Demographic structure in China and its implications to economic growth

Zizhen Guo (95929089)

2022-23, ECON 490 013, The University of British Columbia

Demographic structure in China and its implications to economic growth

Introduction

Over the past several decades, China has undergone a significant demographic transition, with declining fertility rates and a rapidly aging population. This has resulted in a shift in the population's age structure, with a declining share of young people and an increasing share of older adults (Wei & Hao, 2010). This demographic shift has had significant implications for the Chinese economy. With a declining share of young workers, the labor force has become less dynamic and productive, leading to slower economic growth (Luo & Zhang, 2010). Additionally, the increasing share of older adults has put pressure on the social welfare system, with rising healthcare and pension costs (F. Zhao et al., 2019). Furthermore, the declining fertility rates have led to a gender imbalance, with a higher proportion of males in the population. This has led to a rise in unmarried males, contributing to social instability and increased crime rates (Q. Wang, Zou, & Fan, 2019).

This study aims to investigate the consequences of the demographic transition in China for economic growth using stepwise regression models with time-lagged explanatory variables. The literature review summarizes relevant studies of the relationship between economic growth in China and demographic components, including population aging, urbanization, fertility, public health, and the evolution of family planning policy in China. The method section describes the dataset and research methods used in this study. The result and discussion section provides an interpretation of the empirical results and evaluates the models' quality. Limits and potential problems are also addressed.

Literature review

The demographic structure in China has entered a stage of transition and evolution, which raised broad concerns about the impacts on the Chinese economy. Wei and Hao (2010) investigated provincial-level data from 1989 to 2004 and found evidence that the decline in fertility levels and a preferable age structure significantly contributed to economic growth after 1989. However, as China experienced a growth slowdown, population aging, and a low fertility rate in recent decades, the dependent factors of the previous study might have changed. The review primarily focuses on the transition of critical demographic indicators corresponding to population aging, area of residence, mortality, life expectancy, fertility, population policies, and their relationship with economic growth from existing theoretical frameworks and empirical analysis incorporating the Chinese context.

Population aging is the direct consequence of the persisting shrinkage in natural population growth caused by a declining birth rate and longer lifespan. (Attane, 2002) viewed the impact of population aging on economic growth as a negative supply shock and an increase in demand. From the supply side, Fang (2018) discussed the demographic transition towards the exhaustion of the "population dividend," which refers to the diminished advantage in human capital that constituted the necessary condition of China's rapid economic growth. From the micro perspective, Giorgi et al. (2020) reviewed the mental health problem in the aging working force and concluded the decline in cognitive performance of older workers when exposed to occupational stress, which may further imply risk factors to the already overloaded medical system. From the demand side, evidence from Wang et al. (2017) 's study suggests that aging contributes to higher demand for medical services, which corresponds to a decrease in aggregate social saving rates. This coincides with the life cycle hypothesis that explains the saving and consumption pattern of older people (Modigliani, 1986). Although classical economics suggests that low savings could reduce future investments and essentially lead to declining economic growth (Nordhaus, 1989), the high demand among the aging population will shift consumer preferences towards nutrition and health support, bringing new opportunities for the "silver economy" (Attane, 2002).

Urbanization concentrates human populations into discrete areas (US EPA, 2015). Therefore, the population in urban and rural areas has always been a reliable proxy for urbanization. Hoselitz (1953) proposed the connection between urbanization and economic growth in 1952. He stated that “the population growth in urban and industrial centers appears inevitable if there is economic development. “ Henderson (2003) assumed an optimal level of urban concentration exists and examined the impact of deviations from the optimality on the economic growth. The results indicated that insufficient and excessive urban concentration is costly to productivity growth. In the Chinese context, Chen et al. (2013) assessed the evolution of urbanization in China from 1960 to 2010 and discussed three development patterns in different periods: the rapid-decline stage, the stable stage, and the rapid-promotion stage. Moreover, the fast progression of China’s urbanization since 2004 has reflected an excessive focus on quantity over quality in urban planning. Based on Chen et al.’s research, Yang et al. (2017) empirically examined the impact of urbanization on economic growth in China and found such effect only positive and significant for samples excluding the major low-income group, which implied difficulties in China’s urbanization targeting and policies.

Aside from previous studies that implicitly investigated the impact of urbanization on economic growth, Hajkova (2014) explicitly analyzed the knowledge base of cities, which refers to the ability to absorb and exchange knowledge (Cantner, Meder, & ter Wal, 2010). The efficiency of knowledge bases is evaluated concerning the performance in knowledge processes such as R&D and education. The research illustrated that internal knowledge transfer with sales effect contributed to economic growth from 2004 to 2009.

Despite the rapid economic and epidemiological changes over the past few decades, China’s burden on the public health system and medical services is still noticeable. Yang et al. (2013) comprehensively investigated diseases, injuries, deaths, and life expectancy in China with a comparison against other G20 countries in 2010. The modest increase in life expectancy placed China in the 12th position in the G20. The results also indicate a surge in non-communicable diseases (such as heart disease and stroke) driven by the side effects of economic growth. The ambient air pollution index in China is the second highest in the G20, while the household air pollution index ranks third. With a focus on pollution, Ebenstein et al. (2015) explained the insignificant growth of China's life expectancy from the "offset by the cost of pollution exposure" supported by two analyses. First, deaths associated with air pollution are found to decline much slower than other causes of death. Moreover, the study examined a significant positive relationship between cardiorespiratory mortality and exposure to pollutants. Furthermore, Liu (2019) explored the unequal income and wealth accumulation under economic growth in Chinese urban and rural areas, which leads to significantly higher cancer mortality rates and all-causes mortality rates for the rural population. The study implied that health benefits from economic growth might induce more substantial impacts on a group of individuals but have little effect on others. Besides, He et al. (2019) considered different stages of economic growth in China and the implications for mortality rates. The stable economic growth period (2005–2008) corresponds to a significant increase in the population mortality rate. This ratio was significantly reduced during the financial crisis period (2008–2010) and was further decreased in the sustained downward economic period (2011–2013), followed by the new normal economic situation (2013–2017). However, the negative effects of mental illness and cardiovascular diseases also become observable over time, which agrees with the results from Yang et al. (2013) and Ebenstein et al. (2015).

As a consistent extension of growth models, Chatterjee (2018) showed that fertility is procyclical in the short run and decays with long-run economic growth. Using multiple fertility measures, the continuous and steady decline of China's fertility levels is confirmed by Yang et al. (2017). The results demonstrated the significance of marriage and childbearing postponement that accounts for the decrease.

The comparative study of China and India by Bloom et al. (n.d.) stated that a remarkable decline in fertility creates a "bulge" generation that affords the economic takeoff from a higher labor force per capita ratio. This theory is supported by Karra et al. (2017) with in-depth illustrations of increasing female participation in the labor force and higher investments in health care and education due to low fertility levels. However, Wei &

Hao (2010) countered this view from the absence of marketization reform in China that could lead to inefficiencies such as a high unemployment rate or over-staffed that would account for a "deficit" rather than a "dividend" for economic growth. It is also noticeable that the economic environment can affect an individual's decisions for fertility. This decision-making process is explained by Easterlin (1975) using an economic framework. It considered three determinants: demand for children, the potential output of children, and the cost-benefit analysis of fertility regulation. Under this framework, Chen (2022) found a positive relationship between education level and the number of children of married Chinese women. This result raised ambiguities about the impacts of economic status on fertility decisions. On one side, the income effect from economic growth allows females to afford more children. However, parents may choose to spend more on existing children rather than having new children.

Methods

Following the previous demographic research, this study centers on the relationship between demographic structure and economic growth in China. The dataset serving for the empirical analysis is sampled and measured by National Bureau of Statistics of China (NBSC), a government agent responsible for nationwide statistical work, and assessed through the World Bank Open Data portal (*World Bank Open Data*, 2022). Due to the incompleteness of data in 1950s and 2020s, the study will focus on annual national data from 1961 to 2020. The annual population statistics are constituted of yearly sample survey of 1% of the entire population and census on every five-year basis.

The experimental group includes GDP as a measure of economic growth of the entire nation, and growth of GDP per capita that measures the change in output level of individuals. The explanatory variables incorporates the results of former studies and includes a subset of China's annual population indicators with strong evidence of those are either actively affecting the economic growth, or passively affected by the growth. These components are separated into five categories population size factor population aging factors that include age dependency ratio and survival to 65 residential factors that include urban and rural population public health factors that include luife expectancy and mortality rate reproduction factors that includes fertility rate, adolescent fertility rate. It is also notable that the residence's committee or villager's committee is the minimum unit comprising urban classification in China. And the rural area is the rest of region excluding the classified urban area (*Hunan Provincial Bureau of Statistics*, 2006).

The primarily empirical methodology is the stepwise regression model with time-lagged explanatory variables. Stepwise regression is the derivation of the reduced regression model with an optimal feature subset from all original features without losing the goodness of fit (Wieczorek & Lei, 2022). An 8:2 train-test split in time order is applied as a validation procedure. Specifically, the training set contains data from 1961 to 2006, and the testing set includes the later statistics up to 2020. From which the observations in traning set are used to estimate the function with respect to economic growth, and the testing set serves the purpose of evaluating the performance of the models (James, Witten, Hastie, & Tibshirani, 2013). The notations and mathematical expressions in this part use GDP growth as the dependent variable, but the same approach will also be applied with GDP per capita.

We first introduce the notations and the basic OLS model. Equation 1.1 represents the basic ordinary least square model, with GDP growth being the dependent variable and demographic components being the regressors. The classical pre-conditions of linear regression models on the finite sample are assumed, and the residual e_j is assumed identically and independently distributed with zero mean and unknown variance (Wooldridge, 2016).

Notation	Variable
p	Number of regressor
n	Sample size
e	Error term / Residual
D	Demographic indicators (explanatory variables)
GDP	GDP growth (dependent variables)
k	Maximum number of regressors

$$\text{(Equation 1.1)} \quad GDP_j = \beta_0 + \sum_{i=1}^{p} \beta_i D_{ij} + e_j$$

Lagged variables are engineered in this study to preserve the sequential property of the time-series data and to determine the time-dependent causality. A simple approach to lagged features is stated by equation (2.1).

$$\text{(Equation 2.1)} \quad GDP_j = \beta_0 + \{\sum_{i=1}^{p} \beta_{0i} D_{0ij} + e_{0j}\}_{\text{lag}=0} + dots + \{\sum_{i=1}^{p} \beta_{5i} D_{5ij} + e_{5j}\}_{\text{lag}=5}$$

However, as the regression model constructed from Equation 2.1 contains more regressors than observations, the OLS coefficients are no longer unique, and this singularity would lead to invalid results. A time-lagged regression model will be constructed as an alternative approach for each specified year. In particular, a maximum of five-year lagged feature will be considered. Therefore, in addition to the basic OLS model in Equation 1.1, five distinct lag-based models from a one-year shift to a five-year shift are established in Equation 2.2.

(Equation 2.2)

$$GDP_{1j} = \beta_{1,0} + \{\sum_{i=1}^{p} \beta_{0i} D_{0ij} + e_{0j}\}_{\text{lag}=0} + \{\sum_{i=1}^{p} \beta_{1i} D_{1ij} + e_{1j}\}_{\text{lag}=1}$$

$$\vdots$$

$$GDP_{5j} = \beta_{5,0} + \{\sum_{i=1}^{p} \beta_{0i} D_{0ij} + e_{0j}\}_{\text{lag}=0} + \{\sum_{i=1}^{p} \beta_{5i} D_{5ij} + e_{5j}\}_{\text{lag}=5}$$

James (2013) stated that the variable selection algorithms mainly serve two purposes. The first goal is prediction accuracy, provided that with respect to a fixed n, models with smaller p tend to have lower variance and less overfitting. Another is model interpretability, as the algorithm reduced insignificant variables that lead to "unnecessary complexity" in the model. Insignificant variables are either irrelevant or redundant. Although both types are safe to remove from the model, redundant variables could still imply a significant relationship with the response variable Guyon (2003). A stepwise method is chosen for this study over different feature selection algorithms because of its cost-effectiveness in computing resources compared to the best subset selection and its advantages in interpretability compared to the shrinkage approach, such as the ridge and LASSO regression. A brief demonstration is as follows.

Forward selection is initiated with a null model shown in Equation 3.1 and adds the most significant variable on each recurring step until a specific stopping criterion is triggered (Hocking, 1976).

$$\text{(Equation 3.1)} \quad GDP_{lj} = \beta_0$$

Equation 3.2 is the necessary condition for the variable X_j to enter the regression model, where SSE is the sum of squared errors derived using Equation 3.3.

$$\text{(Equation 3.2)} \quad F_j = \max_j \quad \left(\frac{SSE_p - SSE_{p+1}}{SSE_p/(n-p)}\right) > F_{\text{in}}$$

$$\text{(Equation 3.3)} \quad SSE_p = \sum_{j=1}^{p}(GDP_j - G\hat{D}P_{pj})^2$$

Reversely, backward selection starts from a full model and removes the least significant variable at each step (N. Zhao, Liu, Cao, Samson, & Zhang, 2017). Equation 3.4 is the full model for this empirical study. Remarkably, when lag = 0, the equation reverts to the basic OLS model.

$$\text{(Equation 3.4)} \quad GDP_{lj} = \beta_{l,0} + \{\sum_{i=1}^{p}\beta_{0i}D_{0ij} + e_{0j}\}_{\text{lag}=0} + \{\sum_{i=1}^{p}\beta_{li}D_{lij} + e_{lj}\}_{\text{lag}=l}$$

Variable X_j will be eliminated from the regression model if Equation 3.5 is satisfied.

$$\text{(Equation 3.5)} \quad F_j = \min_j \quad (\frac{SSE_{P-1} - SSE_P}{SSE_p/(n-p)}) < F_{\text{out}}$$

The criteria for the algorithm to terminate depends on the local minima of Mallow's C_p to the regression model since a smaller C_p implies a better fit in general. Mallow's C_p, as defined by Equation 3.6, is a metric to evaluate the goodness of fit of OLS models with corrections of the overfitting problem, which is common for a model with many regressors but insufficient observations (Mallows, 1973). Equation 3.7 expands Equation 3.6 regarding the differences between predicted and actual values in the training set.

$$\text{(Equation 3.6)} \quad C_p = \frac{SSE_p}{\hat{\sigma}^2} - n_{\text{train}} + 2(p+1)$$

$$\text{(Equation 3.7)} \quad C_p = \frac{\sum_{i=1}^{p}(GDP_i - G\hat{D}P_{pi})}{\sum_{j=1}^{k}(GDP_j - G\hat{D}P_j)^2} - n_{\text{train}} + 2(p+1)$$

The last stage concentrates on finding a best-fitting model. From the previous steps, two regression models using forward and backward selection will be created for each year of time lagging from 0 to 5. In addition, the full OLS model is adapted as the control group. The performance of the models is measured by root-mean-square error (RMSE), which is the square root of the mean squared difference in the actual GDP growth and the prediction in the testing set, as pronounced in Equation 4.1. Models with lower RMSE commonly have a better fitting to the actual values, which implies a more accurate estimation of the actual relationship between demographic factors and economic growth. Hence the resulting model with minimum RMSE will be the optimal model.

$$\text{(Equation 4.1)} \quad RMSE = \sqrt{\frac{\sum_{j=1}^{n_{\text{test}}}(GDP_{lj} - G\hat{D}P_{lj})^2}{n_{\text{test}}}}$$

Results and discussion

Table 1.

Term	Coefficient	Standard error	Test statistic	P-value
(Intercept)	7.95	1.73	4.58	0.00
Adolescent fertility rate	0.19	0.17	1.11	0.27
Death rate, crude	0.39	0.14	2.78	0.01
Life expectancy at birth, female	34.60	81.02	0.43	0.67
Life expectancy at birth, total	-51.66	165.59	-0.31	0.76
Mortality rate, adult, male	20.12	85.83	0.23	0.82
Population, female	-156.53	95.64	-1.64	0.11

Table 2.

Term	Coefficient	Standard error	Test statistic	p-value
(Intercept)	0.67	5.87	0.11	0.91
Birth rate, crude	-0.13	0.06	-2.08	0.04
Fertility rate, total	-0.05	0.98	-0.05	0.96
Survival to age 65, female	0.38	0.62	0.62	0.54
Adolescent fertility rate, 1-year lag	-0.16	0.28	-0.57	0.57
Mortality rate, adult, female, 1-year lag	-0.23	0.75	-0.30	0.76
Fertility rate, total, 1-year lag	0.25	0.87	0.29	0.78
Age dependency ratio, 1-year lag	-0.78	0.86	-0.91	0.37
Rural population, 1-year lag	1.06	3.61	0.29	0.77
Urban population, 1-year lag	2.08	1.39	1.50	0.14

Table 1 reports the resulting model with an optimal estimation of GDP growth using backward selection with six variables, whereas all lagged features are insignificant in predicting the growth trend. Table 2 represents the optimal growth model in GDP per capita with nine variables, and six are time-lagged by one year. The multicollinearity issue is observed in Table 1 and Table 2 because the variable and model selection process is merely based on the value of C_p and RMSE and omits the internal correspondence between features. The collinear population variables are interpreted as a group to address this problem, and the aggregate relationship with economic growth will be examined.

In summary, reproduction factors have a slightly positive relationship with GDP growth, but the coefficients are negative for GDP per capita. The ambiguous results match previous studies' controversies (Bloom et al., n.d.) (Wei & Hao, 2010). Nevertheless, the opposite correlation suggests that although an increase in fertility levels might positively influence economic growth with an expansion of the "population dividend," the potential cost of raising children still worsens the economic status of individuals.

Population aging factors appear only in the second model for GDP per capita. The absence from the first model implies that population aging factors are insignificant in estimating GDP growth. However, with sufficient evidence that the deformation of the aging structure in China imposes pressures on economic growth (Fang, 2018), we deduce the population aging factors are not irrelevant but redundant in the first model since public health factors may have covered some variations of GDP growth driven by population aging. In the second model, the negative relationship for GDP per capita growth confirms Attane (2002) 's theory in the literature review.

Residential factors are also reduced from the first backward selection model of GDP. The potential reason could be the debatable state of urbanization in China. In the review section, Yang et al. (2017) stated that the concentration of capital and resources would benefit the economy. However, Chen et al. (2022) argued that China is experiencing excessive and inefficient urbanization that could offset the economic growth promoted by urban concentration, making the residential factors a less effective predictor of the GDP growth trend. Nonetheless, in the second model, a 1% increase in the rural population is correlated to a 1% growth in GDP per capita, and the impact is doubled for the same change in the urban population. The results from the second model support a positive relationship between urbanization progression and an individual's production level. However, the substitution effect from excessive urbanization still requires further research.

Public health factors have a doubting relationship in the first model, with a negative coefficient for life expectancy and a positive coefficient for death statistics. This contradicts the "modest increase in life expectancy" concluded in Yang et al. (2013) 's research. However, evidence also supports that pressure in the medical system, pollution exposure, and a surge of non-communicable and chronic diseases may offset the increasing lifespan of Chinese citizens from rapid economic growth (Ebenstein et al., 2015), (Liu, 2019).

Figure 1.

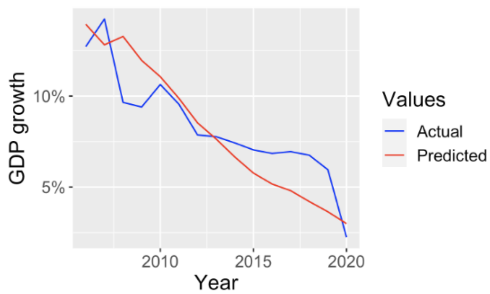

Figure 2.

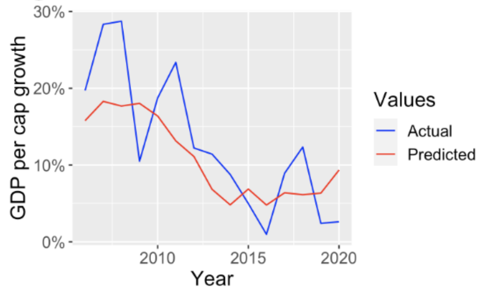

Figure 1 illustrates a summary of the estimation versus the actual values for observations in the testing set using the first model. As the estimated GDP growth (red line) imitates the overall downward trend of the actual GDP growth (blue line), the model failed to capture the major volatility of GDP growth induced by the 2008 financial crisis. The model also underestimates the growth performance in China from 2013 to 2019. This underfitting problem might be attributed to the high-leverage growth shock in 2020 due to the explosion of the COVID-19 pandemic.

Figure 2 is the same representation using the second model. Although the model can estimate the overall trend of the changes in GDP per capita, the estimation of the peaks and troughs around the 2008 financial crisis is inaccurate. Limitations still exist for capturing the major fluctuations over the period.

It is noticeable from both Table 1 and Table 2 that the p-values for both models are commonly large because p-values determine the statistical significance of coefficients under the training set. However, the criterion of model selection is based on the testing set. The validation approach adopted in this study avoids overfitting the training data, but conversely, this may result in overfitting the testing data (Burnham, Anderson, & Burnham, 2002). This leads to the trade-off between prediction accuracy and model interpretability, which refers to the difficulties in inference and interpretation with the increasing complexity of the algorithm aiming for more accurate results (James et al., 2013).

Conclusion

This study examined and interpreted the relationship between demographic structure and economic growth in the Chinese context from 1961 to 2020. The empirical methods produced the best-fitted stepwise regression model with time-lagged variables for GDP growth and GDP per capita growth. The problem of multicollinearity and large p-value were identified and addressed. The results show that reproduction, public health, population aging, and residential factors have significant effects in estimating either one or both GDP growth and GDP per capita growth.

The limitations of this study lie within the incompleteness of empirical strategies and the absence of the government's role. First, the process of feature selection and model selection only depends on the goodness of fit but neglects other factors such as distribution, variance, and statistical significance. This also leads to a regression model with collinear features that impose difficulties for the inferences and interpretations. Second, the inter-correlation between the origin feature and its time-lagged shift is omitted. An autoregressive model should be implemented to address this problem, which requires the author to handle the complexity added to the algorithm. Third, this study does not include the impacts of the family planning policy, which contains two major flaws that have induced serve economic consequences: the harm to the population structure and the intensification of regional population differences (Wu, 2020).

Under the challenges of the deforming demographic structure and decline in economic growth, the Chinese government adopted the policy relaxation to a maximum of two children in 2015 (*Chinese Government Website*, 2015), and further adjustments to the three-child policy were imposed in 2021 (*Xinhua News*, 2021). However, negative influences on labor productivity and social stability due to an unhealthy population structure persist, which urges policymakers to develop a sustainable system for these problems.

References

Attane, I. (2002). China's Family Planning Policy: An Overview of Its Past and Future. *Studies in Family Planning*, *33*(1), 103–113. Retrieved 2022-10-31, from https://www.jstor.org/stable/2696336 (Publisher: [Population Council, Wiley])

Bloom, D. E., Canning, D., Hu, L., Liu, Y., Mahal, A., & Yip, W. (n.d.). Why Has China's Economy Taken Off Faster than India's? , 40.

Burnham, K. P., Anderson, D. R., & Burnham, K. P. (2002). *Model selection and multimodel inference: a practical information-theoretic approach* (2nd ed ed.). New York: Springer. (OCLC: ocm48557578)

Cantner, U., Meder, A., & ter Wal, A. L. J. (2010, September). Innovator networks and regional knowledge base. *Technovation*, *30*(9), 496–507. Retrieved 2022-12-05, from https://www.sciencedirect.com/science/article/pii/S0166497210000507 doi: 10.1016/j.technovation.2010.04.002

Chatterjee, S., & Vogl, T. (2018, June). Escaping Malthus: Economic Growth and Fertility Change in the Developing World. *The American Economic Review*, *108*(6), 1440–1467. Retrieved 2022-12-05, from https://www.proquest.com/docview/2051134125/abstract/25A36AB73C8E411DPQ/1 (Num Pages: 28 Place: Nashville, United States Publisher: American Economic Association) doi: 10.1257/aer.20170748

Chen, M., Liu, W., & Tao, X. (2013, April). Evolution and assessment on China's urbanization 1960–2010: Under-urbanization or over-urbanization? *Habitat International*, *38*, 25–33. Retrieved 2022-12-03, from https://www.sciencedirect.com/science/article/pii/S0197397512000537 doi: 10.1016/j.habitatint.2012.09.007

Chen, S. (2022, March). The Positive Effect of Women's Education on Fertility in Low-Fertility China. *European Journal of Population*, *38*(1), 125–161. Retrieved

2022-12-05, from https://doi.org/10.1007/s10680-021-09603-2 doi: 10.1007/s10680-021-09603-2

Chinese government website. (2015). Retrieved 2022-10-05, from http://www.gov.cn/

Easterlin, R. A. (1975). An Economic Framework for Fertility Analysis. *Studies in Family Planning*, *6*(3), 54–63. Retrieved 2022-12-05, from https://www.jstor.org/stable/1964934 (Publisher: [Population Council, Wiley]) doi: 10.2307/1964934

Ebenstein, A., Fan, M., Greenstone, M., He, G., Yin, P., & Zhou, M. (2015, May). Growth, Pollution, and Life Expectancy: China from 1991-2012. *The American Economic Review*, *105*(5), 226–231. Retrieved 2022-12-05, from https://www.proquest.com/docview/1679723529/abstract/F6D0E906E83F4F26PQ/1 (Num Pages: 6 Place: Nashville, United States Publisher: American Economic Association Section: Papers; High Stakes Energy and Environmental Problems in Developing Countries) doi: 10.1257/aer.p20151094

Fang, C. (2018, September). Population dividend and economic growth in China, 1978–2018. *China Economic Journal*, *11*(3), 243–258. Retrieved 2022-12-02, from https://www.tandfonline.com/doi/full/10.1080/17538963.2018.1509529 doi: 10.1080/17538963.2018.1509529

Giorgi, G., Lecca, L. I., Leon-Perez, J. M., Pignata, S., Topa, G., & Mucci, N. (2020, January). Emerging Issues in Occupational Disease: Mental Health in the Aging Working Population and Cognitive Impairment—A Narrative Review. *BioMed Research International*, *2020*, 1742123. Retrieved 2022-12-03, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7011340/ doi: 10.1155/2020/1742123

Guyon, I., & Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*, *3*(Mar), 1157–1182. Retrieved 2022-12-07, from https://jmlr.csail.mit.edu/papers/v3/guyon03a.html

He, F.-M., Chang, T., Dou, Z.-J., Li, F., & Chang, K.-C. (2019, December). Non-linear Impact of China's Economic Growth on the Health of Residents—An Empirical Study Based on TVP-FAVAR Model. *Frontiers in Public Health*, *7*, 380. Retrieved 2022-12-05, from `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6933769/` doi: 10.3389/fpubh.2019.00380

Henderson, V. (2003, March). The Urbanization Process and Economic Growth: The So-What Question*. *Journal of Economic Growth*, *8*(1), 47–71. Retrieved 2022-12-03, from `https://www.proquest.com/docview/197700297/abstract/CBA2764B53C84423PQ/1` (Num Pages: 47-71 Place: Boston, Netherlands Publisher: Springer Nature B.V.)

Hocking, R. R. (1976). A Biometrics Invited Paper. The Analysis and Selection of Variables in Linear Regression. *Biometrics*, *32*(1), 1–49. Retrieved 2022-12-07, from `https://www.jstor.org/stable/2529336` (Publisher: [Wiley, International Biometric Society]) doi: 10.2307/2529336

Hoselitz, B. F. (1953). The Role of Cities in the Economic Growth of Underdeveloped Countries. *Journal of Political Economy*, *61*(3), 195–208. Retrieved 2022-12-03, from `https://www.jstor.org/stable/1824908` (Publisher: University of Chicago Press)

Hunan provincial bureau of statistics. (2006). Retrieved 2022-10-05, from `http://tjj.hunan.gov.cn/xxgk/tzgg/201507/t20150717_3772301.html`

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning* (Vol. 103). New York, NY: Springer New York. Retrieved 2022-12-07, from `http://link.springer.com/10.1007/978-1-4614-7138-7` doi: 10.1007/978-1-4614-7138-7

Karra, M., Canning, D., & Wilde, J. (2017). The Effect of Fertility Decline on Economic Growth in Africa: A Macrosimulation Model. *Population and Development Review*, *43*(S1), 237–263. Retrieved 2022-12-05, from `https://onlinelibrary.wiley.com/doi/abs/10.1111/padr.12009` (_eprint:

https://onlinelibrary.wiley.com/doi/pdf/10.1111/padr.12009) doi: 10.1111/padr.12009

Liu, L. (2019, January). Rural–urban inequities in deaths and cancer mortality amid rapid economic and environmental changes in China. *International Journal of Public Health*, *64*(1), 39–48. Retrieved 2022-12-05, from https://doi.org/10.1007/s00038-018-1109-3 doi: 10.1007/s00038-018-1109-3

Luo, C., & Zhang, J. (2010). Declining Labor Share: Is China's Case Different? *China & World Economy*, *18*(6), 1–18. Retrieved 2022-12-08, from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1749-124X.2010.01217.x (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1749-124X.2010.01217.x) doi: 10.1111/j.1749-124X.2010.01217.x

Mallows, C. L. (1973). Some Comments on CP. *Technometrics*, *15*(4), 661–675. Retrieved 2022-12-07, from https://www.jstor.org/stable/1267380 (Publisher: [Taylor & Francis, Ltd., American Statistical Association, American Society for Quality]) doi: 10.2307/1267380

Modigliani, F. (1986). Life Cycle, Individual Thrift, and the Wealth of Nations. *Science*, *234*(4777), 704–712. Retrieved 2022-12-02, from https://www.jstor.org/stable/1697934 (Publisher: American Association for the Advancement of Science)

Nordhaus, W. D. (1989). What's Wrong with a Declining National Saving Rate ? *Challenge*, *32*(4), 22–26. Retrieved 2022-12-02, from https://www.jstor.org/stable/40691830 (Publisher: Taylor & Francis, Ltd.)

US EPA, O. (2015, November). *Urbanization - Overview* [Collections and Lists]. Retrieved 2022-12-03, from https://www.epa.gov/caddis-vol2/urbanization-overview

Wang, F., Zhao, L., & Zhao, Z. (2017, January). China's family planning policies and their labor market consequences. *Journal of Population Economics*, *30*(1), 31–68. Retrieved 2022-12-06, from https://www.proquest.com/docview/1833752502/

abstract/C3DAC1D68E2849D6PQ/1 (Num Pages: 31-68 Place: Heidelberg, Netherlands Publisher: Springer Nature B.V.) doi: 10.1007/s00148-016-0613-0

Wang, Q., Zou, Y., & Fan, D. (2019, December). Gender imbalance in China's marriage migration: Quantitative evidence and policy implications. *Economic Modelling*, *83*, 406–414. Retrieved 2022-12-08, from https://www.sciencedirect.com/science/article/pii/S0264999318317565 doi: 10.1016/j.econmod.2019.09.040

Wei, Z., & Hao, R. (2010, December). Demographic structure and economic growth: Evidence from China. *Journal of Comparative Economics*, *38*(4), 472–491. Retrieved 2022-12-05, from https://www.sciencedirect.com/science/article/pii/S014759671000065X doi: 10.1016/j.jce.2010.08.002

Wieczorek, J., & Lei, J. (2022). Model selection properties of forward selection and sequential cross-validation for high-dimensional regression. *Canadian Journal of Statistics*, *50*(2), 454–470. Retrieved 2022-12-07, from https://onlinelibrary.wiley.com/doi/abs/10.1002/cjs.11635 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cjs.11635) doi: 10.1002/cjs.11635

Wooldridge, J. M. (2016). *Introductory econometrics: a modern approach* (Sixth edition ed.). Boston: Cengage Learning. (OCLC: ocn906011217)

World bank open data. (2022). Retrieved 2022-10-05, from https://data.worldbank.org/

Wu, P. (2020). *Population Development Challenges in China: Family Planning Policy and Provincial Population Difference.* Singapore: Springer Singapore. Retrieved 2022-12-06, from https://link.springer.com/10.1007/978-981-15-8010-9 doi: 10.1007/978-981-15-8010-9

Xinhua news. (2021). Retrieved 2022-10-05, from http://m.news.cn/

Yang, G., Wang, Y., Zeng, Y., Gao, G. F., Liang, X., Zhou, M., ... Murray, C. J. (2013, June). Rapid health transition in China, 1990–2010: findings from the Global Burden of Disease Study 2010. *The Lancet*, *381*(9882), 1987–2015. Retrieved 2022-12-05, from https://www.sciencedirect.com/science/article/pii/S0140673613610971 doi: 10.1016/S0140-6736(13)61097-1

Yang, Y., Liu, J., & Zhang, Y. (2017, September). An analysis of the implications of China's urbanization policy for economic growth and energy consumption. *Journal of Cleaner Production*, *161*, 1251–1262. Retrieved 2022-12-03, from https://www.sciencedirect.com/science/article/pii/S0959652617306741 doi: 10.1016/j.jclepro.2017.03.207

Zhao, F., Zhang, M., Xuan, J., Mo, Y., Huang, J., Liu, Z., … Guo, X. (2019, September). Burden of insulin injection-related needlestick injuries in mainland China–prevalence, incidence, and healthcare costs. *International Journal of Nursing Studies*, *97*, 78–83. Retrieved 2022-12-08, from https://www.sciencedirect.com/science/article/pii/S0020748919301294 doi: 10.1016/j.ijnurstu.2019.05.006

Zhao, N., Liu, Y., Cao, G., Samson, E. L., & Zhang, J. (2017, May). Forecasting China's GDP at the pixel level using nighttime lights time series and population images. *GIScience & Remote Sensing*, *54*(3), 407–425. Retrieved 2022-10-31, from https://doi.org/10.1080/15481603.2016.1276705 (Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/15481603.2016.1276705) doi: 10.1080/15481603.2016.1276705