin{pmatrix} \mathbf{0_{np+p,n}} \\ \mathbf{diag}(\sigma_1, \cdots, \sigma_n) \\ \mathbf{0_{1 \times n}} \end{pmatrix}$$

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

where

$$\mathbf{J_p} = \mathbf{diag}(\mathbf{1}, \cdots, \mathbf{p})$$

$$\mathbf{S_0} = (\mathbf{Y_d} - \mathbf{X_d} \times \mathbf{B_0})'(\mathbf{Y_d} - \mathbf{X_d B_0})$$

$$\mathbf{B_0} = (\mathbf{X_d' X_d})^{-1} \mathbf{X_d Y_d}, \ \mathbf{\Omega_0} = (\mathbf{X_d' X_d})^{-1} \text{ and}$$

$$a_0 = T_d - np - p - 1$$

where T_d is the number of rows for both $\mathbf{Y_d}$ and $\mathbf{X_d}$.

We can get

$$\mathbf{Y^*} = \mathbf{X^*} \beta + \mu^* \quad \textbf{where:} \tag{5}$$

$$\mathbf{Y}^* = [\mathbf{Y}', \mathbf{Y}_\mathbf{d}']'; \ \mathbf{X}^* = [\mathbf{X}', \mathbf{X}_\mathbf{d}']'; \ \mu^* = [\mu', \mu_\mathbf{d}']' \tag{6}$$

Then we can estimating the BVAR by conducting an least square regression of Y^* on X^*. The posterior distribution then have the form of

$$vec(\beta)|\Sigma, Y \sim N(vec(\tilde{\beta}), \Sigma \otimes (\mathbf{X}^{*'}\mathbf{X}^*)^{-1}) \ \textbf{and} \tag{7}$$

$$\Sigma|Y \sim \mathbf{IW}(\tilde{\Sigma}, T_d + T - np + 2) \tag{8}$$

where $\tilde{\beta} = (\mathbf{X}*'\mathbf{X}*)^{-1}\mathbf{X}^{*'}\mathbf{Y}^*$ and $\tilde{\Sigma} = (\mathbf{Y}^* - \mathbf{X}^*\tilde{\beta})'(\mathbf{Y}^* - \mathbf{X}^*\tilde{\beta})$
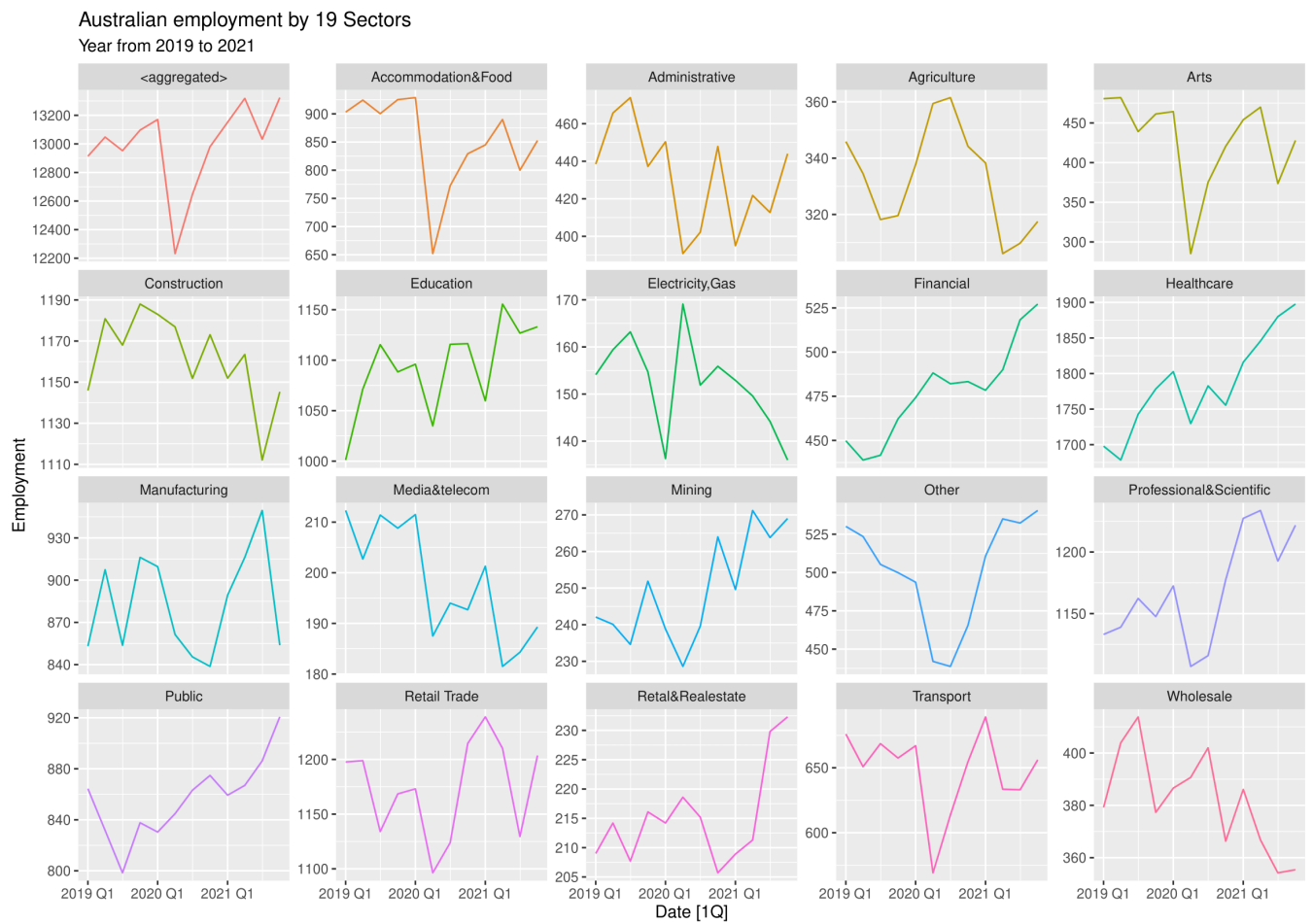
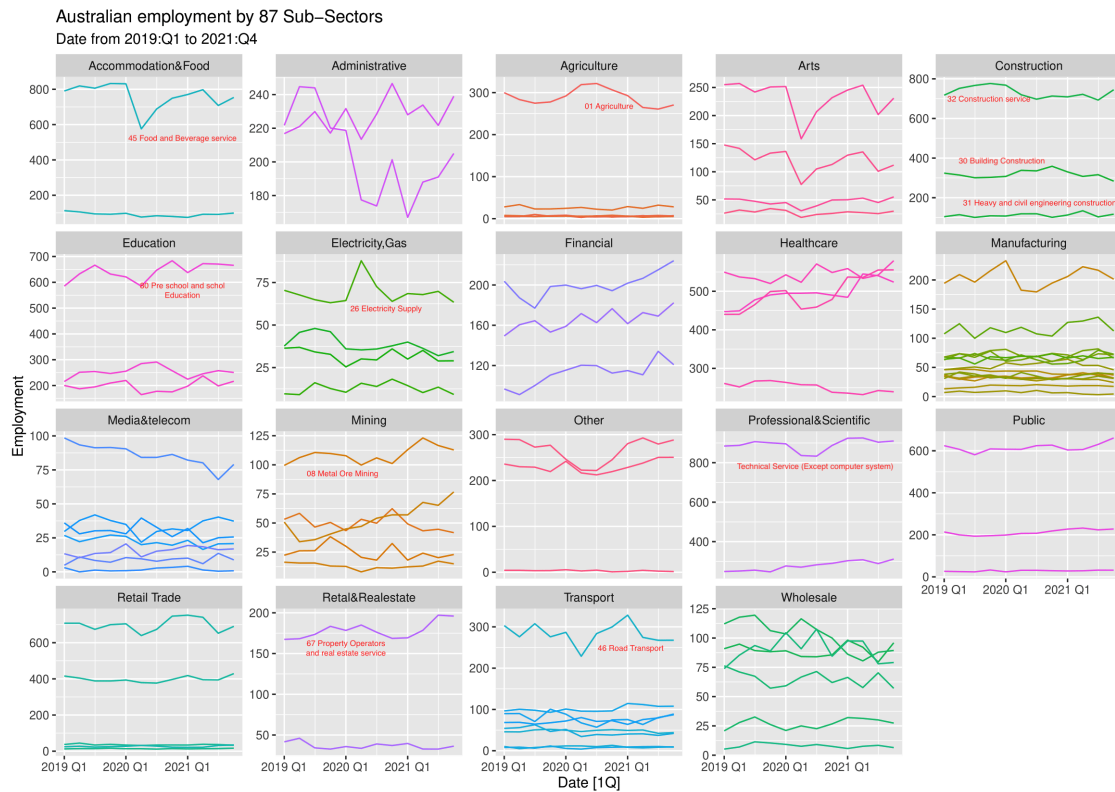**Figure 1:** *Employment('000) of 19 sectors in Australia from 2019:Q1 to 2021:Q4*

Figure 2: *Employment('000) of 87 two-digit subsectors in Australia from 2019:Q1 to 2021:Q4*
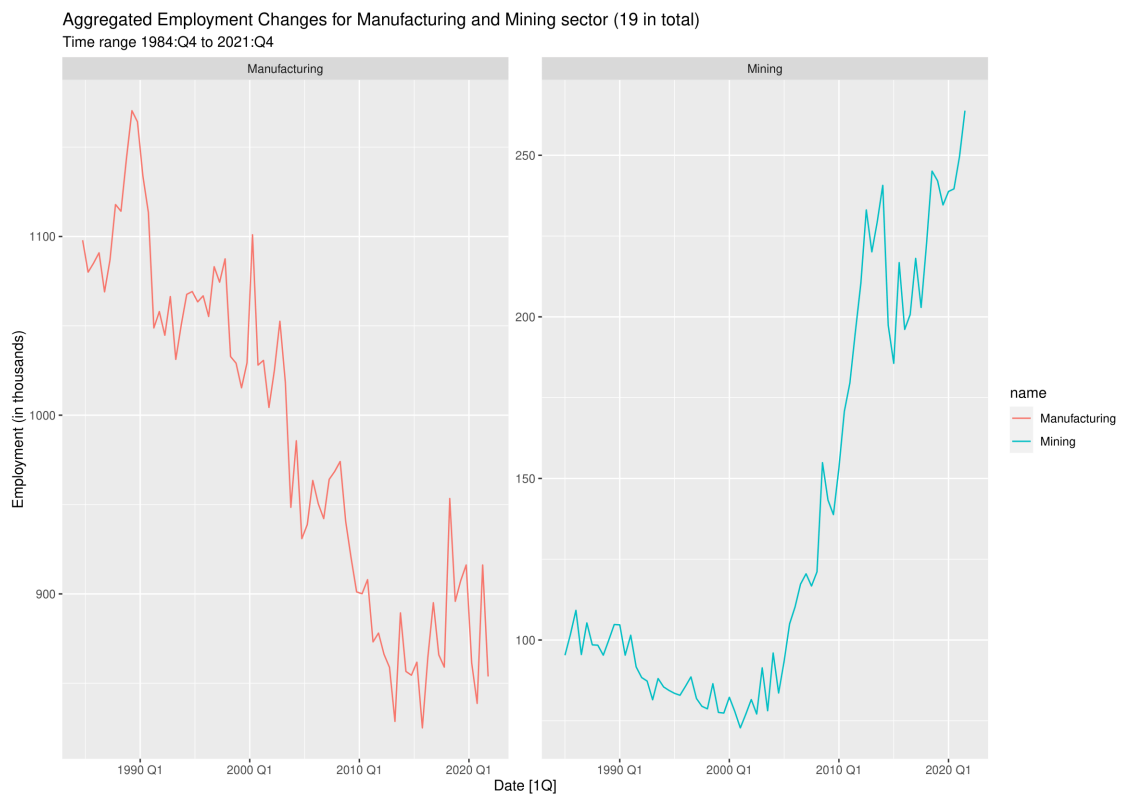

Figure 3: *Aggregated Employment(in thousands) for Manufacturing and Mining sector in Australia from 1984:Q4 to 2021:Q4*
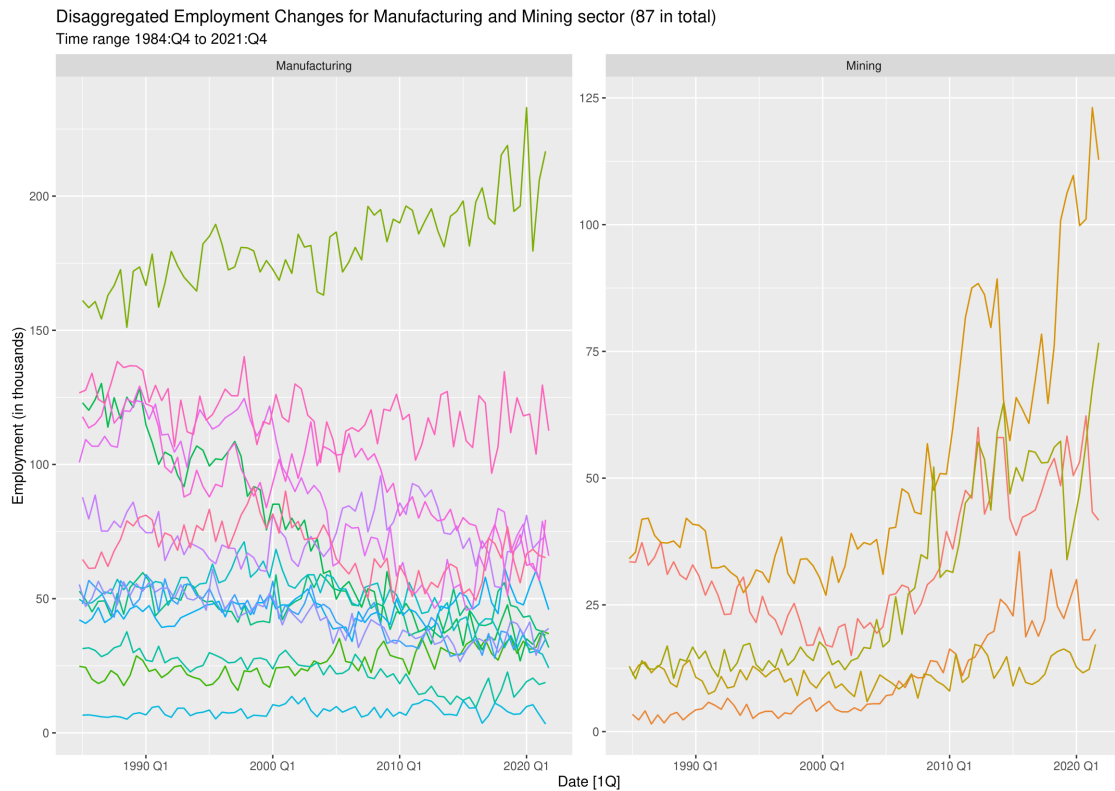
Disaggregated Employment Changes for Manufacturing and Mining sector (87 in total)
Time range 1984:Q4 to 2021:Q4

Figure 4: *Disaaggregate Employment(in thousands) of 87 two-digit subsectors in Manufacturing and Mining sector from 1984:Q4 to 2021:Q4*
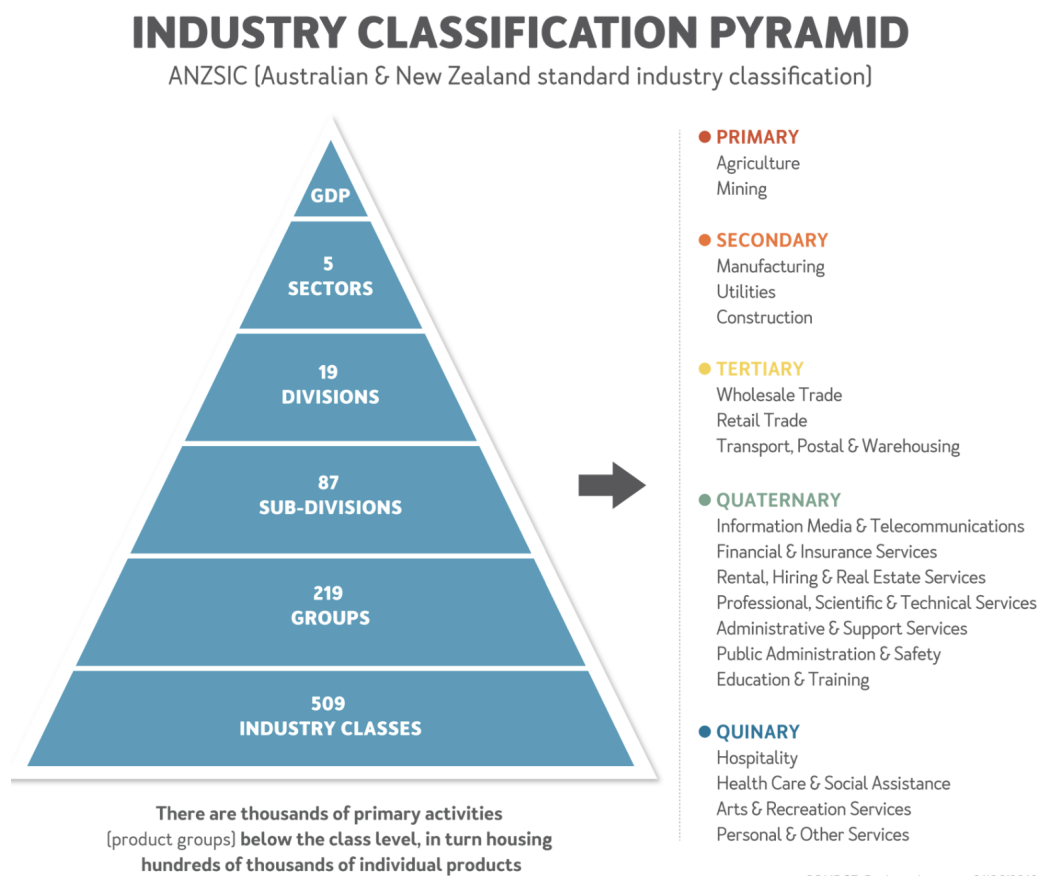

Figure 5: *Australian Industry Pamamid plot by (ANZSIC)*

References

ABS. (2021). *One year of covid-19: Aussie jobs, business and the economy*. Retrieved April 29, 2022, from https://www.abs.gov.au/articles/one-year-covid-19-aussie-jobs-business-and-economy

ABS. (2022). *Labour force, australia, detailed, march 2022*. Retrieved April 29, 2022, from https://www.abs.gov.au/statistics/labour/employment-and-unemployment/labour-force-australia-detailed/latest-release

Anderson, H, Caggiano, G, Vahid, F, & Wong, B. (2020). Sectoral employment dynamics in australia and the covid-19 pandemic. *Australian Economic Review*, **53**(3), 402–414.

Bańbura, M, Giannone, D, & Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of applied Econometrics*, **25**(1), 71–92.

Gregory, V, Menzio, G, & Wiczer, DG. (2020). *Pandemic recession: L or v-shaped?* (Tech. rep.). National Bureau of Economic Research.

Kadiyala, KR, & Karlsson, S. (1997). Numerical methods for estimation and inference in bayesian var-models. *Journal of Applied Econometrics*, **12**(2), 99–132.

Lewis, B, & Hsu, T. (2020). *The collateral damage of coronavirus*. Retrieved April 29, 2022, from https://www.nytimes.com/2020/05/09/business/economy/coronavirus-unemployment.html

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, Ma, S, & Ng, S. (2020). *Covid-19 and the macroeconomic effects of costly disasters* (tech. rep.). National Bureau of Economic Research.

Robertson, JC, & Tallman, EW. (1999). Vector autoregressions: Forecasting and reality. *Economic Review-Federal Reserve Bank of Atlanta*, **84**(1), 4.

Waggoner, DF, & Zha, T. (1999). Conditional forecasts in dynamic multivariate models. *Review of Economics and Statistics*, **81**(4), 639–651.

Woźniak, T. (2016). Bayesian vector autoregressions. *Australian Economic Review*, **49**(3), 365–380.