An Examination of Canada's Service to Good Employment Ratio

Leuven Wang

March 19, 2021

Introduction

For economic planners, understanding the structure of a nation's economy is key to planning for its future. What kind of relief is best delivered in crises? Where is investment most needed in times of growth? It is commonly held that as economies modernize and develop, their focus shifts away from goods to service industries. This investigation will examine Canada's progression in this aspect from January 1976 to February 2021, accounting for almost 5 decades of the greatest economic shocks, developments, revolutions and growth experienced in human history. We will look at how employment in the services and goods sectors has changed during this time, distinguishing between the 10 provinces.

We are not classifying the category of a job based on which industry the worker is employed in but rather, its nature. For example, a secretary within an oil company will have an administrative service job, not a goods job. Very generally speaking, we can define a goods sector job as being one where its holder directly partakes in the production of a good. The industries where these jobs are most commonplace include construction, manufacturing, agriculture, forestry etc... Service sector jobs support the behind the scenes work in those industries, as well as many others, including trade, transportation, finance, business, education, health care, etc...

The main metric of interest in our investigation is the services to goods employment ratio. This is the number of people employed in the services sector of a region divided by the number of people employed in the goods sector in the same region at the same time.

We are interested in achieving two results. The first is an interval which we can say, with a certain level of confidence, contains the mean services to goods employment ratio. This gives us a general idea of where Canada's service to goods balance level has resided in the last 50 years. The second thing we are interested in is a similar interval aiming to identify the locality of the variance of the services to goods employment ratio. This will help us understand how widely the country diverges in its economic structure. Has our economy transformed unitarily? Or are our biggest gaps in economic structure geographical? Which provinces depend most on manufacturing? How much progress have we made in this transformation?

We hypothesize that the mean services to goods employment ratio will be somewhere between 2 to 3 and that the variance of the service to goods employment ratio will be less than 1.

Data

The data we are using for this investigation is provided by Statistics Canada, a Canadian government agency tasked with collecting various data on the country. We are examining a table detailing the number of jobs each month, in each province, from January 1976 to February 2021. This data has been collected through the monthly official Labour Force Survey (LFS). The jobs are classified by which industry they are in and whether it belongs to the goods or services sector through the North American Industry Classification System (NAICS). The LFS itself does not survey the entire population but only a small sample. Thus, the data

gathered is not a population statistic but an estimate made by the agency. For this investigation, we will be using unadjusted projections. That is to say that we will not be offsetting seasonal variations in employment. The reason behind this is that it is important to account for how seasonal changes affect structural features and to offset these factors would be to ignore them and any inefficiencies they bring.

The data is fairly comprehensive. Most of the cleaning revolved around removing extraneous columns unneeded in our investigation. We filtered out all the seasonally adjusted estimates as well as the standard error calculations. We also removed all federal data as one of the main focuses of our investigation is on the variability between provinces. We also discarded any records with missing values. Finally, to calculate the services to goods employment ratio, we took the number of people employed in the service industry each month in each province and divided it by the number of people employed in the goods sector in the same month and province.

Our main important variables are:
Date - The month and year in question. Written in YYYY-MM format.
Geo - The province in question.
S:G Employment Ratio - The services to good employment ratio in that month and province. Calculated by the number of service industry employees divided by the number of goods sector employees.

Summary	**Measurement**
Mean S:G Employment Ratio	2.949947
Standard Deviation of S:G Employment Ratio	0.6866595
Maximum S:G Employment Ratio	5.3603896
Minimum S:G Employment Ratio	1.3045685

Fig 1: Table of Numerical Summaries of Sample Data

Figure 1 illustrates the center and spread of the S:G ratio. This gives us a good general idea of what we can expect in our upcoming investigation. As we can already see, there seems to exist quite a wide difference between the maximum and minimum S:G employment ratio, implying some big disparity, either geographically or historically or both.

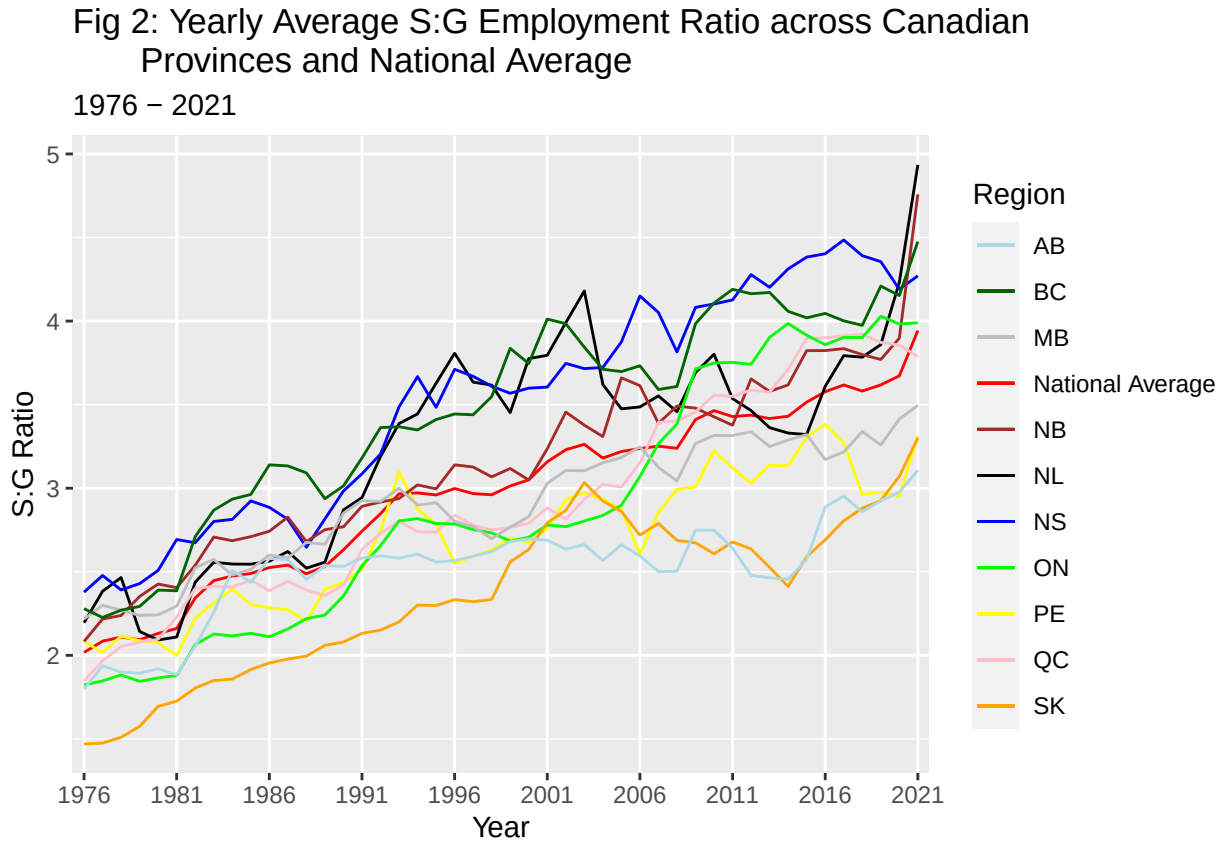

Figure 2 is a line graph that shows the evolution of the S:G employment ratio from 1976 to 2021. It displays the yearly average S:G employment ratio for each province, as well as that of the national average. As we can clearly see, there has been an indisputable increase in the ratio, meaning that the share of service sector jobs has been increasing steadily. This matches with the general perception that as economies modernize and develop, they swap to a focus on services rather than goods.

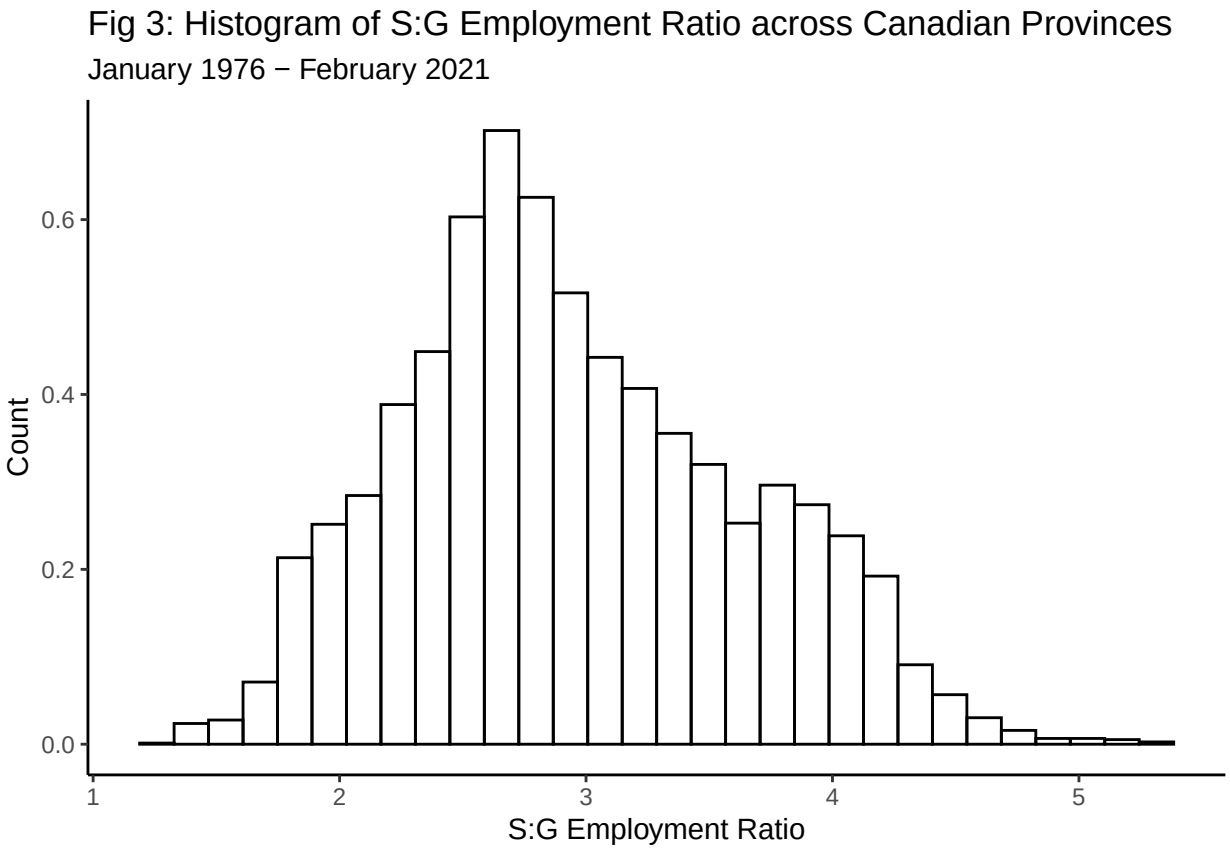

Figure 3 depicts a density histogram for all S:G employment ratios calculated with the given data. This encompasses every single province and every month from January 1976 to February 2021. As can be seen, its shape is reminiscent of a bell curve with a single peak and trailing tails.

All analysis for this report was programmed using `R version 4.0.2`.

Methods

We will assume that the S:G employment ratio follows a normal distribution. To create the two confidence intervals of our investigation, one for the mean of the S:G employment ratios and one for the variance, we will pursue two different methods. First, it is important to understand what a confidence interval is. A confidence interval is a range where we can establish that, to a certain level of confidence, contains the true parameter we are interested in, in this case, the mean and variance. We will be creating intervals with confidence levels of 95%. This means that we are 95% confident that the true parameters lie within those intervals. This gives us a good balance between maintaining the trustworthiness of our intervals and making them precise enough to extract meaningful conclusions. The two methods we will be using to build our confidence intervals are critical values through a t-distribution and bootstrapping.

We will first use critical values and the t-distribution to build the confidence interval of the mean of the S:G employment ratio. What this means is that we will attempt to build a distribution of means of the S:G employment ratio and set a suitable interval on it. Since we do not know the true variance of our distribution of S:G employment ratios, we cannot assume that the distribution of sample means would be normal. Thus, to account for the uncertainty of the variance, we have to widen the interval. This is done by using a t-distribution instead of a normal distribution to model the distribution of the mean. A standardized t-distribution is similarly symmetrically centered at 0 with one peak but has larger tails. This means that it has more maneuverability than the normal distribution. Instead of using the true standard

deviation as we would for the standardized normal distribution to build our model, we will use the standard deviation of the sample.

For a sample $X_1, ..., X_n$ from a $N(\mu, \sigma^2)$ distribution, the studentized mean

$$\frac{\bar{X}_n - \mu}{\frac{S_n}{\sqrt{n}}}$$

has a $t(n-1)$ distribution. Since we are setting a 95% confidence level, that means that we have to cut off 2.5% probability at each end of the distribution to set our interval. This will be done using critical values where z_p represents the point where the right side still contains probability p. For a 95% confidence interval, the critical points will thus be $z_{97.5}$ and $z_{2.5}$. All that is needed now is to calculate the values that correspond to these points and this will be our confidence interval. The mean of the S:G employment ratio simply refers to the mean average of the S:G employment ratio distribution across all provinces throughout the period from 1976 to 2021. The mean S:G employment ratio is a good indicator of where Canada has generally stood over the last 50 years in terms of its transition for a goods to services economy.

To determine a confidence interval for the variance of the S:G employment ratio distribution, we will use another method, bootstrapping. Here, we will attempt to build a distribution of variances of the S:G employment ratio by resampling from our current data sample. We will resample numerous times, each time calculating the variance of that new sample and adding it to our variance distribution. This method depends on the original sample being representative of the larger population. Then we will determine the critical values such that the interval between them makes up 95% of the distribution. This will be our confidence interval. We have chosen to use empirical bootstrapping as it would be difficult to guess the parameters best suited to modeling the S:G employment ratio distribution under a normal distribution. This means we will be relying directly on the sample data to build our bootstrap samples. The variance of the S:G employment ratio measures its spread. This can help us understand how large disparities in the country's economic transition are.

All analysis for this report was programmed using `R version 4.0.2`.

Results

Parameter	**2.5th Percentile**	**97.5th Percentile**
Mean S:G Employment Ratio	2.9316792	2.9682148
Variance S:G Employment Ratio	0.4562181	0.4873569

Fig 4: Table of Confidence Interval Critical Values

Figure 4 is a table that displays the results of the methodology described in the Methods section. These figures form a confidence interval of 95% confidence level for the mean and variance of S:G employment ratios across all provinces from 1976 to 2021. This means that we are 95% confident that the true mean of the S:G employment ratios exists somewhere between 2.9316792 and 2.9682148. Similarly, we are 95% that the true variance of the S:G employment ratios exist somewhere between 0.4562181 and 0.4873569.

These results seem quite reasonable, especially if we compare them to the sample statistics we calculated in the Data section. The sample mean, 2.949947, clearly is within the bounds of the interval and if one squares the sample standard deviation, 0.6866595, then that sample variance, 0.4715013, falls within the calculated confidence interval as well.

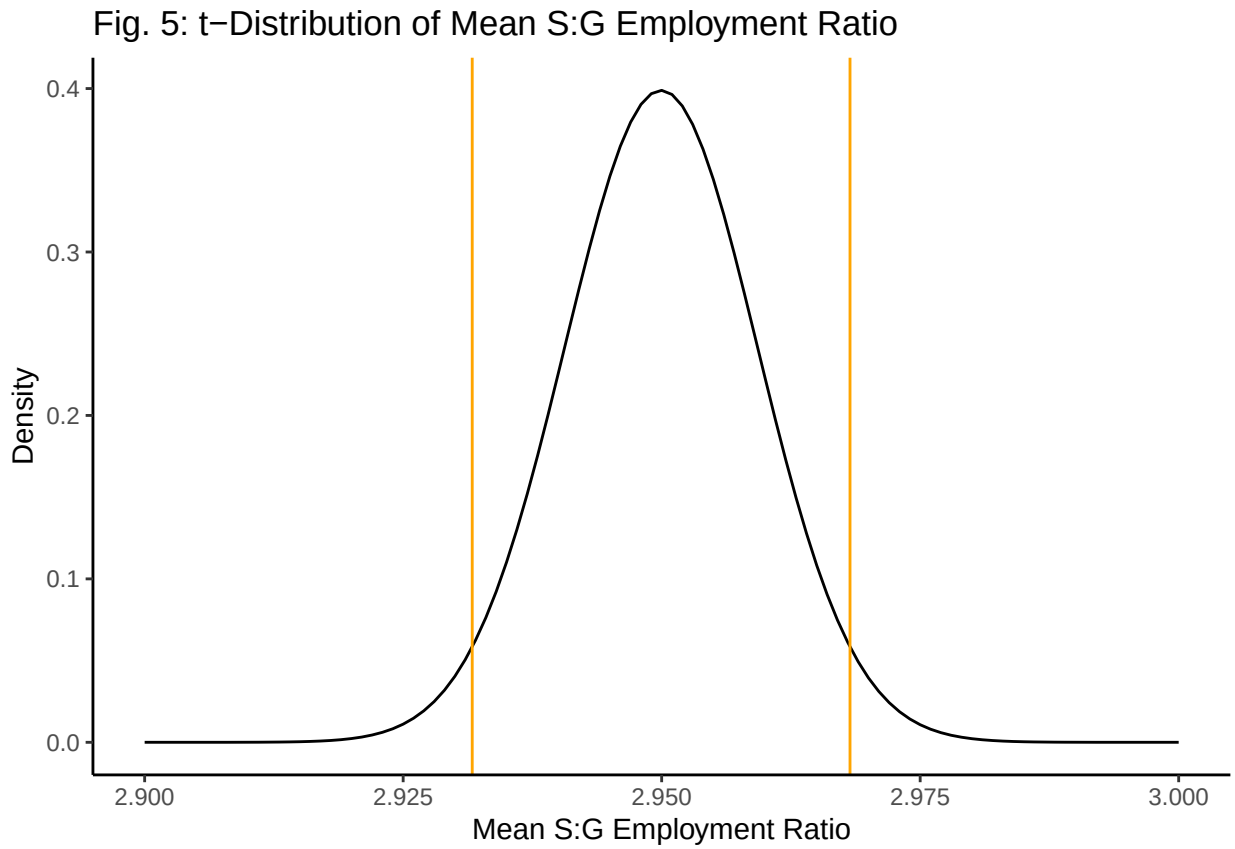

Figure 5 depicts the t-distribution (without centering on 0) that represents the distribution of the mean of S:G employment ratios. The orange line shows the critical points we established in figure 4 as the markers for our 95% confidence interval. As can be seen, the interval in between them makes up for about 95% of the probability of the t-distribution.

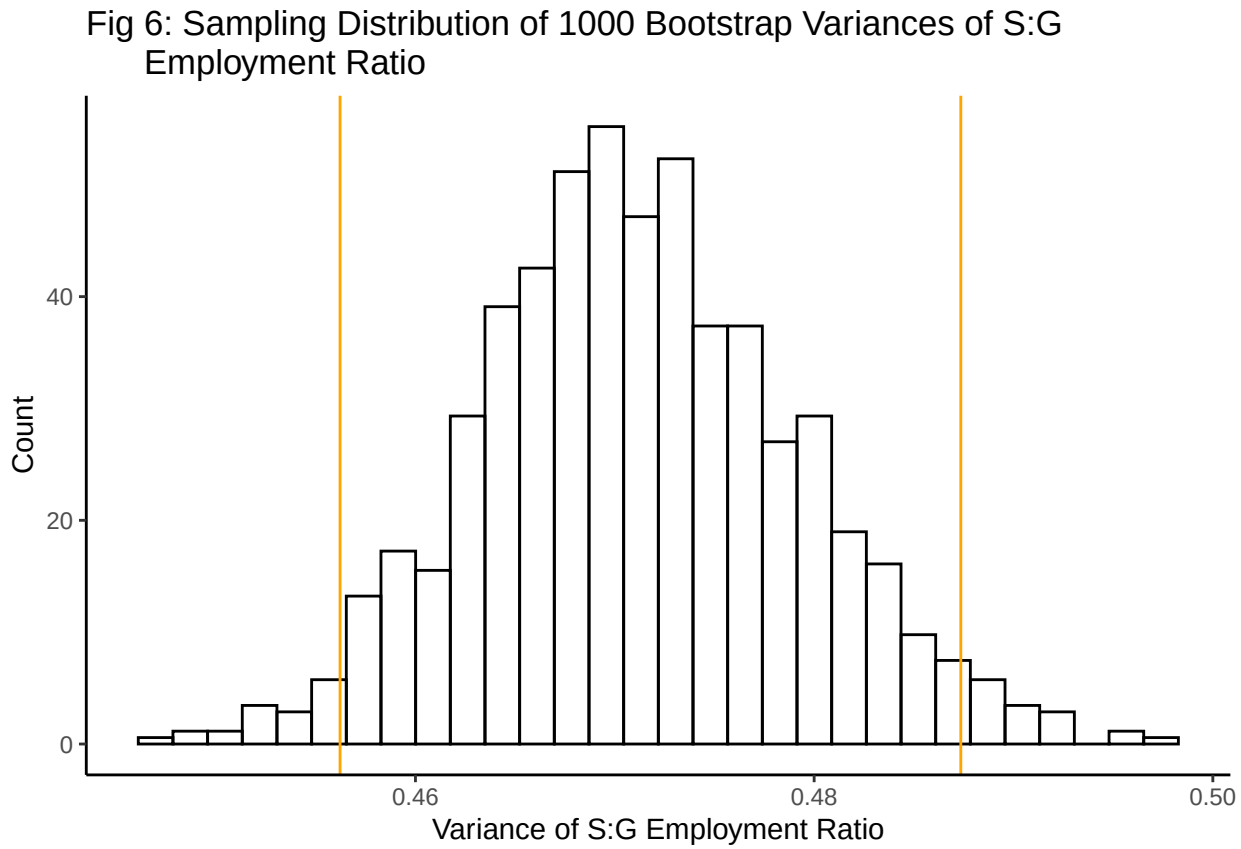

Figure 6 shows the bootstrap sampling distribution of variances of the S:G employment ratio from 1000 bootstrap samples. Again, the orange line represents the endpoints on our confidence interval and it is graphically clear that around 95% of the distribution lies within those two points.

This investigation rests upon the central assumption that the S:G employment ratio is normally distributed. This is not an unreasonable assumption as the sample data seems to possess the bare essence of a normal distribution. This was depicted in figure 3 where we saw it had one peak in the middle. The results we have received seem to be in line with the sample statistics we calculated in figure 1 and thus we may assume that they are reasonable as well. Of course, this entire investigation depends on the original data supplied by Statistics Canada as being largely representative of the whole population. Given the almost 5 decades in which Statistics Canada has engaged in economic estimates and sampling, there is good reason to trust this.

Conclusions

In this investigation, we have attempted to better understand the change in Canada's economic structure from 1976 to 2021 by measuring the share of service sector jobs and the share of goods sector jobs. We did this by measuring the S:G employment ratio, where the number of service sector jobs in a region at a time is divided by the number of goods sector jobs in the same region at the same time.

Our methodology aimed to build 2 intervals that encapsulated the mean and variance of the S:G employment ratio with 95% confidence levels. For the mean, we used the sample mean and standard deviation to generate a t-distribution of the mean S:G employment ratio. We then took the critical points of the t-distribution to determine which values made up the endpoints of this confidence interval. For the variance, we used empirical bootstrapping to resample data, generating 1000 bootstrap sampling distributions of the S:G employment ratio, for which we each calculated a variance. We then put all these

variances together to build a distribution of the variance of S:G employment ratios. Again, we took the critical points of this distribution to determine the values of our confidence interval.

Our initial hypotheses were that the mean S:G employment ratio will be somewhere between 2 to 3 and that the variance of the service to goods employment ratio will be less than 1. To this end, our hypotheses have been rendered correct by the results of our investigation. We can say with a 95% confidence level that the mean S:G employment ratio is between 2.9316792 and 2.9682148 and that the variance S:G employment ratio is between 0.4562181 and 0.4873569. Both of these confidence intervals lie well within the range we have specified.

The relatively low variance suggests a relatively smaller geographical disparity. This seems to imply that Canada as a whole has uniformly evolved its economic structure. Provinces have largely made the transition from goods sector jobs to services sector jobs together and the rate of growth is similar even if they are at different levels. The mean S:G employment ratio is bound to be above 1 which means that Canada, as a developed nation, is very much committed more to the service industry and this is likely to continue and grow.

The shortcoming of this investigation is that the smallest denomination of a region is by province. This means that we are unable to measure the difference between rural and urban areas. Whilst some provinces surely contain more urban powerhouses than others, just looking at provinces doesn't necessarily paint the disparity between rural and urban employment. This is especially important as goods sector jobs are more likely to be in rural areas and service sector jobs are more likely to be in urban centers. We also failed to account for the population growth that has occurred in big cities over the last 5 decades. Whether Canada's transition from a goods economy to a services economy is due to this urban population growth or an actual shrinkage of its goods industry is unclear. Furthermore, the measurement of employment itself might not be wholly representative of the status of industries. Manufacturing jobs are increasingly automated, meaning that there could be an even greater share of output than before but less employment in the goods industry. All of this warrants a closer look.

For the future, statisticians and policymakers eager to address economic issues should look closer at all of the above-mentioned issues. They should focus more on economic output such as tonnage or money. If the issue they are dealing with pertains to Canada's economic self-dependency or autarky, they should particularly pay attention to the agriculture sector and how much of Canada's goods and services are exports or depend on imports. They should start to examine the S:G employment ratio within counties, towns, cities, even neighborhoods maybe. All of this depends on much more exact and representative sampling.

In conclusion, this investigation has examined Canada's economic balance between the services sector and the goods sector. It has discovered that Canada's transition over the last 5 decades to a majority-services country has largely been uniform between provinces. It has found that its dependency on the services sector is strong. In closing, several recommendations have been made for future investigations that can aid in understanding our nation's economic structure.

Bibliography

1. Grolemund, G. (2014, July 16) *Introduction to R Markdown.* RStudio. https://rmarkdown.rstudio.com/articles_intro.html. (Last Accessed: January 15, 2021)

2. Dekking, F. M., et al. (2005) *A Modern Introduction to Probability and Statistics: Understanding why and how.* Springer Science & Business Media.

3. Allaire, J.J., et. el. *References: Introduction to R Markdown.* RStudio. https://rmarkdown.rstudio.com/docs/. (Last Accessed: January 15, 2021)

4. Statistics Canada (2021, March) *Employment by industry, monthly, seasonally adjusted.* [https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410035502] (https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410035502)

All analysis for this report was programmed using `R version 4.0.2`. Libraries used: openintro, tidyverse, statcanR