SPATIAL INEQUALITY AND FERTILITY IN CHINA

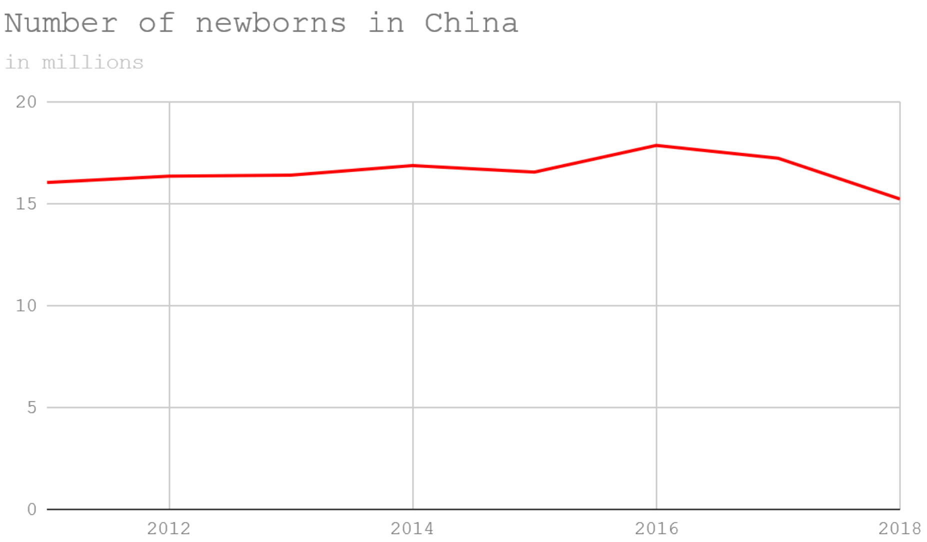
Edwin Mu and Ada Huang[1]
B. Sc Economics
2nd year
University College London

Explore Econ Undergraduate Research Conference
March 2019

[1] Special thanks to Dr Dunli Li and Dr Áureo de Paula for their advice.

Introduction

In 2016, the Chinese government predicted that the abolition of one-child policy would "add an estimated 3 million babies annually in the next five years", so the number of newborns would be 21 million each year (People's Daily Online, 2016). The prediction failed miserably.

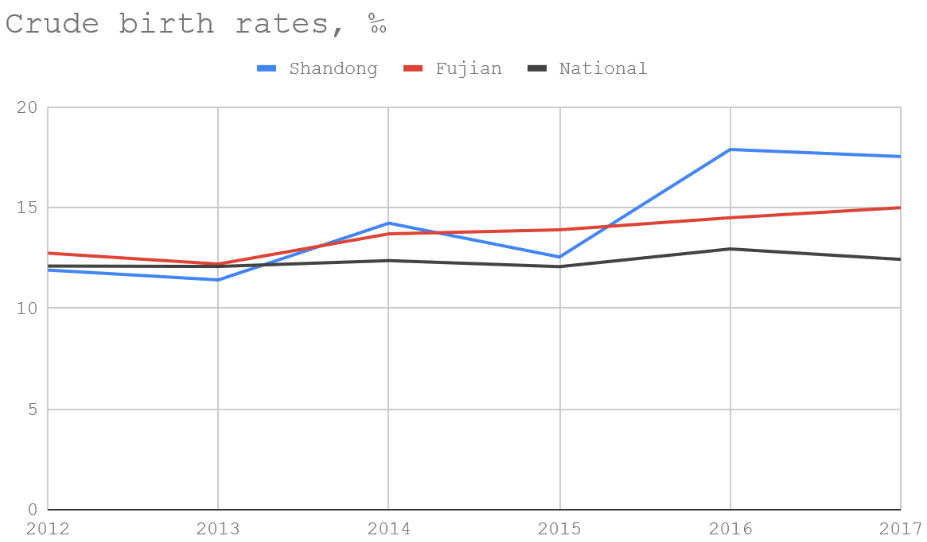


(Source: Global Times, 2019; Xinhua News Agency, 2017)

Figure 1

As shown above, the surge in the number of newborns only occurred in 2016; even in that year, the surge was far smaller than predicted. Why?

We find an explanation offered by Jushishuo, a user of a Chinese online forum, particularly interesting. Jushishuo (2018) argues that the most important factor behind the lacklustre surge is geographically uneven development of cities, which we define as spatial inequality in this article. His evidence is that provinces with more evenly developed cities, such as Shandong and Fujian Province, had seen greater growth in crude birth rates than the country as a whole by 2017.



(Source: National Bureau of Statistics of China)

Figure 2

Yet his evidence is insufficient to validate his argument. Firstly, as commented by another user (Dengbojun, 2018), Jushishuo might have cherry-picked provinces whose data corroborate his argument. Secondly, Jushishuo does not have any controls, such as the number of women of childbearing age and women's education.

The aim of this research is to confirm whether there is a negative correlation between spatial inequality within a province and fertility of that province in China, controlling for usual determinants. Using data from 2000, 2005 and 2010, our results suggest that such a correlation does not exist.

The article is organized as follows. We start with a literature review as to why there might be a correlation between spatial inequality and fertility. Then we introduce our methodology and data. We conclude with results and discussion.

Theoretical background

If there was a negative correlation between spatial inequality within a province and fertility of that province in China, as claimed by Jushishuo (2018), why would that be the case?

One possibility is that discrepancy in prosperity results in migration, which then reduces fertility: “Migratory movements exist only because people perceive some places to be more attractive than others.” (Castro, 2007). Indeed, past literature suggests that rural-to-urban migration results in a decrease in fertility.

Using China’s 1990 Census, You & Poston (2004) argue that rural-to-urban migrants have lower total fertility rates (TFR) compared to non-migrants in their hometowns, after controlling for ethnicity, occupation and education. This phenomenon is possibly due to disruption caused by migration, e.g. separation from family and higher level of stress (known as the disruption hypothesis). Similar results are found by Chen & Wu (2006) and Hao (2015).

Therefore, a possible cause of a negative correlation between spatial inequality and fertility is an increase in within-province migration towards richer areas.

Does inequality itself have an impact on fertility, independent of migration? Repetto (1978) argues that if income affects fertility in a non-linear way, then income distribution also has an impact on fertility, as poor and rich households will react differently when they have more income. However, there is no consensus as to whether income inequality has a positive or negative association with fertility (Chen, 2010; Wright, 1988). In a recent study about mainland China, Chen (2010) argues that there is a negative correlation between them.

Hence another possible reason: if spatial inequality is negatively correlated with fertility, it might be due to higher income inequality.

Empirical Strategy

A primary question is how to measure spatial inequality. We decide to use the standard deviation of prefectural-level real GDP per capita within a province as the measure of spatial inequality of a province, where a greater standard deviation suggests a higher level of spatial inequality.

The prime advantage of GDP per capita is its availability; in comparison, data on average disposable income on the prefectural level is incomplete. Yet our measure is not perfect. Firstly, it does not capture the income households really receive. Secondly, and perhaps most importantly, our measure cannot reflect spatial inequality on the county level (i.e. within a prefecture), and it is possible that spatial inequality across prefectures differs from that across counties. Sadly, data collection for counties requires dramatically more effort that is not viable for this research.

Nevertheless, a low prefectural spatial inequality suggests that income and production are more evenly distributed among major cities of a province, and we believe that it is able to capture the impact of spatial inequality sufficiently. As stated earlier, a major channel through which negative correlation between spatial inequality and fertility might occur is migration. The impact of inter-county migration on fertility should be considerably smaller compared to inter-prefecture migration, given the smaller size of a county.

Another problem regarding measurement is that our measure for within-province migration is the proportion of population whose *hukou* is in a prefecture that is not his or her current place of residence, but within-province migration may induce a change in *hukou*. This is the best we can achieve without using individual-level survey.

Our dependent variable is TFR. In general, it is calculated as

$$\mathrm{TFR} = \frac{1}{\mathrm{N}}\sum_{i}^{N}\frac{\text{number of live births by women in age group i}}{\text{number of women in age group i}}$$

and can be interpreted as the number of children a woman will give birth to over her life given the current age-specific fertility rates (UN, 2003). Notice that the number of women in specific age groups are incorporated in TFR, so it is not additionally controlled in our model.

In terms of model selection, this research is based on a fixed effect model. From the linear model

$$TFR_{it} = \beta_0 + \beta_1 GDPpc\_sd_{it} + \sigma^T X + \delta^T Year + a_i + u_{it}$$

we obtain within estimators

$$\ddot{TFR}_{it} = \beta_1 \ddot{GDPpc\_sd}_{it} + \sigma^T \ddot{X} + \delta^T \ddot{Year} + \ddot{u}_{it}$$

where subscript i is province, subscript t is year, $GDPpc\_sd$ is our measure of spatial inequality, X is a set of control covariates that may also affect TFR identified from past literature (see Pan & Xu, 2012; Yong, 2010), $Year$ is a set of year dummies to control for time fixed effect, a is the province fixed effect and u is the time-variant error.

The pooled OLS model is not adopted since province-specific culture, which is likely to be time-invariant, may affect fertility; the F test for province fixed effect produces a p-value of zero. The random effect model is not chosen since province-specific culture is likely to be correlated with other characteristics of a province; the Hausman test produces a p-value less than 1%.[2]

[2] Results of the F and Hausman test can be provided on request.

Variable	Description
total fertility rate	total fertility rate of a province
s.d. of GDP p.c.	standard deviation of real GDP per capita of prefectures of a province
GDP p.c.	real GDP per capita of prefectures of a province
rural population	population in rural areas as a proportion of total population of a province
Han population	Han population as a proportion of total population of a province
women educated	the number of women who have received high-school education as a proportion of total female population (above age 6) of a province
women employed	the number of women in employment as a proportion of total female population (above age 16) of a province
women married	the number of women married as a proportion of total female population (above age 15) of a province
within province migration	population whose *hukou* is in the same province but not the prefecture of current residence as a proportion of total population of a province
Gini coefficient	Gini coefficient of a province

Table 1

Variable descriptions

Data

Since individual-level data lack key variables, this research is based on aggregate-level data reported by the Chinese government. For instance, region-specific data from China Family Panel Survey (CFPS) are difficult to obtain, and Chinese General Social Survey (CGSS) provides limited data on fertility.

Data covering 27 provincial-level districts in mainland China in 2000, 2005 and 2010 are collected. Municipalities are excluded since they are too small for measurement of spatial inequality.

Provincial-level data on TFR, the proportion of women in employment, the proportion of women with high school education, the proportion of Han population, the proportion of rural population and within-province migration are obtained from the Population Census in 2000 and 2010 and the 1% National Population Sample Survey in 2005. We assume that these data are representative.

Prefectural nominal GDP per capita are obtained from CSMAR Database. Due to CSMAR's limitation, data for some provinces are incomplete, in which case provincial statistical yearbooks are used for the whole province. Provincial nominal GDP per capita and CPI are obtained from China Statistical Yearbooks in 2001, 2006 and 2011. Real GDP per capita is then calculated using CPI.

Estimates of provincial-level Gini coefficients between 2000 and 2010 are obtained from Tian (2012). Note that estimates for Jilin, Shandong, Guangxi, Hainan and Tibet are unavailable.

Results

Assuming homoscedasticity and i.i.d. in time variant error, we obtain the following results:

	(1) total fert~e b/se	(2) total fert~e b/se	(3) total fert~e b/se	(4) total fert~e b/se
s.d. of GDP p.c.	5.56e-06*	4.00e-06	3.94e-06	3.64e-06
	(2.40e-06)	(2.91e-06)	(2.92e-06)	(3.03e-06)
year=2000	0	0	0	0
	(.)	(.)	(.)	(.)
year=2005	.0836652*	-.0263956	-.1108714	.1115171
	(.0360132)	(.0724175)	(.11581)	(.1528885)
year=2010	-.1669785**	-.4138276*	-.5112472*	-.1207263
	(.0480049)	(.1618772)	(.1926657)	(.2541533)
GDP p.c.		5.09e-06	5.93e-06	2.41e-06
		(5.56e-06)	(5.64e-06)	(6.63e-06)
rural population		-1.650877	-1.585283	-.7591354
		(.8788444)	(.8828532)	(1.262757)
Han population		-.2869266	-.1803557	1.535359
		(2.209049)	(2.215047)	(2.978308)
women educated		-.6028501	-.5676745	-.821993
		(2.125045)	(2.128328)	(2.389975)
women employed		.9992569	1.155347	1.143608
		(1.17902)	(1.192388)	(1.68675)
women married		-1.075048	-.4579826	.337159
		(1.612228)	(1.743999)	(2.310394)
within province mi~n			.2965232	-.0681147
			(.3169474)	(.3643455)
Gini coefficient				-.7373665
				(1.452425)
constant	1.293848***	2.937752	2.176243	.0231828
	(.0278979)	(1.878878)	(2.050007)	(3.367766)
R-sqr	0.408	0.480	0.490	0.408

* p<0.05, ** p<0.01, *** p<0.001

(Note: standard errors are reported in parentheses)

Table 2

Impact on TFR, assuming i.i.d. of u and homoskedasticity

Model 1 shows that, without any controls, there is actually a positive correlation between spatial inequality and TFR of a province, with a p-value of 2.4%. The estimate suggests that an increase of 10,000 yuan in the standard deviation is associated with an increase of 0.0556 in TFR. Is this economically significant? The answer is probably normative. For instance, between 2010 and 2015, Zhejiang Province's standard deviation in GDP per capita rose by nearly 20,000 yuan. This is estimated to be associated with a increase of about 111 children born by every 1,000 women, or an increase in TFR by 0.111.

However, as we add more controls, the coefficient on spatial inequality becomes statiscally insignificant. Model 2 produces the lowest p-value, 0.177, which is statiscally insignificant at all conventional levels.

What if we allow for heteroscedasticity and autocorrelation? The results are shown below:

	(1) total fert~e b/se	(2) total fert~e b/se	(3) total fert~e b/se	(4) total fert~e b/se
s.d. of GDP p.c.	5.56e-06**	4.00e-06*	3.94e-06*	3.64e-06
	(1.80e-06)	(1.62e-06)	(1.63e-06)	(1.83e-06)
year=2000	0	0	0	0
	(.)	(.)	(.)	(.)
year=2005	.0836652*	-.0263956	-.1108714	.1115171
	(.0345841)	(.0860428)	(.1241252)	(.1411594)
year=2010	-.1669785*	-.4138276	-.5112472	-.1207263
	(.062414)	(.2421925)	(.2781977)	(.2685425)
GDP p.c.		5.09e-06	5.93e-06	2.41e-06
		(4.16e-06)	(4.29e-06)	(4.11e-06)
rural population		-1.650877	-1.585283	-.7591354
		(1.155868)	(1.13418)	(1.734247)
Han population		-.2869266	-.1803557	1.535359
		(1.670437)	(1.626945)	(2.658651)
women educated		-.6028501	-.5676745	-.821993
		(2.015455)	(2.182214)	(2.043146)
women employed		.9992569	1.155347	1.143608
		(.953293)	(.9654865)	(1.701426)
women married		-1.075048	-.4579826	.337159
		(1.984407)	(1.737392)	(2.047161)
within province mi~n			.2965232	-.0681147
			(.2499497)	(.2225952)
Gini coefficient				-.7373665
				(1.517671)
constant	1.293848***	2.937752	2.176243	.0231828
	(.0195049)	(1.821309)	(1.50191)	(3.582859)
R-sqr	0.408	0.480	0.490	0.408

* p<0.05, ** p<0.01, *** p<0.001

(Note: robust standard errors clustered by provinces are reported in parentheses)

Table 3

Impact on TFR with cluster robust standard errors (CRSEs)

Interestingly, the CRSEs are much smaller: Model 2 and 3 produce a p-value of 0.021 and 0.023 respectively, and Model 4 produces a p-value of 0.059. Does it mean the postive correlation with spatial inequality and TFR is statiscally significant? This is doubtful. Notice that there are only 27 clusters, which may not satisfy the asymptotic justification for CRSE. Cameron, Gelbach, & Miller (2008) argue that, in this case, CRSEs are likely to be downwards biased.

Conclusion

In conclusion, we find that there is no negative correlation between the standard deviation of GDP per capita of prefectural-level districts within a province and the fertility of that province in China. Therefore, the results suggest that Jushishuo's (2018) claim is likely to be wrong.

But there is still some uncertainty left. Firstly, as noted earlier, our measure for inequality does not directly capture spatial inequality across counties. Secondly, when using CRSEs it seems that there is a positive correlation between spatial inequality and fertility. If this was the case (instead of a breakdown of CRSE's asymptotic justification), would it be due to an omitted variable bias, where spatial inequality across prefectures is negatively correlated with that across counties, which is negatively correlated with fertility? We believe that further studies based on individual-level data may address this question, and CFPS may be a particularly useful source of data.

Finally, the relationship between spatial inequality across provinces and fertility of China as a whole is not considered in this study, which may be of greater interest.

References

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, *90*(3), 414–427.

Castro, M. C. de. (2007). Spatial Demography: An Opportunity to Improve Policy Making at Diverse Decision Levels. *Population Research and Policy Review*, *26*(5), 477–509. https://doi.org/10.1007/sllll3-007-9041-x

Chen, W. (2010). 中国的经济平等与生育率 [Equality and Fertility: Evidence from China]. *人口与经济 [Population & Economics]*, *1*(4), 18–23.

Chen, W., & Wu, L. (2006). 中国人口迁移与生育率关系研究 [Research on the Relationships between Migration and Fertility in China]. *人口研究 [Population Research]*, *30*(1), 13–20. Retrieved from http://rkyj.ruc.edu.cn/CN/abstract/abstract449.shtml

Dengbojun. (2018). 如何看待山东省2017年出生人口175万人，一孩出生数量还不到二孩一半？ - 邓铂鋆的回答 [What’s your view on the fact that the number of newborns in Shandong Province was 1.75 million in 2017 and the number of first child was fewer than half of that of second child? -Dengbojun’s answer]. Retrieved February 24, 2019, from https://www.zhihu.com/question/267934281/answer/332460442

Global Times. (2019). China births dip in 2018. Retrieved March 9, 2019, from http://www.globaltimes.cn/content/1134258.shtml

Hao, Z. (2015). 人口流动对生育水平的影响:基于选择性的分析 [Selectivity Bias in the

Effect of Migration on Fertility]. *人口研究[Population Research]*, *39*(1), 14–28.

Jushishuo. (2018). 如何看待山东省2017年出生人口175万人，一孩出生数量还不到二孩一半？ - 居士说的回答 [What’s your view on the fact that the number of newborns in Shandong Province was 1.75 million in 2017 and the number of first child was fewer than half of that of second child? - Jushishuo’s answer]. Retrieved February 24, 2019, from https://www.zhihu.com/question/267934281/answer/330917188

Pan, L., & Xu, J. (2012). Housing price and fertility rate. *China Economic Journal*, *5*(2–3), 97–111. https://doi.org/10.1080/17538963.2013.764675

People’s Daily Online. (2016). Two-child policy creates baby boom in China. Retrieved February 24, 2019, from http://en.people.cn/n3/2016/0803/c90882-9094691.html

Repetto, R. (1978). The Interaction of Fertility and the Size Distribution of Income. *The Journal of Development Studies*, *14*(4), 22–39. https://doi.org/10.1080/00220387808421688

Tian, W. (2012). 省域居民收入基尼系数测算及其变动趋势分析 [Measurement of Gini coefficients of provincial residents’ income and analysis of its trends]. *经济科学 [Economic Science]*, *34*(2), 48–59.

UN. (2003). *Handbook on the Collection of Fertility and Mortality Data*. Retrieved from https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/fertility-and-mortality/SeriesF_92-E.pdf

Wright, R. E. (1988). The Impact of Income Redistribution on Fertility in Canada. *Genus*, *44*(1), 139–156.

Xinhua News Agency. (2017). 2000-2016年我国出生人口数变化情况 [Changes in births in this country between 2000 and 2016]. Retrieved February 24, 2019, from http://www.gov.cn/xinwen/2017-01/22/content_5162356.htm

Yong, C. (2010). China's Below-Replacement Fertility: Government Policy or Socioeconomic Development? *Population and Development Review*, *36*(3), 419–440.

You, H. X., & Poston, D. L. (2004). Are floating migrants in China "child-bearing guerrillas"?: An analysis of floating migration and fertility. *Asian and Pacific Migration Journal*, *13*(4), 405–422. https://doi.org/10.1177/011719680401300401