

1. Introduction (abstract)

The report aims to investigate how the decrease in permanent resident population is associated with the burden of the existing labor force in Heilongjiang, which is measured by dependency ratio. The significance of this investigation is that with a clear association of permanent resident population and dependency ratio, we can further study the causal relationship between those two variables. We will conduct a linear regression between those two variables with data gathered from 2010 to 2019 to get the value of the correlation coefficient, a hypothesis test for slope of the least square regression line to see if the correlation coefficient is statistically significant, and multiple regression, at the end, to refine our current linear model. We expect to find a negative correlation between permanent resident population and dependency ratio. For the hypothesis test, our null hypothesis will be $\beta=0$ and our alternative hypothesis will be $\beta<0$.

2. Background Research

After the reform and opening-up policy has been implemented in China, the overall economic strength of China, especially Southeastern coastal cities like Shenzhen, increased sharply. However, three provinces, Heilongjiang, Jilin, and Liaoning, located in northeast of China seem to be an exception of the economic growth in the 1980s to 1990s. In Heilongjiang province specifically, the reason of the stagnation or even decline of the economy is attributed to the fact that it almost solely depends on heavy industry as the backbone of Heilongjiang's economy. Environmental protection policy, high transportation expenses due to its location and harsh environment, and less demand for its unadvanced product impeded the development of heavy industry in Heilongjiang province.

The decline in economy, job opportunities, infrastructure, and public services make young people from Heilongjiang leave for highly developed cities such as Beijing and Shanghai (Mo Yang, 2022). According to the data of the "Seventh National Population Survey," the overall household population of the three provinces in Northeast China is 103.46 million, which is about 4.46 million less than ten years ago; the resident population is 98.51 million, which is about 11 million less than ten years ago. The major motivation for local people to leave Heilongjiang province is to seek economic gain (Mengyu Shang, 2022). LinXi University researchers suggest that older people left alone in Heilongjiang province require family care and heavy family burden, which leads to a higher dependency ratio. 1% increase in the old-age dependency ratio leading to a decrease of about 0.04% in the likelihood of household entrepreneurship, which demonstrates that decrease in population also negatively impacts economy and that there is a vicious cycle between decrease in population and decline in economy (Xinrui Wang, Rui Hou, Chunyu Qi, 2023).

Previous researchers have also offered methods to solve the problem of population loss in Heilongjiang province, such as upgrade the industrial structure, increase job opportunities, and raise level of payment (Fenggang Lu, 2021).

3. Methods and Procedure:

Data Collection: We collected the raw data of permanent resident population and dependency ratio of Heilongjiang province from year 2010 to 2019. We collected it from National Bureau of Statistics, which offers authoritative data to be used. Our raw data is shown below:

Year	Permanent Resident Population (million)	Dependency Ratio (%)
2010	38.33	25.4
2011	37.82	24.6
2012	37.24	26.5
2013	36.66	26.5
2014	36.08	26.6
2015	35.29	27.3
2016	34.63	28.2
2017	33.99	28.3
2018	33.27	29.5
2019	32.55	31.2

Table 1. Permanent Resident Population and Dependency Ratio of Heilongjiang Province from year 2010 to 2019

Then we will normalize the raw data:

Year	Z-score of Permanent Resident Population	Z-score of Dependency Ratio
2010	1.48266143	-1.323370861
2011	1.371982713	-1.496984618
2012	1.24611274	-1.084651946
2013	1.120242766	-1.084651946
2014	0.994372792	-1.062950226
2015	0.822929207	-0.911038189
2016	0.679697858	-0.715722713
2017	0.540806853	-0.694020993
2018	0.384554471	-0.433600358
2019	0.22830209	-0.064671124

Table 2. Z-score of Permanent Resident Population and Z-score of Dependency Ratio

Then we convert the raw data into a scatterplot:

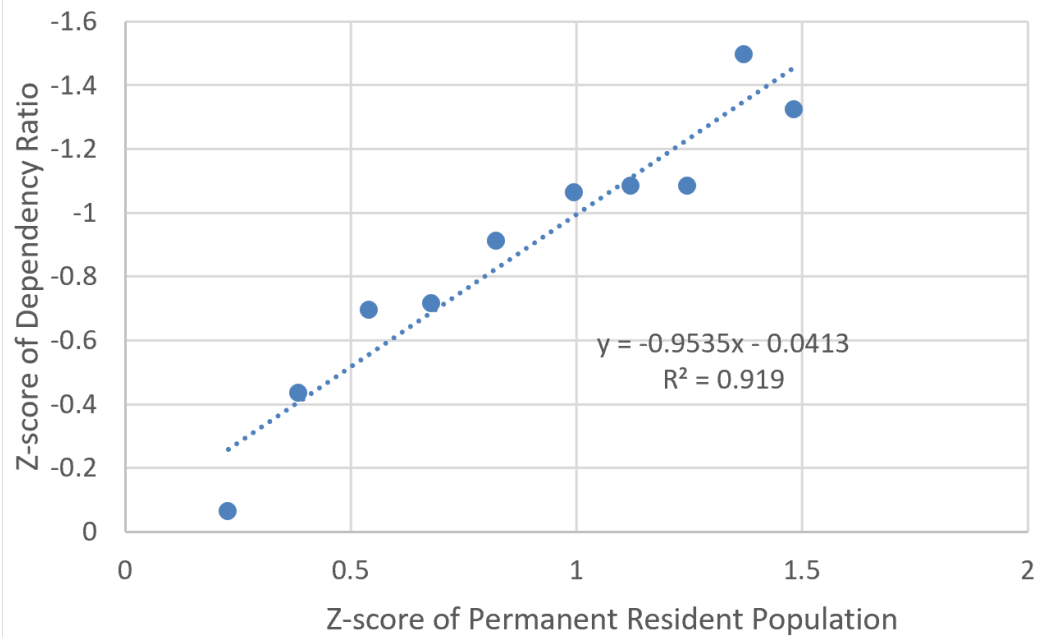


Figure 1. Scatterplot of Z-score of Dependency Ratio and Z-score of Permanent Resident Population

Figure 1. shows that the equation for the least square regression line is $y = -0.9535x + 0.0413$, and the coefficient of determination, R^2 , for the regression line is 0.919, which means that about 91.9% of the variation in dependency ratio can be explained by its linear relationship with permanent resident population. This shows that the linear model is appropriate when modeling the relationship between z-score of permanent resident population and z-score of dependency ratio, although it may not be the best fit.

To see if the slope of the least square regression line is statistically significant, we need to conduct a hypothesis test for β . We need to check conditions that need to be met first in order to conduct the test. Our explanatory variable,