The Evolution of Employment: A 40-Year Analysis of Canada's Labor Market

Gabrielle Friedman, Yang Gao, Ivan Misic, Lydia Tesconi, Mike Wang
Professor Bhaduri, MA 611 – Section 1, 16 December 2024

Abstract—The "Employment Rate in Canada" dataset reflects monthly provincial employment measurements by the Labor Market Council of Canada from years 1976 to 2019. The employment rate is crucial to analyze because it reflects the local economy's health. Historically we have seen that when there is a higher percentage of people employed in the country, the local economy tends to perform better. This dataset also provides criteria about full-time and part-time employment for each sex within each province that we can cross-analyze. These variables could highlight local economic weaknesses in particular provinces or larger social issues that take place in different parts of Canada. This study could provide insight into the demographic shifts and technological changes over time reflecting Canada's economic strength.

While there have been a variety of other studies related to Canada's employment rate, few analyze the differences between provinces. Additionally, most lack the substantial historical data needed for significant findings related to employment trends over an extended period. The other studies tend to focus on short-term changes in employment or provincial differences that are not backed by years of monthly change, especially not over a range of forty years. We have seen an emergence of analyses post-pandemic to determine the lasting effects of Covid-19 on Canadian employment. In our study, we are looking at a longer-range analysis to focus on the evolution of the Canadian labor market through various economic downturns and significant globalization.

We conducted a thorough analysis of this data set by implementing a variety of statistical techniques. After implementing an informative exploratory analysis of our variables, we executed a time series analysis using a variety of models to determine the best-performing model that is suitable for forecasting changes in employment count over time. In addition to sectioning by province and clustering across sexes in the time series model, we will also seek to understand the relationship between the employment counts and the other variables in the dataset through different subsetted time series analyses. Our goal will be to define the best forecasting model for each of our main analyses: total employment counts in all of Canada, provincial models, sex-based models, and a model that highlights a more applicable topic of employment rates rather than counts.

Our study's outcomes can provide strategic insights into the Canadian labor market. We aim to uncover interesting patterns relating to how demographics could be correlated to employment rates. These analyses should produce broader findings about trend and seasonality of employment rate in Canada and significant differences in patterns between provinces across sexes over time.

Index Terms— Employment trends, Canadian labor market, unemployment patterns, provincial disparities, employment rate predictors, sex-based differences, time-series analysis

1. Introduction

The employment rate is a key determinant of the local economy's health. “Employment rate is the extent to which available labor resources (people available to work) are being used” (“Employment Rate”, 2023). This measurement can be used to inspect the relationship between employment, inflation, and economic downturns (Suthar, 2022). The employment and unemployment rates have also provided insight into a country’s Gross Domestic Product (GDP). Many economists and statisticians have been able to utilize the employment rate to gain deeper insights into economic cycles, such as periods of recessions and periods of growth (Suthar, 2022). It is essential to measure the employment rate because it usually provides context to the well-being of the individuals in the area being measured and the overall economic productivity in the region.

There are a variety of studies relating to the topic of Canada’s labor market. The Canadian Labour Market and Skills Researcher Network (CLSRN) has deployed various analyses of economic shifts and demographic trends in employment (“CLSRN”, 2015). There is a plethora of studies in CLSRN that aim to address employment trends over time. These studies seem to focus on one specific attribute at a time rather than a comprehensive dataset that includes longitudinal information with demographics and province segments. Statistics Canada is another comprehensive source that releases publications that look into Canada’s economy and factors that have impacted the country’s economic performance over time (“Economic Insights”, 2020). These studies include both short-term and long-term analyses, even providing insights into Covid-19’s impact on employment.

While there are key insights in these reports about the country’s overall health, there are few findings on the differences between provinces. The “Canadian City Unemployment Rates and the Impact of Economic Diversity” study provides greater depth into the disparities between Canadian provinces, but it does not include a longitudinal analysis of trends and seasonality over time (Tarzwell, 1997). Another article, “Seasonality of Labour Markets” compiled in 2000, details interesting findings on employment seasonality in Canada (Guillemette et. al., 2000). It employs data from the years 1976 to 1997 and compares seasonality in Canada to that in the United States.

The dataset being analyzed in this study, “Employment Rate in Canada” departs from previous analyses that have been conducted surrounding Canada’s employment rates. This dataset includes information surrounding different types of employment and different sexes in a variety of Canadian provinces. This is a shift from surrounding research on the topic because these variables can provide insights into economic disparities within various demographics. The breakdown of employment by province could also uncover some industry differences and their ups and downs over the past forty years.

To further explain, each province in the dataset seems to have slightly varied important businesses that make up employment within the region. In Alberta energy, agriculture, and forestry are dominant areas while in Ontario there is more of a focus on information and technology, with many individuals working within automobile manufacturers (“Immigration to Canada’s

Provinces", 2021). The employment trends uncovered in this analysis by province can enable a deeper exploration into broader implications for industries.

In the following portion of the report, section two, we will provide additional background information on the dataset and our initial exploratory data analysis. Beyond highlighting the scope and data characteristics, we will also list any initial challenges that we faced. This investigation stage helps to provide a framework for the following time series analyses that we conducted. In section three we will describe the various time series models that we have chosen to apply and examine further.

We will explain any model assumptions and how we set up these models for future analysis. In section four we supply various statistical tests, our model fitting steps, and the interpretations of these models. This portion of the report will highlight our statistical analysis of employment trends over time in different time series models and explain how we decided to segment the data. Finally, section five will summarize our key takeaways from the analyses. We make sure to touch on how these models provide insight into future employment trends and also list several areas for further analysis.

2. The Data Set

The "Employment Rate in Canada" dataset was collected from the Labour Market Council of Canada and published on Kaggle in 2020 (Ortiz-MacLeod, 2020). This dataset provides statistics from ten provinces in Canada; Alberta, British Columbia, Manitoba, New Brunswick, Newfoundland and Labrador, Nova Scotia, Ontario, Prince Edward Island, Quebec, and Saskatchewan. Within each of these provinces full-time and part-time employment measures were collected as well as female and male sex demographics. For each of these variables, there are monthly collections of data spanning the years 1976 to 2019, representing almost forty years of analysis. The dataset includes 4,743 observations with 13 total variables. There are no missing values, but there are several outliers from the date range of the months between July 2018 and November 2019.

After removing some of our outliers and creating the appropriate subsets of our data, we sought to understand some of the variables in our dataset.

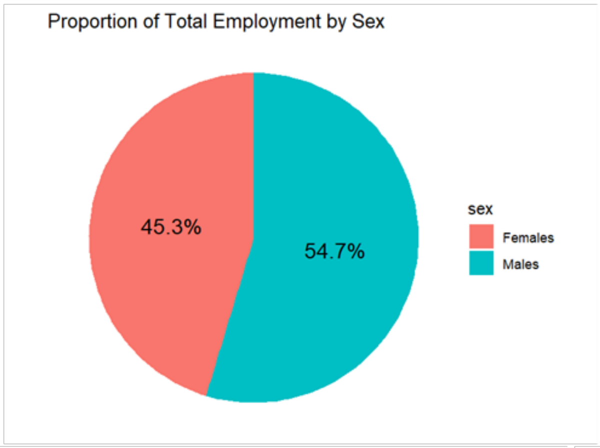

Figure 1: Sex Proportion

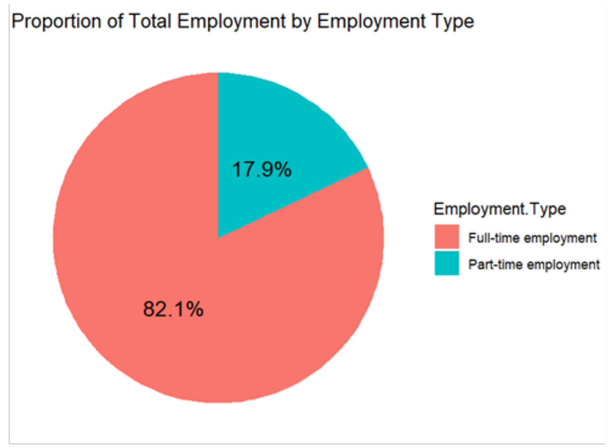

Figure 2: Employment Type Proportion

Figure 1 and Figure 2 explore our sex and employment type variables and their proportions in our dataset. There are slightly more males in our dataset compared to females with them representing roughly 55% of the dataset and females representing 45% of the dataset. In comparison, for the employment type, we can see largely unequal proportions. Around 82% of the data is full-time employment and 18% is part-time employment. This is useful information to know going into our specific analyses as a favoring toward one sex or employment type could skew our findings if we were to think that they were equal.

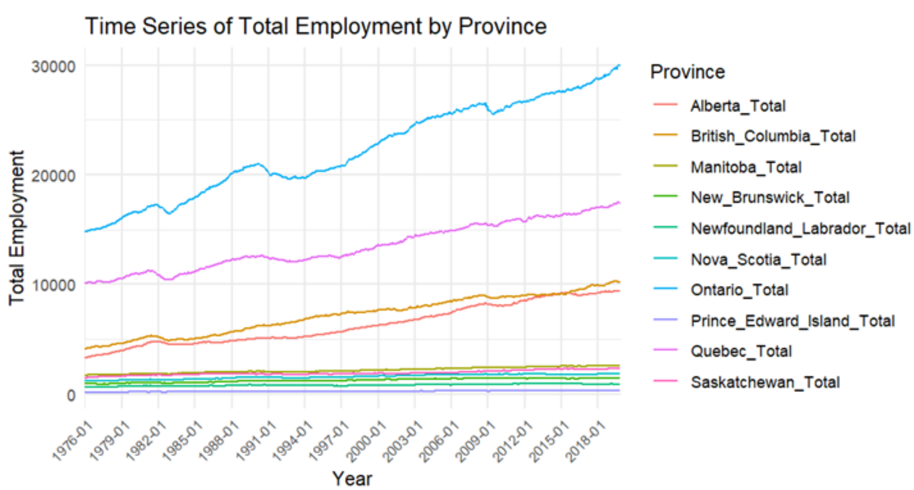

Figure 3: Time Series of Total Employment by Province

Figure 3 represents a time series of total employment by province. This is a helpful graphic to help us understand which provinces may be more interesting to explore further. The larger provinces with a greater population have more substantial increases in total employment over time. Ontario leads by showcasing the largest upward increase in total employment over the last forty years. Quebec is not far behind, and Alberta and British Columbia have very similar trends.

This report is focused on presenting a detailed analysis of the development of the Canadian labor market over the past forty years. We utilize a time series analysis to explore significant employment trends across different provinces.

3. Theory and Methods

We employed ETS, ARIMA, Neural Network, and Bagged Rate time series models to predict future values in our time series analysis. The ETS models helped us to capture the trends and seasonality in the Canadian employment dataset. This model operates on weighted averages of previous data and has weights that deteriorate exponentially. In comparison, ARIMA combines past value differences and moving averages in its time series models. Neural network models can portray nonlinear relationships in the dataset that ETS and ARIMA models cannot. Neural methods can showcase patterns that linear models are not able to reveal. The Bagged Rate time series model helped us forecast the employment rate that our previous models were unable to capture.

ETS Model: $Y_t = E_t + T_t + S_t + \epsilon_t$

The ETS model entails error, trend, seasonality, and random error components. For each of the main elements of ETS, there can be different levels; additive, damped additive, multiplicative, and no. The fitted values are generated using the optimal combination of error, trend, and seasonality through the model fitting process with the lowest AIC values. The ETS model assumes additive or multiplicative effects on errors, trends, and seasonality based on how the data performs.

ARIMA Model: $(1-\phi 1\ B)(1-\Phi 1\ B^\wedge s)(1-B)(1-B^\wedge s)\mathrm{X}t = (1+\theta 1\ B)(1+\Theta 1\ B^\wedge s)\varepsilon t$

The SARIMA model is a seasonal autoregressive integrated moving average. The predicted values of an ARIMA model are calculated using estimated ARIMA parameters. This model assumes stationarity after differencing and that there are no correlations between the error terms. Like an ETS model, the parameters are estimated in R using likelihood metrics like the lowest AIC values.

- SARIMA(p, d, q)(P, D, Q, s):
    - AR(p): Autoregressive component of order p (ϕ)
    - MA(q): Moving average component of order q (θ)
    - I(d): Integrated component of order d
    - Seasonal AR(P): Seasonal autoregressive component of order P (Φ)
    - MA(Q): Seasonal moving average component of order Q (Θ)
    - Seasonal I(D): Seasonal integrated component of order D
    - s: Seasonal period
    - Error Terms: ϵt

Neural Network Model: NNAR(# of ordinary lags, # of seasonal lags, # of hidden nodes) [seasonality]

To figure out if a Neural Network model may be a better model for the data, we employ various nonlinearity tests. We assume that we should use a linear model until we reject the null hypothesis that the mean function is linear. If we find that in multiple of the nonlinearity tests, we reject the null hypothesis, we know that a neural model may capture any additional intricacies and patterns that a linear model would be unable to uncover.

A Neural Network model investigates different layers of nodes that can capture non-linear patterns in data. The fitted model is generated through different inputs and backward integration for training. A large assumption of a Neural Network model is that it requires a large dataset, and that standard linear relationships cannot capture all the appropriate patterns. Linearity is contained and the effects of outliers deteriorate quickly. In comparison to ETS and ARIMA, the estimates are fluid and can change from run to run.

Rate Time Series & Nonlinear Model:

The rate time series was included because of the connections within this dataset, and the ability to visualize change within context (see Appendix B1). Every possible subgroup combination of Canadian workers by sex, employment type, and province had an employee count that could be compared to that of one or more greater workforce totals.

Time series rates were created to compare employment per province to total Canadian employment per workforce type (see Table 1 and referenced appendices below):

Table 1	Provincial Employment & Corresponding Canadian Workforce Totals				
	Canadian Workforce Totals				
Subgroup	All	Part-Time	Full-Time	Female	Male
Provincial	**App. B1a**	**App. B1b**	**App. B1c**	**App. B1d**	**App. B1e**

More granular time series rates were created to compare the same subgroup to the multiple greater populations to which it belonged (see Tables 2 and 3 with appendices below):

Table 2	Female PT Employment & Types of Workforce Totals in Ontario		
	Ontario Workforce Totals (App. B3a)		
Subgroup	All	Female	Part-Time
Female PT	Plot 1	Plot 2	Plot 3

Table 3	Female FT Employment & Types of Workforce Totals in Ontario		
	Ontario Workforce Totals (App. B3b)		
Subgroup	All	Female	Full-Time
Female FT	Plot 1	Plot 2	Plot 3

Based on initial analyses, including assessing time series entropy, one combination of subgroup, workforce, and province was chosen to model and forecast future employment. By harnessing the greatest possible levels of granularity available in our dataset, we compared the employment rate of female part-time (PT) workers in Ontario against the province's total PT employment. This prepared us to explore the following questions:

1) What proportion of Ontario's part-time workforce has historically been female?
2) What proportion of Ontario's part-time workforce will likely be female in the future?

To accomplish these goals, models were created. The granular rate-based time series was modeled by minimizing AIC to determine the best ARIMA and the best exponential smoothing models (ETS) for non-transformed, Box-Cox transformed, and log-transformed data.

With regards to assumptions, after we considered the null hypothesis of linearity, we created two nonlinear models: a neural network and a bagged (using 15 bootstraps) model. Accuracy was compared using test MASE and retrospective analysis.

The bagged model was expected to have wider forecasting intervals and higher MASE due to its use of 15 bootstrapped time series of simulated data. Model-dependent versus data-dependent approaches were considered logically in the context of our findings.

4. Data Analyses

In our data analysis, we decided to split our study into multiple different components and attributes to analyze. We looked at four main areas. First, we looked at a broader analysis of employment counts in all of Canada. Next, we analyzed each of the ten provinces and plotted ETS, ARIMA, and Neural models to see which the best-performing model was based on accuracy metrics.

We aimed to discover differences and interesting findings between these different areas of Canada. We also wanted to analyze one of the demographic variables, sex, to see if there were fascinating discrepancies between male and female employment and forecasting models. Lastly, we investigated more granular time series models in the form of proportion-based time series and comparative methods to model nonlinearity.

<u>Model 1: Canada Employment Count Time Series</u>

At the beginning of our analysis, we created a time series to represent all of Canada's employment data. First, we summed up all the employment data from each of the ten provinces into one column and subsetted our data to look at only total employment and both sexes so that we did not have repetitive data in our models.

We fitted a time series model to our data and graphed the original model. As shown in Figure 4 below, Canadian employment has steadily increased over the last forty years. The data seems relatively consistent, with only minor dips during economic downturns, like the 2008 recession. This also seems to align with the steady increase in Canada's population.

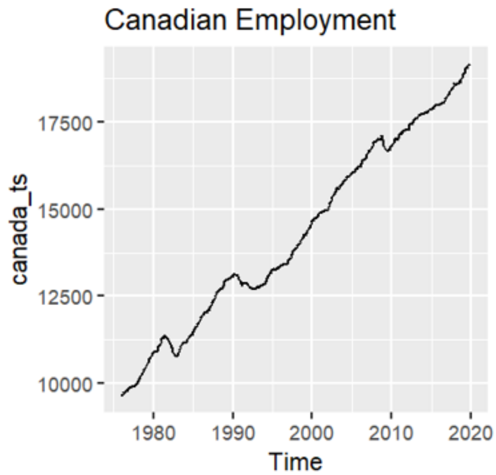

Figure 4: Canada Time Series Plot

We then needed to split our data into training and testing splits. We decided to use the standard format of 80% in our training set and 20% in our testing set which meant that we withheld the last eight years of data for testing. This withholding process allowed us to better understand how accurate our models were at forecasting future employment values. In our nonlinearity test, the training dataset in Canada demonstrated some potential linearity issues.

We have to reject the null hypothesis on three out of five of these tests that the mean function is linear. Therefore, we analyzed all three models of ETS, ARIMA, and Neural on our total Canadian employment time series. After fitting these models, we made sure to check the residuals to see if there may be any dependency issues.

The ETS and ARIMA models seemed to pass the residual check for a variety of reasons. Figures 5 and 6 portray the ARIMA model residual check. In the time plot, the residuals seem to have patterns removed and they are mostly random. There are only minor normality deviations in the histogram and the residuals seem to be normally distributed.

Additionally, the Ljung-Box test outputs a large p-value which means that the residuals are uncorrelated and independent. Unfortunately, in the Neural Network model residual check, we saw a very tiny p-value output from the Ljung-Box test. A smaller p-value means that the residuals may have some dependencies.

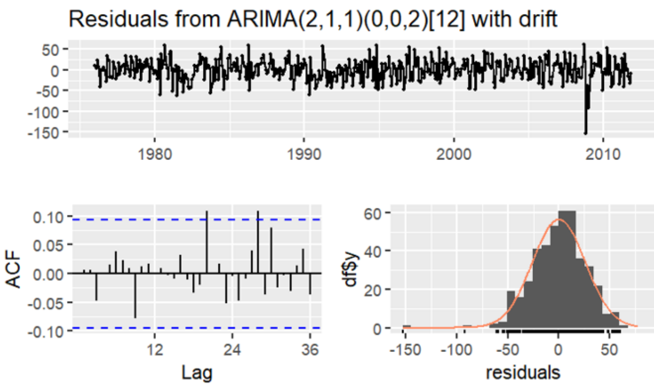

Figure 5: Canadian Total ARIMA Residual Check

```
> checkresiduals(canada_arima)

        Ljung-Box test

data:  Residuals from ARIMA(2,1,1)(0,0,2)[12] with drift
Q* = 13.068, df = 19, p-value = 0.8351

Model df: 5.   Total lags used: 24
```

Figure 6: Residual Check Continued

Lastly, we checked the accuracy measures on our training and testing sets to understand which of the three models should be used for forecasting. The ARIMA model had the lowest MAPE and MASE and did not show signs of overfitting on the test set. The ETS and Neural Network models both showed some overfitting as the test set performance was dramatically worse than the training set performance in each of these models.

For all these reasons, we have concluded that the ARIMA model is the best to forecast future Canadian employment trends. R concluded that the best ARIMA model was ARIMA(2, 1, 1)(0, 0, 2)[12] with drift. The equation that we fitted based on our summary findings is as follows:

$$(1-0.56B\ -0.26B^2)(1-B)Xt$$
$$= c + (1-0.42B)(1\ -0.20B^{12}-0.15B^{24})e_t + 17.70t$$

Our fitted model demonstrates two non-seasonal components in the form of autoregressive terms and one moving average term. We also have two seasonal moving average terms with a period of twelve months. Additionally, there is a drift term that encapsulates a constant trend over time.

The presence of non-seasonal and seasonal components in our model means that the employment count data showcase patterns across and within individual years. The drift terms help to capture a long-term increase in employment count which makes sense given the demographic changes and economic growth over the last forty years. The small standard errors in comparison to the coefficients showcase a relatively well-fitted model.

We wanted to be able to utilize this model for some baseline forecasting for future Canadian employment predictions. It is important to note that this model represents counts of employment rather than the rate itself. We have a future model in this analysis that incorporates specific employment rate calculations. Figure 7 below represents our three models employed in this introductory analysis with their fitted and forecasted representations.

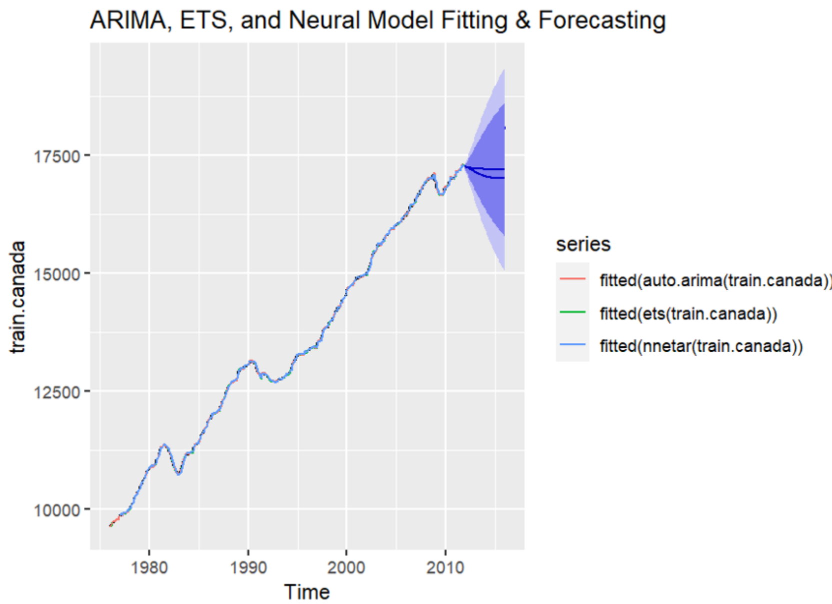

Figure 7: Canadian Totals Model Forecasting

Each model fitted seems to closely resemble the original data and appropriately captures the trend over time. Of all the models, ARIMA seems to be the most appropriate at forecasting and the least overfit. The shaded area indicates the confidence in our predictions over the future which becomes more uncertain as more time goes on.

Nationally, we observed a consistent increase in employment across Canada, even following significant global disruptions like the 2008 financial crisis and the recessions of 1982 and 1991, reflecting the country's economic resilience and population growth. Our ARIMA model excelled in predicting future employment trends at the national level, effectively capturing both long-term patterns and seasonal fluctuations.

<u>Model 2: Canada Province Time Series</u>

The next challenge that we wanted to tackle was looking at different provinces and what model would be best for forecasting employment counts for each one. Our analysis shows that four provinces had ARIMA with the best forecasting accuracy, four provinces had ETS with the best forecasting accuracy, and two provinces had a Neural Network model with the best forecasting accuracy. We thought this finding was interesting as it emulates the dramatic differences between the complex employment structures in each of the provinces and potential differences in their trend and seasonality even though most of the provinces are close to one another.

Table 1 in Appendix A shows the best-fitting models for the employment count at the province level in Canada, with ARIMA, ETS, and NNAR models being chosen based on their performance. Provinces with larger populations, such as Ontario, Quebec, and British Columbia, tend to have models that account for both trend and seasonality, indicating complex dynamics in employment data that likely reflect the provinces' economic diversity and size.

For Ontario and Quebec, ARIMA models with multiple seasonal and non-seasonal components were selected, suggesting that employment data for these provinces follow systematic patterns with dependencies on prior values and seasonal trends. For example, Ontario's ARIMA (1,1,3)(0,0,2)[12] includes seasonal components ([12] for monthly seasonality) and a moving average to smooth out irregularities, indicating high variability in employment trends. Similarly, Quebec’s ARIMA (2,1,1)(2,0,1)[12] reflects a slightly more complex model, likely due to pronounced seasonal effects and autocorrelation, representing a more intricate structure in its employment dynamics.

For smaller provinces, such as New Brunswick and Newfoundland and Labrador, the NNAR model emerges as the best fit, highlighting the value of neural networks in capturing non-linear relationships in smaller datasets where traditional statistical models might struggle. The NNAR (1,1,2)[12] for New Brunswick and NNAR (3,1,2)[12] for Newfoundland and Labrador incorporate lagged observations and a seasonal lag structure ([12]), effectively learning patterns without assuming a specific data-generating process.

In contrast, provinces like Nova Scotia and Prince Edward Island rely on ETS models, which excel at modeling data with multiplicative errors and trends while accounting for additive seasonal variations. This reflects consistent yet simpler patterns in their employment data compared to larger provinces.

Model 3: Sex-Based Time Series

Looking at the Canadian employment stratified by sex can show whether there are some disparities between the two sexes. We decided to look at male and female employment on a national level continuing with our 80-20 train-test methodology by leaving the last eight years of our data out for investigation. Figure 8 shows how the two sides have become closer than ever through time as male and female employment nears. There are about 2% more females nationally so we can see how proportionally more males employed than women, but that gap will look to even out the coming years if the trends continue.

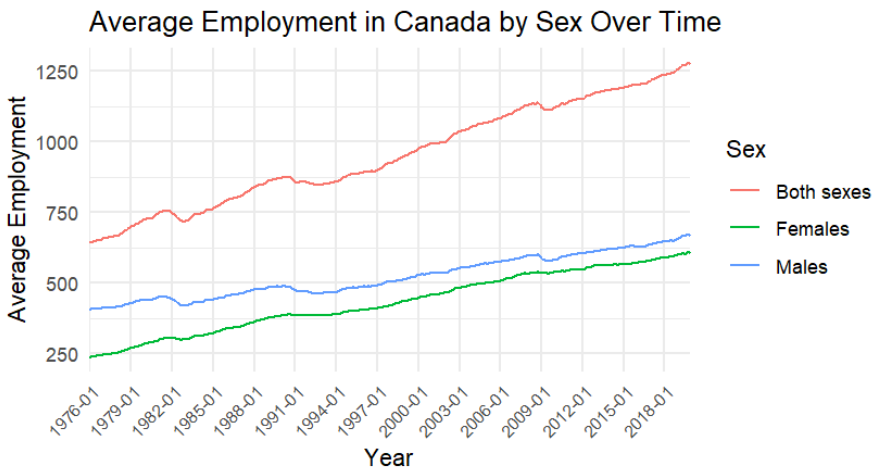

Figure 8: Average Employment in Canada by Sex

We can also investigate the trends based on different provinces to see where the employment population is growing proportionate to the national numbers. Figure B2c and Figure B2d in the Appendix look at the employment proportion of the province compared to the national employment to see the relative growth in each province by sex. When we zoom in using Figure 9 below, we can see that sexes follow the same general pattern between each other for all the provinces except for Ontario where the men are proportionally growing while the female employment has slightly diminished through time.

Figure 9: Ontario Employment Rates by Sex

This could indicate that a lot of men are moving to Ontario for job opportunity whether it is from other countries or from other provinces where proportional employment has been decreasing. Ontario has the strongest job potential as they have the largest economy in Canada. Our EDA provides evidence that men are looking to go there rather than women when analyzing the graphs from the sex point of view.

Outside of EDA, we ran linear models for each sex on a national level as both passed the linearity tests which ruled out the need to use the neural models. Before looking into results, all four models passed the Ljung-Box test by showing insignificant p-values, proving that the models are suitable by avoiding issues of autocorrelation.

The ARIMA models proved to be superior to the ETS models by a decisive margin for both male and female. They scored far better in the MAPE and MASE categories in both sections, which played the biggest role in their selection. R concluded that the best function for both models was the ARIMA(1,1,2)(0,0,2)[12] with drift. This means we must use the one autoregressive term, a first differencing, two moving average terms, two seasonal moving averages, and a drift term.

The female ARIMA model shown in Figure 10 below, shows the graphical performance on the test set displaying forecast and actual. Linear does a very good job predicting the general trend of employment over the eight years and the MAPE score was about 1.009 showing strong promise. The forecast continues the strong linear increase of female employment with a steep slope that shows promise towards eventual alignment with the males assuming their growth continues as it has.

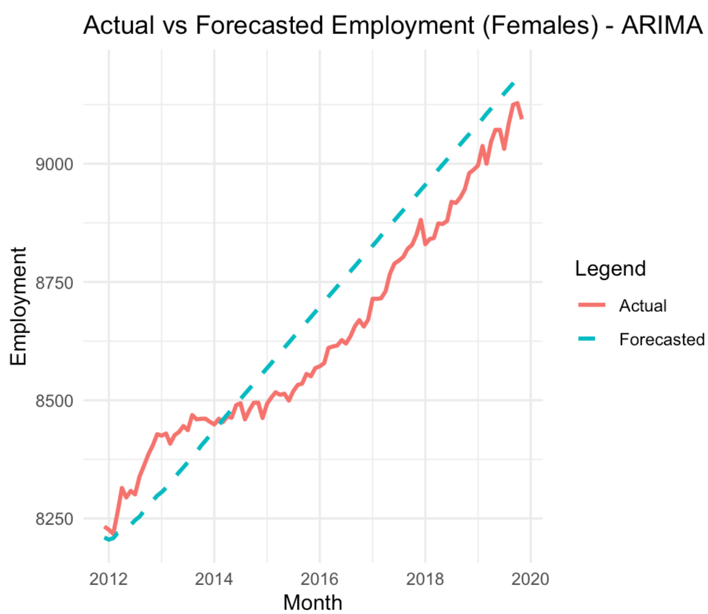

Figure 10: Actual vs. Forecasted Female Employment

The male ARIMA model in Figure 11 below shows the graphical performance on the test set displaying forecast and actual. Linear does a very good job predicting the general trend of employment over the 8 years and the MAPE score was about 1.2779 showing strong promise. The forecast shows how the men were not necessarily projected to grow as heavily as the women but exceed the expectations which means that the alignment between the two sexes is potentially growing further apart as the women performed under their projections.

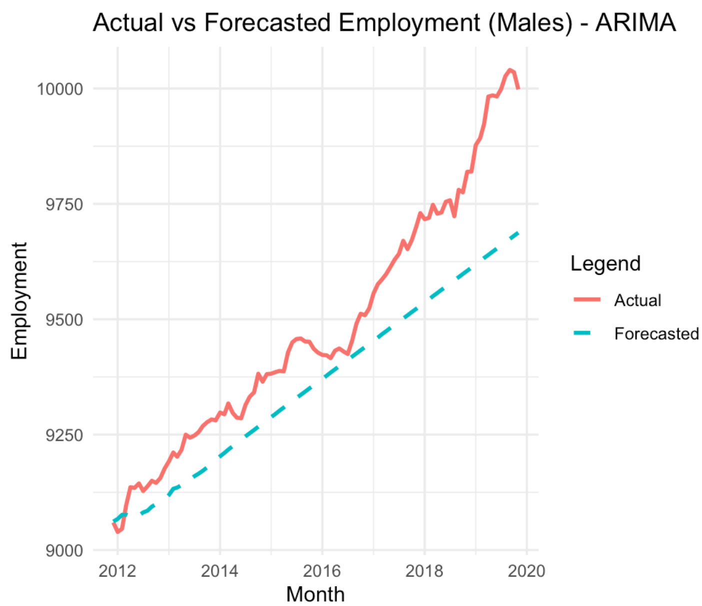

Figure 11: Actual vs. Forecasted Male Employment

To conclude the sex section, ARIMA(1,1,2)(0,0,2)[12] with drift was selected as the strongest model as it managed to reduce the error more than ETS as only linear models were in consideration as they passed linearity checks. Female employment was expected to grow at a steeper slope than men when looking at the ARIMA forecast, but men exceeded their projections while women underperformed. The important finding for Canada is that employment is growing at generally the same rate at present day with the men having a slight edge proportionally at the national level.

<u>Model 4: Rate Time Series & Nonlinear Model:</u>

The decision not to use linear models for forecasting was made using these findings:

1) Although the trend strength was high (0.8959), the best ETS was MNN and the best ARIMA was ARI(0,0), indicating that linear-based decomposition may not be correct.
2) The training and testing MASE for ETS of untransformed and transformed data were all almost identical, for both BoxCox and log transformation, and approaching the naive.
3) The best ARIMA training and testing MASE were the highest of all.

The determination to use nonlinear models was made given these factors:

- 3 out of 6 nonlinearity testing p-values were significant at $p<0.05$: 0.0357, 0, 0.0035.
- The best training MASE was by neural net, at 0.033.
- The model with best performance by residuals was also neural net, with $p = 0.8415$

- Neural net and bagged dominated the retrospective errors approaching zero in Fig. 12:

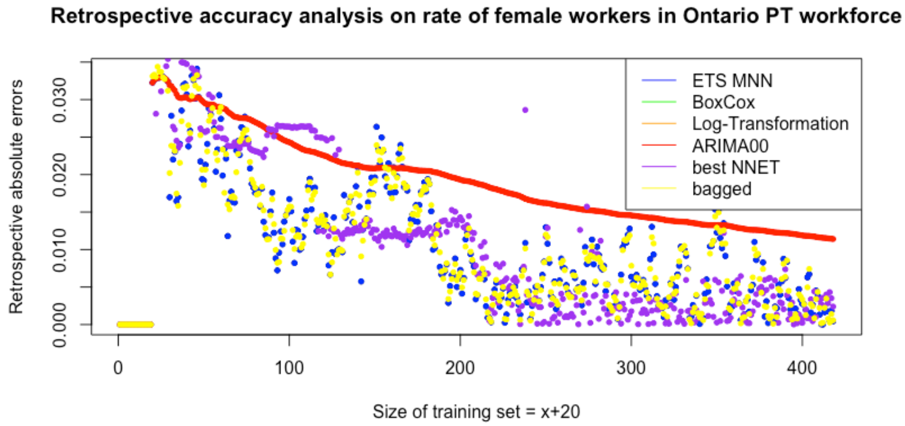

Figure 12: Comparative Retrospective Absolute Errors on Granular Time Series

The determination to choose bagged model over neural network NNAR(25,1,13)[12]:

- The bagged out-of-sample MASE (0.7304) was only slightly higher than those of the transformed (0.729) and untransformed (0.7294) ETS models, despite using multiple time series (bootstrapped simulations) versus their only using one (the original).
- The neural network out-of-sample MASE was approaching 1 at 0.9087 (see Fig. 13).
- Forecast intervals are reasonable although unreliable with a bagged model (see Fig 14)
- Although the neural net fit may seem superior, it assumes a decreasing trend (see Fig 15)

	[,1]	[,2]	[,3]
[1,]	"model"	"train MASE"	"test MASE"
[2,]	"naive"	"0.5799"	"1.0281"
[3,]	"randomwalk"	"0.5437"	"0.8084"
[4,]	"best ETS"	"0.5478"	"0.7294"
[5,]	"BoxCox ETS"	"0.5476"	"0.7291"
[6,]	"log-trans ETS"	"0.5479"	"0.729"
[7,]	"best ARIMA"	"1.3209"	"2.3238"
[8,]	"best neural"	"0.033"	"0.9087"
[9,]	"bagged"	"0.5264"	"0.7304"

Figure 13: Training & Testing MASE Comparisons by Model

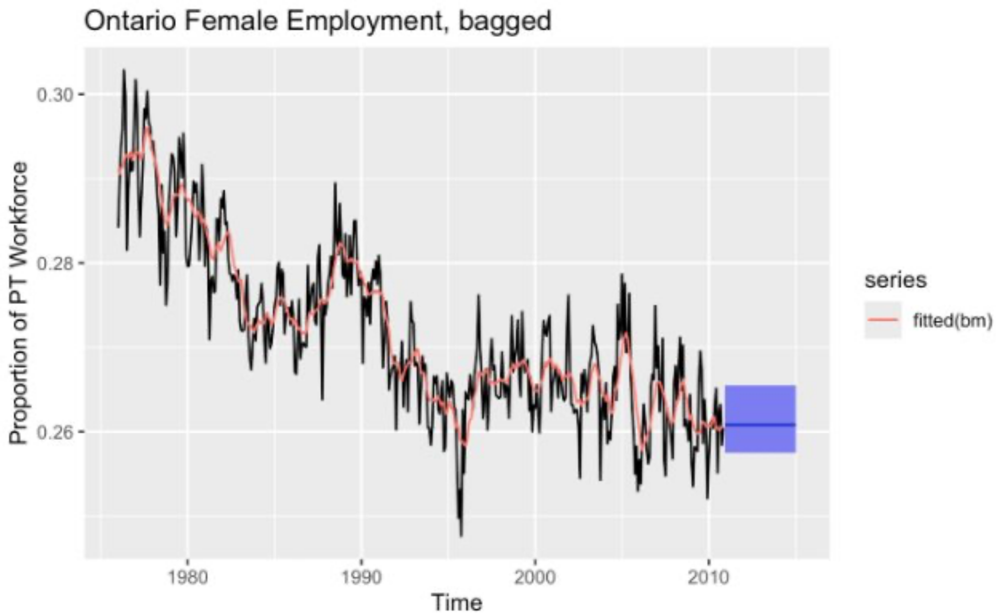

Figure 14: Bagged Model Forecasts

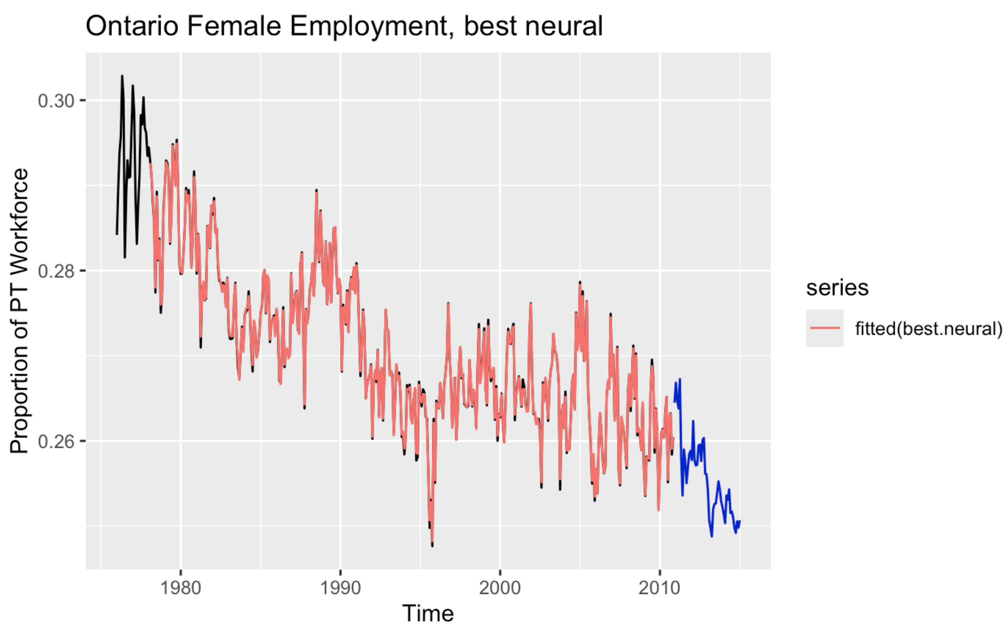

Figure 15: Neural Net Forecasts

Finally, for us what it came down to when choosing between the relative strengths and weaknesses of the neural and the bagged models was based on logic. Choosing a neural net is limiting because the results are model-dependent – and if the model is weak (see bad test MASE), so is our analysis.

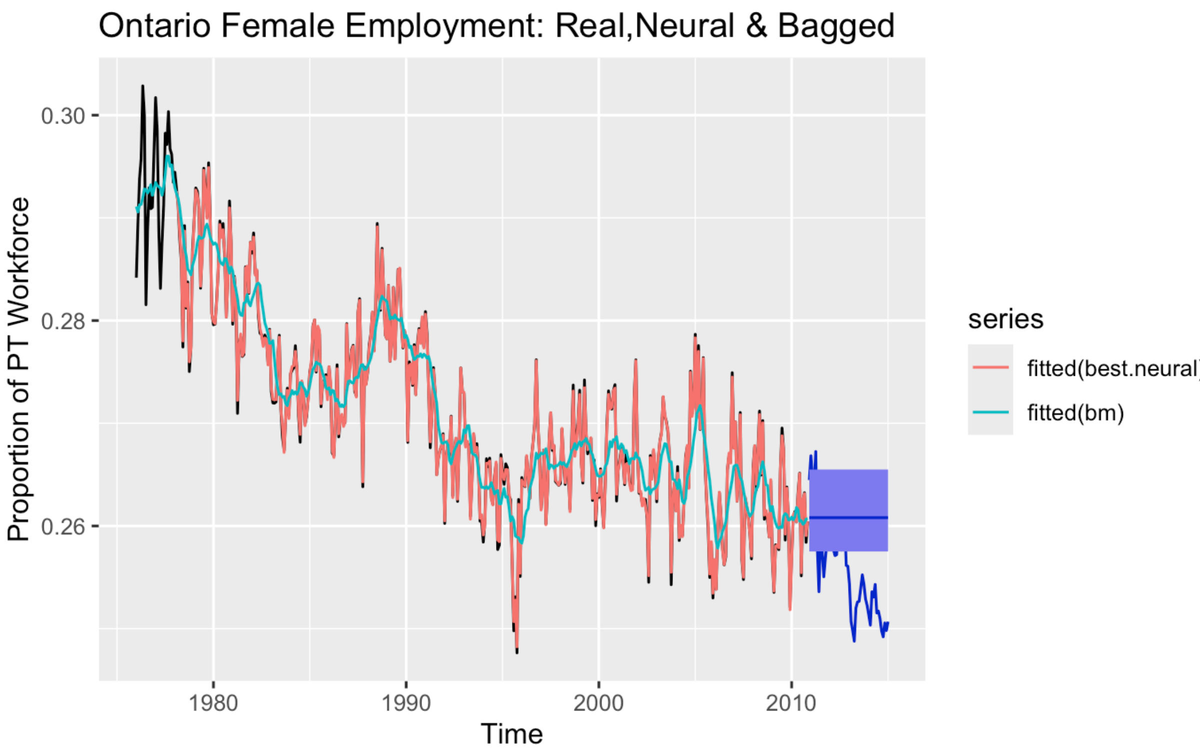

Figure 16: Forecast Comparison for Neural Net & Bagged Models

The bagged model is preferable because it is a data-dependent form of modeling and relies on the creation of many bootstrapped simulations as many as we desire. There is more flexibility and more likelihood of accuracy over time. For example, in Fig. 16 above we are reminded that the bagged model differed in not assuming a decreasing trend – based on the data.

The results of the bagged model, in contrast to the neural net, also appear to suggest that the level of the time series (ie. the proportion of Ontario part-time workers who are female) is unlikely to change greatly in the coming decade (remaining about 26 - 27%). Given the high level of variability and unpredictability in the granular time series, as evidenced in the attempted moving average decomposition of Fig, 17 below, bagged is best as the most conservative option.

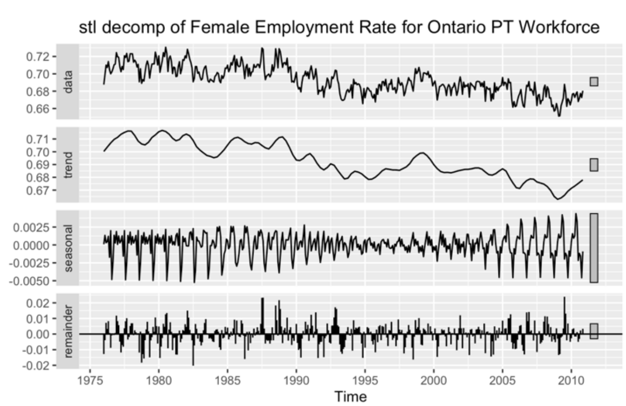

Figure 17: Illogical STL with nonperiodic seasonality

5. Conclusions and Future Work

This research project focused on the evolution of the Canadian labor market over the last forty years, between 1976 and 2019. Our dataset includes the employed population across each Canadian province, employment type, and sexes over time. By using time series modeling methods including ARIMA, ETS, and neural network models, we explored the employed population trend, seasonality, and forecast nationwide, across each province, and sex. Our findings indicate major employed population differences in regional population, provincial economy, and sexes.

At a national level, we found that the employed population has been steadily growing in Canada including right after major world events such as the 2008 financial crisis and recessions of 1982 and 1991. This indicated economic resilience and Canada's growing population over time. At a national level, our ARIMA model performed best for forecasting future employment, being able to capture both the long-term trend and seasonality. On the provincial level, different provinces have different optimal models. In the larger provinces, Ontario and Quebec, the ARIMA model performed the best. For the 3rd and 4th largest provinces in employed population, the ETS model performed the best.

For employment trend based on sex, we see a narrowing trend over time—indicating gender disparity in employment is becoming less and less as time goes by. At the provincial level, strong regional disparities in employment based on sex still exists. An interesting example is Ontario. Its male workforce shows an increasing trend while the female workforce shows a slight decreasing trend. Meanwhile, when we took a closer look, we found high variability and nonlinear patterns despite the appearance of a general downwards trend. It may be challenging to predict change in the proportion of part-time workers in Ontario who are female over time, and more achievable to forecast just the average level. These forms of abnormality may be explained by the provincial labor and employment policy, economic status and industry distribution.

Our project places high importance on understanding employment trends and seasonality within economic and political patterns. While ARIMA is the most reliable model for forecasting national employment, additional exploring for non-linear models and more micro variables such as employment count by industry, GDP per captia, significant labor policy published nationally or provincially could provider further insights into our dataset and help us understand the Canadian employment trend and seasonality better. Additionally, the effect the COVID pandemic has on employment could be further studied. It would provide valuable insights for the Candian and other administrative entities on how employment would fluctuate in a regional/global pandemic in the future. However, a limitation is that our dataset only has the employment data until late 2019, right before the pandemic.

For future research, we would like to see the ratio-based granular approach continue with other combinations of province, sex, and employment type given the richness of the available data and the diversity within Canada as a whole. It could be valuable to ask these further questions:

- Could sex contribute to the apparent great differences in trend and seasonality across employment types within each province of Prince Edward Island and Quebec?
- Could sex also help to explain the relative lack of variation for each of the Saskatchewan, Manitoba, and British Columbia provinces?
- Could employment type help to explain the apparent similarity in trend and seasonality between male and female for these provinces: Newfoundland, New Brunswick, and Prince Edward Island?
- Could the employment type be involved in the apparent greater variability between male and female for the three provinces above, Quebec, and Nova Scotia?

In conclusion, our research project gives a comprehensive understanding of Canadian employment, providing policy makers, economists, entrepreneurs and stakeholders valid statistical findings to help solve provincial disparities in employment and advance economic growth for Canada, while also providing clear indications of avenues for future meaningful investigation.

References

Economic Insights. Statistics Canada. (2020). https://www150.statcan.gc.ca/n1/en/catalogue/11-626-X#wb-auto-2

Employment Rate. OECD. (2023). https://www.oecd.org/en/data/indicators/employment-rate.html

Guillemette, R., L'Italien, F., & Grey, A. (2000). Seasonality of Labour Markets Comparison of Canada, the U.S. and the Provinces. *Government of Canada*, (R-00-8E). https://doi.org/https://publications.gc.ca/collections/Collection/MP32-29-00-8E.pdf

Immigration to Canada's Provinces and Major Industries. Winny Immigration. (2021). https://winnyimmigration.com/immigration-to-canadas-provinces-and-major-industries

Ortiz-MacLeod, D. (2020). Employment Rate in Canada, Version 1. Retrieved November 21, 2024, from https://www.kaggle.com/datasets/ortizmacleod/employment-rate-canada/data

Suthar, J. (2022). *Employment Rate.* WallStreetMojo. https://www.wallstreetmojo.com/employment-rate/

Tarzwell, G. (1997). Canadian City Unemployment Rates and the Impact of Economic Diversity. *The Canadian Journal of Regional Science*, *20*(3), 389-.

Appendix A

Province	**Population**	**Best Model**
Ontario	14,223,942	ARIMA (1,1,3) (0,0,2) [12]
Quebec	8,501,833	ARIMA (2,1,1) (2,0,1) [12]
British Columbia	5,000,879	ETS (M, Ad, N)
Alberta	4,262,635	ETS (M, Ad, N)
Manitoba	1,342,153	ARIMA (2,1,2) (2,0,1) [12]
Saskatchewan*	1,132,505	ARIMA (3,1,2) (2,0,1) [12]
Nova Scotia	969,383	ETS (M, A, N)
New Brunswick	775,610	NNAR (1,1,2) [12]
Newfoundland and Labrador	510,550	NNAR (3,1,2) [12]
Prince Edward Island	154,331	ETS (M, A, N)

Table 1: Province Findings

Appendix B1: Count vs. Proportion Time Series

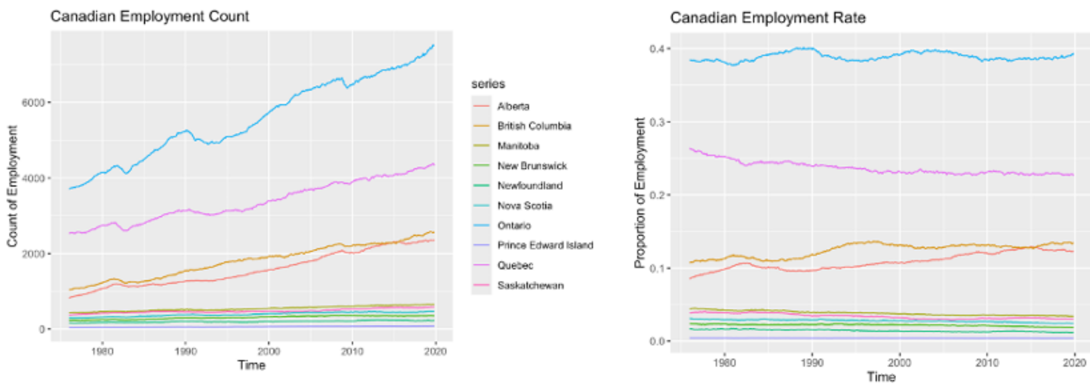

Fig. B1a. Canadian employment by province as count or proportion

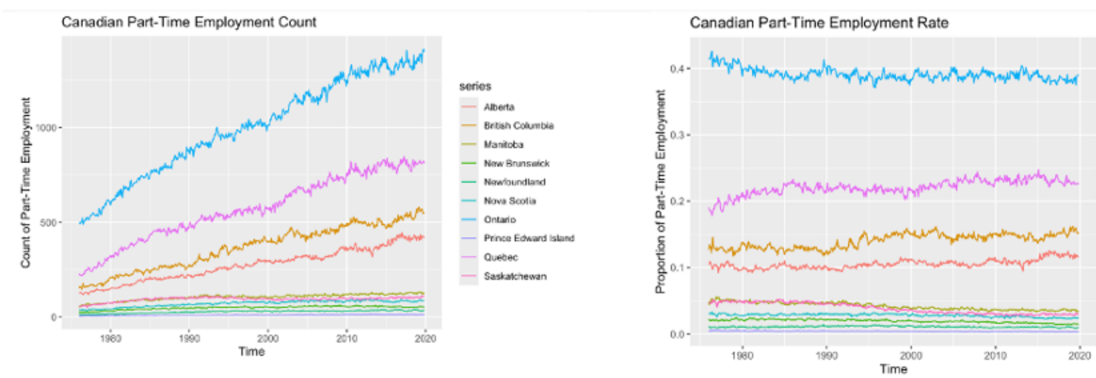

Fig. B1b. Canadian PT employment by province as part/whole or as a proportion (rate)

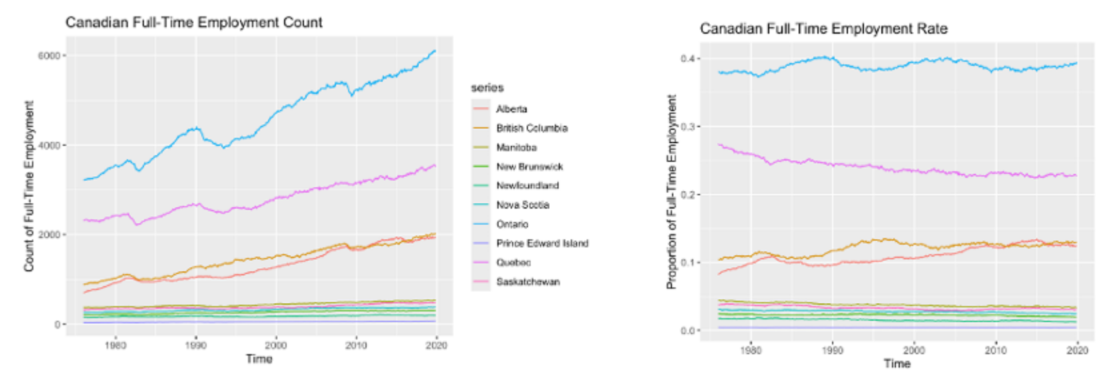

Fig. B1c. Canadian FT employment by province as part/whole or as a proportion (rate)

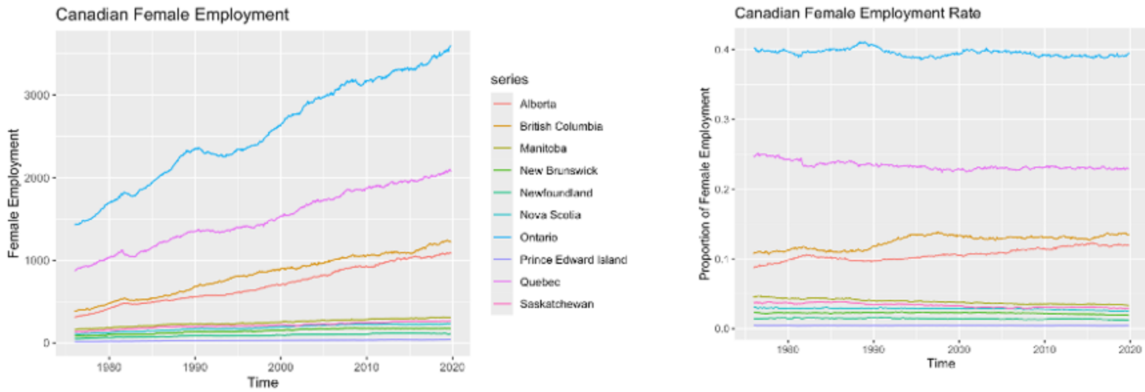

Fig. B1d. Canadian female employment by province as part/whole or as a proportion (rate)

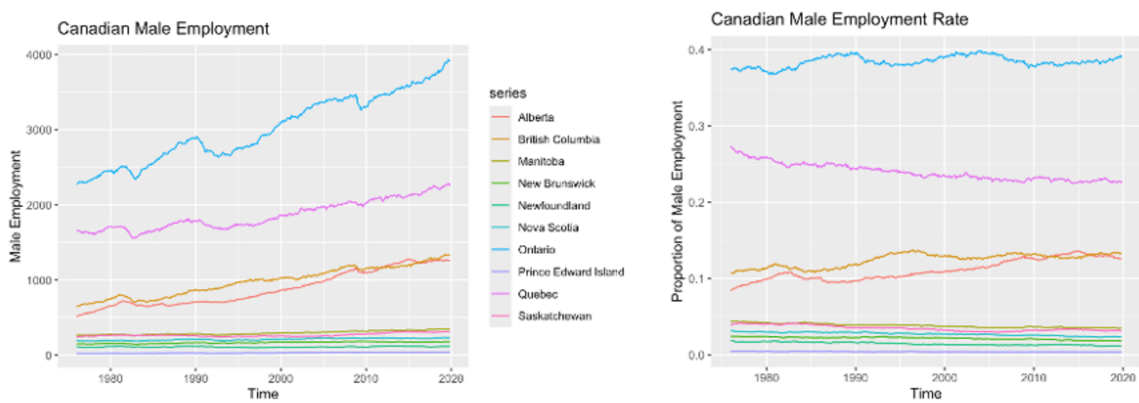

Fig. B1e. Canadian male employment by province as part/whole or as a proportion (rate)

Appendix B2: Proportion Time Series

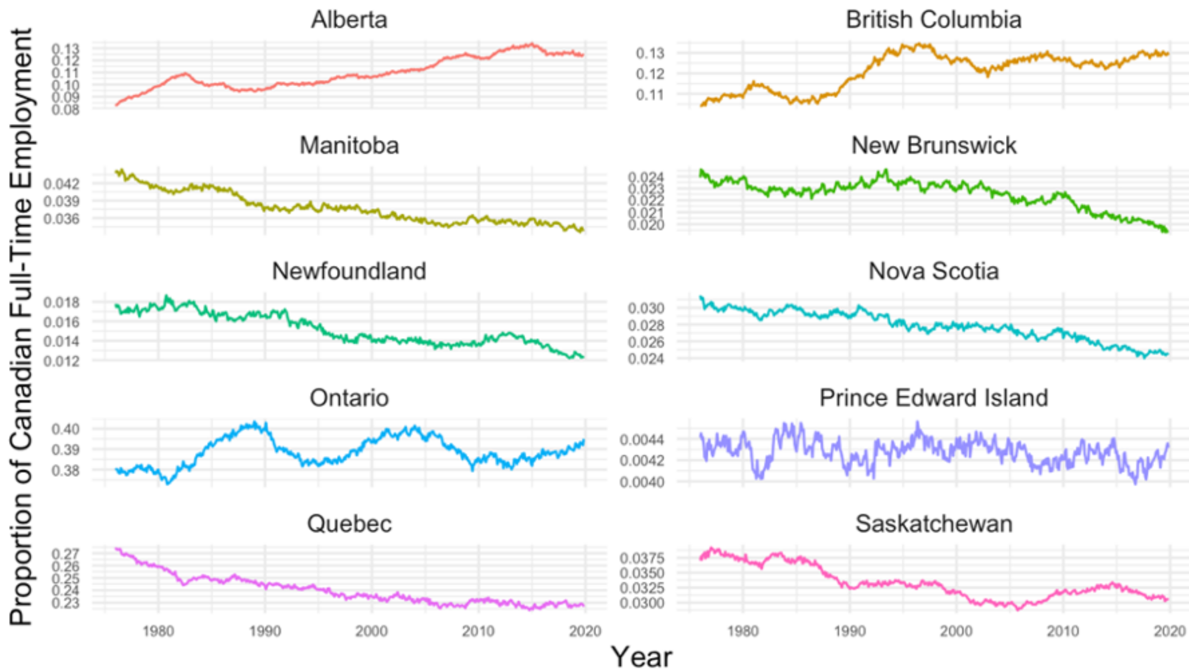

Fig. B2a. Per-Province Proportions of Canadian FT Employment

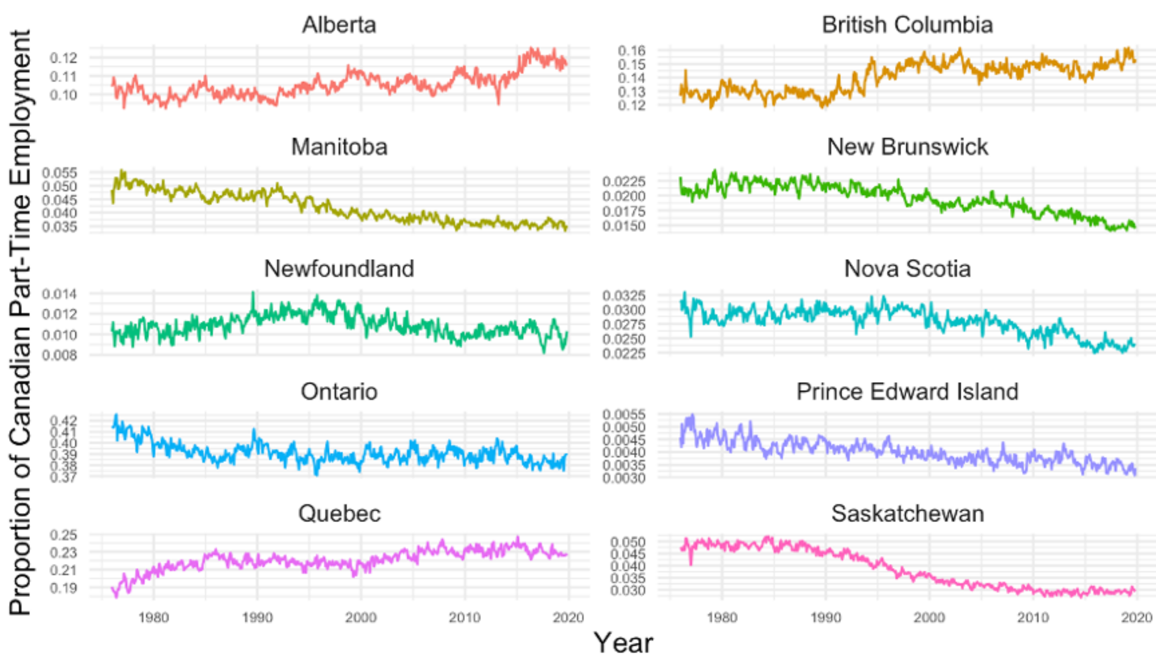

Fig. B2b. Per-Province Proportions of Canadian PT Employment

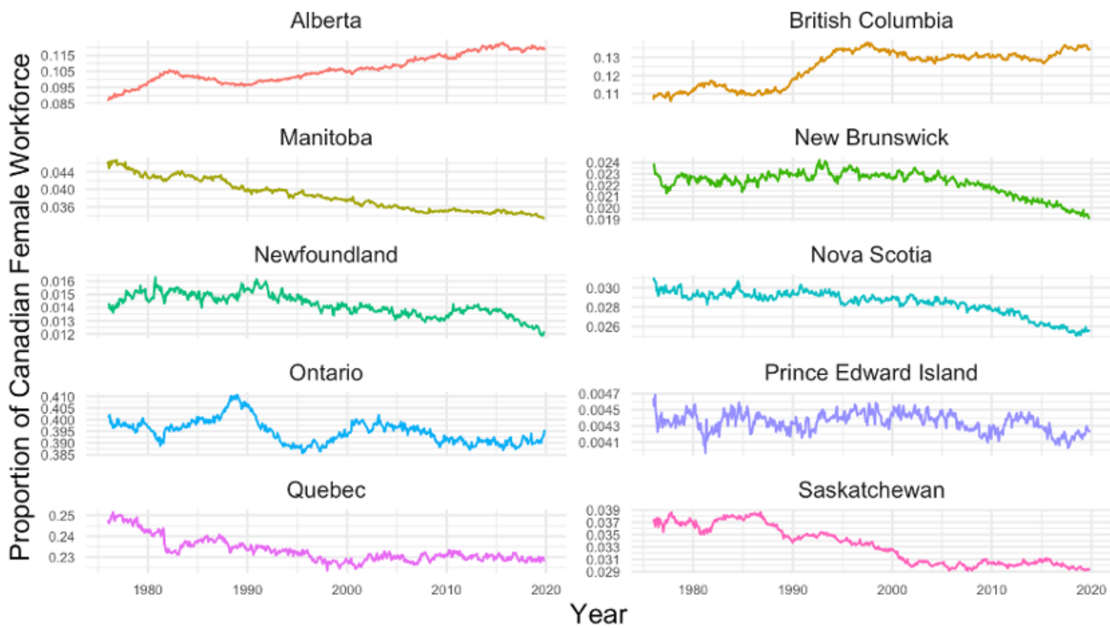

Fig. B2c. Per-Province Proportions of Canadian Female Workforce

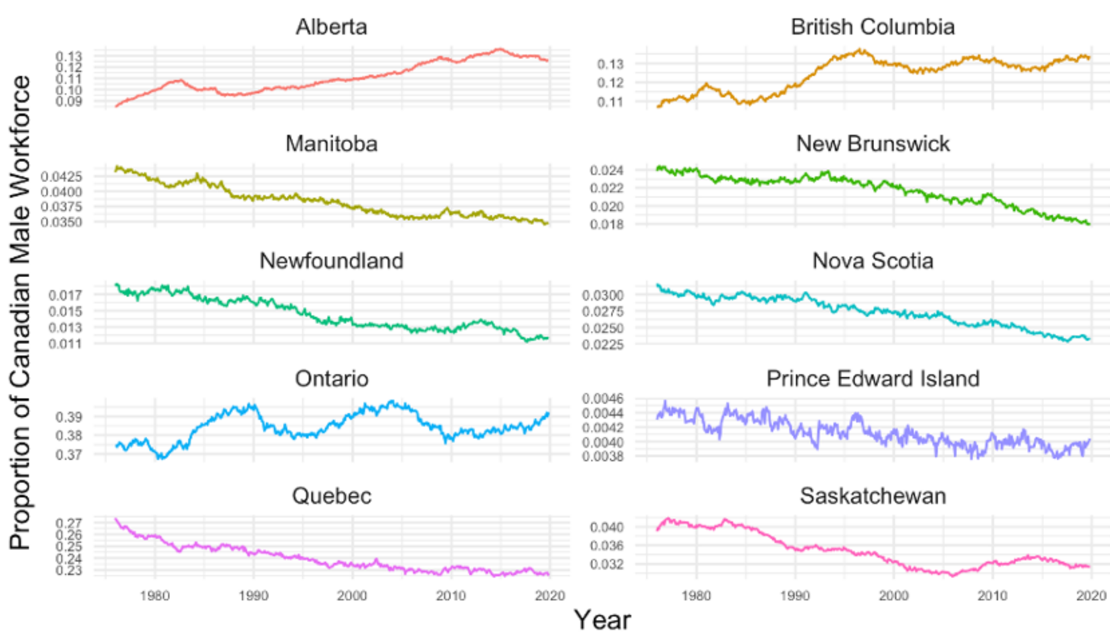

Fig. B2d. Per-Province Proportions of Canadian Male Workforce

Appendix B3: Granular Proportions

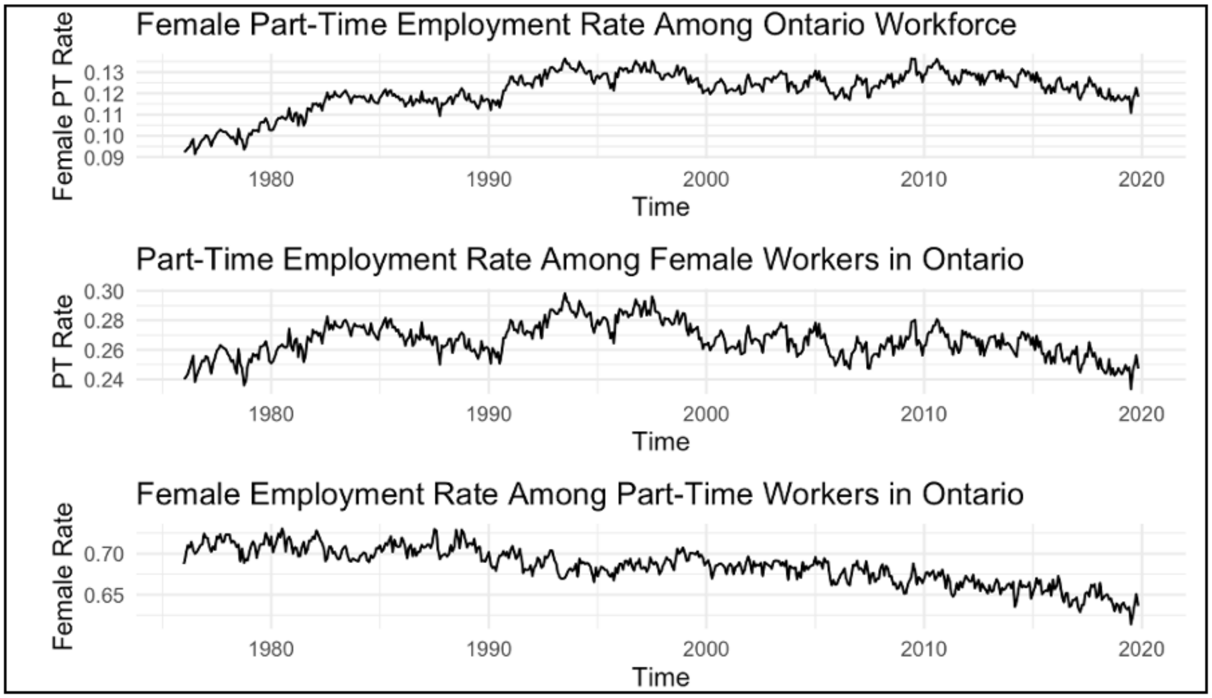

Fig. B3a. Ontario: Female PT compared to 1) Total, 2) Total Female, & 3) Total PT Workforces

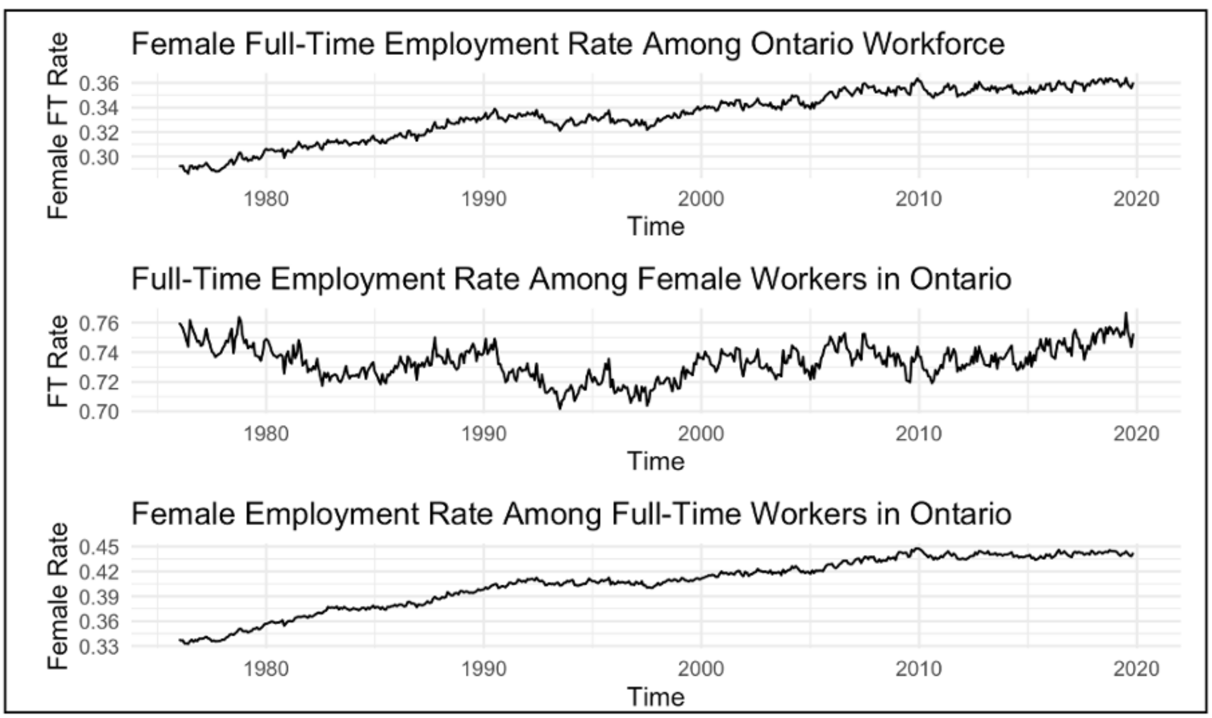

Fig. B3b. Ontario: Female FT compared to 1) Total, 2) Total Female, & 3) Total FT Workforces