Mathematical Interpretation of China's Economic Outputs

The Role of Demographics in Economic Growth

Presenters

Ye Liu
Major in Data Science
yl6408@nyu.edu

Yichen Qian
Major in Mathematics
yq801@nyu.edu

Qihang Xu
Major in Mathematics
qx512@nyu.edu

Jianan Liao
Major in Economics
jl10558@nyu.edu

Introduction

Background

- Population
  - Labor Force
  - Economic Growth
- Family Planning Policy
  - Moderate growth of population growth
  - Aging population and shrinking workforce
- Analyze the impact of the size of Labor Force on economic development

General Assumption

Assumption I

- Adam Smith's statement
  - population generates greater labor force
  - larger aggregate production
  - demand of labor → favoring population growth
- Population would produce a considerable human capital
  - positive impact on production growth based on Adam Smith's statements

Assumption II

- Structure of the Economy
- Labor-intensive industries
  - Benefit from a large labor force
  - More sensitive to the change in labor supply
- Capital-intensive industries
  - Less labor-intensive
  - Higher leverage on the capital

Data and Methods

WE Collected the data from China's National Bureau of Statistics (NBS)

- Finance*
- Birth Rate
- Urbanization
- Labor
- College** (number of college students after high school)

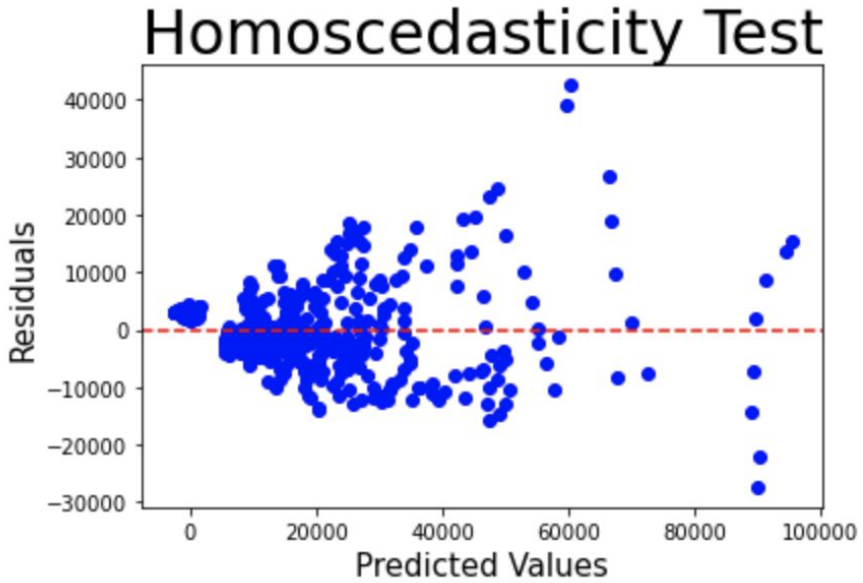
City	Year	Urbanization	GDP	Labor	Finance	Birth_rate	College
北京市	2020	0.875286	36102.6	739.9	0.195473		15.47
北京市	2019	0.873516	35445.1	791.3	0.184629	8.12	15.22
北京市	2018	0.870894	33106	819.3	0.179765	8.24	15.2
北京市	2017	0.869189	29883	812.9	0.177345	9.06	15.05
北京市	2016	0.867426	27041.2	791.5	0.176941	9.32	15.12
北京市	2015	0.867002	24779.1	777.3	0.176177	7.96	15.35
北京市	2014	0.865039	22926	755.9	0.162994	9.75	15.69
北京市	2013	0.864	21134.6	742.3	0.153658	8.93	15.98
北京市	2012	0.862849	19024.7	717.4	0.146294	9.05	15.86
北京市	2011	0.862154	17188.8	685.9	0.140882	8.29	15.69
北京市	2010	0.859327	14964	646.6	0.1362	7.48	15.51
北京市	2009	0.85	12900.9	619.3	0.135215	8.06	15.98
北京市	2008	0.849238	11813.1	570.3	0.139548	8.17	15.61

Panel Data Regression
Fixed Effect Model

Panel Data Regression

Test for Homoscedasticity

Durbin-Watson Test to detect autocorrelations




```
from statsmodels.stats.stattools import durbin_watson

durbin_watson_test_results = durbin_watson(pooled_OLS_dataset["residual"])
print(durbin_watson_test_results)
```

```
0.5136184573873889
```

Strong positive autocorreltion indicated, thus Fixed Effect or Random Effect models are suitable for this dataset analysis. IV = Labor, DV = GDP.

Fixed Effect Model

$$Y_{it} = \beta_1 X_{1,it} + \ldots + \beta_k X_{k,it} + \underbrace{\gamma_1 Z_{1i} + \ldots + \gamma_m Z_{mi}}_{=\alpha_i} + u_{it}$$

where the Z1, . . . ,Zm variables do not change for individual i over time but may be important in explaining Yit

1. Straight-up pooled regression
2. With region fixed effects
3. With region fixed effects but estimate in first differences
4. With region and time fixed effects
5. With time fixed effects

Interpretation

Sample rows of data

Description: df [10 × 7]

	Urbanization <dbl>	GDP <dbl>	Labor <dbl>	Finance <dbl>	Birth_rate <dbl>	College <dbl>	College_ratio <S3: AsIs>
黑龙江省-2016	0.6110309	11895.0	424.9	0.07268600	6.12	19.78	0.046552....
甘肃省-2014	0.4227578	6518.4	264.7	0.06891262	12.21	12.93	0.048847....
贵州省-2010	0.3380282	4519.0	224.3	0.05713654	13.96	9.93	0.044271....
湖南省-2002	*NA*	4151.5	*NA*	*NA*	11.56	15.69	*NA*
内蒙古自治区-2006	0.4865424	4161.8	*NA*	0.02724782	9.87	7.98	*NA*
北京市-2008	0.8492380	11813.1	570.3	0.13954800	8.17	15.61	0.027371....
青海省-2007	0.4003623	720.1	*NA*	0.03763366	14.93	1.11	*NA*
新疆维吾尔自治区-2008	0.3965275	4142.5	248.2	0.04482800	14.31	6.53	0.026309....
内蒙古自治区-2016	0.6338259	13789.3	293.2	0.05294685	9.03	12.19	0.041575....
广西壮族自治区-2019	0.5297069	21237.1	404.1	0.06919495	13.31	35.98	0.089037....

1–10 of 10 rows

Straight-up pooled regression

Straight-up pooled regression

```
m1 <- lm(log(GDP) ~ Urbanization + log(Labor) + Finance + Birth_rate + College_ratio,
         data = data)
coeftest(m1, df = Inf, vcov = vcovHC(m1, type = "HC1"))
```

```
##
## z test of coefficients:
##
##                Estimate Std. Error z value  Pr(>|z|)
## (Intercept)   1.0131777  0.1586472  6.3864 1.699e-10 ***
## Urbanization  1.2064188  0.1709118  7.0587 1.680e-12 ***
## log(Labor)    1.0945919  0.0152898 71.5899 < 2.2e-16 ***
## Finance       2.8583162  0.6742181  4.2395 2.241e-05 ***
## Birth_rate    0.0583962  0.0048592 12.0176 < 2.2e-16 ***
## College_ratio 9.4148418  0.9473645  9.9379 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With region fixed effects

Region fixed effects regression

```
m2 <- plm(log(GDP) ~ Urbanization + log(Labor) + Finance + Birth_rate + College_ratio,
          data = data, model = "within")
coeftest(m2, df = Inf, vcov = vcovHC(m2, type = "HC1"))
```

```
##
## z test of coefficients:
##
##                Estimate Std. Error z value  Pr(>|z|)
## Urbanization  4.3924284  0.5463130  8.0401 8.974e-16 ***
## log(Labor)    0.8657813  0.1366146  6.3374 2.337e-10 ***
## Finance       3.0840124  1.5428487  1.9989 0.0456183 *
## Birth_rate    0.0050896  0.0104284  0.4880 0.6255156
## College_ratio 6.1050108  1.8010435  3.3897 0.0006997 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With region fixed effects but estimate in first differences

Region fixed effects but estimate in first differences regression

```
m3 <- plm(log(GDP) ~ Urbanization + log(Labor) + Finance + Birth_rate + College_ratio,
          data = data, model = "fd")
coeftest(m3, df = Inf, vcov = vcovHC(m3, type = "HC1"))
```

```
##
## z test of coefficients:
##
##                  Estimate  Std. Error z value  Pr(>|z|)
## (Intercept)    0.08797485  0.00623837 14.1022 < 2.2e-16 ***
## Urbanization   1.40137073  0.31491194  4.4500 8.585e-06 ***
## log(Labor)     0.20719298  0.07261328  2.8534  0.004326 **
## Finance       -2.34247232  0.81831642 -2.8626  0.004202 **
## Birth_rate    -0.00083921  0.00238197 -0.3523  0.724598
## College_ratio -0.62878896  0.61235363 -1.0268  0.304496
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With region and time fixed effects

Region and time fixed effects regression

```
m4 <- plm(log(GDP) ~ Urbanization + log(Labor) + Finance + Birth_rate + College_ratio,
          data = data, model = "within", effect = "twoways")
coeftest(m4, df = Inf, vcov = vcovHC(m4, type = "HC1"))
```

```
##
## z test of coefficients:
##
##                 Estimate Std. Error z value  Pr(>|z|)
## Urbanization   1.2277355  0.4484721  2.7376 0.0061890 **
## log(Labor)     0.5289553  0.1450987  3.6455 0.0002669 ***
## Finance       -1.5154153  0.8119288 -1.8664 0.0619800 .
## Birth_rate     0.0090604  0.0100527  0.9013 0.3674355
## College_ratio  3.2401107  1.6932715  1.9135 0.0556814 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With time fixed effects

Only time fixed effects regression

```
m5 <- plm(log(GDP) ~ Urbanization + log(Labor) + Finance + Birth_rate + College_ratio,
          data = data, model = "within", effect = "time")
coeftest(m5, df = Inf, vcov = vcovHC(m5, type = "HC1"))
```

```
##
## z test of coefficients:
##
##                Estimate Std. Error z value  Pr(>|z|)
## Urbanization  0.832742   0.370352  2.2485 0.0245433 *
## log(Labor)    1.081261   0.031863 33.9347 < 2.2e-16 ***
## Finance       0.335597   2.151472  0.1560 0.8760451
## Birth_rate    0.036157   0.010811  3.3443 0.0008249 ***
## College_ratio 5.384494   1.940591  2.7747 0.0055258 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Limitations

Limitations

- Data
  - Limited sample size
  - Difficulty obtaining certain types of data
  - Reliability of data
- Methodology
  - Insufficient consideration of the interaction effects of variables
  - Potential multicollinearity among variables
  - Inability to the effectiveness of change in birth policies and the lagged effect related to the change in birth rate

Acknowledgement

Acknowledgement

DURF GRANTS

Dean's Undergraduate Research Fund

Reference

- Brugger, B. (2021, February 14). A guide to panel data regression: Theoretics and implementation with python. Medium. Retrieved April 28, 2022, from https://towardsdatascience.com/a-guide-to-panel-data-regression-theoretics-and-implementation-with-python-4c84c5055cf8
- National Bureau of Statistics. (n.d.). Regional Data. Retrieved April 28, 2022, from https://data.stats.gov.cn/easyquery.htm?cn=C01

Q&A

We welcome any further questions and concerns about our research

Please contact us via yq801@nyu.edu