Does Housing Price Impede the Rising Birth Rate?

Yao Weitong

According to the Malthusian theory of population, it' s the idea that population growth is potentially exponential while the growth of the food supply is linear. Malthusianism has concluded two "checks" to impede the population growth: "preventive checks" (moral restraints) and "positive checks" (disease, starvation, war.etc.). Obviously, technology development hasn' t been taken into account by Malthusian; in another words, these "checks" affect little in modern society, and the population of the world will rise dramatically (and it does!)[1]. However, some people study the relationship between high housing price and declining birth rate in Japan and they consider the rising housing price, as cost of raising babies, is a kind of "check" of Malthusiam. Is it true and general? Based on the dramatically increasing housing price in China since 2008, I would like to check whether housing prices do reduce fertility in the long run.

1 Regression Analysis

Fertility rates are linked to many factors, such as the cost of raising children – rising house prices, education – especially the ratio of female educated, women' s labor participation rate, the idea of "rearing kids for old age" – reflected in pension expenditures and the entire society' s elderly dependency rate. This passage collects data from the National Bureau of Statistics in mainland China from 2000 to 2017, and supplements it with data from the China Statistical Yearbook to make a simple multiple linear regression and to make a simple diagnosis of the model.

1.1 Variable Descriptions and Multiple Linear Regression Analysis

Variable	**Description**
birthrate	yearly fertility rate
femalegrow	growth rate of female labor participation rate
houseprice_rate	growth rate of residential house prices (rmb per square)
pensionout	proportion of social pension spending
devocerate	crude devorce rate

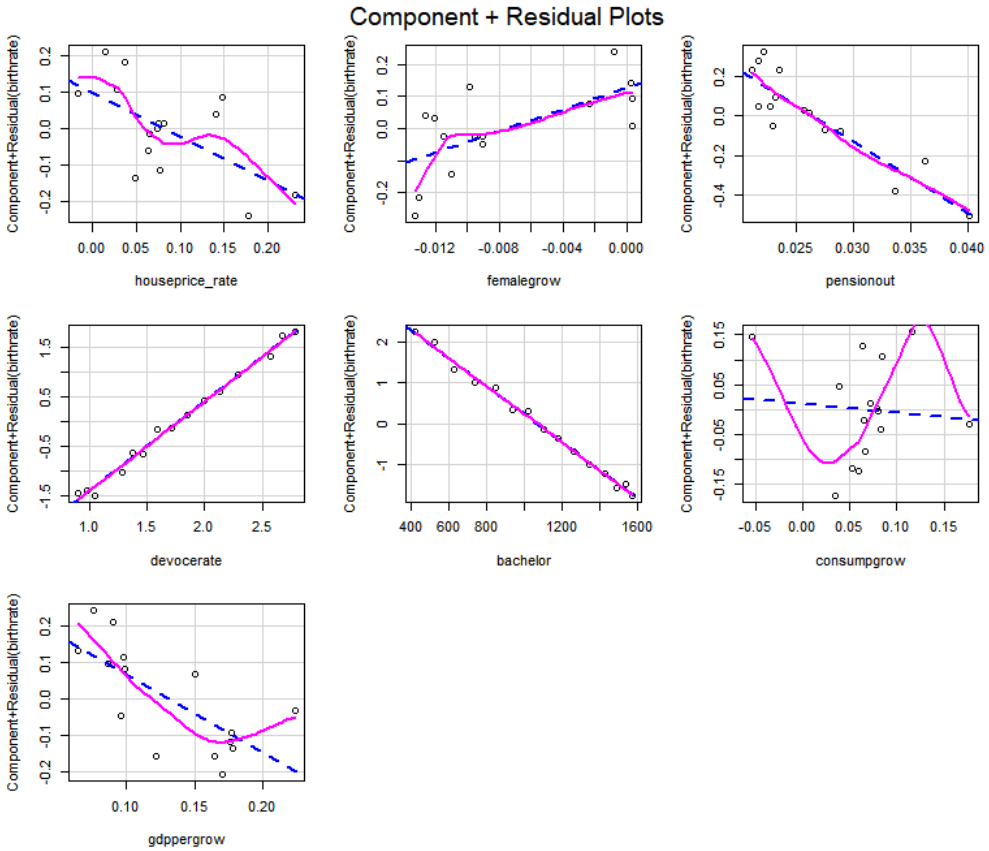
bachelor	number of undergraduates
consumpgrow	growth rate of consumer price index
gdppergrow	growth rate of gdp per capita

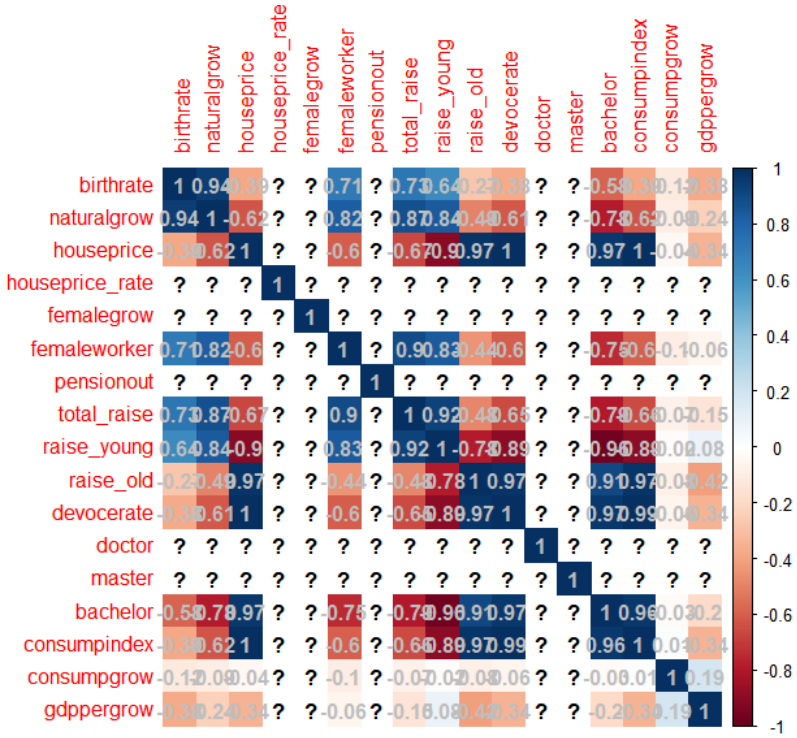
$$Birthrate_t = \beta_0 + \beta_1 housepricerate_{t-1} + \beta_2 femalegrow_{t-1} + \beta_3 pensionout_{t-1}$$
$$+\beta_4 devocerate_{t-1} + \beta_5 bachelor_{t-1} + \beta_6 consumpgrow_{t-1} + \beta_7 gdppergrow_{t-1} + e_t$$

The regression result and assessment of multiple linear model are as follows:

```
Call:
lm(formula = birthrate ~ houseprice_rate + femalegrow + pensionout +
    devocerate + bachelor + consumpgrow + gdppergrow, data = reg)

Residuals:
      Min        1Q    Median        3Q       Max
-0.178105 -0.061134 -0.000508  0.075046  0.165295

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     1.425e+01  5.811e-01  24.522 4.78e-08 ***
houseprice_rate -1.213e+00  7.428e-01  -1.633  0.14643
femalegrow       1.657e+01  1.739e+01   0.953  0.37240
pensionout      -3.660e+01  3.034e+01  -1.206  0.26693
devocerate       1.796e+00  6.529e-01   2.751  0.02845 *
bachelor        -3.436e-03  7.809e-04  -4.400  0.00316 **
consumpgrow     -1.766e-01  9.505e-01  -0.186  0.85787
gdppergrow      -2.129e+00  1.474e+00  -1.444  0.19186
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1435 on 7 degrees of freedom
  (3 observations deleted due to missingness)
Multiple R-squared:  0.9335,    Adjusted R-squared:  0.8669
F-statistic: 14.03 on 7 and 7 DF,  p-value: 0.001239


    ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
    USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
    Level of Significance =  0.05

    Call:
     gvlma(x = fm_birthrate)

                         Value p-value                Decision
    Global Stat        3.342654  0.5022 Assumptions acceptable.
    Skewness           0.001041  0.9743 Assumptions acceptable.
    Kurtosis           0.501445  0.4789 Assumptions acceptable.
    Link Function      1.811430  0.1783 Assumptions acceptable.
    Heteroscedasticity 1.028738  0.3105 Assumptions acceptable.
```

Next, check the multicollinearity of this model:


Component + Residual Plots
Component+Residual(birthrate)
houseprice_rate
femalegrow
pensionout
devocerate
bachelor
consumpgrow
gdppergrow


birthrate
naturalgrow
houseprice
houseprice_rate
femalegrow
femaleworker
pensionout
total_raise
raise_young
raise_old
devocerate
doctor
master
bachelor
consumpindex
consumpgrow
gdppergrow

Obviously, there is a problem of multicollinearity in this linear regression. To improve this modle, the passage uses ridge regression instead of simple multiple linear regression.

1.2 Ridge Regression[2,3]

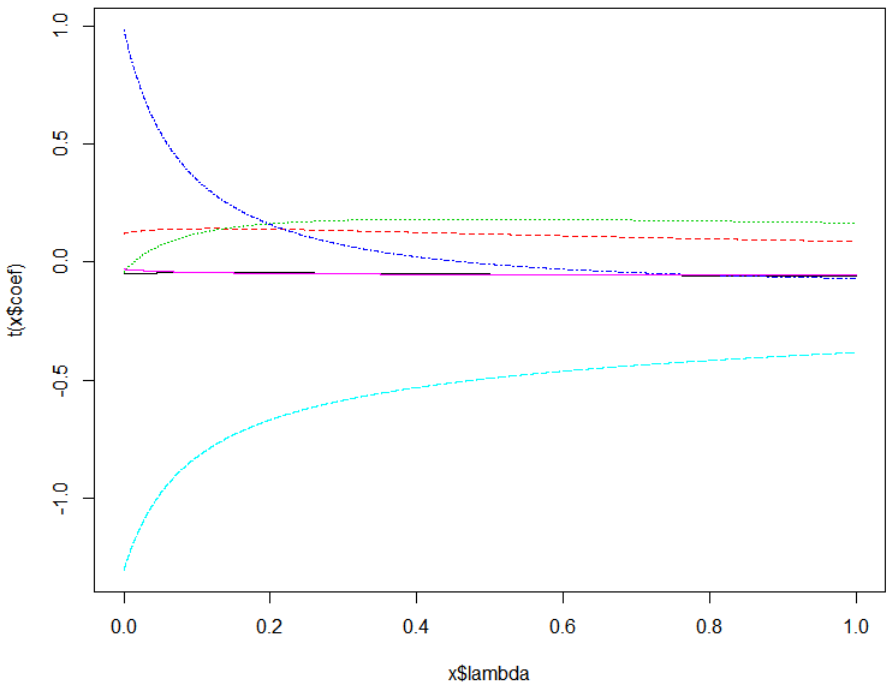
The logit that ridge regression can solve multicollinearity problem is to reduce the variance of parameter estimators, at the cost of sacrifying unbiasness. The general principle of ridge regression is to add a constraint to the parameter after the cost function, based on the MLE. This constraint is called a **regularizer**. The cost function of ridge regression is

$$J(\theta) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - BX_i)^2 + \lambda||B||^2 = MSE(\theta) + \lambda\sum_{i=1}^{m}\theta_i^2$$

where B is a coefficient vector excluding intercept term; θ is a n+1 vector including intercept term θ_0; n is the sample size and m is the number of independent variables. By making the gradient of $J(\theta)$ equal to 0, the optimal solution to minimize $J(\theta)$ is

$$\theta = (X^TX + \lambda I)^{-1}(X^TY)$$

Here are the ridge regression results; the significance of regression coefficient is significantly higher than that of multiple linear regression.



```
Call:
linearRidge(formula = birthrate ~ houseprice_rate + femalegrow +
    pensionout + devocerate + bachelor + consumpgrow, data = reg)


Coefficients:
                Estimate Scaled estimate Std. Error (scaled) t value (scaled)
(Intercept)     12.940912              NA                  NA               NA
houseprice_rate -1.173839       -0.288902            0.195188            1.480
femalegrow       1.065748        0.021537            0.188230            0.114
pensionout      11.121559        0.242107            0.191606            1.264
devocerate      -0.135003       -0.313159            0.101387            3.089
bachelor        -0.000501       -0.708423            0.147880            4.791
consumpgrow     -1.219595       -0.219044            0.195992            1.118
                Pr(>|t|)
(Intercept)           NA
houseprice_rate  0.13884
femalegrow       0.90891
pensionout       0.20639
devocerate       0.00201 **
bachelor        1.66e-06 ***
consumpgrow      0.26373
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Ridge parameter: 0.3375779, chosen automatically, computed using 2 PCs

Degrees of freedom: model 3.029 , variance 2.122 , residual 3.936
```

1.3 Explanation of Results

The results show that in the long run, the feritility rate is only negatively correlated with the divorce rate and education. The influence of devorce rate is greater in an economical sense, and the statistical results are more significant, too. It has nothing to do with the rate of house price appreciation. Further, assuming that there is a strong correlation between the increase in house prices and the divorce rate, I further run two regression and find that the results are either insignificant or significant but weakly correlated. Although the cost of raising children is rising when house prices rise, the regression shows that women's labor participation rate and GDP growth explain the decline in child dependency ratio more. Therefore, the decline in fertility in the long-term cannot be attributed to the rise in housing prices; or said, the house as a fixed asset is the main collateral for developers and corporate borrowing. In order to maintain the normal economy operation, the government would like maintain stability rather than suppress it. Directly speaking, it is an inevitable trend that house prices are at high prices, but people wouldn't choose to be DINK because of this. It is obvious that the house only temporarily sway the decision of fertility, but it cannot affect fertility in the long run.

2 Conclusion and Other Evidences

Other evidences can show why they are not exactly negatively correlated. First is about the two-sided impacts. According to *Lisa J.Dettling and Melissa S.Kearney*[4],

current house prices can lead to a negative effect on the fertility rates among renters but positive effect on births among homeowners during the current period and these two effects can offset each other. Second, children-rearing cost is a bridge linking housing prices and fertility rate, however, the young dependency rate is falling in recent year, from the data collected from National Bureau of Statistics. In other words, if having a children but not raising them, then the impact of housing prices on fertility rate will not be established. Third one is the elasticity issue. By *Creina Day and Ross Guest*[5], the key is the level of substitution effect of rising female wages on fertility depends on how important housing is as a cost of children; price elasticity of housing supply is low in Asia, so the increasing wage leads to decreasing birth rate. Things are different in western countries, where housing supply is relatively price elastic.

References

[1] *Malthusiam*, wikipedia.

[2] *M. El-Dereny and N.I. Rashwan.* 2011. Solving Multicollinearity Problem Using Ridge Regression Models. Int J. Contemp. Math. Sciences, Vol.6,2001, no.12,585-600.

[3] 【机器学习】正则化的线性回归 —— 岭回归与Lasso回归

[4] *Lisa J.Dettling and Melissa S.Kearney*.2016. Fertility and female wages: A new link via house prices. Economic Modelling.

[5] *Creina Day and Ross Guest.* 2014. House prices and birth rates: The impact of the real estate market on the decision to have a baby. Journal of Public Economics