DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 53, ARTICLE 18, PAGES 525–568
PUBLISHED 18 SEPTEMBER 2025
https://www.demographic-research.org/Volumes/Vol53/18
DOI: 10.4054/DemRes.2025.53.18

Research Article

Universal yet local: Estimating county-level fertility ideals and intentions in China

Donghui Wang

Yongai Jin

Tao Liu

Contents

Demographic Research: Volume 53, Article 18
Research Article

Universal yet local: Estimating county-level fertility ideals and intentions in China

Donghui Wang[1]

Yongai Jin[2]

Tao Liu[3]

Abstract

BACKGROUND
Understanding China's persistent low fertility requires detailed information regarding fertility attitudes at a finer geographic scale. However, data on fertility preferences at appropriate spatial resolutions are often unavailable.

OBJECTIVE
This study aims to estimate county-level fertility ideals and intentions in China.

METHODS
This study employs the multilevel regression and post-stratification method to estimate county-level fertility ideals and intentions. Fertility ideals and intentions data are drawn from a large national fertility survey, while post-stratification data come from the 2020 population census. The estimates are internally validated using a split sample approach and externally validated against independent national and regional surveys.

RESULTS
The estimates reveal that the county-level average ideal number of children for women of reproductive age is 1.98 (ranging from 1.29 to 3.78), while the average for the intended number of children is 1.81, with a broader range (1.02 to 3.96). The spatial distribution of fertility ideals exhibits a north–south contrast, suggesting cultural influences on family norms. Fertility intentions show coastal–inland disparities, underscoring socioeconomic conditions. Within-province variations are no less than between-province variations.

[1] Department of Sociology, Faculty of Social Sciences, University of Macau, Macau SAR, China. Email: donghuiwang@um.edu.mo.
[2] Corresponding author. Center for Population and Development Studies, School of Population and Health, Renmin University of China, Beijing, China. Email: jinyongai0416@ruc.edu.cn.
[3] Corresponding author. College of Urban and Environmental Sciences, Center for Urban Future Research, Peking University, Beijing, China. Email: liutao@pku.edu.cn.

CONTRIBUTION
These findings highlight the complexity of the fertility attitudes landscape in China. The estimates also serve as an important data source for predicting future fertility and designing place-based policies.

1. Introduction

Fertility attitudes – including the ideal number of children for a family and the number of children a person wants – have been the cornerstone of fertility research (Bongaarts 2001; Casterline and Sinding 2000; Lee 1980). With sustained low fertility rates now a major demographic reality in many societies, there is a rising scholarly interest in studying fertility attitudes in low fertility contexts. China contributes a compelling case for such studies. The country has experienced a rapid fertility decline over the last several decades. As China's fertility rate declines, people's fertility attitudes also change (Hou 2015). To date, China stands out among the low fertility countries, with fertility ideals falling below the replacement level, whereas other nations still maintain fertility ideals above this threshold (Chen and Gietel-Basten 2024).

Chinese people's fertility attitudes have been closely measured, analyzed, and debated over the past several decades (Attané 2016; Chen and Gietel-Basten 2024; Hermalin and Liu 1990; Hou 2015; Jiang, Li, and Sánchez-Barricarte 2016; Merli and Morgan 2011; Nie and Wyman 2005; Whyte and Gu 1987; Zheng et al. 2009). Despite such intensive investigations, our understanding of fertility attitudes in China remains incomplete. Most existing studies tend to draw conclusions at the national level and disregard large subnational variations. This gap exists because survey data on fertility attitudes at fine-grained geographic scales are scarce. Yet this gap is significant, and theoretical developments of fertility behaviors have suggested extending the discussion to the new frontier of geographic space (Lesthaeghe 2010). People live and socialize in local places. They form their childbearing attitudes and expectations in relation to their local environments (Bachrach and Morgan 2013). Ignoring such variations limits our understanding of the nature of China's low fertility. From a policy-making perspective, local governments also need fertility attitudes data at fine-grained geographic scales to allocate resources and make predictions of future fertility levels.

Furthermore, with a few exceptions (e.g., Chen and Yip 2017), few empirical studies have made a clear distinction between the two concepts: fertility ideals and intentions. Even if such distinctions are made, few studies have examined both. Fertility ideals reflect childbearing desires under ideal conditions for a typical family, while fertility intentions reflect individuals' actional fertility plans (Miller 2011; Philipov and Bernardi

2011). Although the two concepts are interconnected, they serve different analytical purposes. Fertility ideals are often used to describe the fertility norms prevalent in a place (Goldstein, Lutz, and Testa 2003; Sobotka and Beaujouan 2014). Fertility intentions provide insights into childbearing plans and readiness to have children; thus they are used to predict the future (Hartnett and Gemmill 2020; Yeatman, Trinitapoli, and Garver 2020). Distinguishing the two helps to clarify different contributing factors underpinning the fertility decision-making process. This distinction is also crucial for policy-makers and researchers aiming to design family support programs that align with both the aspirations and practical realities of individuals and families.

This study takes the first step to deal with these knowledge gaps and estimates fertility ideals and intentions for detailed demographic subgroups at the county level in China. We employ multilevel regression and post-stratification (MRP), a method that combines the strengths of model- and design-based small-area estimation techniques (Gelman and Little 1997). The method has a proven record of producing reliable subnational-level attitudes (Buttice and Highton 2013; Lax and Phillips 2009; Pacheco 2011; Warshaw and Rodden 2012). We model individuals' fertility ideals and intentions based on a large nationally representative survey sample. These model estimates are then post-stratified using detailed county-by-age-group female population counts from the national census.

Our county-level estimates contribute to the existing knowledge of fertility attitudes in China in several ways. First, they allow us to investigate the prevalence of high or low fertility ideals and intentions, thus effectively evaluating the prevalence of the new fertility norms and the persistence of the old ones. Second, they allow us to identify spatial patterns that are otherwise masked by the national averages. Third, they allow us to inspect the geographic scales at which the underlying mechanisms operate. Altogether, our study contributes to a broad scholarly effort that emphasizes the importance of space for demographic and sociological behaviors (Entwisle 2007; Lobao 2004).

Our study also advances the methodology of small-area estimations of demographic attitudes. Despite the growing need to understand demographic behaviors based on individual perceptions, existing small-area estimations in population research focus mainly on behavioral measures. We apply MRP to the study of fertility attitudes, thus extending the applicability of this approach.

2. Fertility ideals, intentions, and their spatial variations

Fertility ideals and intentions are important concepts that capture individuals' childbearing attitudes. While the two concepts are closely related, they have different theoretical underpinnings and serve different analytical purposes (Hagewen and Morgan

2005; Philipov and Bernardi 2011; Trent 1980). Fertility ideal is the "preferred number of children for some typical family" (Hagewen and Morgan 2005: 511). Despite scholarly debates on its constructive validity (Blake 1966; Philipov and Bernardi 2011), ideal family size still plays a fundamental role in human reproductive decision-making (Hagewen and Morgan 2005; Trent 1980). In comparison, fertility intention is conceptualized as individuals' actionable plans regarding their reproductive goals (Ajzen and Klobas 2013; Miller 2011). Compared to the fertility ideal, fertility intention is a better predictor of fertility behavior (Bongaarts 2001; Hartnett and Gemmill 2020; Schoen et al. 1999). But intentions do not always translate to behaviors; a large strand of literature takes intention as the starting point and strives to understand what facilitates or hinders the realization of fertility intentions (Beaujouan and Berghammer 2019; Hagewen and Morgan 2005; Schoen et al. 1999).

Prior studies on fertility ideals and intentions place greater emphasis on temporal variations over spatial variations. Yet understanding spatial variations of fertility ideals and intentions is important, not only because such variations exist but also because existing fertility theories point to such direction. First, people make fertility decisions in response to local structural constraints. Long-standing research in demography suggests that while reproductive goals have a personal dimension, they also respond to local socioeconomic conditions beyond personal characteristics (Bachrach and Morgan 2013; Behrman and Weitzman 2024; Billingsley and Ferrarini 2014; Blake 1966). Therefore fertility attitudes are expected to co-vary with local socioeconomic and policy conditions.

Second, cultural models of fertility change also foreshadow geographic variations of fertility attitudes (Lesthaeghe 2010; Thornton et al. 2012). For example, Goldstein, Lutz, and Testa (2003) identified the emergence of a sub-replacement-level fertility ideal in the German-speaking part of Europe and attributed this trend to a societal shift toward low fertility norms. Sobotka and Beaujouan (2014) examined the same question, and using more comprehensive datasets they arrived at somewhat different conclusions compared to Goldstein, Lutz, and Testa (2003). They noted that the two-child ideal still prevails across European countries. Despite the difference in findings, both studies use geographic in(variation) to discuss whether there is an emergence of new fertility norms in low fertility contexts.

To sum up what we learned from the existing literature, fertility ideals and intentions are expected to vary across space because they are shaped by local material structures and cultural contexts. However, the conceptual differences between the two lead us to formulate distinctive expectations regarding their spatial variations. First, previous studies suggest that fertility ideals tend to remain relatively stable over time (Sobotka and Beaujouan 2014). We infer that they should also be relatively stable across space because the concept reflects the collective understanding of what constitutes an ideal family size in a place. On the other hand, fertility intentions represent individuals' childbearing plans and thus are expected to have greater geographic variations. Second, while both fertility ideals and intentions are shaped by local contexts, their primary driving forces may differ. The ideal family size, which reflects dominant cultural norms regarding childbearing, is expected to diffuse readily along cultural boundaries. In contrast, socioeconomic factors, such as levels of economic development, are likely to exert a stronger influence on the geographic distribution of people's intended family size.

3. The changing landscape of fertility ideals and intentions in China

During the one child policy era (1980s to the 2000s), having a firm grasp of the Chinese people's true fertility preferences was challenging. Survey responses on intended family size were deemed unreliable, as individuals tended to provide answers that aligned with policy-permitted birth quotas (Hermalin and Liu 1990; Merli and Smith 2002). This led researchers to focus on the ideal family size, as the concept presents a hypothetical scenario that allows respondents to express preferences that are free from birth restrictions (Hou 2015). Thus, despite its limitations in predicting actual behavior, ideal family size became the primary, albeit imperfect, indicator of people's genuine fertility preferences at that time.

China gradually phased out its restrictive fertility policies in the 2010s, moving from a selective two child policy in 2013 to a universal two child policy in 2016 and ultimately to a three child policy in 2021 (Zhai and Jin 2023). A crucial question remains: Has this shift in policy triggered a substantial increase in fertility intentions and, furthermore, altered fertility ideals? Chen and Gietel-Basten (2024) identified a causal relationship between the removal of the one-child policy and an increase in the ideal number of children. However, the effects appear to be modest, and the authors conclude that people in urban China have genuinely embraced sub-replacement fertility ideals. Their findings mirrored other recent regional studies (Jin, Song, and Chen 2016), pointing toward the persistence of low fertility ideals despite policy changes.

To date, it becomes clear that with the removal of the restrictive fertility policy, fertility ideals and intentions remain low in China. However, our understanding of China's low fertility ideal and intention in the post–one child policy era remains underdeveloped. Most prior studies are at the national level, which unfortunately overlooks large subnational heterogeneities. While some acknowledge subnational differences, they often focus on broad regions or provinces, overlooking differences within those areas. We argue from three aspects that fertility ideals and intentions at small geographic scales merit scholarly attention in contemporary China.

First, demographers have long sought to understand the link between inequality and fertility. How resources are allocated across social groups and geographic locations greatly shapes people's childbearing decisions. In contemporary China, it is the geographic disparities, rather than disparities across social groups, that stand out as the primary source of inequality (Xie and Zhou 2014). Because we take the county as our primary geographic unit of analysis, our study sheds light on the importance of geographic scales: If variations in fertility attitudes occur primarily at the provincial or broader regional level, this suggests that the underlying mechanisms – be they economic, cultural, or policy related – are at play at such broader geographic scales. Future studies may safely investigate regional patterns or trends and get an accurate picture of fertility attitudes in China. On the contrary, if most variations occur at smaller geographic scales, this may indicate that local environments are more influential. This would suggest that detailed, localized studies are necessary.

Second, people form fertility attitudes by socializing with others, presumably those who live within close geographic proximity (Cleland and Wilson 1987). The cultural models of fertility change suggest that a shift in fertility preferences can occur independent of economic development. For example, the developmental idealism framework suggests that beliefs in the benefits of low fertility can be endorsed in places with low development levels (Thornton et al. 2012; Thornton and Xie 2016). Past empirical studies have shown that fertility ideals are below replacement level at the national level, suggesting a shift toward small-family norms. However, the prevalence and geographic distributions of these new norms remain unclear. With county-level estimates, we can better identify the prevalences and spatial patterns of high and low fertility ideals and intentions, thus gaining a deeper understanding of how new norms spread and old ones persist.

Third, from a policy-making perspective, the demand for small-area data on fertility attitudes is also growing in China. This demand is driven by the increasing need for place-based family policies. China's approach to fertility policy has always involved a degree of decentralization. Even back in the 1980s, the strengths and incentives of fertility policy were known to vary considerably across space (Short and Fengying 1998). When a shift toward a more lenient policy was under way, local governments were given more freedom to design their own policies (Basten and Jiang 2014). To date, many local governments are now considering taking active roles in implementing family support policies in response to low fertility (China Daily 2021). While guidelines for family support policies may be made from higher administrative levels, it is municipal and county-level authorities who are responsible for putting these programs into action and managing their budgets. Data on fertility ideals and intentions can be crucial for predicting future fertility trends and ensuring that resources are allocated effectively.

4. Our approach to small-area estimation of fertility attitudes

This study aims to estimate county-level fertility ideals and intentions. A major challenge is the dearth of fine-grained data. While population censuses can provide reliable information for behavioral indicators, estimating attitudes relies mostly on surveys. Such surveys are often designed for national representativeness and often lack sufficient data points for reliable subnational estimates.

Scholars have developed two main small-area estimation approaches: design-based and model-based (Pfeffermann 2013). The design-based approach adjusts estimates using sampling weights based on the survey's sampling design (Pfeffermann 2013). The model-based approach use statistical models and auxiliary information to predict variables of interest (Chi 2009). Both approaches have limitations. The design-based approach relies heavily on the sampling design and its assumptions, which can be restrictive when dealing with complex small-area estimation problems. The model-based approach assumes that the constructed models apply to the entire target population, which is often not the case (Zhang et al. 2015).

We adopt MRP, a hybrid approach that leverages the advantages of both design- and model-based small-area estimation techniques. MRP combines the strengths of two methods: multilevel regression, which considers the hierarchical structure of data, and post-stratification, which ensures the representativeness of estimates (Gelman and Little 1997). Thus MRP enables researchers to model relationships between individual- and group-level variables while adjusting for non-representativeness and other biases using post-stratification techniques. MRP has proven to be particularly useful for estimating preferences and behaviors at the subnational level or for subgroup estimates for surveys with limited sample sizes or non-representative samples (Buttice and Highton 2013; Lax and Phillips 2009; Pacheco 2011; Wang et al. 2015).

MRP involves two steps. First, multilevel regression models individual responses as a function of individual and geographic characteristics. Unlike the traditional model-based approach, MRP extensively uses random effects, which allows for the partial pooling of information across units, especially units with smaller sample sizes (Gelman and Little 1997). The estimated coefficients are used to predict detailed cross-tabulated geographic and demographic subcategories. In the second post-stratification step, the estimated subcategories are weighted by their true population sizes. A post-stratification framework is usually obtained from a cross-tabulated population census. This step adjusts the estimates to match the targeted population to reduce any biases resulting from survey-related issues, such as sample sparseness, non-representativeness, or under-coverage of certain demographic groups (Wang et al. 2015).

While we build on earlier studies, our analysis departs from previous MRP studies in two ways. First, while most prior MRP research estimates public opinion, we estimate preferences closely aligned with demographic behaviors. With this work, we extend the application of MRP to the field of demographic attitudes. Second, while scholars have extensively verified that MRP can be applied to estimate state-level public opinion (Buttice and Highton 2013; Lax and Phillips 2009; Warshaw and Rodden 2012) and have situated the method as the gold standard for state-level public opinion estimation (Selb and Munzert 2017), MRP's validity at smaller scales remains uncertain. In response to this gap, we also validate the method's effectiveness at the county level.

5. The survey

We use the 2017 National Fertility Survey as our primary data source. The survey is administered by the National Health and Family Planning Commission of China. This comprehensive survey collects information on reproductive knowledge, attitudes, and practices among women aged 15–60 in mainland China (Zhuang et al. 2019). It uses cluster sampling and probability proportional to sample size (PPS) sampling. To ensure accuracy, face-to-face interviews were conducted with rigorously trained participants. A post-survey analysis confirmed that the survey was of high quality (Zhuang et al. 2019).

The survey's notable strength lies in its extensive sample size, which comprises 249,946 women. Although only women of reproductive age (15–49 years) were asked about their fertility attitudes, which limited our analytical sample to 160,494 observations, the sample still covered 2,439 out of 2,821 counties in mainland China (Figure A-1). With county-level identifiers, we can link the survey with auxiliary data sources, such as the population census.[4] Although the analytical sample size seems substantial, the survey was not representative at the county level. The average sample size per county was 63, with substantial variation (13–833 observations per county). Therefore the survey cannot be used directly to generate county-level averages of fertility ideals and intentions, particularly for counties with limited sample sizes. Thus a small-area estimation is necessary.

The survey included two questions. First, "What do you think is the ideal number of children for a family to have?" Second, "How many children do you want to have?" The first question asked respondents about their beliefs regarding the ideal family size in general rather than for themselves; thus we use it as a measure of fertility ideal. The second question asked respondents about the size of the family they intended to have, and

[4] We removed observations from 222 counties in the survey that could not be merged with 2010 county-level population census data. The detailed data-matching procedure can be found in Section A of the online supplemental materials. The four directly administered municipalities, Beijing, Tianjin, Shanghai, and Chongqing, are treated as equivalent to the provincial-level units.

we use it as a measure of fertility intention.[5] Detailed summary statistics of the two variables can be found in Table A-1.

6. Multilevel model specification

MRP involves predicting and post-stratifying survey responses within demographic and geographic subgroups. As the first step of the analysis, we model respondents' ideal and intended family sizes as a function of individual demographics and geographic covariates. Following Gelman and Hill's notation (2007), Equation (1) summarizes the basic model. In particular, y_i (ideal or intended number of children)[6] for individual *i* is expressed as follows:

$$y_i = \beta^0 + \alpha_{j[i]}^{age} + \alpha_{l[i]}^{county} + \alpha_{m[i]}^{prefecture} + \alpha_{n[i]}^{province} + \beta^1 livebirth_{2010}, \qquad (1)$$

where

$\alpha_{j[i]}^{age} \sim N\left(0,\ \sigma_{age}^2\right)$ for j= 1,2,…,7

$\alpha_{l[i]}^{county} \sim N(0,\ \sigma_{county}^2)$ for l= 1, 2, …,2,439

$\alpha_{m[i]}^{prefecture} \sim N\left(0, \sigma_{prefecture}^2\right)$ for m= 1,2, …,349

$\alpha_{n[i]}^{province} \sim N\left(0,\ \sigma_{province}^2\right)$ for n =1,2…,32

In Equation (1), β^0 is the overall intercept, and $\alpha_{j[i]}^{age}$, $\alpha_{l[i]}^{county}$, $\alpha_{m[i]}^{prefecture}$, and $\alpha_{n[i]}^{province}$ denote random intercepts for age, county, prefecture, and province categories, respectively. The subscripts *j, l, m,* and *n* indicate category membership for age, county, prefecture, and province, respectively. For example, $\alpha_{j[i]}^{age}$ corresponds to seven age categories (j = 1,2,…7): 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49. We focus on five-year age groups as random intercepts for two reasons. First, five-year age groups

[5] We excluded non-numerical responses, including "don't know" or "not sure" (0.6% of the sample), from questions about fertility ideals and intentions.

[6] Although MRP is typically used with binary outcomes, it is not limited to them and can be applied to other forms of measurement. We treat the ideal and intended numbers of children as continuous variables and model them using a Gaussian link to save computing time and avoid model convergence failure in the subsequent cross-validation analysis.

are consistently available in both the survey data and the population census, and this allows us to post-stratify the predicted cells effectively. Second, age is a well-established predictor of individual fertility preferences (Balbo, Billari, and Mills 2013; Liefbroer 2009) and a critical dimension in measuring actual fertility rates. Most of the fertility indicators, such as the total fertility rate, are age-graded. Therefore our age-stratified estimates enable scholars and practitioners to directly compare age-specific fertility preferences with actual age-graded fertility rates at small geographic scales.

We further included three geographic variables (county, prefecture, and province) as random intercepts. All random intercepts are drawn from a normal distribution with unknown variance and a mean of 0. The inclusion of demographic and geographic variables as random effects, rather than fixed effects, is crucial in MRP. This approach leverages the "partial pooling" effect of the multilevel model (Gelman and Hill 2007), which is particularly valuable when working with limited sample sizes in specific locations or demographic groups.

Prior studies suggest that including important contextual variables can improve MRP model fit and estimation accuracy (Buttice and Highton 2013; Lax and Phillips 2009). For our analysis, we added a fixed effect, livebirth2010, as a proxy for past fertility. The variable measures the county's average number of live births per woman of reproductive age in the 2010 census. Past fertility matters, as individuals form their fertility attitudes through social learning by observing real fertility behaviors around them (Lutz, Skirbekk, and Testa 2006). Thus places with low fertility rates may further reinforce low fertility norms, whereas places with high fertility rates may reinforce high fertility norms. The analysis is performed using R (version 4.4.1).

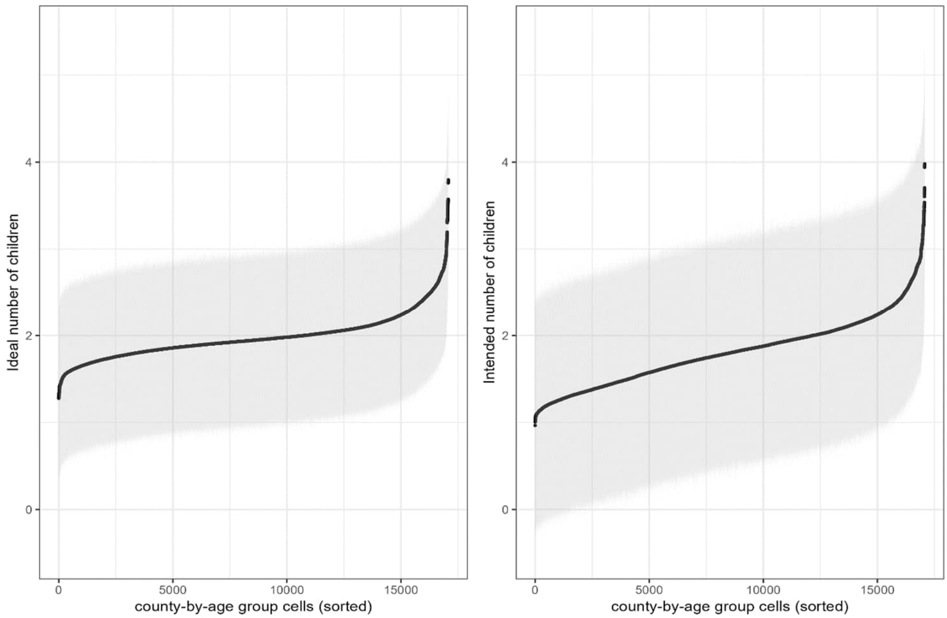
Table 1 presents the summary of the multilevel model. The fixed effect livebirth2010 was positively correlated with the respondents' ideal and intended numbers of children. This finding supports previous research indicating that past fertility contexts influence fertility ideals and intentions. After controlling for livebirth2010, we find that substantial between-group variations persist. Geographic random effects, such as county, prefecture, and province, show relatively high between-group variance, whereas the two demographic random effects show relatively low between-group variance. The only exception is education, which exhibits the largest variance (0.04 in the fertility intention model). We also experiment with alternative model specifications. Detailed discussions regarding alternative model specifications can be found in Table A-2.

Table 1: Multilevel model results for ideal and intended number of children

	Fertility ideal		Fertility intention	
Fixed Effect	Estimates	S. E	Estimates	S. E
Intercept	1.44	0.07	0.68	0.13
Livebirth2010	0.39	0.01	0.80	0.02
Random Effect	Variance		Variance	
σ^2	0.24		0.43	
County	0.02		0.03	
Prefecture	0.01		0.02	
Province	0.02		0.03	
Age group	0.01		0.01	
N				
County	2439		2439	
Prefecture	338		338	
Province	32		32	
Age group	7		7	
Observations	160,494		160,494	

Figure 1 presents a scatter plot of the predicted values of the 17,073 demography-by-geography subgroups (7 age groups × 2,439 counties) based on the estimated fixed and random effects from the multilevel models. The figure provides the first impression of varying fertility ideals and intentions across geographic and demographic subgroups. Fertility ideals are generally higher than intentions, which is expected from their definitions. However, fertility intentions have a wider confidence interval than ideals, suggesting that greater uncertainty is associated with predicting realistic fertility plans compared to fertility ideals.

Figure 1: Predicted cell values based on multilevel models



Notes: The graphs show the scatter plot of the predicted values of 17,073 cells (7 age groups × 2,439 counties), based on the multilevel models presented in Table 3. The estimates were sorted from the smallest to the largest estimated values. Shaded areas indicate 95% confidence intervals produced from 999 simulations using the predictInterval function in R.

7. The post-stratification framework

The second step of MRP is post-stratification, where we adjust the estimates in each demography-by-geography cell to match actual counts. Specifically, our model requires us to post-stratify the estimates with detailed age group counts at the county level. We first obtain fully enumerated female population counts by seven age groups (15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49) at the county level, which yields 17,073 cell counts (7 age groups × 2,439 counties). Next we weigh the cells according to their population percentages. For each estimate, θ_j, in each county, s, the weighted predictive mean ideal and intended family size $\hat{y}$ can be expressed as:

$$\hat{y} = \frac{\sum_{j \in s} N_j \theta_j}{\sum_{j \in s} N_j}$$

We further evaluate the performance of our estimates through both internal and external validation methods.

8. Results

8.1 County averages and prevalences

Table 2 summarizes MRP estimates. Panel A shows county average estimates. Panel B shows the estimates by age subgroup within each county. Panel C describes the prevalence of high and low fertility ideals and intentions, dividing county average estimates into four categories. As Panel A reveals, the mean estimated ideal and intended numbers of children are 1.98 and 1.81, respectively, which is consistent with the raw survey data. The statistics vary greatly by county, however, with the estimated ideal number of children ranging from 1.29 to 3.78 and the intended number of children ranging from 1.02 to 3.96. When broken down further by age subgroup (Panel B), the range widens, with the ideal number of children spanning from 1.28 to 3.8 and the intended number of children ranging from 0.96 to 3.98.

Panel C further breaks down the prevalence of high and low estimates, categorizing them into four discrete categories: 0.9–1.5, 1.5–2, 2–2.5, and 2.5 and above. Of the 2,439 counties, 63% have fertility ideals below replacement level, while over 71% have below-replacement fertility intentions. This highlights the level of prevalence of below-replacement fertility ideals and intentions. Among these counties, around 24% have fertility intentions below 1.5. If these intentions translate into actual fertility behavior, the realized fertility rates could end up even lower.

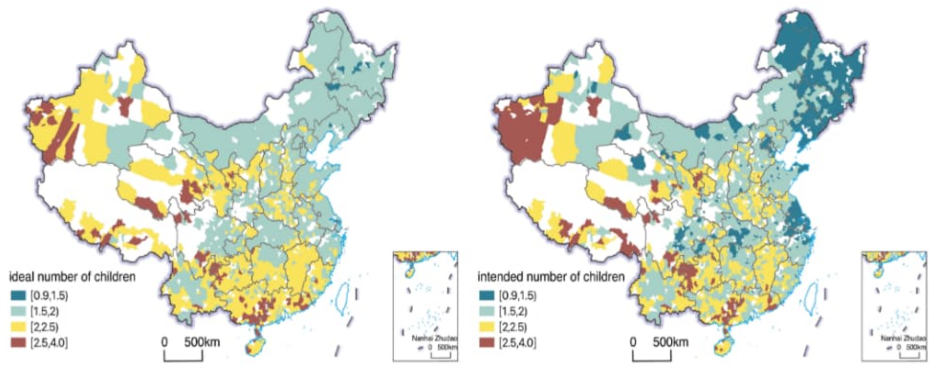
Table 2: Summary statistics of MRP estimates

Panel A: County Average						
	Min	1st quantile	Median	Mean	3rd quantile	Max
Ideal	1.29	1.84	1.94	1.98	2.07	3.78
Intention	1.02	1.51	1.80	1.81	2.03	3.96
Panel B: County Average of Age Subgroups						
	Min	1st quantile	Median	Mean	3rd quantile	Max
Ideal	1.28	1.83	1.94	1.98	2.07	3.80
Intention	0.96	1.51	1.80	1.81	2.04	3.98
Panel C: Percentage of Counties That Fall into Each Category						
		0.9–1.5	1.5–2.0		2.0–2.5	2.5–4.0
Ideal		0.87%	62.22%		32.44%	4.47%
Intention		24.08%	46.96%		23.95%	5.62%

8.2 Spatial patterns

Figure 2 presents maps of the aggregated county average estimates of fertility ideals and intentions. The corresponding local spatial cluster maps are in Figure A-2, while the detailed age-by-education subgroup maps are in Figures A-3 and A-4.

Figure 2: Geographic variations of the MRP estimates



Note: Counties with no survey data are left blank.

At first glance, the spatial distributions of fertility ideal and intention appear similar, which is expected, since the two concepts are closely related. However, a closer look reveals notable differences. The fertility ideal generally exhibits a north–south difference. Northern counties (north of the Yangtze River)[7] tend to have below-replacement fertility ideals, while southern counties tend to have above-replacement fertility ideals. This pattern aligns with the cultural explanation of the fertility behaviors (Peng 2010; Zhang and Li 2017; Zou et al. 2024). When considering the ideal number of children in a family, people often take cues from local family size norms. Therefore fertility ideals are more likely to diffuse along cultural boundaries. In China this cultural influence is particularly evident in the contrast between lineage culture (or clan culture) in the south and culture in other regions. As early anthropological work documents, lineage culture has been deeply entrenched in southern China for centuries (Fei 1946; Freedman 1958; Hu 1948). This tradition is less preserved in the north because of warfare eruptions (Peng 2010). The lineage system places great importance on the continuities of the family bloodline through male descendants, which in turn gives rise to a pronatalist culture. To ensure at least one son, families faced great social pressure to have more children, which further

[7] The location of the Yangtze River is presented in Figure A-1.

elevated overall fertility levels as a form of biological and social insurance for the patriline. Recent empirical studies further confirmed this strong relationship between the presence of lineage culture and fertility levels in China (Zhang and Li 2017; Zhang 2022; Zou et al. 2024). One notable example of this effect is Guangdong Province, located in the southernmost part of mainland China.[8] As one of the country's most economically developed regions, Guangdong might be expected to follow the trend of other affluent areas, preferring smaller families. However, residents of Guangdong tend to prefer larger families,[9] likely due to the region's strong pronatalist norms (Attané 2001).

There are exceptions to this north–south difference. Places with large ethnic minority populations, such as the northwest, have high fertility ideals. The northwest has a strong presence of Turkic and Tibetan populations, whose childbearing norms differ from those of the Han Chinese (Yang and Schieman 2024). The practice of decades-long lenient fertility policy further reinforces such difference (Jin, Hu, and Wang 2023). The small-area estimates allow us to see much more refined cultural boundaries than what is typically revealed by regional or provincial statistics.

Meanwhile, although culture continues to shape the distribution of fertility intentions, economic forces play a greater role. The spatial distribution of fertility intention shows a contrast between coastal and inland areas. Residents in economically developed coastal areas and large cities like Beijing and Shanghai generally intend to have fewer children. In contrast, those in less developed inland areas intend to have more children. When individuals make fertility decisions, they carefully consider a range of social and economic constraints. High living standards and the associated costs of childbearing often lead individuals in economically developed areas to have fewer children. Beyond financial concerns, the opportunity costs of childbearing – such as career advancement and education – are more pronounced in coastal areas. This further discourages women of reproductive age in coastal areas from having children.

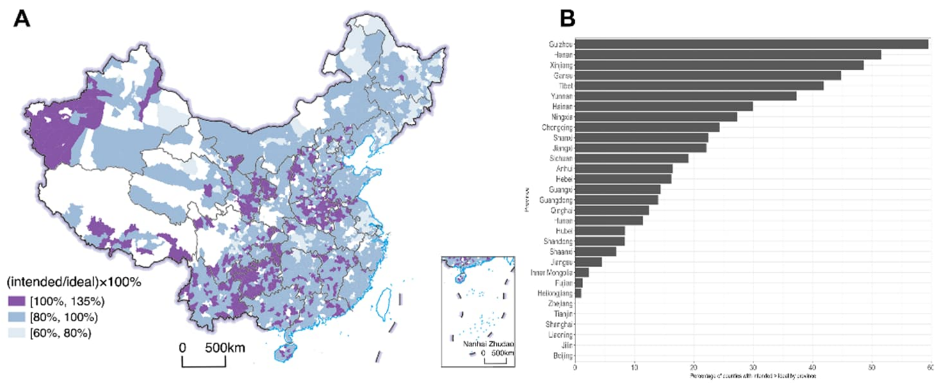
Among the areas with low fertility intentions, the northeast stands out with the lowest fertility intention, averaging less than 1.5 children per female. This low fertility intention may be explained by the feedback loop between past fertility levels and future fertility intentions proposed by Lutz, Skirbekk, and Testa (2006). As one of the earliest industrialized areas, the northeast exhibits some of the lowest observed fertility rates (Wang and Chi 2017). Consequently, younger cohorts growing up in this low fertility setting may internalize low fertility norms, which reinforces the intention to have fewer children. Our county-level estimates confirm the very low fertility intentions in this region.

To evaluate discrepancies between fertility ideals and intentions, we computed the ratio between the two. We divided county-level fertility intention estimates by fertility

[8] The locations of the provinces are shown in Figure A-1.
[9] Refer to the local cluster map in Figure A-2 for the spatial "hot spot" of high fertility ideals.

ideal estimates and multiplied the ratio by 100% (Figure 3). Ratios close to 100% indicate a convergence between ideals and intentions. For most counties (N = 1,542), this ratio is below 100%, indicating that fertility intention is lower than the ideal. Approximately 18% (N = 443) of these counties fall into the lowest ratio category (below 80%). These counties are concentrated in the most industrialized and urbanized places in China, such as large metropolitan areas like Shanghai, or in the northeast, where the actual fertility rate is also low. In a low fertility context, fertility ideal is considered as the upper limit of fertility (Hagewen and Morgan 2005). A gap between intention and ideal suggests there are constraints preventing people from translating their ideals into intentions (Brinton et al. 2018). In places with such gaps, family policies to eliminate childbearing barriers may be especially effective.

Figure 3: Ideal-to-intention ratios and distributions by province



Notes: The left-hand side of Figure 3 presents the spatial distribution of counties with ideal-to-intention ratios. The right-hand side of Figure 3 presents the percentages of counties where the intended number of children exceeds the ideal number of children, by province.

Around 19% of the counties (N = 454) exhibited higher estimated fertility intentions than ideals. Figure 3 (on the left) shows the spatial distributions of these counties, while Figure 3 (on the right) further breaks down their proportions by province. Together these two figures reveal that these counties tend to be concentrated in less economically developed regions with historically high fertility levels. Notable examples include regions such as Xinjiang and Guizhou, which have large ethnic minority populations, and central provinces like Henan, characterized by a high concentration of rural residents. A supplemental analysis (Table A-3) examining the socioeconomic characteristics and fertility levels of these counties supports our visual inspection. Counties where fertility intentions exceed ideals tend to have higher actual fertility rates and lower levels of socioeconomic development. These areas are marked by lower average educational attainment, larger ethnic minority populations, higher sex ratios at birth, and a smaller proportion of residents with nonagricultural *hukou* status. Additionally, they show lower percentages of employment in industrial and service sectors, and higher deprivation scores. Within this high fertility context with lower socioeconomic development levels, people may intend to have more children than the popular ideal number, a phenomenon that is little discussed in existing literature. These findings can be made only when both fertility ideals and intentions are examined side by side at fine-grained geographic scales.

8.3 Within- and between-province heterogeneity

To illustrate how geographic scales matter in understanding fertility attitude variations, we constructed the Theil index to quantify between- and within-province variations in fertility ideals and intentions.[10] As shown in Table 3, the national-level Theil index of fertility intention is more diversified (0.022) compared to fertility ideals (0.007). We further decomposed the index into within- and between-province differences. The results suggest that within-province differences contribute no less than between-province differences. For fertility intention, within-province differences account for 52.7% of the total variation, compared to 47.3% for between-province differences. Similarly, for the fertility ideal, within-province variation contributes nearly half (48.4%) of the total variation. Past studies primarily describe fertility variations at broader regional levels. Our findings highlight that while regional differences are prominent, socioeconomic and cultural influences also operate at a smaller scale within provincial boundaries.

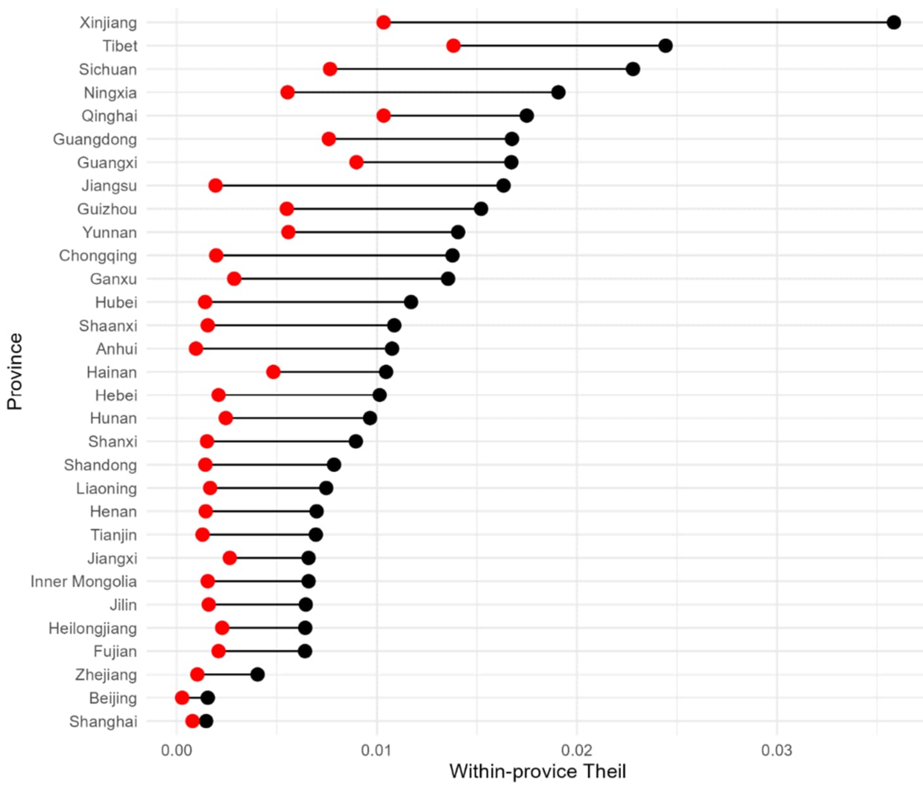
Table 3: Theil index and contribution rates of intra-province and inter-province differences

	Theil	$T_{between}$	T_{within}	Between contribution rate	Within contribution rate
Ideal	0.007	0.004	0.003	0.516	0.484
Intended	0.022	0.011	0.012	0.473	0.527

[10] Following the notation of OECD (2016), the Theil T index is defined as Theil = $\frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{\bar{y}}\ln(\frac{y_i}{\bar{y}})$, where N is the number of counties, y_i is the variable of interest (ideal or intended fertility) in the *i*th county, and $\bar{y}$ is the mean fertility ideal or intention cross all counties. The overall Theil index can be further decomposed into within- and between-province differences: Theil = $\frac{1}{N}\sum_{i=1}^{N} s_j \frac{y_{ij}}{\bar{y_j}}\ln(\frac{y_i}{\bar{y}}) + \frac{1}{M}\sum_{j=1}^{M} s_j \ln(\frac{y_j}{\bar{y}})$, where M is the number of provinces and s_j is the ratio between province *j*’s average and the national average of fertility ideals or intentions.

We further plotted the between-province Theil index by province (Figure 4). We find that provinces with high within-province Theil, such as Xinjiang, Tibet, Sichuan, and Ningxia, also have high fertility intentions and ideals. Places with low within-province Theil, such as Shanghai, Beijing, and Zhejiang, are also places with low fertility ideals and intentions. This may suggest a geographic convergence of fertility attitudes in China, with places transitioning from heterogeneously high to uniformly low fertility attitudes.

Figure 4: Within-province Theil index, by province



Note: The red dots denote fertility ideals; the black dots denote fertility intentions.

8.4 The validity of the MRP estimates

We evaluate the accuracy of the MRP estimates through internal and external validation. First, we evaluate the estimation accuracy and precision of the estimates with a split sample validation approach. Second, we compare the MRP estimates with independent national and regional surveys. For internal validation, we randomly split the data into testing and training sets. We calculate county-level averages for fertility ideals and intentions in the testing set, treating these as the ground truth. For the training set, we draw subsamples of varying sizes (10%, 25%, 50%, and 100%) to generate estimates. This process is repeated 200 times for each sample size, resulting in 800 random samples. Similar practices have been used in other studies to evaluate the performance of MRP estimates (Claassen and Traunmüller 2020; Lax and Phillips 2009).

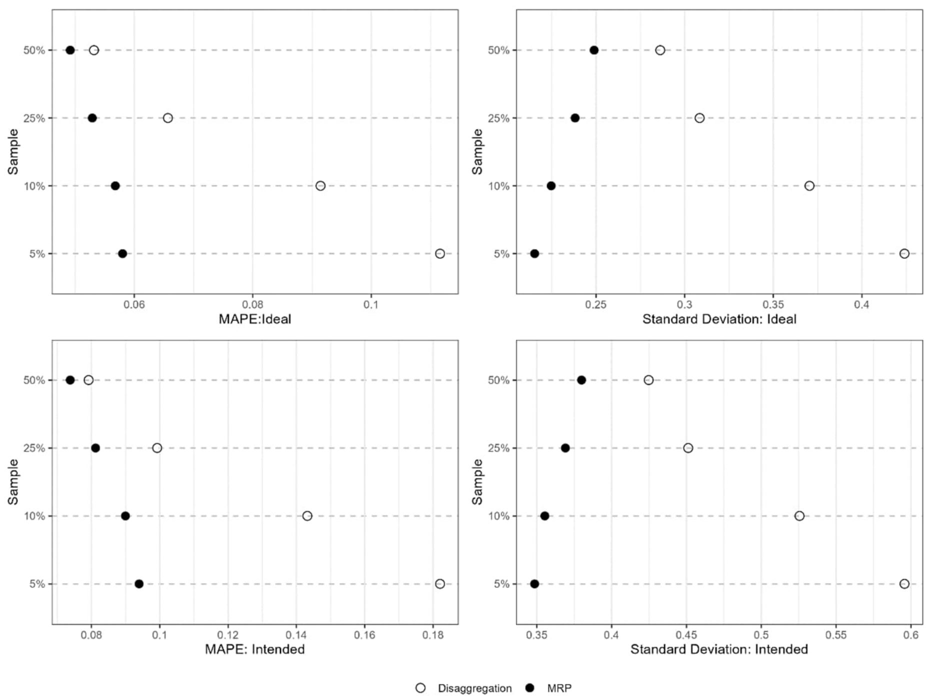
Using these random samples, we estimate fertility ideals and intentions using both MRP and simple disaggregation methods. We assess model performance primarily using the mean absolute percentage error (MAPE), as it is widely used and rescales errors to a common metric, making it more effective than other measures, such as mean absolute error (MAE) (Smith et al. 2006; Claassen and Traunmüller 2020).[11] The left panel in Figure 5 presents the MAPE for the different sampling scenarios and fertility attitude measures. The upper left panel displays the MAPE for ideal family size, whereas the lower left panel shows the MAPE for intended family size. It is no surprise that estimation accuracy increases as we increase the sample size in the training set. But more importantly, the results clearly show that MRP achieves better performance than simple disaggregation across all four sampling scenarios for both fertility ideal and intention, especially with smaller sample sizes. For example, MRP estimates for fertility ideal achieve 5% for MAPE when we use only 10% of the sample as the training data, whereas the MAPE for disaggregation estimates is as high as 11.3%. For fertility intention, the MAPE is only 9% for MRP estimates when we use only 10% of the sample as our training data, whereas the MAPE is as high as 18% for disaggregation. This suggests that using the full dataset in our final estimation will yield even lower estimation errors. Even when considering the absolute value of the MAPE, the MRP estimates remain within an acceptable error range of less than 10% in all validation scenarios, which is comparable to other small-area estimation methods (Chi and Voss 2011) and MRP-based attitude estimates (Claassen and Traunmüller 2020; Wiertz and Lim 2021).

---

[11] We calculate MAPE following the equation $MAPE = \frac{\sum_{l,q} \frac{|true - estimates|}{true}}{200 \times 2{,}439}$. The equation shows that in each sample, q, we calculate the absolute difference between the ground truth and the estimates (using both MRP and simple disaggregation) for each county, l, and then divide this difference by the ground truth value. For each sampling scenario, we calculate the average of these accuracy measures across 2,439 counties and 200 simulations.

We further calculate the standard errors of the estimates for both approaches (right panel of Figure 5). Both the upper and lower right panels show that the MRP estimates have lower variances across the 200 simulations in each sampling scenario. For fertility ideal estimates, increasing the sample size from the smallest to the largest raises the standard deviation of the MRP estimates only from 0.21 to 0.24, while the standard deviation of simple disaggregation estimates nearly doubles. A similar pattern is observed for fertility intention estimates. These findings suggest that MRP estimates are less sensitive to sample size and yield more stable estimates.

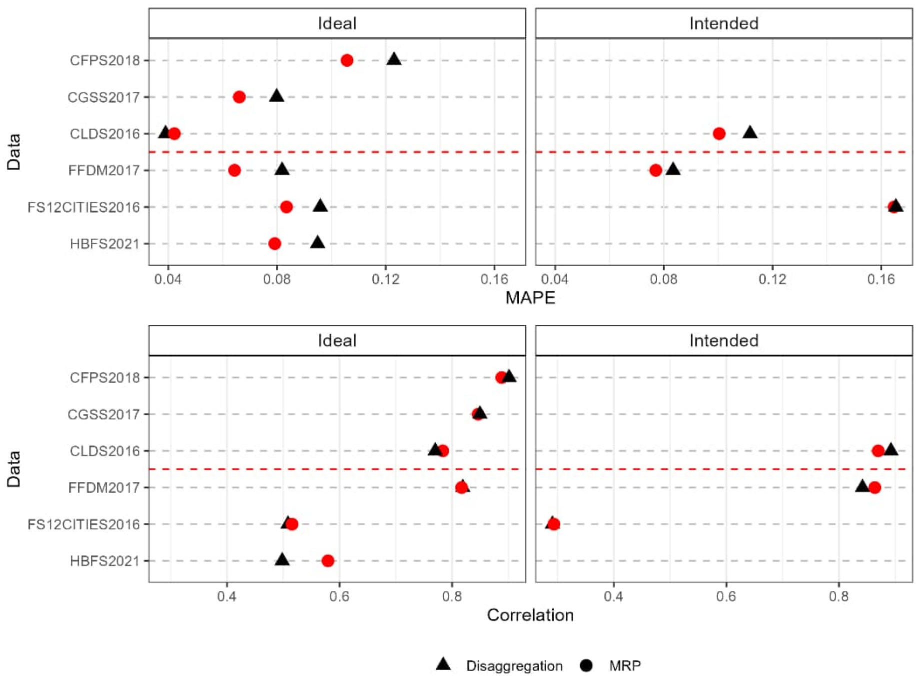
Figure 5: Internal validation: Mean absolute percentage error and standard deviation



Notes: The MAPE was calculated as the average absolute difference between the estimates (MRP and disaggregation) and testing data across 200 samples for each sample size. The standard deviation was calculated as the standard deviation of the estimates across counties and was averaged across 200 simulations.

For external validation, we benchmarked the MRP estimates and simple disaggregation estimates against six independent national and regional surveys. The detailed descriptions of these six surveys are in Section B of the appendix. For national surveys, we evaluate estimates at the provincial level. For regional surveys, comparisons were made at each survey's claimed representative level (prefecture or county). Figure 6 summarizes the results of the external validation. Overall, our estimation consistently outperforms simple disaggregation across all benchmarks. For example, when validated against three national surveys at the provincial level, the MRP estimates exhibit high correlation coefficients exceeding 0.8. This is comparable to the result of Lax and Phillips (2009), who also validate the MRP results at the state level. The MAPEs for our MRP estimates are all less than 10%, regardless of which external survey is used for validation.

Figure 6: External validation: Mean absolute percentage error and correlation against independent surveys



Notes: The dashed red horizontal line divides national surveys (above the line) from regional surveys (below the line). CFPS2018 = China Family Panel Studies (2018). CGSS2017 = Chinese General Social Survey (2017). CLDS2016 = China Labor Force Dynamics Survey (2016). FFDM2017 = Family Fertility Decision Making in China (2017). FS12CITES2016 = Fertility Survey of Twelve Cities in China (2016). HBFS2021 = Hubei Fertility Survey 2021. Detailed descriptions of these surveys are in Section B of the appendix. Correlation = Pearson correlation coefficient.

As expected, the estimation accuracy declines at smaller geographic scales, due to the inherent uncertainties of small-area estimation (Zhang and Chambers 2004), and the indicators of estimation accuracy vary across the external surveys. For example, the MAPE for MRP fertility intention estimates is 6% when benchmarked against Family Fertility Decision Making in China 2017 (FFDM2017) but rises to 16% for the Fertility Survey of Twelve Cities in China 2016 (FS12CITIES2016). We speculate that the notably lower performance relative to FS12CITIES2016 may stem from its data collection and cleaning procedures: Respondents self-reported their addresses, requiring researchers to manually identify their provinces and prefectures of residence, which could have potentially introduced additional errors. Despite this, the MRP estimates still outperform simple disaggregation at sub-provincial levels. In sum, both internal and external validation demonstrate that our MRP estimates reliably outperform benchmarks in nearly all scenarios.

9. Discussion and conclusion

Using data from the 2017 National Fertility Survey of China and the population census, we applied multilevel regression and post-stratification (MRP) to produce detailed county average as well as county-specific five-year age group fertility ideals and intentions in China. While individual fertility attitudes can change over the life course (Hayford 2009), county-level estimates allow us to understand how contextual factors such as culture, institutions, economic development, and policies shape reproductive attitudes beyond individual volatility. We discuss insights drawn from these estimates.

First, the estimates enable us to investigate the prevalence of high and low fertility ideals and intentions, thus helping address important questions regarding the nature of Chinese fertility attitudes. For example, past studies suggest that sub-replacement fertility ideals appear to be genuine in urban China (Chen and Gietel-Basten 2024). Our estimates suggest that this phenomenon might be geographically more widespread. Shifting to fertility intentions, around two-thirds of the counties have below-replacement fertility intentions, and 16% of counties have extremely low fertility intentions (less than 1.5 children). Concerns regarding Chinese people's dwindling fertility intentions have been growing. Our small-area estimates provide the first comprehensive confirmation of such concerns.

With the Theil index, we further identified a geographic convergence to low fertility preferences. Provinces with high average fertility ideals and intentions are driven by several outlier counties with exceptionally high values, while provinces with low fertility preferences are concentrated with counties with uniformly low values. It remains to be seen what will happen after the entire nation arrives at a low level. However, we emphasize that only detailed county average statistics allow us to uncover this pattern.

Second, moving beyond prevalences, the estimates also enable us to identify spatial patterns of fertility ideals and intentions that are otherwise dismissed by national statistics. We find that while the two concepts are closely related, fertility intentions tend to display coastal–inland gradients, and fertility ideals follow a north–south divide. While the coastal–inland differences in fertility intention are well-known, the north–south divide in fertility ideals is much less discussed in the existing literature. Yet this pattern underscores the role of culture in influencing fertility attitudes. Prior studies in China have underscored cultural influences on fertility behavior and attitudes (Zhang 2022; Zhang and Li 2017; Zou et al. 2024). Our small-area estimates offer further compelling evidence of how geographically structured cultural factors shape fertility attitudes.

Another unexpected spatial pattern emerges when we examine the gap between fertility ideals and intentions. We find that 19% of counties have fertility intentions that exceed ideals. These counties are mainly located in less socioeconomically developed places with historically high fertility levels. Existing research has mostly focused on cases where fertility ideals exceed intentions (Brinton et al. 2018). However, our estimates suggest that even in a commonly known low fertility setting such as China, there are still places where people intend to have more children than the social norm suggests. In China, worldviews promoting small families are widespread (Thornton and Xie 2016). Prior empirical studies indicate that, even in the most impoverished rural areas of China, people still consider small families a good thing (Lai and Thornton 2015). Our estimates show that in some places, residents may feel pressure to conform to their surroundings, even amid the weight of development idealism on expressed fertility ideals.

Last, our study demonstrates how geographic scale matters in understanding variations of fertility attitudes. We find that within-province heterogeneity contributes as much to the overall variations in fertility attitudes as between-province heterogeneity. This finding corresponds with existing research that argues that spatial disparities in China are particularly stark between urban centers and their periphery counties within provinces rather than across broader provincial or regional lines (Liu et al. 2015; Wei and Ye 2009). This has important implications for future research, indicating that studies at small geographic scales are just as crucial as broader regional analyses.

In addition to our substantive findings, our study also demonstrates the utility of MRP for small-area estimations of demographic attitudes. Despite its popularity in other fields, demographers have not fully embraced its potential. In demonstrating the effectiveness of MRP for estimating fertility attitudes, we advance small-area estimation techniques. In particular, while demographers engaging small-area estimation have long speculated that sophisticated model-based estimations may not perform better than simple ones (Chi 2009; Keyfitz 1982), studies that evaluate the performance of MRP highlight the value of local knowledge (Buttice and Highton 2013; Downes and Carlin 2020; Warshaw and Rodden 2012). Although our analysis does not comprehensively evaluate the factors contributing to MRP estimation accuracy, it strongly suggests MRP's superior performance, particularly in areas with a limited sample size.

This study has several limitations. First, our survey data focus only on women of reproductive age. As a result, our estimates are restricted to this specific subgroup. This reflects a broader issue in fertility research in China, as most fertility surveys focus exclusively on women of reproductive age. The rationale is that women are the primary actors in childbearing, so their attitudes were considered most consequential. However, growing evidence indicates that childbearing decisions are complex. The fertility preferences of other family members also influence fertility decisions (Qian and Jin 2018; Testa 2012; Thomson 1997). Studies have shown that in places where women have less autonomy, they may not have the final say about a fertility decision; their husbands' or other relatives' fertility preferences may exert a stronger influence on actual fertility behavior (Doepke and Tertilt 2018; Morgan and Niraula 1995). So our estimates may be less predictive indicators of actual fertility behavior in places where females have lower reproductive autonomy. Future studies should include a broader demographic group. Second, existing research suggests that non-numerical responses to fertility attitudes, such as "don't know" or "not sure," hold important value, as they may reflect societal uncertainty (Hayford and Agadjanian 2011). In our study, we excluded these non-numerical responses (0.6% of the sample) to focus only on numerical expressions of fertility ideals and intentions. Exploring the meanings and implications of the non-numerical responses to fertility attitudes is beyond the scope of this study. Future studies could investigate the meanings, predictors, and behavioral correlates of non-numerical fertility responses to deepen understanding of fertility-related ambivalence, particularly in low fertility contexts such as China. Third, although we estimate fertility ideals and intentions by age group at the county level, we do not consider other important variations. For example, it would be ideal to estimate county-level fertility ideals and intentions by birth parity. However, this requires data on county-level parity distribution for post-stratification purposes, which are currently unavailable. Accordingly, future studies may expand upon the current one if such data become available.

10. Acknowledgments

We thank Tianji Cai, Hao Dong, Ze Hong, Airan Liu, Jilei Wu, Yu Xie, Feng Yang, Jia Yu, and Yi Zhou, for their valuable feedback at various stages of this project. We also extend our sincere thanks to Zhilei Shi for generously sharing the Hubei Fertility Survey with us, and to Yun Wang for his exceptional data analysis on this survey. Additionally, we thank Wanchen Xiao for her excellent mapping assistance. Our thanks extend to the participants of the seminars at the following institutions and conferences: the Paul and Marcia Wythes Center on Contemporary China at Princeton University, the Department of Sociology at the University of Macau, the Center for Social Research at Peking University, and the RC28 2024 spring meeting. Donghui Wang receives support from the National Social Science Foundation of China (Grant Number: 21CRK001).

References

Ajzen, I. and Klobas, J. (2013). Fertility intentions: An approach based on the theory of planned behavior. *Demographic Research* 29(8): 203–232. doi:10.4054/DemRes.2013.29.8.

Attané, I. (2001). Chinese fertility on the eve of the 21st century: Fact and uncertainty. *Population: An English Selection* 56(2): 71–100. doi:10.3917/popu.p2001.13n2.0100.

Attané, I. (2016). Second child decisions in China. *Population and Development Review* 42(3): 519–536. doi:10.1111/j.1728-4457.2016.00151.x.

Bachrach, C.A. and Morgan, S.P. (2013). A cognitive–social model of fertility intentions. *Population and Development Review* 39(3): 459–485. doi:10.1111/j.1728-4457.2013.00612.x.

Balbo, N., Billari, F.C., and Mills, M. (2013). Fertility in advanced societies: A review of research. *European Journal of Population* 29(1): 1–38. doi:10.1007/s10680-012-9277-y.

Basten, S. and Jiang, Q. (2014). China's family planning policies: Recent reforms and future prospects. *Studies in Family Planning* 45(4): 493–509. doi:10.1111/j.1728-4465.2014.00003.x.

Beaujouan, E. and Berghammer, C. (2019). The gap between lifetime fertility intentions and completed fertility in Europe and the United States: A cohort approach. *Population Research and Policy Review* 38: 507–535. doi:10.1007/s11113-019-09516-3.

Behrman, J. and Weitzman, A. (2024). State-level immigrant policies and ideal family size in the United States. *Population and Development Review* 50(2): 375–401. doi:10.1111/padr.12621.

Billingsley, S. and Ferrarini, T. (2014). Family policy and fertility intentions in 21 European countries. *Journal of Marriage and Family* 76(2): 428–445. doi:10.1111/jomf.12097.

Blake, J. (1966). Ideal family size among white Americans: A quarter of a century's evidence. *Demography* 3(1): 154–173. doi:10.2307/2060069.

Bongaarts, J. (2001). Fertility and reproductive preferences in post-transitional societies. *Population and Development Review* 27: 260–281.

Brinton, M.C., Bueno, X., Oláh, L., and Hellum, M. (2018). Postindustrial fertility ideals, intentions, and gender inequality: A comparative qualitative analysis. *Population and Development Review* 44(2): 281–309. doi:10.1111/padr.12128.

Buttice, M.K. and Highton, B. (2013). How does multilevel regression and poststratification perform with conventional national surveys? *Political Analysis* 21(4): 449–467. doi:10.1093/pan/mpt017.

Casterline, J.B. and Sinding, S.W. (2000). Unmet need for family planning in developing countries and implications for population policy. *Population and Development Review* 26(4): 691–723. doi:10.1111/j.1728-4457.2000.00691.x.

Chen, M. and Yip, P.S.F. (2017). The discrepancy between ideal and actual parity in Hong Kong: Fertility desire, intention, and behavior. *Population Research and Policy Review* 36: 583–605. doi:10.1007/s11113-017-9433-5.

Chen, S. and Gietel-Basten, S. (2024). How genuine are sub-replacement ideal family sizes in urban China? *Population Studies* 78(2): 305–324. doi:10.1080/00324728.2023.2194670.

Chi, G. (2009). Can knowledge improve population forecasts at subcounty levels? *Demography* 46(2): 405–427. doi:10.1353/dem.0.0059.

Chi, G. and Voss, P.R. (2011). Small-area population forecasting: Borrowing strength across space and time. *Population, Space and Place* 17(5): 505–520. doi:10.1002/psp.617.

China Daily (2021). A laudable local effort to raise the fertility rate. Beijing, China https://www.chinadaily.com.cn/a/202107/30/WS61033740a310efa1bd66561b.html.

Claassen, C. and Traunmüller, R. (2020). Improving and validating survey estimates of religious demography using Bayesian multilevel models and poststratification. *Sociological Methods and Research* 49(3): 603–636. doi:10.1177/0049124118769086.

Cleland, J. and Wilson, C. (1987). Demand theories of the fertility transition: An iconoclastic view. *Population Studies* 41(1): 5–30. doi:10.1080/0032472031000142516.

Doepke, M. and Tertilt, M. (2018). Women's empowerment, the gender gap in desired fertility, and fertility outcomes in developing countries. *AEA Papers and Proceedings* 108: 358–362. doi:10.1257/pandp.20181085.

Downes, M. and Carlin, J.B. (2020). Multilevel regression and poststratification as a modeling approach for estimating population quantities in large population health studies: A simulation study. *Biometrical Journal* 62(2): 479–491. doi:10.1002/bimj.201900023.

Entwisle, B. (2007). Putting people into place. *Demography* 44(4): 687–703. doi:10.1353/dem.2007.0045.

Fei, H.-T. (1946). Peasantry and gentry: An interpretation of Chinese social structure and its changes. *American Journal of Sociology* 52(1): 1–17. doi:10.1086/219914.

Freedman, M. (1958). *Lineage organisation in South-Eastern China*. London: The Athlone Press. doi:10.4324/9781003135296.

Gelman, A. and Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press. doi:10.32614/CRAN.package.arm.

Gelman, A. and Little, T.C. (1997). Poststratification into many categories using hierarchical logistic regression. *Survey Methodology* 23: 127.

Goldstein, J., Lutz, W., and Testa, M.R. (2003). The emergence of sub-replacement family size ideals in Europe. *Population Research and Policy Review* 22: 479–496. doi:10.1023/B:POPU.0000020962.80895.4a.

Hagewen, K.J. and Morgan, S.P. (2005). Intended and ideal family size in the United States, 1970–2002. *Population and Development Review* 31(3): 507–527. doi:10.1111/j.1728-4457.2005.00081.x.

Hartnett, C.S. and Gemmill, A. (2020). Recent trends in U.S. childbearing intentions. *Demography* 57(6): 2035–2045. doi:10.1007/s13524-020-00929-w.

Hayford, S.R. (2009). The evolution of fertility expectations over the life course. *Demography* 46(4): 765–783. doi:10.1353/dem.0.0073.

Hayford, S.R. and Agadjanian, V. (2011). Uncertain future, non-numeric preferences, and the fertility transition: A case study of rural Mozambique. *Etude de La Population Africaine / African Population Studies* 25(2): 419–439. doi:10.11564/25-2-239.

Hermalin, A.I. and Liu, X. (1990). Gauging the validity of responses to questions on family size preferences in China. *Population and Development Review* 16(2): 337–354. doi:10.2307/1971594.

Hou, J. (2015). Changes in the Chinese population's fertility intentions: 1980–2011. *Social Sciences in China* 36(1): 46–63. doi:10.1080/02529203.2015.1001482.

Hu, H. C. (1948). The Common Descent Group in China and its functions. The Viking Fund, Inc.

Jiang, Q., Li, Y., and Sánchez-Barricarte, J.J. (2016). Fertility intention, son preference, and second childbirth: Survey findings from Shaanxi Province of China. *Social Indicators Research* 125: 935–953. doi:10.1007/s11205-015-0875-z.

Jin, Y., Hu, W., and Wang, D. (2023). Fertility transition of Han and ethnic minorities in China: A tale of convergence and variation. *Chinese Journal of Sociology* 9(4): 522–552. doi:10.1177/2057150X231207916.

Jin, Y., Song, J., and Chen, W. (2016). Women's fertility preference and intention in urban China: An empirical study on the nationwide two-child policy (In Chinese). *Population Research* 40(6): 22–37.

Keyfitz, N. (1982). Can knowledge improve forecasts? *Population and Development Review* 8(4): 729–751. doi:10.2307/1972470.

Lai, Q. and Thornton, A. (2015). The making of family values: Developmental idealism in Gansu, China. *Social Science Research* 51: 174–188. doi:10.1016/j.ssresearch.2014.09.012.

Lax, J. R. and Phillips, J. H. (2009). How should we estimate public opinion in the states? *American Journal of Political Science* 53(1): 107–121. doi:10.1111/j.1540-5907.2008.00360.x.

Lee, R.D. (1980). Aiming at a moving target: Period fertility and changing reproductive goals. *Population Studies* 34(2): 205–226. doi:10.1080/00324728.1980.10410385.

Lesthaeghe, R. (2010). The unfolding story of the second demographic transition. *Population and Development Review* 36(2): 211–251. doi:10.1111/j.1728-4457.2010.00328.x.

Liefbroer, A.C. (2009). Changes in family size intentions across young adulthood: A life-course perspective. *European Journal of Population* 25: 363–386. doi:10.1007/s10680-008-9173-7.

Liu, T., Qi, Y., Cao, G., and Liu, H. (2015). Spatial patterns, driving forces, and urbanization effects of China's internal migration: County-level analysis based on the 2000 and 2010 censuses. *Journal of Geographical Sciences* 25: 236–256. doi:10.1007/s11442-015-1165-z.

Lobao, L. (2004). Continuity and change in place stratification: Spatial inequality and middle-range territorial units. *Rural Sociology* 69(1): 1–30. doi:10.1526/003601104322919883.

Lutz, W., Skirbekk, V., and Testa, M. R. (2006). The low-fertility trap hypothesis: Forces that may lead to further postponement and fewer births in Europe. *Vienna Yearbook of Population Research* 4 167–192. doi:10.1553/populationyearbook2006s167.

Merli, M.G. and Morgan, S.P. (2011). Below replacement fertility preferences in Shanghai. *Population* 66(3–4): 519–542. doi:10.3917/pope.1103.0519.

Merli, M.G. and Smith, H.L. (2002). Has the Chinese family planning policy been successful in changing fertility preferences? *Demography* 39(3): 557–572. doi:10.1353/dem.2002.0029.

Miller, W.B. (2011). Differences between fertility desires and intentions: Implications for theory, research and policy. *Vienna Yearbook of Population Research* SI Reproductive decision-making: 75–98. doi:10.1553/populationyearbook2011s75.

Morgan, S.P. and Niraula, B.B. (1995). Gender inequality and fertility in two Nepali villages. *Population and Development Review* 21(3): 541–561. doi:10.2307/2137749.

Nie, Y. and Wyman, R. (2005). The one-child policy in Shanghai: Acceptance and internalization. *Population and Development Review* 31(2): 313–336. doi:10.1111/j.1728-4457.2005.00067.x.

OECD (2016). Indexes and estimation techniques. In: *OECD regions at a glance*. Paris: OECD Publishing: 174–177. doi:10.1787/reg_glance-2016-50-en.

Pacheco, J. (2011). Using national surveys to measure dynamic U.S. state public opinion: A guideline for scholars and an application. *State Politics and Policy Quarterly* 11(4): 415–439. doi:10.1177/1532440011419287.

Peng, Y. (2010). When formal laws and informal norms collide: Lineage networks versus birth control policy in China. *American Journal of Sociology* 116(3): 770–805. doi:10.1086/657102.

Pfeffermann, D. (2013). New important developments in small area estimation. *Statistical Science* 28(1): 40–68. doi:10.1214/12-STS395.

Philipov, D. and Bernardi, L. (2011). Concepts and operationalisation of reproductive decisions implementation in Austria, Germany and Switzerland. *Comparative Population Studies* 36(2–3): 495–530. doi:10.12765/CPoS-2011-14.

Qian, Y. and Jin, Y. (2018). Women's fertility autonomy in urban China: The role of couple dynamics under the universal two-child policy. *Chinese Sociological Review* 50(3): 275–309. doi:10.1080/21620555.2018.1428895.

Schoen, R., Astone, N.M., Kim, Y.J., Nathanson, C.A., and Fields, J.M. (1999). Do fertility intentions affect fertility behavior? *Journal of Marriage and Family* 61(3): 790–799. doi:10.2307/353578.

Selb, P. and Munzert, S. (2017). Estimating constituency preferences from sparse survey data using auxiliary geographic information. *Political Analysis* 19(4): 455–470. doi:10.1093/pan/mpr034.

Short, S.E. and Fengying, Z. (1998). Looking locally at China's one-child policy. *Studies in Family Planning* 29(4): 373. doi:10.2307/172250.

Sobotka, T. and Beaujouan, É. (2014). Two is best? The persistence of a two-child family ideal in Europe. *Population and Development Review* 40(3): 391–419. doi:10.1111/j.1728-4457.2014.00691.x.

Testa, M.R. (2012). Couple disagreement about short-term fertility desires in Austria: Effects on intentions and contraceptive behaviour. *Demographic Research* 26(3): 63–98. doi:10.4054/DemRes.2012.26.3.

Thomson, E. (1997). Couple childbearing desires, intentions, and births. *Demography* 34(3): 343–354. doi:10.2307/3038288.

Thornton, A. and Xie, Y. (2016). Developmental idealism in China. *Chinese Journal of Sociology* 2(4): 483–496. doi:10.1177/2057150X16670835.

Thornton, A., Binstock, G., Yount, K.M., Abbsai-Shavazi, M., Ghimire, D., and Xie, Y. (2012). International fertility change: New data and insights from the developmental idealism framework. *Demography* 49(2): 677–698. doi:10.1007/s13524-012-0097-9.

Trent, R.B. (1980). Evidence bearing on the construct validity of 'ideal family size.' *Population and Environment* 3(3–4): 309–327. doi:10.1007/BF01255345.

Wang, D. and Chi, G. (2017). Different places, different stories: A study of the spatial heterogeneity of county-level fertility in China. *Demographic Research* 37(16): 493–526. doi:10.4054/DemRes.2017.37.16.

Wang, W., Rothschild, D., Goel, S., and Gelman, A. (2015). Forecasting elections with non-representative polls. *International Journal of Forecasting* 31(3): 980–991. doi:10.1016/j.ijforecast.2014.06.001.

Warshaw, C. and Rodden, J. (2012). How should we measure district-level public opinion on individual issues? *Journal of Politics* 74(1): 203–219. doi:10.1017/S0022381611001204.

Wei, Y.H.D. and Ye, X. (2009). Beyond convergence: Space, scale, and regional inequality in China. *Tijdschrift voor Economische en Sociale Geografie* 100(1): 59–80. doi:10.1111/j.1467-9663.2009.00507.x.

Whyte, M.K. and Gu, S.Z. (1987). Popular response to China's fertility transition. *Population and Development Review* 13(3): 471–493. doi:10.2307/1973135.

Wiertz, D. and Lim, C. (2021). The rise of the Nones across the United States, 1973 to 2018: State-level trends of religious affiliation and participation in the General Social Survey. *Sociological Science* 8(21): 429–454. doi:10.15195/v8.a21.

Xie, Y. and Zhou, X. (2014). Income inequality in today's China. *Proceedings of the National Academy of Sciences* 111(19): 6928–6933. doi:10.1073/pnas.1403158111.

Yang, X.Y. and Schieman, S. (2024). Racial disparities in death rates and death incidences in Xinjiang: A study of multilevel ecological mechanisms. Social Science and Medicine 340: 116405. doi:10.1016/j.socscimed.2023.116405.

Yeatman, S., Trinitapoli, J., and Garver, S. (2020). The enduring case for fertility desires. *Demography* 57(6): 2047–2056. doi:10.1007/s13524-020-00921-4.

Zhai, Z. and Jin, G. (2023). China's family planning policy and fertility transition. *Chinese Journal of Sociology* 9(4): 479–496. doi:10.1177/2057150X231205773.

Zhang, C. and Li, T. (2017). Culture, fertility, and the socioeconomic status of women. China *Economic Review* 45: 279–288. doi:10.1016/j.chieco.2016.07.012.

Zhang, L. (2022). Patrilineality, fertility, and women's income: Evidence from family lineage in China. *China Economic Review* 74: 101805. doi:10.1016/j.chieco.2022.101805.

Zhang, L.-C. and Chambers, R.L. (2004). Small area estimates for cross-classifications. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 66(2): 479–496. doi:10.1111/j.1369-7412.2004.05266.x.

Zhang, X., Holt, J.B., Yun, S., Lu, H., Greenlund, K.J., and Croft, J.B. (2015). Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the behavioral risk factor surveillance system. *American Journal of Epidemiology* 182(2): 127–137. doi:10.1093/aje/kwv002.

Zheng, Z., Cai, Y., Wang, F., and Gu, B. (2009). Below-replacement fertility and childbearing intention in Jiangsu Province, China. *Asian Population Studies* 5(3): 329–347. doi:10.1080/17441730903351701.

Zhuang, Y., Yang, S., Qi, J., Li, B., Zhao, X., and Wang, Z. (2019). China fertility survey 2017: Design and implementation. *China Population and Development Studies* 2: 259–271. doi:10.1007/s42379-018-0018-8.

Zou, W., Ma, R., Ma, Z., and Zheng, P. (2024). How informal institutions matter: Clan culture and fertility in China. *Review of Economics of the Household*. doi:10.1007/s11150-024-09737-2.

Appendix

Section A. Additional tables and figures

Table A-1: Summary statistics of the ideal and intended number of children

		As a categorical variable (%)			
	Mean	Zero	One	Two	Three and above
Ideal number of children					
Total	2.00	0.5%	11.2%	78.7%	9.6%
By age group					
15–19	1.86	2.2%	14.8%	78.6%	4.4%
20–24	1.9	1.4%	13.6%	79.1%	5.9%
25–29	1.96	0.4%	11.9%	80.6%	7.2%
30–34	1.99	0.3%	11.3%	79.4%	8.9%
34–39	2.00	0.3%	11.8%	78.4%	9.6%
40–44	2.03	0.3%	10.7%	78.2%	10.9%
45–49	2.08	0.2%	9.1%	77.7%	13.0%
Intended number of children					
Total	1.81	1.4%	31.8%	54.9%	11.9%
By age group					
15–19	1.66	6.9%	23.5%	66.8%	2.8%
20–24	1.77	3.7%	22.6%	67.4%	6.4%
25–29	1.85	1.0%	24.0%	65.3%	9.7%
30–34	1.87	0.6%	26.7%	60.4%	12.2%
34–39	1.83	0.6%	33.2%	53.2%	13.0%
40–44	1.78	0.9%	38.9%	46.5%	13.8%
45–49	1.80	0.8%	38.7%	45.4%	15.1%

Notes: This table provides summary statistics of the survey. The summary statistics are weighted with the original survey weights. For women of reproductive age, the national averages for the ideal and intended number of children are 2 and 1.8, respectively. These statistics are higher than those in previous studies (e.g., Chen and Basten 2023; Feng 2017), possibly due to differences in survey wording and target populations. When viewing fertility ideals and intention as categorical variables, we found that a large proportion of respondents (79%) considered two children per family to be ideal; however, only approximately half of the respondents intended to have two children. Meanwhile, a small percentage of respondents considered having no children to be ideal (0.5%) or intended not to have children (1.4%), suggesting a prevailing social norm of universal childbearing. Further breaking the data down by age group reveals a slight upward trend in the ideal and intended number of children with increasing age.

Table A-2: Model selection procedure

Model	Model description	Ideal number of children AIC	Ideal number of children BIC	Ideal number of children RML	Intended number of children AIC	Intended number of children BIC	Intended number of children RML
M1	county-level random intercept only	236783.8	236813.8	236777.8	329540.9	329570.8	329534.9
M2	M1 + age group random intercept	233316.7	233356.7	233308.7	327709.9	327749.8	327701.9
M3	M2 + prefecture and province-level random intercept	231642.9	231702.8	231630.9	325936.3	325996.2	325924.3
M4	**M3 + one fixed effect (livebirth2010)**	231105.5	231175.4	231091.5	324762.4	324832.3	324748.4
M5	M4 + other seven contextual fixed effects[a]	231062.6	231202.4	231034.6	324700.4	324840.2	324672.4
M6	M3 + other seven contextual fixed effects	231213.6	231323.4	231191.6	324960.5	325070.4	324938.5

Notes: AIC = Akaike information criterion, BIC = Bayesian information criterion, RML = restricted maximum likelihood. We experiment with different model specifications and select the final model based on the best fit and interpretation. While Model 5 appears to have a slightly better performance, it is at the cost of adding seven additional contextual variables. Therefore we choose Model 4, with only one contextual variable, livebirth2010. We fit all models using the restricted maximum likelihood implemented in the lmer function of the lme4 package in R.

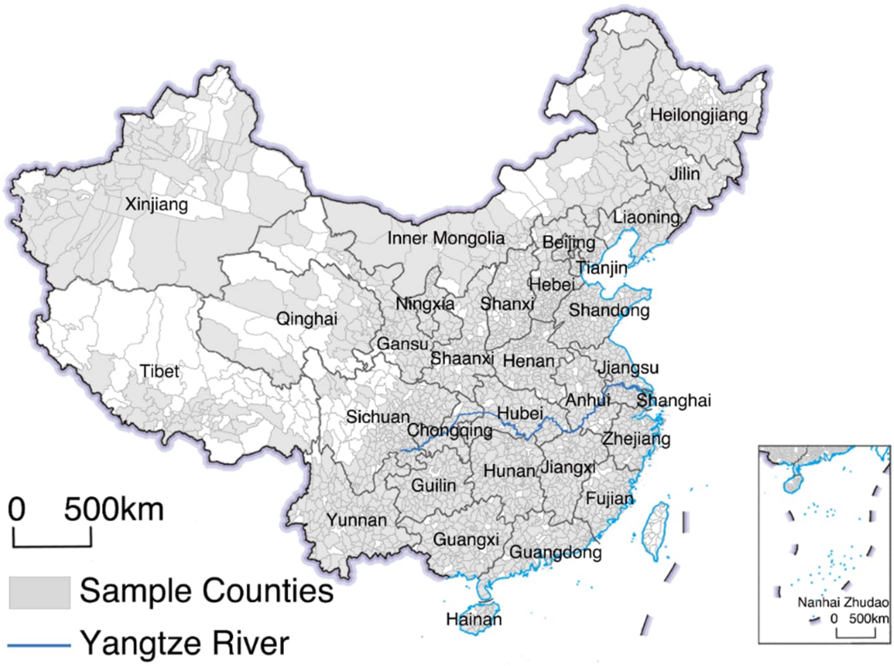
[a]Other fixed effects include average educational level, percentage of ethnic minority groups, sex ratio at birth, percentage of nonagricultural *hukou* registration, deprivation score, percentage of industrial sector employment, and percentage of service sector employment. All these variables are measured for 2010 at the county level.

Table A-3: Comparison of 2010 county-level socioeconomic and demographic characteristics between the two types of counties

	Type A: Intended < = Ideal (N = 1985)	**Type B: Intended > Ideal (N = 454)**	**T test p-value**
Average live birth	1.35 (0.282)	1.70 (0.203)	< 0.001
Average education	9.00 (1.30)	7.79 (1.08)	< 0.001
% ethnic population	11.6 (23.4)	26.4 (35.9)	< 0.001
Sex ratio at birth (boys/girls)×100	115 (10.5)	118 (10.5)	< 0.001
% nonagricultural *hukou*	32.0 (23.8)	13.9 (8.37)	< 0.001
% industrial occupation	22.9 (14.9)	11.7 (8.76)	< 0.001
% service occupation	30.1 (18.2)	15.3 (7.34)	< 0.001
Deprivation score	–0.367 (1.46)	1.17 (1.45)	< 0.001

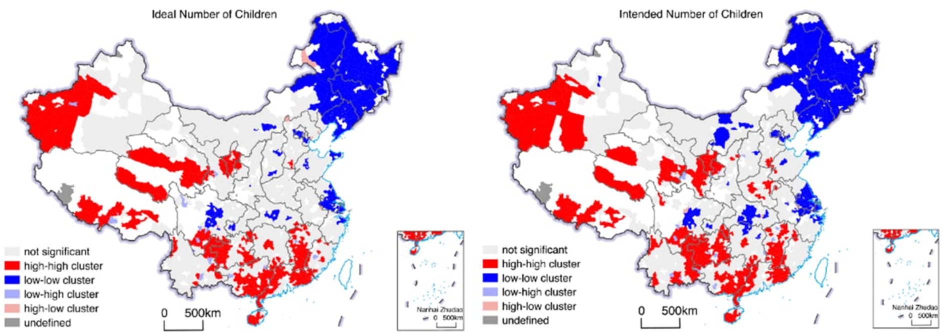
Notes: The deprivation score is a composite indicator reflecting the percentage of households lacking basic amenities (kitchen, tap water, toilet, or shower) within each county. This score was constructed using principal component analysis and demonstrates good internal consistency (Cronbach’s alpha = 0.838). All variables are measured based on the 2010 county-level population census.

Figure A-1: Map of the People's Republic of China



Note: Shaded areas denote counties where survey data are available. The Yangtze River is the traditional north–south divide in China.

Figure A-2: Local cluster map



Note: The spatial weights used to identify local clusters and outliers are first-order Queen's continuities. High-high clusters indicate places where high values of ideal fertility (intention) are concentrated. Low-low clusters indicate places where low values of ideal fertility (intention) are concentrated. High-low outliers indicate places with high values of ideal fertility (intention) surrounded by neighboring areas with low values. Low-high outliers indicate places with low ideal fertility (intention) values surrounded by neighboring areas with high values. NA denotes spatially isolated counties (where no spatial weights are assigned).

Figure A-3: County-level estimates of the ideal number of children, by age subgroup

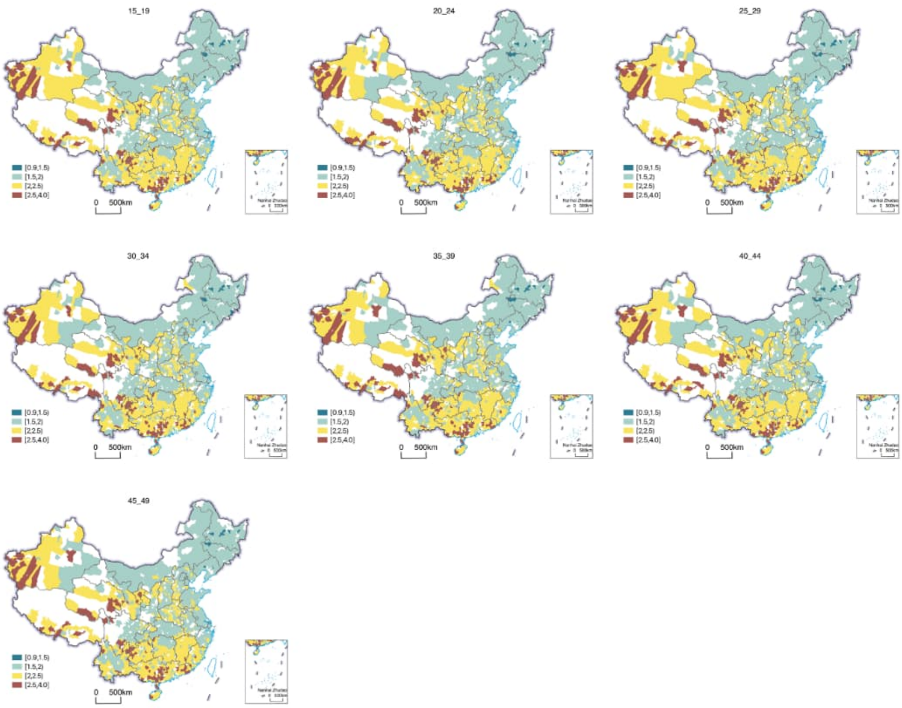
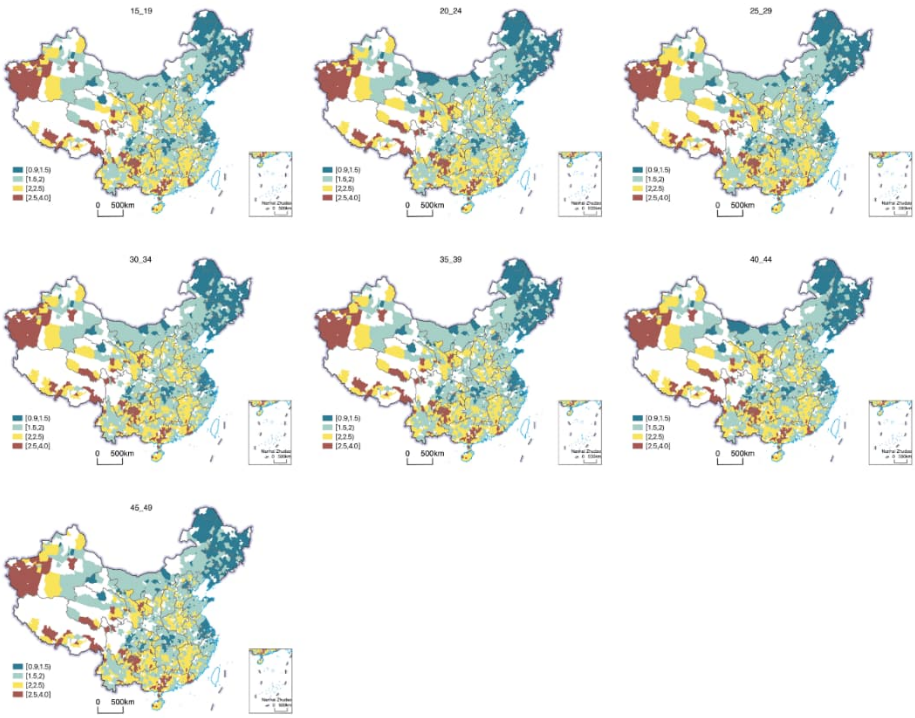


Figure A-4: County-level estimates of the intended number of children, by age subgroup



Section B. Description of Independent Surveys for External Validation Purposes

China Family Panel Studies (2018)

China Family Panel Studies (CFPS) is a nationally representative, biennial longitudinal survey of Chinese communities, families, and individuals launched in 2010 by the Institute of Social Science Survey (ISSS) of Peking University. The CFPS is designed to collect individual-, family-, and community-level longitudinal data in contemporary China. The survey adopts a multi-stage probability sampling framework.

The analytical sample for validation purposes is drawn from the 2018 wave. The sample included women of reproductive age (15–49), which yielded 8,389 observations. The survey asked respondents, “How many children do you think is ideal”? We treat this variable as our measurement of the fertility ideal. Individual cross-sectional weights were applied when estimating the province-level fertility ideal.

Official website: https://www.isss.pku.edu.cn/cfps/index.htm

Reference: Xie, Y. and Lu, X. (2015). The sampling design of the China Family Panel Studies (CFPS). *Chinese Journal of Sociology* 1(4): 471–484.

Chinese General Social Survey (2017)

The Chinese General Social Survey (CGSS) is the earliest nationally representative survey project of China that aims to systematically track social change, public attitudes, and demographic trends across the country. The project is administered by the National Survey Research Center at Renmin University. The survey adopted a multi-stage stratified sampling framework.

The analytical sample is drawn from the 2017 survey, which included women of reproductive age (15–49). The final data yielded 3,007 observations. The survey asked respondents, “If there were no policy restrictions, how many children would you like to have?” We treat this variable as a measurement of fertility intention. Individual weights were applied when estimating province-level fertility attitudes.

Official website: http://cgss.ruc.edu.cn/English/Home.htm

China Labor Force Dynamic Survey (2016)

The China Labor Force Dynamic Survey has established a comprehensive database, with the labor force as the survey object, through the biennial follow-up survey of Chinese urban and rural villages. It contains the tracking and cross-sectional data of the labor force at three levels, individual, family and community, to provide high-quality basic data for evidence-oriented theoretical and policy research.

The analytical sample is drawn from the 2016 survey and includes women of reproductive age (15–49; N = 5,170). The survey asked respondents, "How many children do you want to have" and "How many children do you think is ideal for a family?" We treat the first question as a measurement of fertility intention and the second as a measurement of fertility ideal. Individual weights were applied to ensure representativeness.

Data access website:
https://www.scidb.cn/en/detail?dataSetId=36d6d9d24afc4bc8a5ce0da5feaf22ff

Reference: Wang, J., Zhou, Y., and Liu, S. (2017). China labor-force dynamics survey: Design and practice. *Chinese Sociological Dialogue* 2(3–4): 83–97.

Fertility Survey of Twelve Cities in China (2016)

The Fertility Survey of Twelve Cities in China (2016) (FS12CITIES2016) was conducted in April 2016. The primary goal of the survey was to understand the fertility decision making of women of reproductive age as well as their spouses in urban China. The survey was part of the large research project "Study on the Fertility Decision-Making Mechanisms of Chinese Families," initiated by the Center for Population and Development Studies at Renmin University. The sampling cities were selected based on regional fertility levels, sex ratio at birth, geographical location, population size, and economic development. The survey ultimately included 12 cities across six provinces. These cities are Guangzhou and Jieyang in Guangdong Province; Chengdu and Luzhou in Sichuan Province; Wuhan and Jingzhou in Hubei Province; Jinan and Jining in Shandong Province; Hangzhou and Lishui in Zhejiang Province; and Shenyang and Chaoyang in Liaoning Province.

For each sampled city, a multi-stage PPS sampling method was employed, identifying three districts or counties in each city, two streets or towns in each district or county, four to ten neighborhood or village committees in each street/town, and eight to ten households randomly selected from each neighborhood/village committee. Five hundred households in each city were surveyed, resulting in a final sample of 6,000 households. Considering that urban areas are the primary regions where the target population eligible for the universal two child policy resides, urban residents were over-sampled, with 86% of the final sample being urban and 14% rural. The research subjects were married women and their spouses who had lived in the selected neighborhood or village committee for more than six months prior to the survey and were born between March 1, 1966, and March 1, 1996. The final valid sample size for women was 5,972.

Reference: Jin, Y., Jian, S., and Wei, C. (2016). Comprehensive two-child policy background under China Urban Women's Fertility Preference and Fertility Plan. (In Chinese) *Population Research* 40(6): 22.

Family Fertility Decision Making in China (2017)

Family Fertility Decision Making in China (FFDM2017) was conducted in 2017 and aimed to understand households' fertility decision-making process after the implementation of the universal two child policy. The survey was a follow-up to the Fertility Survey of Twelve Cities in China (2016), thus adopting similar survey question designs. Compared to FS12CITIES2016, FFDM2017 had much smaller geographic coverage: It was conducted in five cities across three provinces: Beijing; Xi'an and Yan'an in Shaanxi Province; and Zhengzhou and Zhumadian in Henan Province.

The study employs a PPS sampling method. Three districts or counties were identified in two cities, Shaanxi and Henan, alongside the four districts in Beijing. In each district or county, two streets or towns were selected, and in each street/town, four to ten neighborhood or village committees were determined, with eight to ten households randomly selected from each neighborhood/village committee. The subjects of the survey were married women of reproductive age born between January 1973 and January 1997. The final sample contains 3,415 respondents, with 1,143 from Beijing, 1,131 from Shaanxi, and 1,142 from Henan.

Reference: Jin, Y. and Shen, X. (2022). Socioeconomic status, fertility motivations, and second-child plans among urban women in China. (In Chinese). *Population Research* 46(6):88.

Hubei Fertility Study (2021)

The Hubei Fertility Study (2021) was conducted at Hubei Province in July 2021. The target population for the survey consists of families with no children, one child, and two children within the permanent resident population of Hubei Province as of July 31, 2021. The survey employed a stratified, two-stage, equal-size random sampling method. This approach avoided the issue of insufficient sample size in certain areas that can arise from PPS sampling, ensuring the sample was representative of the entire province. The research team first conducted random sampling across all 125 counties in Hubei Province, selecting 400 households from each district to form the preliminary sampling frame. Second, within each of the 125 county-level districts, the survey team randomly selected

100 samples from the preliminary list of 400 households for telephone surveys. The survey ultimately yielded 12,014 valid responses.

The analytical sample comprises women of reproductive age (15–49). The survey asked respondents, “How many children do you think is ideal for yourself?” We treat this variable as a measure of fertility ideal. Individual weights are applied when estimating county-level fertility attitudes.

References: Shi, Z, Xi, S., Zhang, W. et al. (2022). Fertility intentions of urban and rural residents under the three-child policy. (In Chinese). *Chinese Journal of Population Science* 44(3): 1–18. doi:10.16405/j.cnki.1004-129X.2022.03.001.

Shi, Z. and Xi, S. (2023). The impact of COVID-19 on fertility intentions under the three-child policy. (In Chinese). *Population Research* 47(2): 78–95.