Discussion Paper Series
RIEB
Kobe University

DP2022-13

A Spatial Panel Data Analysis of Fertility Rates: Unraveling Two Myths*

Yosuke SASAKI
Takashi KAMIHIGASHI

March 30, 2022

Research Institute for Economics and Business Administration
Kobe University
2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

A Spatial Panel Data Analysis of Fertility Rates: Unraveling Two Myths

Yosuke Sasaki* Takashi Kamihigashi†

2022/03/30

Abstract

This paper analyzes the relationship between total fertility rates and other socio-economic variables. The main results are against two myths: rural areas have higher total fertility rates than urban ones, and economic inequality harms total fertility rates. The key is a spatial panel data analysis with the lag between inception and birth. Here, spatial panel data analysis improves the efficiency of explanatory variables, and the lag produces a different result from former papers.

1 Introduction

In 2019, only 864 thousand babies were born in Japan, with a 126 million population, the number is the lowest in the statistics, and the percentage of the total population is only 0.69%. Several academics and journalists cite the chronically low birth rates as one of Japan's crucial challenges. Because East Asian countries have similar problems, and even the United States now records the lowest birth rates in its history, the mechanism between the total fertility rates (TFRs) and other socio-economic variables are supposed to be vital to find the solution and evaluate the policy on low total fertility rates.

Numerous people suggest that income inequality and educational costs have adverse effects on TFRs. One may question if TFRs relate to the Gini index and the admission rates to colleges and the age cohort, and one may also be interested in the time series structure of TFRs. Moreover, several think rural areas have higher TFRs than cities, and economic inequality affects TFRs negatively.

Several papers suggest that the factors affect TFRs, such as female employment, male and female wage, wealth transfer between the generations, parental care costs, the number of children in one household. Moreover, in Japan, the policies to promote parental care should be evaluated. However, researchers have not estimated the relationship between the fertility rates, the Gini index, and the admission rates after high school.

The spatial panel data analyses show that the increase of the Gini index has adverse effects on TFRs. Also, the relationship between the population density and TFRs is the inverted U-shape. Furthermore, real expenditure and TFRs have an inverted J-shape relationship. Then, the impulse response of the panel vector autoregression (Sigmund & Ferstl 2019[22]) indicates that even a one-shot policy on TFRs has favorable effects for about 20 years and adverse economic shocks have opposite effects. This paper's methods also validate the Easterlin hypothesis on fertility (Macunovich & Easterlin 2018[15]). This paper shows that the negative relationship between TFRs and the population density is inverted-U-shaped, using the simple unbalanced panel data with large cross-sectional dimensions and the balanced

*Center for Computational Social Science, Kobe University

†Center for Computational Social Science, Kobe University

panel data with smaller cross-sectional dimensions denser time-series ones. However, this analysis faults that one does not consider the lag between inception and birth. With this lag, there is no statistically significant relationship between population density and TFRs. The relationship between economic inequality and TFRs is statistically false with this lag. The bivariate nonparametric regression and time series of Moran I and Theil indices indicate the negative relationship between TFRs and colleges' admission rates. However, the panel non-Granger causality test and the spatial panel regression do not show a clear relationship between TFRs and the admission rates.

2 The Literature and Data Characteristics

2.1 The Literature

2.1.1 Economic and Analytical Studies on Total Fertility Rates

Several social scientists work on human fertility. However, here we focus on socio-economic approaches.

Becker (2009)[4] considers the interaction between the quantity and quality of children to explain why education per child tends to be lower in families having more children.

Willis (1973)[23] points out that the interaction model captures an essential empirical regularity in the cross-section relationship between fertility and measures of husband's income and wife's education that has become apparent in the emergence of a U-shaped relationship between fertility and income. The predictions of the theoretical model of fertility demand developed in his paper fit well with this paper's estimation result.

Prettner et al. (2013)[19] theoretically and empirically show that if individuals put more weight on education, they reduce fertility, increase educational investments, and hold consumption, savings, and health investments constant.

Lutz et al. (2006)[14] empirically find a consistent and significant negative relationship between human fertility and population density, using fixed-effects models on the time series of 145 countries and controlling for critical social and economic variables such as GDP per capita, infant mortality, female labor force participation, and female literacy. In the empirical study, Kulu (2013)[12] shows that the desired family size in small towns and rural areas is larger than that in urban areas.

Myrskylä et al. (2009)[17] find that TFRs in highly developed countries have inverted J-shape curves over the Human Development Index (HDI) except Japan, Canada, and South Korea. The HDI is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable, and having a decent living(United Nations Development Programme's Human Development Report[20]). However, their finding is not robust to the redefined HDP or the decomposition of the HDI into its subindices of education, the standard of living, and health (Harttgen & Vollmer 2014[9]).

Aldieri & Vinci (2012)[1] estimate the correlation between the female's education and the number of children in Italy and use the partner's education to take into account the family dimension. They use a zero-inflated Poisson regression and observe a negative correlation between the number of children born and the educational level.

2.1.2 Sociological and Demographic Studies on Total Fertility Rates

As Pampel & Peters (1995)[18] review the Easterlin hypothesis, Easterlin presented his basic argument that swings in relative cohort size and resulting levels of relative income among cohorts of child-bearing age produced the baby boom and then the baby bust. Initially, his discussion is concerned with describing demographic cycles (Easterlin 1968[7]). Moreover, the linkage between higher birth rates and adverse social-economic effects arises because large cohorts face crowding problems in three major social institutions such as the family, education, and labor markets (Macunovich & Easterlin 2018[15]).

The second demographic transition (SDT) explains the inverted J-shape curve in sociology and demography. Zaidi & Morgan (2017)[24] summarize the SDT as three arguments: the shift from king-child to king-couple, the Maslowian drift, the rise of individualism, and pushback against economic explanations.

2.1.3 Studies on Total Fertility Rates in Japan

Gini Index

Few researchers have the Gini index as the explanatory variable for total fertility rates. The panel data in this study shed light on the Gini index's effects on TFRs.

Educational level

Some papers regress the female educational levels to TFRs.

2.2 Data

Total Fertility Rates (TFR) and the population of each cohort of every prefecture ($p0to4$, $p5to9$, $p10to14$, $p15to19$, $p20to24$, $p25to29$, $p30to34$, $p35to39$, $p40to44$, $p45to49$, $p50to54$, $p55to59$, $p60to64$, $p65to69$, $p70to74$, $p75to79$ and $pover80$) are from the National Institute of Population and Social Security Research (Japan) and its former organization. Municipal data on TFR is available online from the Statistical Bureau of Japan.

Figure. 1 depicts the movements of total fertility rates of each prefecture in Japan.

Data on the Gini index ($gini$) and mean real expenditure $real\_exp\_mean$ are estimated originally from the microdata courtesy of the Statistic Bureau, Ministry of Internal Affairs and Communications.

3 Results of the Econometric Approaches

3.1 Nonparametric Analysis and the Panel Non-Granger Causality

At first, one can nonparametrically regress the bivariate relationship between the total fertility rates (TFRs) and other socio-economic variables. Figure 2 shows the relationship between TFR and $pop\_density$ and one between TFR and ad_rate appear inverse corrletions.

The relationship between TFR and $pop\_density$ and one between TFR and $real\_exp\_mean\_100$ have the panel Granger causality in both directions. These tests are based on Dumitrescu & Hurlin (2012)[6]. They note that the null of Homogeneous Non-Causality does not provide any guidance concerning the number or the identity of the particular panel units for which the null of non-causality is rejected. In other words, only non-Granger causality is meaningful. However, $gini$ and ad_rate have the panel Granger causality from TFR, and TFR has the panel non-Granger causality from $gini$ and ad_rate. This $gini$ and ad_rate do not determine TFR in the non-Granger causality standpoint. Then, the panel non-Granger Causality test assumes that the individuals of the panel have no cross-sectional dependence.

3.2 The Relationship Between the Total Fertility Rates in the Municipal Level and the Population Density

Because municipal data on TFR have relatively large cross-sections (at least two thousand) and seven terms, panel data analysis is available. The estimation is based on the fixed-effects model, and the Least

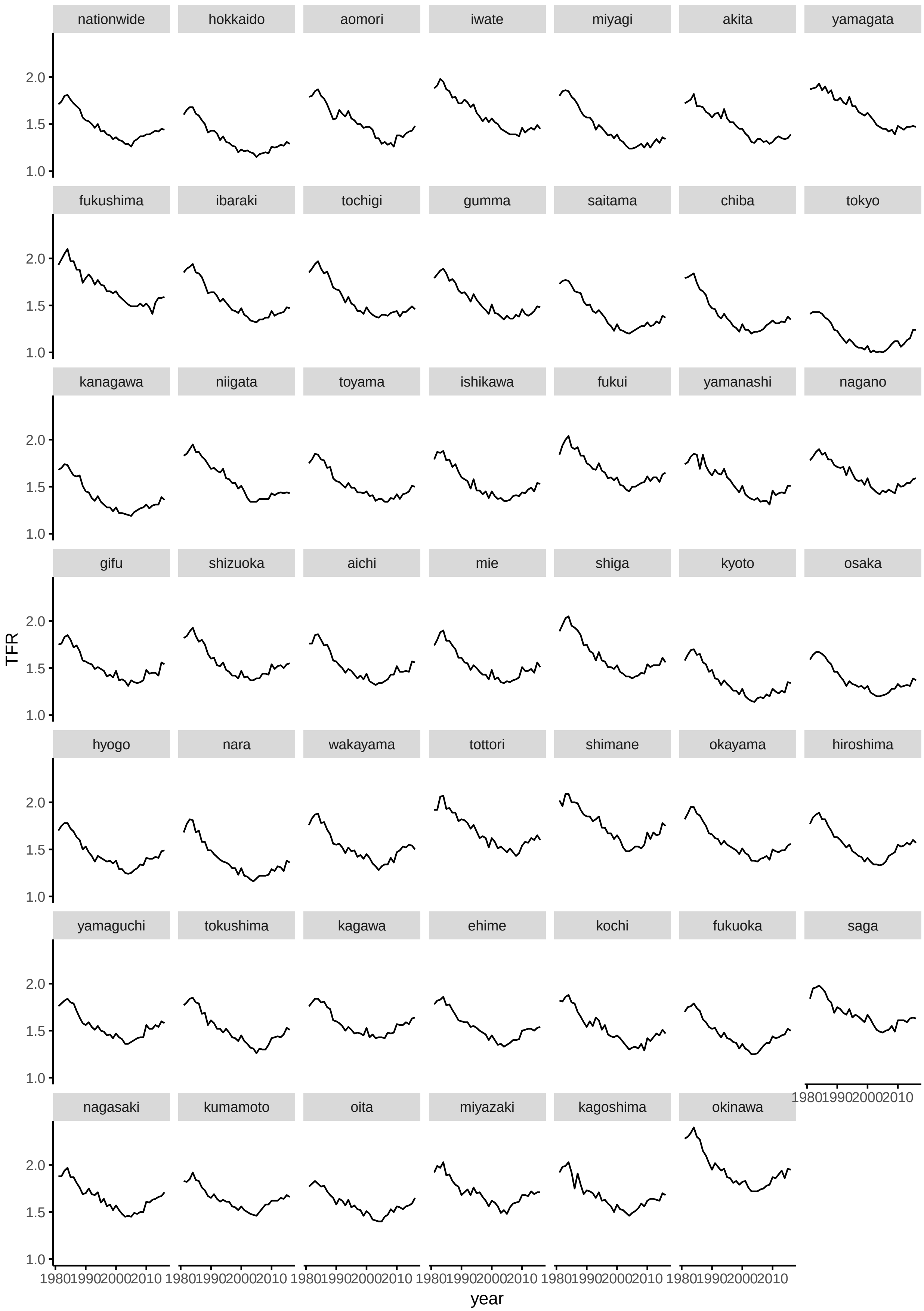

Figure 1: Total Fertility Rates in Prefectures

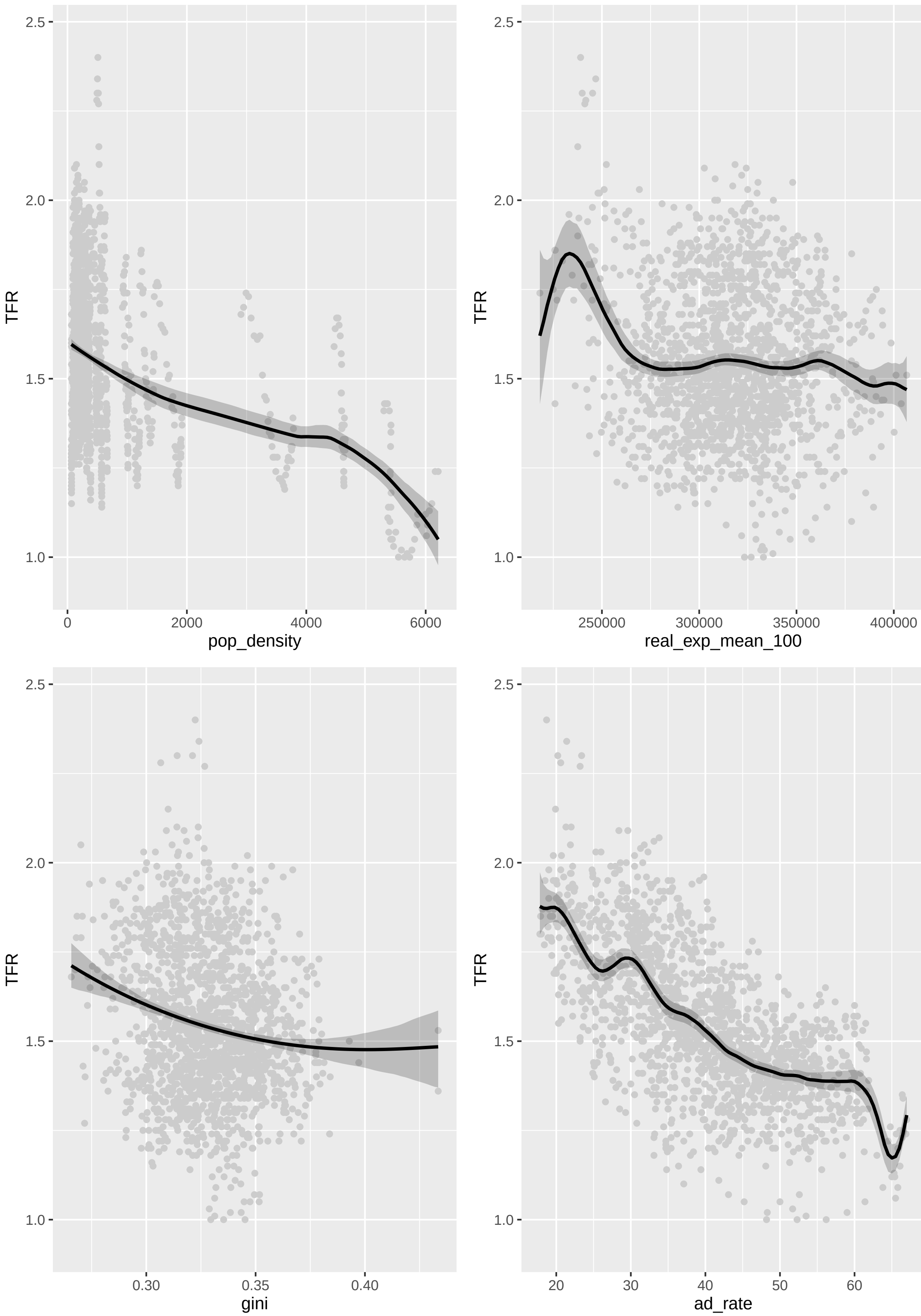

Figure 2: Nonparametric Regression Results

Table 1: Estimation Results: The Large Cross-Sectional Panel ($n = 3755$, $T = 1 - 7$, and $N = 19235$)

$\log(TFR)$	Fixed Effects	Random Effects	Pooling
Intercept		0.2332224 (0.0154558) ***	0.36260635 (0.01481855) ***
$\log(pop\_density)$	0.6627606 (0.0654288) ***	0.1294545 (0.0057179) ***	0.08314563 (0.00562807) ***
$(\log(pop\_density))^2$	-0.0504881 (0.0056033) ***	-0.0140688 (0.0005074) ***	-0.01043975 (0.00051242) ***
$Adj.R^2$	-0.11254	0.16	0.19968

$. < 0.10 \quad * < 0.05 \quad ** < 0.01 \quad *** < 0.001$

Double clustered standard deviations are in parentheses.

Square Dummy Variable as explanatory variables are few, and the omitted variables bias is severe. The Hausman test between the fixed effects and the random effects supports this reasoning.

$$exp(0.6627606/2(0.0504881)) = 708.7712 \tag{1}$$

This subsection's relationship between the TFR and $pop\_density$ is spurious because it is natural that we consider the lag between inception and birth.

3.3 The Spatial Autocorrelation and the Inequality Index

Moran's I in the spatial autocorrelation of cross-sectional spatial data has an analogy to Durbin-Watson in the serial autocorrelation of time-series data (Li et al., 2007[13]). Then, the higher Moran's I is, the higher the degree of spatial autocorrelation is. The spatial dependence (global spatial autocorrelation) measure of Moran's I is represented by equation 2 (the formulation follows Rey 2004[21]):

$$Moran\_I_t = \frac{n}{s} \frac{\sum_i \sum_j w_{ij} z_{it} z_{jt}}{\sum z_{it}^2} \tag{2}$$

$$\tag{3}$$

where n is the number of prefectures (i.e.$n = 47$), z_{it} and z_{jt} are the deviation of each variable from the mean of each prefecture in each year (the prefecture i, j and the year t), w_{ij} are the elements of weight matrix W $(n \times n)$ and it is equal to 1 if i and j are neighbors and 0 if they are not; s is the sum of all elements of W (spatial weights). One note that n, s and w_{ij} are invariant over time.

Figure 3 plots Moran's I of each variable. The critical value, which rejects at the significant level 0.05 null hypothesis that there is no spatial correlation, is about 0.14 in the case of this dataset and is added as *The Critical Value*(horizontal lines) in line plots of Figure 3.

According Rey (2004)[21], Figure 4 plots Moran's I and Theil index of each variable. The Theil index of inter-prefecutral inequalities is presented as follows(the formulation follows Cowell 2011[5]):

$$Theil_t = \frac{1}{n} \sum_{i=1}^{n} \frac{y_{it}}{\mu_t} \log(\frac{y_{it}}{\mu_t}) \tag{4}$$

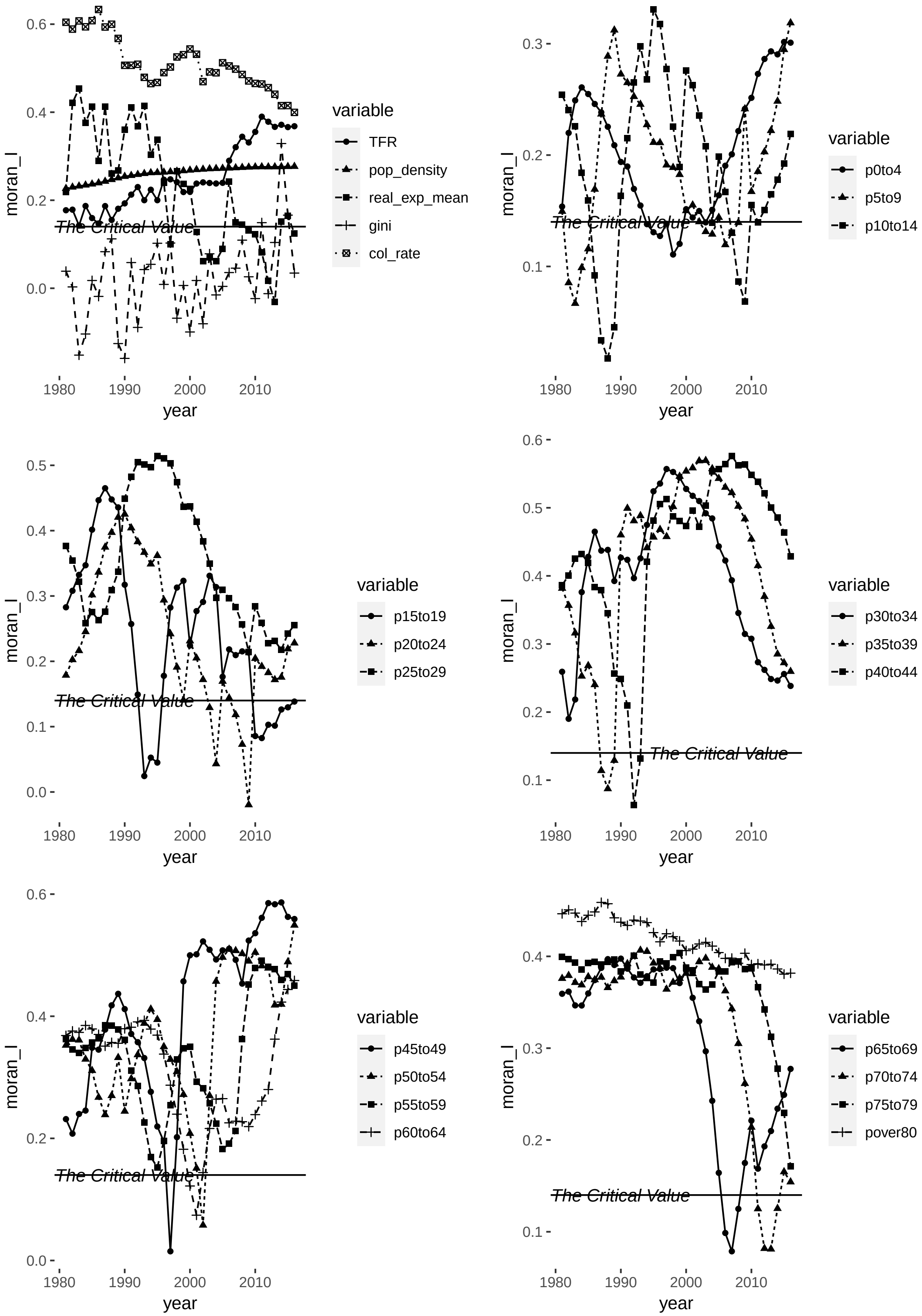

Figure 3: Moran's I of Each Variable in This Study

where y_{it} is the value of each variable in each prefecuture and each year, μ_t is the mean over Japan in each year and n is the number of prefectures(i.e.$n = 47$). Rey (2004)[21] empirically shows higher Moran's Is tend to be higher Theil indices. Theil indices have the drawback that they are affected by the sample size. With higher inequality of TFR, *ad_rate* has lower inequality over 36 years in this paper's dataset. Figure 4 indicates that inequality of TFR and that of *ad_rate* have a negative relationship, and one must consider geographical inequality to analyze the relationship of both. Next, we focus on spatial panel regression as the model of spatial autocorrelation.

3.4 The Spatial Panel Data Analysis

Hsiao (2014)[10] suggests that panel data analysis could control the impact of omitted variables (or individual or time heterogeneity) and generate more accurate predictions for individual outcomes than the cross-sectional analysis and time-series analysis.

One can apply the spatial panel data analysis to the analysis of Total Fertility Rates (TFRs) to gain more efficiency than the static panel analysis such as the Least Squares Dummy Variable (LSDV) or the dynamic panel analysis such as the dynamic panel Generalized Method of Moments (GMM) (Arellano & Bond, 1991[2]).

One can conduct the Maximum Likelihood (ML) (Baltagi et al. 2007[3]) and the Generalized Moments (GM) estimation (Kapoor et al. 2007[11]) to estimate the spatial panel. In this paper, the fixed effects spatial error model is estimated because the fixed effects model is generally more appropriate than the random effects model since adjacent spatial units' space-time data are located in unbroken study areas (Elhorst 2014[8]). Furthermore, as the spatial error model's parameters are considered marginal effects themselves (Elhorst 2014[8]), these parameters are easily interpretable. Then, other models using the spatial weight matrix are similar to the spatial error model. This paper's notations are based on Millo and Piras (2012)[16], and we use their package. This paper utilizes the spatial error model to control spatial autocorrelation. In the spatial error model, spatial error terms and the spatial lag terms are considered the spatial autocorrelation terms in the error terms.

The fixed effects spatial error panel data model estimated by the ML and the GM in this paper is as follows:

$$\mathbf{y} = (\boldsymbol{\iota}_T \otimes \mathbf{I}_N)\,\boldsymbol{\mu} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{5}$$

$$\boldsymbol{\varepsilon} = \rho\,(\mathbf{I}_T \otimes \mathbf{W}_N)\,\boldsymbol{\varepsilon} + \boldsymbol{\nu} \tag{6}$$

$$\tag{7}$$

$\mathbf{y}$ is the $NT \times 1$ vector of the dependent variable (here $\log(TFR)$), $\boldsymbol{\mu}$ the vector of the time-invariant fixed effect (not spatially autocorrelated), and $\mathbf{X}$ the $NT \times k$ of independent variables. Also, $\boldsymbol{\iota}_T$ is the vector of $T \times 1$ ones, $\mathbf{I}_N$ the identity matrix with $N \times N$ dimensions, $\mathbf{I}_T$ the identity matrix with $T \times T$ dimensions, and $\mathbf{W}_N$ the spatial weight matrix with $N \times N$ dimensions. $\boldsymbol{\varepsilon}$ is $\varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2)$ error terms, and $\boldsymbol{\nu}$ is $\nu_{it} \sim IID(0, \sigma_\nu^2)$. Moreover, $\boldsymbol{\beta}$ is the parameter vector, and ρ ($|\rho| < 1$) is the spatial autocorrelation coefficient. $\otimes$ represents the Kronecker product.

To control age structure, variables such as age-specific proportions (from $p0to4$ to $pover80$) imply young population proportions have negative impacts on the TFR, and this estimation results fit well the Easterlin hypothesis that young cohort sizes have adverse effects on their fertility (Easterlin 1968[7]).

$$\mathbf{y}_t = (\boldsymbol{\iota}_T \otimes \mathbf{W}_N)\,\boldsymbol{\mu} + \mathbf{X}_{t-1}\boldsymbol{\beta} + \boldsymbol{u}_t \tag{8}$$

$$\boldsymbol{u}_t = \rho\,(\mathbf{I}_T \otimes \mathbf{W}_N)\,\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t \tag{9}$$

$$t = 1, \ldots, T \tag{10}$$

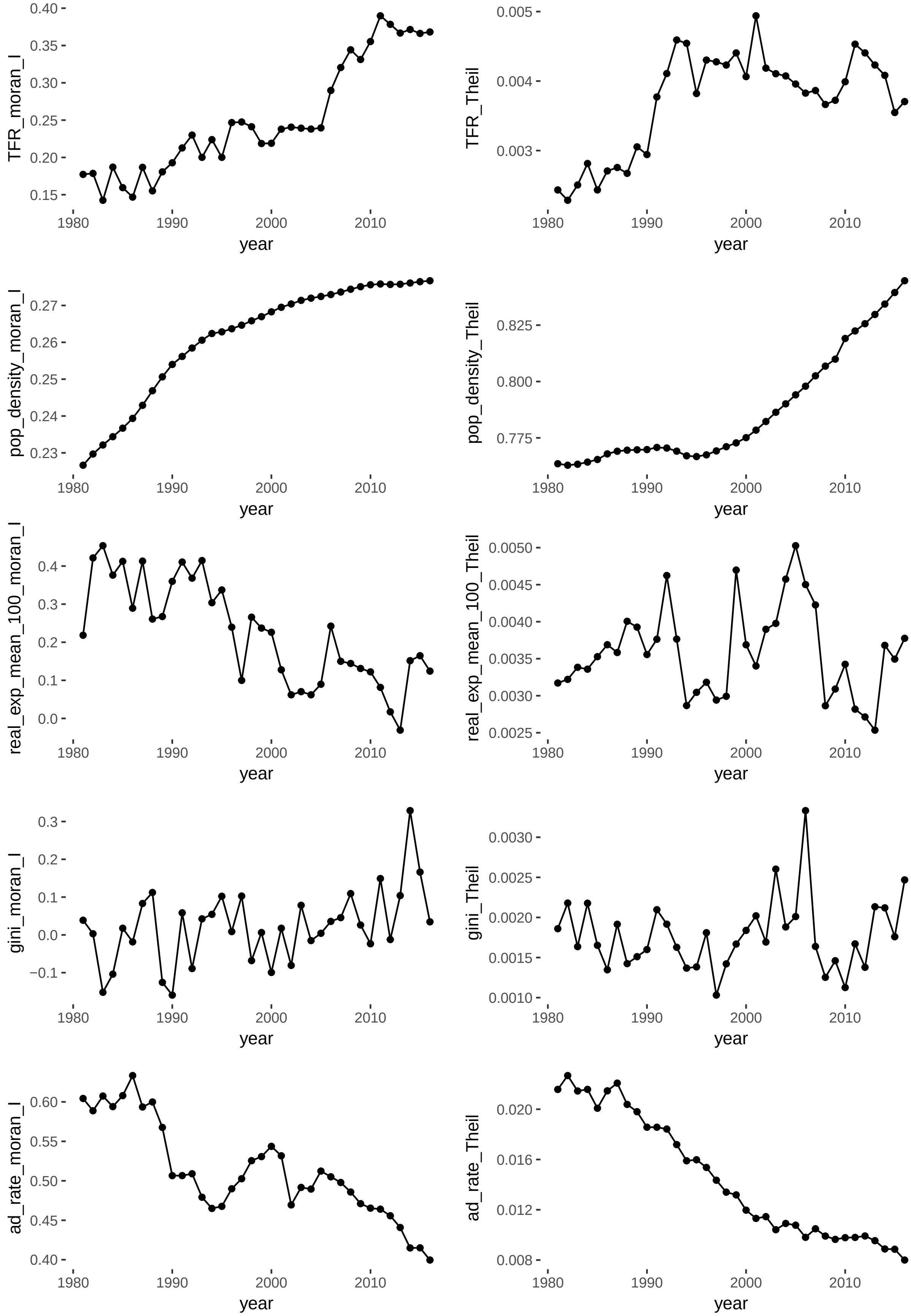

Figure 4: Moran's I and Theil Index of Each Variable in This Study

Table 2: Estimation Results: Spatial Panel Maximum Likelihood (SPML), Spatial Panel Generalized Moments (SPGM), Least Squares Dummy Variable (LSDV) and Dynamic Panel Generalized Methods of Moments (Arellano & Bond 1991[2])(DPGMM(AB)) ($n = 47$, $T = 35$, and $N = 1645$)

$\log(TFR)$	SPML	SPGM	LSDV	DPGMM(AB)
ρ	0.674254 (0.022147) ***	0.54946607		
σ_ν^2		0.00082763		
lag				0.353606 (0.037013) ***
$\log(lag(pop\_density, 1))$	-0.0163771 (0.0174095)	-0.017743 (0.015476)	0.0776786 (0.2496147)	0.485169 (0.324223)
$(\log(lag(pop\_density, 1)))^2$	0.0016538 (0.0013904)	0.0018175 (0.0012316)	0.0080262 (0.0190958)	-0.030917 (0.025209)
$\log(lag(real\_exp\_mean, 1))$	-5.7234362 (1.7326863) ***	-6.2441 (1.7248) ***	-8.1587635 (2.0155213) ***	-5.089522 (2.337559) *
$(\log(lag(real\_exp\_mean, 1)))^2$	0.2270749 (0.0685412) ***	0.24826 (0.068263) ***	0.3227035 (0.0796065) ***	0.202124 (0.092200) *
$\log(lag(gini, 1))$	0.0118918 (0.0137283)	0.014496 (0.013994)	-0.0124552 (0.0159547)	0.011498 (0.017563)
$\log(lag(ad\_rate, 1))$	-0.0360324 (0.0087634) ***	-0.031813 (0.0081342) ***	-0.0291019 (0.0168526) .	0.049324 (0.024595) *
$\log(lag(p0to4, 1))$	0.6246373 (0.0315136) ***	0.67079 (0.031383) ***	0.6306060 (0.0550940) ***	0.272078 (0.050675) ***
$\log(lag(p5to9, 1))$	0.2359812 (0.0399109) ***	0.21131 (0.040257) ***	0.0257211 (0.0621721)	0.103450 (0.051732) *
$\log(lag(p10to14, 1))$	0.0467749 (0.0313585)	0.041251 (0.031855)	0.0229333 (0.0479956)	-0.091458 (0.054953) .
$\log(lag(p15to19, 1))$	-0.1808072 (0.0205411) ***	-0.18262 (0.020838) ***	-0.2903170 (0.0307829) ***	-0.298755 (0.031187) ***
$\log(lag(p20to24, 1))$	-0.0957044 (0.0138215) ***	-0.10322 (0.013685) ***	-0.1192106 (0.0236034) ***	-0.129304 (0.021896) ***
$\log(lag(p25to29, 1))$	-0.2683597 (0.0209416) ***	-0.27073 (0.021170) ***	-0.4103448 (0.0330438) ***	-0.267500 (0.034114) ***

$\log(lag(p30to34, 1))$	-0.2831767 (0.0250284) ***	-0.28320 (0.025049) ***	-0.3239812 (0.0428268) ***	-0.289960 (0.045548) ***
$\log(lag(p35to39, 1))$	-0.0768934 (0.0156915) ***	-0.083393 (0.015441) ***	-0.1834299 (0.0328672) ***	-0.169882 (0.030038) ***
$\log(lag(p40to44, 1))$	-0.0481813 (0.0148061) **	-0.044103 (0.014591) **	-0.0777104 (0.0301375) *	-0.061538 (0.031882) .
$\log(lag(p45to49, 1))$	0.0037819 (0.0157401)	0.0088936 (0.015535)	-0.0343759 (0.0254181)	-0.061216 (0.034499) .
$\log(lag(p50to54, 1))$	0.0093755 (0.0149187)	0.000045573 (0.014862)	-0.0367954 (0.0203576) .	-0.076422 (0.027751) **
$\log(lag(p55to59, 1))$	0.0206318 (0.0112965) .	0.017104 (0.011299)	-0.0110085 (0.0111422)	-0.019842 (0.023959)
$\log(lag(p60to64, 1))$	-0.0528298 (0.0122143) ***	-0.059268 (0.012431) ***	-0.0676346 (0.0146975) ***	-0.058528 (0.020833) **
$\log(lag(p65to69, 1))$	-0.0041745 (0.0162989)	-0.0069989 (0.016680)	-0.0512689 (0.0199666) *	-0.032209 (0.020328)
$\log(lag(p70to74, 1))$	0.0247657 (0.0212499)	0.016970 (0.021771)	0.0124371 (0.0196962)	-0.037297 (0.027946)
$\log(lag(p75to79, 1))$	-0.1701261 (0.0213852) ***	-0.15362 (0.021546) ***	-0.1367153 (0.0244545) ***	-0.085345 (0.024770) ***
$\log(lag(pover80, 1))$	0.1153499 (0.0165672) ***	0.11265 (0.015793) ***	-0.0489186 (0.0364597)	-0.149514 (0.044055) ***
$adj.R^2$			0.92753	

$. < 0.10 \quad * < 0.05 \quad ** < 0.01 \quad *** < 0.001$
Double clustered standard deviations are in parentheses.

Table.2 does not indicate that $\log(pop\_density)$ and $\log(gini)$ in the model with the lag have the statistical significance in the regression to $\log(TFR)$. This result is consistent with the fact that TFR has panel non-Granger causality from $gini$ in Section 3.1. In short, $gini$ has no impact on TFR in the viewpoint of panel data. Moreover, $pop\_density$'s effect on TFR remains unclear.

4 Conclusion

Several people think rural areas have higher TFR than cities, and economic inequality negatively correlates with TFR. However, the negative relationship between TFR and $pop\_density$ and the positive relationship between TFR and $gini$ is false if one considers the lag between inception and birth. The spatial panel data analysis validates the Easterlin hypothesis. ad_rate has adverse effects on TFR in the spatial panel data model. However, the panel non-Granger causality shows *ad_rate* does not determine TFR in the non-Granger causality standpoint.

Compliance with Ethical Standards

Funding

This study was financially supported by the JSPS KAKENHI Grant Number 20H05633.

Conflict of interest

The author has no conflicts of interest to declare.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

References

[1] Luigi Aldieri and Concetto Paolo Vinci. Education and fertility: An investigation into italian families. *International Journal of Social Economics*, 2012.

[2] Manuel Arellano and Stephen Bond. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297, 1991.

[3] Badi H Baltagi, Seuck Heun Song, Byoung Cheol Jung, and Won Koh. Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics*, 140(1):5–51, 2007.

[4] Gary S Becker. *A Treatise on the Family*. Harvard university press, 2009.

[5] Frank Cowell. *Measuring inequality*. Oxford University Press, 2011.

[6] Elena-Ivona Dumitrescu and Christophe Hurlin. Testing for granger non-causality in heterogeneous panels. *Economic modelling*, 29(4):1450–1460, 2012.

[7] Richard A Easterlin. The american baby boom in historical perspective. In *Population, labor force, and long swings in economic growth: the American experience*, pages 77–110. NBER, 1968.

[8] J Paul Elhorst. *Spatial econometrics: from cross-sectional data to spatial panels*, volume 479. Springer, 2014.

[9] Kenneth Harttgen and Sebastian Vollmer. A reversal in the relationship of human development with fertility? *Demography*, 51(1):173–184, 2014.

[10] Cheng Hsiao. *Analysis of panel data*. Number 54. Cambridge university press, 2014.

[11] Mudit Kapoor, Harry H Kelejian, and Ingmar R Prucha. Panel data models with spatially correlated error components. *Journal of econometrics*, 140(1):97–130, 2007.

[12] Hill Kulu. Why do fertility levels vary between urban and rural areas? *Regional studies*, 47(6):895–912, 2013.

[13] Hongfei Li, Catherine A Calder, and Noel Cressie. Beyond moran's i: testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis*, 39(4):357–375, 2007.

[14] Wolfgang Lutz, Maria Rita Testa, and Dustin J Penn. Population density is a key factor in declining human fertility. *Population and Environment*, 28(2):69–81, 2006.

[15] Diane J. Macunovich and Richard A. Easterlin. *Easterlin Hypothesis*, pages 3165–3171. Palgrave Macmillan UK, London, 2018.

[16] Giovanni Millo, Gianfranco Piras, et al. splm: Spatial panel data models in r. *Journal of Statistical Software*, 47(1):1–38, 2012.

[17] Mikko Myrskylä, Hans-Peter Kohler, and Francesco C Billari. Advances in development reverse fertility declines. *Nature*, 460(7256):741–743, 2009.

[18] Fred C Pampel and H Elizabeth Peters. The easterlin effect. *Annual review of sociology*, 21(1):163–194, 1995.

[19] Klaus Prettner, David E Bloom, and Holger Strulik. Declining fertility and economic well-being: Do education and health ride to the rescue? *Labour economics*, 22:70–79, 2013.

[20] United Nations Development Programme. Human development index (hdi)—human development reports, 2020. `http://hdr.undp.org/en/content/human-development-index-hdi`.

[21] Sergio J Rey. Spatial analysis of regional income inequality. *Spatially integrated social science*, 1:280–299, 2004.

[22] Michael Sigmund and Robert Ferstl. Panel vector autoregression in R with the package panelvar. *The Quarterly Review of Economics and Finance*, 2019.

[23] Robert J Willis. A new approach to the economic theory of fertility behavior. *Journal of political Economy*, 81(2, Part 2):S14–S64, 1973.

[24] Batool Zaidi and S Philip Morgan. The second demographic transition theory: A review and appraisal. *Annual review of sociology*, 43:473–492, 2017.