NBER WORKING PAPER SERIES

DOES FERTILITY STABILIZE AT LOW LEVELS? EVIDENCE FROM PERIOD, COHORT, AND SUBNATIONAL DATA

Diane Coffey
Michael Geruso
Dean Spears
Sangita Vyas

Working Paper 35824
http://www.nber.org/papers/w35824

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2026

This research was supported by grant, P2CHD042849, Population Research Center, awarded to the Population Research Center at The University of Texas at Austin by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). Diane Coffey acknowledges support from a Mentored Research Scientist Development Award from the NICHD, grant number K01HD103999. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w35824

Does Fertility Stabilize At Low Levels? Evidence from Period, Cohort, and Subnational Data
Diane Coffey, Michael Geruso, Dean Spears, and Sangita Vyas
NBER Working Paper No. 35824
September 2026
JEL No. C53, J11, J13

ABSTRACT

Many population projections used by international and government agencies assume that, in today's low-fertility societies, future fertility rates will stop declining and roughly stabilize. This paper assesses the evidence for an equilibrium point or rebound from low levels. We show that there is no robust evidence for a systematic fertility rebound, or for a stable equilibrium point at any fertility level conventionally assumed. Instead, where and when fertility has fallen low, it has tended to keep falling. This is true across a range of data sources, time periods, and fertility measures: United Nations data on period total fertility rates for all country-years from 1950 to 2023; Human Fertility Database data on completed fertility for cohorts born between 1935 and 1975 in 31 countries; and fertility rates in subnational geographies in Japan, India, and the United States. We reconcile these results with a prior literature that documented fertility reversals among some low-fertility countries in the early 2000s. Our findings provide an empirical basis for re-examining the equilibrium and rebound assumptions embedded in widely used population projections.

Diane Coffey
University of Texas at Austin
coffey@utexas.edu

Michael Geruso
University of Texas at Austin
Department of Economics
and NBER
mike.geruso@utexas.edu

Dean Spears
University of Texas at Austin
Department of Economics
dspears@utexas.edu

Sangita Vyas
Hunter College of the City University
of New York
sangita.vyas@hunter.cuny.edu

1 Introduction

Scholarship on the demographic transition recognizes three stages of population-level fertility: first high fertility, then falling fertility, and then low fertility in a "post-transition" stage (Notestein, 1945; Kirk, 1996). By 2025, a demographic transition from high mortality and high fertility to low mortality and low fertility had occurred for most of the world: Over two-thirds of the world's population lived in countries with below-replacement fertility rates (UN DESA, 2024). Now, in an era of global low fertility, a focus of population science has been to characterize the contemporary "post-transition" stage.[1] In particular, a key question in contemporary population science is whether fertility decline will end or reverse in low-fertility societies—and if it does, at what levels will fertility stabilize?

Influentially, the National Research Council (2000) report *Beyond Six Billion* by the National Academies proposed that "homeostatic mechanisms have the potential, with considerable lag, to counteract very low fertility" (p. 84), restoring fertility at or near replacement levels. But as recognized in both academic and public scholarship, the experience of the past quarter century has not revealed a homeostatic tendency to stabilize around replacement-level fertility (Lesthaeghe, 2014; United Nations, Department of Economic and Social Affairs, Population Division, 2024). Indeed, from the earliest examinations of low fertility, expert disagreements have endured on whether an assumption of eventual reversal of downward trends or stabilization of fertility rates (at any level) was supported by theory (Caldwell, 1976; Easterlin, 1978; Demeny, 1986; Lee, 1987; Kirk, 1996).

A body of empirical research in the 2000s and early 2010s revised the idea of fertility stabilization, this time around stable rates below replacement. Period rates had increased in a number of low-fertility countries during the early 2000s, which suggested that the "lowest-lows" of fertility had perhaps been reached (Goldstein et al., 2009). It was argued

[1] See, e.g., Goujon (2025) "[W]e are at an 'end of history' moment for classic demographic transition theory. Most of the world has moved through the demographic transition; the focus now lies in a heterogeneous post-transition landscape where fertility is low, mortality gains are slowing or stalling in some places, and migration is central yet hard to predict." See also Morgan (2003), Balbo et al. (2013), and Graham (2021).

that fertility rates in low-fertility societies were on the path to rebound, or at least to stabilization, following a period of birth postponement that was only temporarily depressing period rates (Bongaarts and Sobotka, 2012). Examinations of cohort fertility measures in data from this period supported that conclusion (Myrskylä et al., 2013; Caltabiano et al., 2009), leading researchers to project that fertility would stabilize or recover in a number of countries (Schmertmann et al., 2014). Formal methods that would form a new basis for population projection at the United Nations were developed for the 2010 Revision of the World Population Prospects (WPP), building in assumptions that country-level fertility rates were converging to stable equilibria.[2,3]

Yet, in the years since, fertility rates have fallen to ever new lows around the world—including in the particular societies in Northern Europe and elsewhere that once appeared to demonstrate evidence of fertility stabilization or rebound (Hellstrand et al., 2021; Ohlsson-Wijk and Andersson, 2022). So a central question of fertility dynamics remains whether fertility decline has an identifiable endpoint. This paper addresses that question by analyzing the past progression of period and cohort fertility rates in low-fertility societies.

To inform this question, we estimate and analyze phase diagrams for period and cohort fertility. Phase diagrams, as we describe in detail below, summarize the expected magnitude and direction of change in a variable for a given level of that same variable. We follow Alkema et al. (2011) in applying the phase diagram approach to fertility dynamics in period rates, and build upon them in applying it to cohort rates, in each case plotting growth rates in fertility against the level of these variables.

We apply this methodology to rates of change in period total fertility rates (TFRs)

[2]These new population projections also incorporated probabilistic approaches to handle uncertainty in the future trajectories of demographic rates (Alkema et al., 2011), though our paper does not touch on these aspects of the projection method concerned with probabilistic, uncertain projections.

[3]Prior revisions of the World Population Prospects also assumed that low fertility societies' total fertility rates would converge to a stable equilibrium (e.g. in 1996, below replacement countries were assumed to rebound to replacement, and in the 4 revisions between 2002 and 2008, below replacement countries were assumed to rebound to 1.85).

for all country-years since 1950 in the United Nations World Population Prospects, in completed cohort fertility for country-cohorts in the Human Fertility Database, and in fertility measures from subnational geographies in India, Japan, and the United States. In all cases, the core question we analyze is whether there is any level of (period or cohort) fertility that, once reached, has predicted a subsequent stabilization or rebound.

Across all data sources and fertility outcomes, we find no robust evidence of systematic fertility stabilization (i.e., no robust evidence of zero subsequent fertility change) at any level of period or cohort fertility, and no robust evidence for systematic rebound (positive fertility change). For country-years with period TFRs starting in the range 1.45 to 2.05, subsequent trajectories have been negative on average. Mean annualized growth rates were -0.6% (CI: -0.9% to -0.3%) for period TFRs starting in the 1.45–1.55 range and were -1.1% (CI: -1.4% to -0.8%) for TFRs starting in the 1.95–2.05 range. Below 1.45, the evidence is less clear, as there are fewer contributing observations and these span a shorter window. The analysis of completed cohort fertility is consistent with the period TFR results: For cohort fertility rates as low as 1.45 to 1.55—the lowest levels for which there is sufficient data to generate a precise estimate of completed fertility for subsequent cohorts—the average trajectory of completed cohort fertility among subsequent cohorts has been negative.

These results contribute to answering a core question of contemporary population science: whether history provides evidence robustly supporting the assumption of a stable endpoint at which fertility decline ends—reaching "homeostasis" in the language of Kirk (1996) and National Research Council (2000), or "Phase III" convergence in the language of Alkema et al. (2011)[4] and UN WPP methodology reports published since 2010. We find no robust evidence for stabilization or rebound—neither in the pooled global country-year data, nor, as we show in Section 4.4, for those countries identified in the prior demographic literature as examples of rebound from low fertility.

[4]The 2010 "Phase III" countries are defined by Alkema et al. (2011) as follows: "The countries that have entered Phase III are defined as the countries in which two subsequent five-year increases below a TFR of 2 children have been observed" (p. 820). We discuss this designation in more detail below.

Our findings indicate a different conclusion than earlier work that argued that lowest low fertility was a temporary or transitory state (Goldstein et al., 2009; Sobotka, 2011; Bongaarts and Sobotka, 2012). In Section 4.5, we compare our conclusions with the prior literature, showing that our differing results stem from the extended time window available to researchers today and from our broader sample inclusion criteria for low-fertility country-years.

This paper's findings provide compelling evidence for reassessing the set of stylized facts and assumptions that underpin several important projections. Recent projections from the UN WPP are governed by models that project long-term fertility stabilization around, or rebound to, levels with a global average near 1.6.[5] These assumptions, motivated by prior empirical research, are not supported by the latest data. Fertility rates may someday stabilize around a low level, but to project they will do so is to project that the future of fertility dynamics will, on average, be unlike its past.

2 Background

2.1 Theory and evidence on fertility trends in low-fertility societies

Early descriptions of demographic transition focused on the decline from high fertility to and through falling fertility—addressing important explanatory questions about the social, economic, and political forces initiating and driving these stages. This body of theory largely has not sought precise or quantitative predictions on *endpoints* of fertility decline in low-fertility societies. There is no theoretical consensus on the question of where any fertility floor may lie, or even on the higher-order question of whether a floor for fertility

[5]We discuss the UN WPP model of TFR equilibrium near a global average of 1.6 extensively in Section 2.2. The Wittgenstein Center projections, though different, make a similar projection of equilibrium: "Beyond 2050, we follow the method outlined in KC et al. (2014) for both low- and high-fertility countries, with the TFR converging slowly to 1.75 either by decreasing (for countries having not reached a TFR of 1.6 before 2100) or by increasing (for countries having reached a TFR of 1.6 or below before 2100) until 2200." (KC et al., 2024, p. 23).

exists at all. Some scholars (e.g., Notestein, 1945; Caldwell, 1976; Kirk, 1996) suggest the possibility of a long-run lower bound or convergence to some equilibria, though others (e.g., Demeny, 1986; Lee, 1987, 2003) leave open the question of whether any endpoint could be expected on the basis of theory.[6]

Bongaarts (2001) observed that "[c]onventional demographic theories have little to say about the level at which fertility will stabilize at the end of the transition." Neither does Second Demographic Transition theory identify a quantitative fertility floor or stable endpoint to which fertility may rebound (Van de Kaa, 1987; Lesthaeghe, 2014). Nor does McDonald (2000b)'s theory of how gender equity, inside and outside of the home, shapes fertility outcomes offer an answer.[7] Other observers of Europe's low fertility, which was below-replacement in many countries by the mid-1970s, were similarly skeptical of stabilization or rebound, and called for public policies to support families in an effort to increase fertility rates (Lutz et al., 2006). These families of theories did not aim to estimate or forecast fertility floors or stabilization points.

It was consequential, then, when a body of empirical work in the 2000s and early 2010s emerged which documented and commented upon fertility rebound in some low fertility countries, many of which were in Northern Europe (Kohler et al., 2002; Goldstein et al.,

[6]Kirk (1996) suggests the possibility of convergence to an equilibrium below, or up to, replacement: "Perhaps most promising for transition theory is an equilibrium or homeostatic framework... Application of the homeostasis principle could lead to a better understanding of what post-transition levels of fertility are likely to be. I venture to guess that the present level of fertility in Europe, below and in some cases well below replacement level, is an over-correction that will be modified." Lee (1987, p. 459) makes a contrasting case that highlights how the theoretical literature has not produced quantitative expectations of where a floor would be found: "Ordinary homeostatic tendencies essentially vanish in the course of economic development. ... Current theories of fertility give little insight into how equilibration would then occur." O'Neill et al. (2001, p. 229) agree: "Demographic transition theory provides the theoretical basis for the expectation that currently high fertility countries will experience, or continue to experience, falling fertility rates in the future. It provides little guidance, however, on the long-term average level they might eventually reach. It also has little to offer demographers grappling with the question of future fertility trends in countries that have already completed the transition to low fertility." Similarly, Lee (2003, p. 176) explains, "Most of the theories of fertility, as well as the experience of the More Developed Countries, imply that the demand for children will continue to decline in the future. But these theories point to no natural lower bound for fertility."

[7]McDonald's framework does include the possibility of avoiding very low fertility should gender equity in family-oriented social institutions advance to match that of individual-oriented institutions. However, it does not imply any *particular fertility rate* toward which post-transition societies tend (McDonald, 2000a).

2009). Figure 1 plots the fertility patterns that motivated this literature: it shows an increase in TFR in the 2000s in Northern Europe.[8] Considering the importance of the apparent rebound from low fertility in some countries, demographers sought to understand these patterns. Bongaarts and Sobotka (2012) argued that the observed increase in fertility could largely be attributed to a slowing pace of the postponement of childbearing. Myrskylä et al. (2009) posited that at the highest levels of the human development index (HDI), government policy might reverse fertility decline "by implementing policies that improve gender equality or the compatibility between economic success, including labor market participation, and family life" (p. 742).[9] Goldstein et al. (2009) focused on the contribution of economic growth and declining unemployment to the fertility increase in this period, and concluded "that formerly lowest-low-fertility countries will continue to see increases in fertility as the transitory effects of shifts to later childbearing become less important."

But Figure 1 shows that after the increase of the early 2000s, fertility in Northern Europe stalled and then turned downward again. The historical pattern revealed itself to be similar to prior episodes in the same set of countries—visible in the same figure in the late 1950s and the late 1980s—in which temporary upturns were followed by reversions to an overall downward trajectory.

2.2 The evolution of United Nations' fertility projections for low-fertility settings

Since the 1990s, the United Nations (UN) medium-variant population projections have assumed rebound and long-run stabilization of fertility rates for low-fertility countries (O'Neill et al., 2001, §5.2.1.3). The level of stabilization and the assumed trajectories

[8] This figure uses 2024 country-level data from the UN WPP, which we use in our main analyses and describe in greater detail in Section 3.

[9] The reversal reported by Myrskylä et al. (2009) was later re-evaluated and revised (Harttgen and Vollmer, 2014; Gaddy, 2021; Schubert et al., 2024). In the most recent global data, there is no evidence of a reversal of the HDI–fertility relationship at very high levels of human development (Spears et al., 2026; Marois and Lutz, 2026). In Section 4.4 we report estimates separately by level of HDI.

towards stabilization have varied,[10] influenced by successive waves of population science. But the central claims of rebound and stabilization have been consistent. For example, critiques of the assumption that fertility rates in below replacement countries would rebound to replacement (≈ 2.1) appear to have influenced the 1998 World Population Prospects (WPP) revision of the 1996 report's assumptions (Lutz et al., 1996, 2006). In the 1998 revision, below-replacement countries were assumed to rebound to and stabilize at heterogeneous below-replacement rates depending on the fertility levels at the time. In later revisions, all low fertility countries were assumed to rebound to a TFR of 1.85.

Although the rate and patterns of convergence to 1.85 varied in subsequent reports, the basic assumption remained the same until 2010, when the WPP made two changes to its methodology. One change, which is not our focus, but which has received more attention from the demographic community, was the move away from deterministic modeling of future fertility rates to Bayesian probabilistic projection models (Alkema et al., 2011).[11] The other change, which is our focus, was a return to the assumption that low-fertility countries would eventually rebound to replacement fertility. As United Nations Population Division (2013) explains, this decision was taken on the basis of the research on low fertility at the time—the research we discussed above in Section 2.1.[12] The most recent UN WPP Report, published in 2024, no longer assumes a rebound to replacement but nevertheless maintains the assumption of convergence to each country's stable rate (averaging around 1.6). One contribution of this paper's analysis that follows is to test whether such an assumption is supported by the most recent data.

[10] To draw this conclusion, we examined all UN World Population Prospects Reports published between 1990 and 2024. Earlier reports can be found in the UN Digital Archive (United Nations, Population Division, 1991, 1995, 1998, 2004, 2006, 2010, 2014, 2015, 2017, 2022; Gerland and United Nations, Population Division, 2019). The most recent World Population Prospects Report publishes a separate Methodology Report (United Nations, Department of Economic and Social Affairs, Population Division, 2024).

[11] Our analysis is concerned only with the assumption that fertility will converge towards stable equilibria, which has been present in WPP projections both before and after the adoption of a Bayesian framework, and which is separable from considerations of deterministic versus probabilistic forecasting. Similar assumptions underpin other institutions' projections, including those produced by the Wittgenstein Center (KC et al., 2024).

[12] Goldstein et al. (2009), Caltabiano et al. (2009), Myrskylä et al. (2009), and Sobotka (2011) were cited to justify the assumption that fertility rates in below-replacement countries would rebound to replacement.

2.3 The phase diagram approach to modeling fertility change

Phase diagrams are used in mathematics, ecology, economic growth theory, and elsewhere, to visually represent the laws of motion in a dynamic system by plotting the rate of change in a variable as a function of the level of that same variable (Strogatz, 2015, Ch. 2). We follow Alkema et al. (2011) in applying the phase diagram approach to fertility dynamics, plotting growth rates in fertility against the level of these variables.

A phase diagram provides a natural mathematical analog of the empirical question that we study: Conditional on some level of fertility at a point in time, should we expect fertility to subsequently decline or increase in the following years, and at what rate? For example, among countries that have reached a period total fertility of 1.5 in some year, what has been the average rate of change in period total fertility within those countries over the subsequent years? Has it been positive, negative, or approximately zero (so that TFR is approximately unchanged from 1.5)? One conceptual purpose of the phase diagram approach is to identify any steady-state "attractor" or equilibrium point as a crossing point, above which fertility falls, on average, and below which fertility rises.

To introduce notation, let $f_{c,t}$ be the period total fertility rate of country c in year t. The continuous, exponential growth rate of TFR for country c over the n years after t is

$$g_{c,t} = \frac{\ln(f_{c,t+n}) - \ln(f_{c,t})}{n}. \quad (1)$$

Figure 2 introduces a univariate phase diagram for TFR. Panel (a) is a conceptual illustration.[13] The horizontal axis represents the *level* of TFR, $f_{c,t}$. The vertical axis represents the subsequent *growth* in TFR, $g_{c,t}$, as a function of $f_{c,t}$. If an interval of $f_{c,t}$ predicts negative $g_{c,t}$, then the time-path of fertility would be declining in that range. If an interval of $f_{c,t}$ predicts positive $g_{c,t}$, then the time-path of fertility would be increasing in that range. And if a point or interval of $f_{c,t}$ predicts zero $g_{c,t}$, then reaching such an f_c^* would make fertility

[13]This figure is comparable with Figure 6 of Alkema et al. (2011, p. 827).

stable: TFR would remain at the equilibrium point f^*_c, neither increasing nor decreasing.

The model outlined in Alkema et al. (2011) and used in projections by the UN WPP amounts to a hypothesis that some steady-state f^* exists robustly at low fertility levels—either at replacement fertility ($f^* \approx 2.1$) or at some lower level ($f^* < 2.1$), either shared at one global value ($f^*_c = f^*_{c'} = f^*$ for different countries c, c') or heterogeneously across countries or other populations ($f^*_c \neq f^*_{c'}$). In the Alkema et al. (2011) phase diagram, $g = 0$ at $f = 2.1$, meaning that at any below-replacement fertility level, $g > 0$, so f is increasing.

To illustrate a phase diagram in practice, panel (b) of Figure 2 considers period TFR projections from 2027 to 2100 in the 2024 WPP Medium projection (UN DESA, 2024). (This is the only exhibit in this paper that uses WPP *projections*, rather than WPP *historical estimates*.) The figure shows the relationship between subsequent TFR growth rates on the vertical axis and TFR levels on the horizontal axis. The growth rate data are plotted as binned means.[14] The figure shows that the 2024 WPP Medium projection closely tracks an implied phase diagram that has an average $f^* \approx 1.6$. Below a TFR of 1.6, g tends to be positive and TFR tends to be projected to rise. Above a TFR of 1.6, g tends to be negative and TFR tends to be projected to fall towards about 1.6. Although the WPP does not frame its projection methodology in the language of phase diagrams, this accords with the WPP's own assessment of its assumed equilibrium or attractor point.[15,16]

[14] We compute five year growth rates for each country-year. For example, for the subsequent five-year growth rate for Pakistan following index year 2036, we compute $g_{c,t} = [\ln(f_{\text{Pakistan},2041}) - \ln(f_{\text{Pakistan},2036})]/5$. We then collapse country-year growth rates and fertility levels into 15 bins based on the level of fertility in the index year. Figure 2b plots the means for each bin. This strategy follows the methods we use for our main results, which are described in greater detail in Section 3. In appendix Figure A1, we plot each of the 16,284 country-year observations contained in the medium projection (rather than as binned means, which we do in Figure 2b). The finer presentation of the data in Figure A1 shows the extent of the spread around the mean value.

[15] From United Nations, Department of Economic and Social Affairs, Population Division (2024, p. 31): "The unweighted mean value for the world distribution based on the 39 countries or areas that have entered phase III is 1.62 children per woman in 2100." This is reported as a mean because the WPP Phase III model is hierarchical, allowing for country-specific levels. The intercept that is applied to the majority of countries that were not yet classified as already in Phase III is approximately 1.6.

[16] A note on *phase diagram* versus *Phase III*: Alkema et al. (2011) recognize three stages of the demographic transition. In the third, termed Phase III, fertility is low and trending towards a stable equilibrium. Alkema et al. (2011) model Phase III with a relationship called a *phase diagram*. It is only a coincidence of language that such a figure is known in the literature as a phase diagram; it is a distinct usage from Phase III.

To quantify how well one, log-linear, univariate phase diagram describes the WPP projection, we note that the fit of an OLS regression on the underlying growth rates by country-year is strikingly high: $R^2 = 91\%$. This indicates that a simple phase diagram rule that is applied identically to all countries in the WPP (without any regard to differences across countries) would come close to replicating the WPP's 2024 total-world projection output.[17] In short, countries' future TFR paths in the 2024 WPP Medium projection are well approximated, on average, by log-linear mean reversion toward a stable equilibrium near 1.6 births per woman.[18]

In this paper, we evaluate whether *historical* fertility trajectories, pooled across countries,[19] exhibit a stable low-fertility equilibrium as in the WPP 2024 Medium projections, or show continued decline after crossing different sub-replacement levels. Specifically, we test whether the empirical phase diagram contains a crossing-point attractor, where the rate of change in fertility, g, crosses zero from above (as does the solid line in Figure 2a), or whether g remains negative throughout the observed domain (as does the dashed line in Figure 2a).

Phase diagrams have been influential in demographic theory and population modeling to conceptualize transition dynamics. We apply them empirically to directly isolate our primary object of interest: whether stable low-fertility equilibria can be found robustly in the pooled, global historical experience and in subnational histories.

[17] This observed fit is high even though the 2024 WPP in fact permits different f_c^* values for 39 countries, and even though it models an AR(1) process in levels rather than a log-linear functional form. Country-by-country projection output would be replicated less well for the 39 countries that do not receive the global, pooled model.

[18] These paths close roughly 16% of the log distance to the f^* equilibrium per decade, which implies a half-life of about four decades.

[19] We do not consider the question, for which we suspect the historical record would be underpowered, of whether some country may have an idiosyncratic country-specific equilibrium, f_c^*.

3 Data and methods

3.1 Data: Periods and cohorts; countries, states, districts, and prefectures

Our primary data source is the set of historical country-by-year estimates of period TFR for all countries and UN-recognized territories spanning 1950 to 2023 from the 2024 World Population Prospects (WPP) (United Nations, Department of Economic and Social Affairs, Population Division, 2024). As described above, generating phase diagrams requires information on fertility rates (levels) at a point in time, $f_{c,t}$, and the annualized fertility rate changes (growth rates) that followed, $g_{c,t}$. For each country-year observation we calculate g in the WPP data as $g_{c,t} = (\ln(f_{c,t+n}) - \ln(f_{c,t}))/n$. We primarily focus on growth rates in five year windows ($n = 5$) because this length is a focus of the prior literature, but in the Appendix, we report on three- and ten-year windows as well (see Figure A2).

In separate analyses, we draw on the Human Fertility Database (HFD) (Human Fertility Database, 2026) to generate phase diagrams for completed cohort fertility rather than period TFR. Using all available country-cohorts in the HFD, which span 1935 to 1980, we replicate the procedure described above substituting country-cohorts for country-years, so that the subscript t indexes birth cohort, and the change $g_{c,t}$ is calculated as the average change in cohort TFR after five consecutive single-year-of-birth birth cohorts (so the last cohort with subsequent 5-year growth observed is 1975).

In further analyses, we draw on several other datasets to generate phase diagrams at the subnational level following the same pattern but replacing country-year observations of period fertility with state-year and prefecture-year observations in the United States and Japan, respectively. For India, the structure of the available data allows us to better examine cohort fertility, which we analyze at the cohort-district level. Data for India come from the fifth round of the National Family Health Survey, India's Demographic and Health Survey, which collected data between 2019 and 2021 (International Institute for Population Sciences (IIPS) and ICF, 2021). Data for Japanese prefectures come from the

2024 Vital Statistics of Japan (Ministry of Health, Labour and Welfare, 2025). Data for U.S. states come from Galofré-Vilà and Gómez-León (2025).

3.2 Method: Non-parametric, binned estimation of phase diagrams of period and cohort fertility

Panel (a) of Figure 2 is a conceptual illustration of a fertility phase diagram. It contrasts two possibilities: First, a phase diagram that crosses zero, establishing an equilibrium level at which fertility stabilizes, and second, a phase diagram in which fertility growth is negative throughout the plotted range of fertility levels. The research question of our paper is: Which of these possibilities is better supported by the historical data? More precisely, does the historical record provide robust evidence that countries' fertility rates rebound ($\mathbb{E}[g|f] > 0$) below any observed, low f_c^*? Is there a fertility equilibrium point at 1.6 or any other level?

As a purely conceptual tool, panel (a) of Figure 2 describes a simplified and noiseless deterministic law of motion. Estimating a phase diagram empirically amounts to finding the *average* tendency of fertility rate changes, conditional on fertility levels. We construct an empirical phase diagram by computing a series of non-parametric means. This approach does not presuppose any functional form for the phase diagram. In various samples, we compute the average rate of growth in period TFR for a series of neighborhoods around f: $\mathbb{E}[g|f]$.

Formally, we compute a sample mean of g to estimate $\mathbb{E}[g \,|\, f]$, which is necessarily calculated in some radius h around f:

$$\widehat{\mathbb{E}[g|f]} \equiv \frac{1}{|\{f_{c,t} : f_{c,t} \in [f-h, f+h)\}|} \sum_{f_{c,t} \in [f-h, f+h)} g_{c,t}, \tag{2}$$

with radius $h = 0.05$ in our main results, chosen to balance statistical precision and sample size against the desire for a resolution fine enough to be of practical interest. Standard errors are clustered at the relevant geographic level (country, Japanese prefecture, US state, or Indian state, depending on the data source), if the same unit appears multiple times per estimate.[20] Otherwise, heteroskedasticity-robust standard errors are used.

Our central results include analyses of national period TFRs, national completed cohort fertility rates, subnational period fertility, and subnational cohort fertility. We also present non-parametric estimates of $\mathbb{E}[g \mid f]$ for subsamples divided by historical period and by theoretically or statistically motivated subgroups of country-years, such as according to human development, age at childbearing, unemployment, geographic region, or population size.

One important difference between our empirical estimation of the phase diagram and its implementation by Alkema et al. (2011)—beyond our non-parametric approach—is that when analyzing fertility rate growth or decline starting from some given level of period TFR, we include *all observations* at that fertility rate in the estimation sample. In contrast, Alkema et al. (2011) limit their analysis to the country-years they designate as "Phase III." Specifically, a country is classified as having entered Phase III if two things are true: (1) the country has experienced three consecutive five-year periods in which the mean of TFR over those five years was below two, and (2) the country exhibits increasing fertility across the three consecutive five-year means covering a 15-year window.[21] We discuss this sample restriction and its implications in Section 4.5 and Appendix B.

[20]For example, of the 412 historical country-years in the WPP with a TFR between 1.55 and 1.65 and with observed five year growth, Canada is 21 of them, Germany is 5, and Chile is 1. Clustering recognizes that the 21 Canadian observations are not statistically independent.

[21]From Alkema et al. (2011, p. 820): "...we use a deterministic rule to identify the start of Phase III during the observation period for countries in which the fertility transition has ended. The countries that have entered Phase III are defined as the countries in which two subsequent five-year increases below a TFR of 2 children have been observed. By this definition, in the period 2005–2010 the start of Phase III was observed in 20 countries (Belgium, Bulgaria, Channel Islands, Czech Republic, Denmark, Estonia, Finland, France, Germany, Italy, Latvia, Luxembourg, Netherlands, Norway, Russian Federation, Singapore, Spain, Sweden, United Kingdom, and the United States)."

4 Results

4.1 National period TFRs: 1950-present

Figure 3 plots the first main result: a phase diagram for $\mathbb{E}[g|f]$ using all available country-years in the WPP historical estimates with below-replacement fertility. Each country×year observation with below-replacement fertility forms an index observation ($f_{c,t}$), for which an annualized growth rate over the subsequent five years ($g_{c,t}$) is calculated. Points on the phase diagram are estimated over the indicated ranges for f according to Eq. (2). Countries may contribute to different ranges of f in different years as they pass through different fertility levels. Country×years contribute to estimates in the figure only if there are at least five years of subsequent recorded period fertility rates following the index observation.[22] 95% confidence intervals are plotted with the means, and are computed using standard errors clustered by country. Marker sizes are proportional to the sample size available to compute each estimate.

Figure 3 shows that across the world, period TFRs have, on average, fallen after starting from 1.45 or higher. For country×years with TFRs from 1.45 to 2.05, subsequent trajectories have been negative on average, with mean, annualized growth rates ranging from -0.6% for period TFRs starting in the 1.45–1.55 range to -1.1% for TFRs starting in the 1.95–2.05 range. The means reported in the figure do not imply uniform trajectories. For example, in the sample that identifies growth rates beginning from TFRs in the 1.45 to 1.55 range, 174 out of 411 country×years had subsequent five year growth rates that were positive. What the data show is that the *mean* growth rates have been negative.

The historical record includes enough data points to generate precision in these parameter estimates: 95% confidence intervals rule out zero or positive mean growth rates after fertility rates of 1.45 or higher—that is, they rule out equilibration or rebound, on average. Even at the lower end of this range (meaning, 1.45 to 1.55), the mean growth rate estimate

[22]For example, Japan contributes to the estimate for $[1.15, 1.25)$ in 2005, to the estimate for $[1.25, 1.35)$ in years 1999-2004 and 2006-2009, and so on.

is informed by many data points spanning across 72 countries.

Below 1.45, the evidence is less clear, as there are fewer contributing observations and these span a shorter time window.[23] Growth rate estimates were negative following country-years in which TFRs started from 0.95 to 1.15 and from 1.35 to 1.45, and were positive for country-years from 1.15 to 1.35. But for all period TFRs below 1.45, the results are not statistically significant and have large confidence intervals, and so are consistent with positive, negative, or zero expected growth rates.

The pattern of these results is robust to using 3-year or 10-year growth windows, instead of 5-year growth windows (Figure A2); to using linear rather than non-parametric estimation (Figure A3); to averaging growth in five-year geometric means, to reduce the influence of transitory noise (Figure A4); to splitting the sample roughly in half by population size so that larger- and smaller-population countries are analyzed separately (Figure A5); to collapsing means of g by country to weight each country equally within a TFR category (Figure A6);[24] and to excluding years affected by COVID (Figure A7).

In sum, although assumptions of rebound or stabilization at levels above TFR=1.45 were influentially advanced by the prior literature and adopted by various population projections, the evidence of Figure 3 indicates that assumptions of a rebound ($\mathbb{E}[g|f] > 0$) or stabilization ($\mathbb{E}[g|f] = 0$), at sub-replacement levels of fertility above $f = 1.45$, are not supported by this evidence. Where country-period TFR has fallen to these ranges, TFR has tended to subsequently *decline*—not to stabilize, and not to rebound. In Section 4.5, we reconcile these results with a prior literature that has come to conflicting conclusions, and with the UN WPP methods, which project ($\mathbb{E}[g|f] = 0$) at approximately $f = 1.6$.

[23]For example, the estimate for 0.95 to 1.05 is informed by 23 country-years, for 1.15 to 1.25 by 133 country-years, and 1.35 to 1.45 by 328 country-years.

[24]Figure 3's main result takes an average of (f, g) observations, and therefore weights a country more heavily within a bin if it has had more country-years at that TFR.

4.2 Completed cohort fertility

Period fertility can be affected by shifts in the age of childbearing that might not reflect changes in fertility desires, or in completed cohort fertility rates that eventually obtain (Bongaarts and Feeney, 1998). The literature that highlights this phenomenon (Bongaarts, 2001; Kohler et al., 2002; Goldstein et al., 2009; Bongaarts and Sobotka, 2012) offers a structural reason why declines in period TFR in earlier periods could be consistent with a Phase III rebound or equilibration in later periods. Figure 4 broadens our analysis by investigating the change in completed cohort fertility, using all observations in the Human Fertility Database that have cohort fertility in the relevant range and have subsequent five-year changes observed. This is 890 country-cohorts, covering 31 countries, with an average of 28.7 cohorts observed per country.

As in Figure 3, the horizontal axis in Figure 4 is the level of fertility. The vertical axis is the subsequent annualized five-year change expressed as a growth rate. In this case, the growth rate measures the average change over successive *cohorts* separated by one birth year, rather than successive one-year periods. 95% confidence intervals are calculated using standard errors clustered by country.

The range of fertility rates for which it is possible to generate estimates of completed cohort fertility does not extend as low as for period rates in Figure 3, as the historical experience in these ranges has not resolved yet. For example, the 1977 cohort of Spain, which turned 45 in 2022, had completed rates below 1.35, but there have not yet been five years of subsequent cohorts to evaluate the changes from that starting point.

The figure shows that over the range of available data, which spans cohorts born between 1935 and 1975, all point estimates in the phase diagram are negative: For *any* level of observed completed fertility, the average value of completed fertility over the subsequent five cohorts was lower. These negative growth rates are statistically significant in all but the lowest range (1.35–1.45), where the observations (eight observations from two countries) are too few to generate an informative estimate. Results using 3-year or

10-year growth rates in completed cohort fertility (see Figure A8) are similar to the Figure 4 results using 5-year growth rates. Although, in principle, the cohort results might have revealed fertility trajectories quite different from the period analysis, these cohort results in fact strengthen the statistical conclusion from the period data: For all rates above $f = 1.45$, the growth rate of fertility is negative ($\mathbb{E}[g|f] < 0$) and statistically significant.

4.3 Subnational fertility in Japan, the United States, and India

We next turn to subnational evidence. In tests for convergence in growth rates, comparing states within a country can be an informative method for holding constant variables that a cross-country comparison may not (Barro and Sala-i Martin, 1991, 1992; Sala-i Martin, 1996). In that tradition, we next plot fertility dynamics for subnational units of three countries: Japan, where TFR has remained below 2 since 1975; the United States, where national TFR fluctuated near replacement levels for decades; and India, where some states have long had below-replacement TFR, but country-level TFR has only recently fallen below replacement. Together, these three cases capture dynamics across diverse baseline fertility levels, stages of demographic transition, and economic contexts.

Panel (a) of Figure 5 shows fertility trajectories for Japanese prefectures, which are Japan's major subnational divisions.[25] The figure plots a phase diagram in period rates, restricting to data from 2010 onward, in order to focus on the most recent fertility declines.[26] Most point estimates are negative, indicating negative trajectories, even starting from fertility rates below 1.3, a rate that Kohler et al. (2002) called "lowest-low" in the European

[25]Because some bins have few clusters, standard errors for analyses of Japan, the United States, and India are the larger of heteroskedasticity robust standard errors or standard errors clustered by prefecture (Japan) or state (United States and India). Choosing the larger of two alternative standard errors is a conservative practice common in similar situations, such as a small number of clusters (Angrist and Pischke, 2009, §8.1). For an example in the demography literature, see Kneip et al. (2014).

[26]Because the follow-up period for measuring the mean growth rate is 5 years, the latest date for establishing starting period TFRs ($f_{s,t}$ in $\mathbb{E}[g|f]$) is 2019 for Japan. In panel (b) of Figure 5, it is 2015 for the U.S. (where the state-level data end at 2020). Thus, for neither country is $f_{s,t}$ influenced by Covid-19, but subsequent fertility trajectories in Japan may incorporate impacts of Covid-19. Figure 7 shows estimates for both countries over an extended historical period, including pre-2010 data.

context. For fertility levels with the greatest number of prefecture×year observations—TFRs between 1.15 and 1.75—subsequent five-year growth rates have been statistically significantly negative. Note that there are no observations on the rightmost range of the horizontal axis: No Japanese prefecture had a TFR as great as 2 in 2010 or after. The median 2015 prefecture-level TFR was 1.53. From that starting point, TFR growth was negative over the subsequent five years in each of Japan's 47 prefectures.

Results for U.S. states are shown in panel (b). Though the U.S. has had relatively higher birth rates compared to other high-income countries, mean annualized growth rates have been statistically significantly negative at every fertility level since 2010. Despite starting from higher fertility levels, results for the U.S. are similar to those for Japan.

Figure 6 shows results for India. Due to incomplete vital registration, we rely on birth histories from India's fifth Demographic and Health Survey (DHS), collected between 2019 and 2021. Because the sample of older women in the DHS is not sufficiently large to generate reliable estimates for completed cohort fertility, we construct measures of cohort parity at age 30 for five-year birth cohorts born 1981-1985 and 1986-1990.[27] Thus, this figure gives a phase diagram in cohorts, rather than periods, indicating how cohort fertility—not completed cohort fertility, but cohort parity at age 30—evolved for later birth cohorts, for each initial level of parity at age 30.

For understanding the full life course, this age-30 restriction is less quantitatively important in the modern Indian context than, for example, it would be in the United States or Japan. This is because marriages and births occur at significantly younger ages in India (see, e.g., Park et al., 2023; Geruso and Spears, 2026).[28]

We organize the analysis at the level of Indian districts, which are administrative subdivisions of states. The average Indian district population in 2021 was about two

[27]Studying parity at age 30 enables us to include more recent birth cohorts. This is especially important in the modern Indian context where fertility has fallen fast in recent years. Our analysis plots the growth rate of parity at age 30 between the 1981-1985 cohort and the 1986-1990 cohort against the 1981-1985 cohort parity at age 30.

[28]Among women aged 40 and older in districts where mean children ever born was no greater than 2.1, 86 percent of their births reported at the interview had occurred by age 30.

million people—comparable to a larger Japanese prefecture or a smaller U.S. state.

We restrict attention in Figure 6 to the 351 of 707 observed Indian districts where estimated parity at age 30 for the 1981-1985 birth cohort was below 2.25 and at least 1.25. Even among these, there is a wide range of cohort parity across districts. For the entire range of indexed parities (1.25 through 2.25), trajectories for subsequent cohorts have been negative, and are statistically significantly negative at fertility levels above 1.35.

Thus the subnational evidence—covering the U.S., India (two of the world's three largest countries), and Japan (which has experienced one of the longest historical spells of subreplacement fertility)—points to the same robust conclusion as the country-level analyses using period and cohort fertility data. The statistical evidence is inconsistent with an expectation of fertility rebound or stabilization.

4.4 Is f^* robust? Estimates from many sources and subsamples

We have demonstrated the robustness of our main result by showing that according to several fertility measures, in both national and subnational data, the historical evidence is inconsistent with rebound from or stabilization at low fertility levels. In Figure 7, we show that our results are also robust to re-estimating fertility trajectories in subsamples of the global data.

In each panel, the top set of estimates (above the horizontal black line) uses 2024 UN WPP TFRs. The bottom set of estimates summarizes the findings from cohort and subnational analyses. The figure assembles sample means and confidence intervals for $\mathbb{E}[g|f \approx 1.6]$ in panel (a) and $\mathbb{E}[g|f \approx 1.5]$ in panel (b).[29]

Moving from top to bottom in Figure 7: there is no evidence for fertility rebound or stabilization whether we analyze data from the earlier half of the sample or the later half of the sample. This conclusion holds when examining regions separately. It holds when splitting the sample by the population sizes of country-years. Although international

[29]The bandwidth around the central f in these analyses is 0.1.

demographic analyses often treat evidence from larger and smaller countries symmetrically, WPP estimates include many countries with very small populations: The 25th percentile population size among country-years with TFR below replacement and an observed five-year growth rate is about 250,000 people. Among the 797 observations informing the main estimate of panel (a), the largest population size is over 10^5 times as large as the smallest.[30]

Other subsamples we analyze in Figure 7 are motivated by splits or hypotheses in the prior literature. As mentioned above, Alkema et al. (2011) classified 20 countries as being "post-transition" or in "Phase III" based on having had "two subsequent five-year increases below a TFR of 2 children" observed by 2010 (p. 820). Lesthaeghe (2014) lists "increasing mean ages at parenthood" as a feature that distinguishes the first and second demographic transitions, and Bongaarts and Sobotka (2012) argue that a reduced pace of change in the mean age of childbearing could raise fertility rates. Myrskylä et al. (2009) and others propose fertility rebound at high levels of the Human Development Index (UNDP, 2025). Adserà (2005) and Goldstein et al. (2009) study relationships between fertility and unemployment. Rows 11 through 21 in each panel of Figure 7 stratify the analyses along each of these dimensions.

The evidence presented in Figure 7 is consistent: Across all 54 estimates, $\mathbb{E}[g|f \approx 1.6]$ and $\mathbb{E}[g|f \approx 1.5]$ are negative. This evidence, along with the evidence from the full-sample period and cohort rates presented in Sections 4.1 and 4.2, does not support the claim that fertility shows a tendency to rebound or stabilize at low levels of fertility. A notable outlier

[30] Until this point, we have followed Alkema et al. (2011) and the WPP 2024 Methodology in averaging country-periods' TFRs without weighting by population size. This treats each country-period as an equally-weighted data point to learn about any country-year-level law of motion for fertility, and has the practical consequence of reducing standard errors in this case. If we depart from their methods to weight by population size, the point estimates are very similar (changing from -0.6% to -0.5% in panel (a) and from -0.4% to -0.6% in panel (b)) but the standard errors multiply almost by ten (from 0.0007 to 0.0047 and from 0.0007 to 0.0069, respectively). This loss of precision is expected, given the heterogeneity in population size: Population weighting concentrates much of the influence on a small number of large countries. But the stability of the point estimates and the robustness to splitting the sample by population size offer reassurance. Ultimately, it is a substantive question of demography whether a shared law of motion should describe the typical country-period or instead give countries influence in proportion to their populations. The much wider confidence intervals generated under population weighting are further reason for uncertainty and agnosticism about exactly how low fertility rates change.

in panel (a) of the figure—which, like all other estimates, is negative, but which generates the smallest expected fertility decline—is the set of countries identified in Alkema et al. (2011) as "Phase III." That is plausible, considering this subset was purposefully selected due to having experienced at least some periods of fertility rebound. In the next section and Appendix B, we compare this study's findings, and findings from several other important papers from the same time period, with our results.

4.5 Comparison with prior findings on low-fertility trajectories

Several influential studies—including Myrskylä et al. (2009), Goldstein et al. (2009), and Bongaarts and Sobotka (2012)—studied fertility trends in the 2000s, documenting a notable recovery in period TFR across several low-fertility countries. Evaluated from the empirical vantage point of that period, these patterns offered grounds to hypothesize that persistent fertility declines were ending or reversing in some countries that had reached low fertility levels.

Relative to these earlier studies, our analysis uses 15 additional years of observation. Figure 1 illustrates the reversal of fertility decline in Northern Europe during the early 2000s that motivated earlier arguments for a rebound. However, fertility in Northern Europe reversed course once again in the following decade, resuming its long-term decline. To illustrate how the observation window influences the conclusions, Figure 8 recomputes $\mathbb{E}[g \mid f]$ separately across five-year periods spanning the period 1973 to 2018. Here, the horizontal axis indicates years rather than initial TFR ranges (which are represented by marker shapes). Plotted 95% confidence intervals for mean annual growth rates are heteroskedasticity-robust.[31] Appendix Figure A3 presents an alternative visualization fitting a parametric log-linear law of motion to country-level TFR change over each five-year period in a separate panel.

Across nearly half a century of data, every average growth rate $\mathbb{E}[g \mid f]$ has been

[31]Country-clustering is equivalent in this case because each country appears only once per mean.

negative, with one exception: the five-year window starting in 2003. This episode is central to reconciling our findings with prior research. Analyzing data from the 2000s, researchers were indeed observing an uptick in fertility. The fact that TFR increased among low-fertility countries in the early-2000s could have suggested a fertility rebound or convergence. With 15 additional years of evidence, however, it became clear that the early-2000s episode was historically unique, rather than a broadly-shared, persistent pattern. Why fertility trajectories in the 2000s diverged so markedly from both preceding and subsequent periods remains an open question for social science. Regardless of the underlying reasons, however, this transient period does not provide robust evidence of long-term TFR convergence.

Our findings can also be compared to the assumptions embedded in leading population projection methodologies. The framework in Alkema et al. (2011) and adopted by the UN WPP assumes post-transition fertility stabilization. The quantification of stabilization, in those methodologies, is based on estimates using a “Phase III” subsample of the available data. That subsample is purposively selected to include country-years that have shown a pattern of increasing fertility.

We examine this endogenously selected subsample further in Appendix B, where we step away from our main research question to consider the assumptions and methodological implications of using the Phase III subsample as an input to projections of future fertility rates. Despite the selection of a Phase III sample on the basis of an interval of observed rebound, average fertility has fallen in these countries (Figures B2 and B3). In the appendix, we show that even small changes to the samples, methods, or assumptions around Phase III convergence change its empirical implications (Figure B4).

5 Discussion

Whether low fertility rates have shown a tendency to converge to a stable point is a substantive, empirical question. Our data include the full historical record of period total fertility rates since 1950 from the WPP; all available completed cohort fertility rates in the HFD; and granular subnational data from India and the United States (two of the world's three most populous countries) and from Japanese prefectures (in a country whose national TFR has remained below two continuously since 1975). We show that these histories have not robustly supported a hypothesis of convergence to stable fertility rates. Fertility rates in the range of 1.45 to 2.0 have tended subsequently to decline, not to equilibrate. For fertility rates that begin below 1.45, subsequent trajectories have followed no precisely-estimated, robust course to stabilization or rebound, though for rates in this range the data are too sparse to make precise statistical statements. Historical data cannot yet inform on scenarios in which a large number of diverse countries experience very low TFRs near, say, 1.0.

One limitation of any cross-country analysis is that conclusions are based on a small sample of unalike country-years. Our study—like the literature we respond to, engage with, and/or replicate portions of—uses samples that are small by conventional statistical standards (because there is only a small number of countries in the world). We have shown that some periods were unlike others and that some countries' trajectories were unlike others. Further research may find heterogeneities and sensitivities that we have not.

Some leading national and international population projections assume an equilibrium point to which fertility rates rebound and at which they stabilize. Such projections embed an assumption of stabilization and should be understood as hypothetical or conditional scenarios *predicated on the assumption of an equilibrium point*, rather than empirical extrapolations of past dynamics. Population projections may usefully explore the conditional hypothesis of future fertility equilibria, but that hypothesis is not established by past experience. While unknown homeostatic mechanisms, policy interventions, or other shifts might eventually arrest further decline, the absence of historical evidence for a floor indicates that population projection models without a fertility floor, or with very low floors, may also be useful to explore.

This paper's results do not imply that declining fertility rates will never halt at any level, but they do suggest that halting would be a departure from the average trends that have been observed. We do not propose an alternative assumption. Instead, we conclude that there is not yet enough data to draw any firm conclusion about how low future fertility rates may fall below two.

References

Adserà, A., 2005. Vanishing children: From high unemployment to low fertility in developed countries. American Economic Review 95, 189–193.

Alkema, L., Raftery, A.E., Gerland, P., Clark, S.J., Pelletier, F., Buettner, T., Heilig, G.K., 2011. Probabilistic projections of the total fertility rate for all countries. Demography 48, 815–839.

Angrist, J.D., Pischke, J.S., 2009. Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Balbo, N., Billari, F.C., Mills, M., 2013. Fertility in advanced societies: A review of research. European Journal of Population 29, 1–38.

Barro, R.J., Sala-i Martin, X., 1991. Convergence across states and regions. Brookings Papers on Economic Activity 1991, 107–182.

Barro, R.J., Sala-i Martin, X., 1992. Regional growth and migration: A Japan–United States comparison. Journal of the Japanese and International Economies 6, 312–346.

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

Bongaarts, J., Feeney, G., 1998. On the quantum and tempo of fertility. Population and development review , 271–291.

Bongaarts, J., Sobotka, T., 2012. A demographic explanation for the recent rise in European fertility. Population and Development Review 38, 83–120.

Caldwell, J.C., 1976. Toward a restatement of demographic transition theory. Population and Development Review 2, 321–366.

Caltabiano, M., Castiglioni, M., Rosina, A., 2009. Lowest-low fertility: Signs of a recovery in Italy? Demographic research 21, 681–718.

Demeny, P., 1986. Population and the invisible hand. Demography 23, 473–487.

Easterlin, R.A., 1978. What will 1984 be like? socioeconomic implications of recent twists in age structure. Demography 15, 397–432.

Gaddy, H., 2021. A decade of TFR declines suggests no relationship between development and sub-replacement fertility rebounds. Demographic Research 44, 125–142.

Galofré-Vilà, G., Gómez-León, M., 2025. Annual and state-level data on U.S. total fertility rates, 1931–present. Data in Brief .

Gerland, P., United Nations, Population Division, 2019. World population prospects 2019 : methodology of the United Nations population estimates and projections. Technical Report. United Nations. New York. URL: `https://digitallibrary.un.org/record/3908121?ln=en&v=pdf`.

Geruso, M., Spears, D., 2026. The likelihood of persistently low global fertility. Journal of Economic Perspectives 40, 3–26.

Goldstein, J.R., Sobotka, T., Jasilioniene, A., 2009. The end of lowest-low fertility? Population and Development Review 35, 663–699.

Goujon, A., 2025. The demographic future that we do not know about. Science 390, eaed1652.

Graham, E., 2021. Theory and explanation in demography: The case of low fertility in Europe. Population Studies 75, 133–155.

Harttgen, K., Vollmer, S., 2014. A reversal in the relationship of human development with fertility? Demography 51, 173–184.

Hellstrand, J., Nisén, J., Miranda, V., Fallesen, P., Dommermuth, L., Myrskylä, M., 2021. Not just later, but fewer: Novel trends in cohort fertility in the Nordic countries. Demography 58, 1373–1399.

Human Fertility Database, 2026. Human fertility database. Max Planck Institute for Demographic Research and Vienna Institute of Demography. Country-level fertility series and cohort fertility measures.

International Institute for Population Sciences (IIPS), ICF, 2021. India National Family Health Survey (NFHS-5), 2019–2021. URL: https://dhsprogram.com/data/dataset/India_Standard-DHS_2020.cfm. dHS-VII.

Van de Kaa, D.J., 1987. Europe's second demographic transition. Population bulletin 42, 1–59.

KC, S., Dhakad, M., Potančoková, M., Adhikari, S., Yildiz, D., Mamolo, M., Sobotka, T., Zeman, K., Abel, G., Lutz, W., et al., 2024. Updating the Shared Socioeconomic Pathways (SSPs) global population and human capital projections .

Kirk, D., 1996. Demographic transition theory. Population Studies 50, 361–387.

Kneip, T., Bauer, G., Reinhold, S., 2014. Direct and indirect effects of unilateral divorce law on marital stability. Demography 51, 2103–2126. doi:10.1007/s13524-014-0337-2.

Kohler, H.P., Billari, F.C., Ortega, J.A., 2002. The emergence of lowest-low fertility in Europe during the 1990s. Population and Development Review 28, 641–680.

Lee, R., 2003. The demographic transition: Three centuries of fundamental change. Journal of Economic Perspectives 17, 167–190.

Lee, R.D., 1987. Population dynamics of humans and other animals. Demography 24, 443–465.

Lesthaeghe, R., 2014. The second demographic transition: A concise overview of its development. Proceedings of the National Academy of Sciences 111, 18112–18115.

Lutz, W., Sanderson, W., Scherbov, S., Goujon, A., 1996. World population scenarios for the 21st century. The future population of the world. What can we assume today , 361–396.

Lutz, W., Skirbekk, V., 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 , 167–192.

Marois, G., Lutz, W., 2026. Low fertility may persist and could be good for the economy. Nature Human Behaviour 10, 633–634.

Sala-i Martin, X., 1996. Regional cohesion: Evidence and theories of regional growth and convergence. European Economic Review 40, 1325–1352.

McDonald, P., 2000a. Gender equity in theories of fertility transition. Population and development review 26, 427–439.

McDonald, P., 2000b. Gender equity, social institutions and the future of fertility. Journal of the Australian Population Association 17, 1–16.

Ministry of Health, Labour and Welfare, 2025. Vital statistics of Japan, 2024: Trends in total fertility rates by each prefecture. URL: `https://www.e-stat.go.jp/en/stat-search/files?stat_infid=000040320380`. vital Statistics, Final Data, Natality, Volume 1, Table 4-5.

Morgan, S.P., 2003. Is low fertility a twenty-first-century demographic crisis? Demography 40, 589–603.

Myrskylä, M., Goldstein, J.R., Cheng, Y.h.A., 2013. New cohort fertility forecasts for the developed world: Rises, falls, and reversals. Population and Development Review 39, 31–56.

Myrskylä, M., Kohler, H.P., Billari, F.C., 2009. Advances in development reverse fertility declines. Nature 460, 741–743.

National Research Council, 2000. Beyond Six Billion: Forecasting the World's Population. National Academy Press, Washington, DC.

Notestein, F.W., 1945. Population-the long view. Food for the World , 36–57.

Ohlsson-Wijk, S., Andersson, G., 2022. Disentangling the Swedish fertility decline of the 2010s. Demographic Research 47, 345–358.

O'Neill, B.C., Balk, D., Brickman, M., Ezra, M., 2001. A guide to global population projections. Demographic research 4, 203–288.

Park, N., Vyas, S., Broussard, K., Spears, D., 2023. Near-universal marriage, early childbearing, and low fertility: India's alternative fertility transition. Demographic research 48, 945.

Provencher, C., Galbraith, N., 2024. Fertility in Canada, 1921 to 2022. Demographic Documents 91F0015M2024001. Statistics Canada. URL: https://www150.statcan.gc.ca/n1/pub/91f0015m/91f0015m2024001-eng.htm.

Schmertmann, C., Zagheni, E., Goldstein, J.R., Myrskylä, M., 2014. Bayesian forecasting of cohort fertility. Journal of the American Statistical Association 109, 500–513.

Schubert, H.A., Dudel, C., Kolobova, M., Myrskylä, M., 2024. Revisiting the J-shape: human development and fertility in the united states. Demography 61, 1949–1973.

Sobotka, T., 2011. Fertility in central and eastern europe after 1989: Collapse and gradual recovery. Historical Social Research/Historische Sozialforschung , 246–296.

Sobotka, T., 2026. The global fertility freefall: trends, drivers and uncertainties. Reproductive BioMedicine Online , 105831.

Spears, D., Hsu, Y.C., Coffey, D., Conceição, P., Geruso, M., Pasanen, J., 2026. Low Fertility Persists Despite High Levels of Human Development and Gender Equity. Technical Report.

Strogatz, S.H., 2015. Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering. CRC press.

UN DESA, 2024. World population prospects 2024. URL: https://population.un.org/wpp/.

UNDP, 2025. Human Development Report 2025: A Matter of Choice: People and Possibilities in the Age of AI. Human Development Report Office, New York. URL: https://hdr.undp.org/content/human-development-report-2025.

United Nations, Department of Economic and Social Affairs, Population Division, 2024. World Population Prospects 2024: Methodology of the United Nations Population Estimates and Projections. Technical Report. United Nations. New York. URL: https://population.un.org/wpp/assets/Files/WPP2024_Methodology-Report_Final.pdf.

United Nations, Population Division, 1991. World population prospects, 1990. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/109656?v=pdf.

United Nations, Population Division, 1995. World population prospects: the 1994 revision. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/207239?ln=en&v=pdf.

United Nations, Population Division, 1998. World population prospects, the 1996 revision. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/261934?ln=en&v=pdf.

United Nations, Population Division, 2004. World population prospects, the 2002 revision. : Vol. 3, Analytical report. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/540509?ln=en&v=pdf.

United Nations, Population Division, 2006. World population prospects : the 2004 revision. Vol. 3, Analytical report. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/586158?ln=en&v=pdf.

United Nations, Population Division, 2010. World population prospects : the 2006 revision. Vol. 3, Analytical report. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/3924313?ln=en&v=pdf.

United Nations Population Division, 2013. World Population Prospects: The 2010 Revision, Volume I-Comprehensive Tables. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/3922214?ln=en&v=pdf.

United Nations, Population Division, 2014. World population prospects : the 2012 revision : methodology of the United Nations population estimates and projections. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/3921799?ln=en&v=pdf.

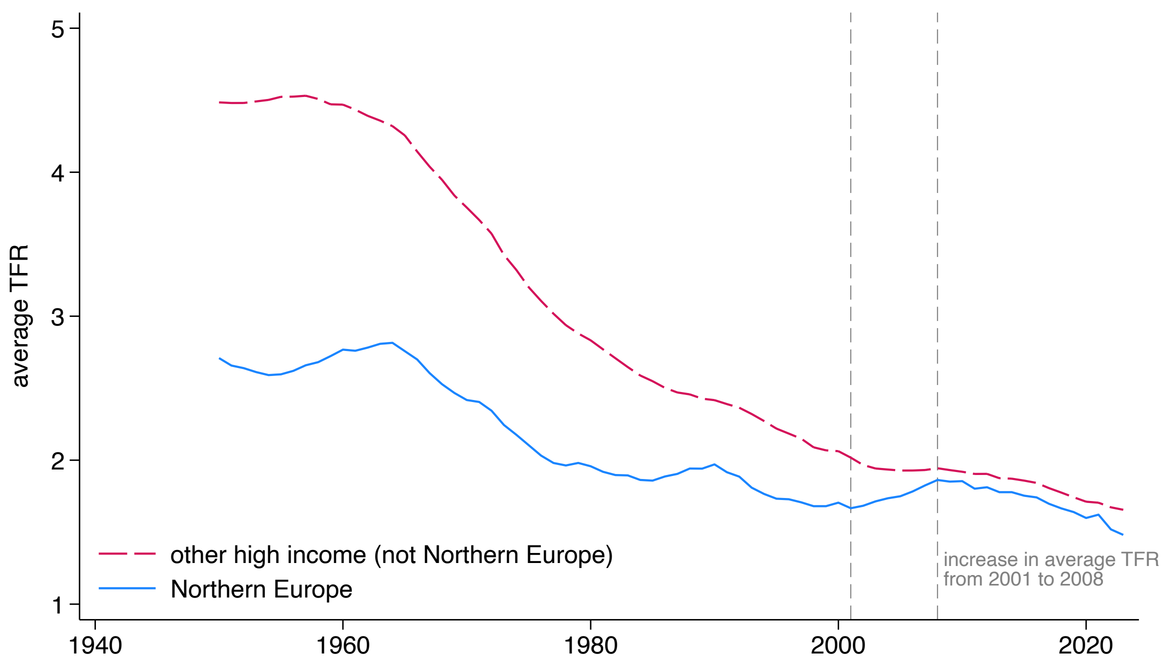
United Nations, Population Division, 2015. World population prospects : the 2015 revision : methodology of the United Nations population estimates and projections. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/3921805?ln=en&v=pdf.

United Nations, Population Division, 2017. World population prospects : 2017 revision : methodology of the United Nations population estimates and projections. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/3969860?ln=en&v=pdf.

United Nations, Population Division, 2022. World Population Prospects 2022 : Methodology of the United Nations population estimates and projections. Technical Report. United Nations. New York. URL: https://digitallibrary.un.org/record/3992029?ln=en&v=pdf.

World Bank, 2026. Unemployment, total (% of total labor force) (modeled ILO estimate). World Development Indicators, series SL.UEM.TOTL.ZS. URL: https://data.worldbank.org/indicator/SL.UEM.TOTL.ZS. source: ILO Modelled Estimates database (ILOEST), ILOSTAT. Accessed September 2, 2026.

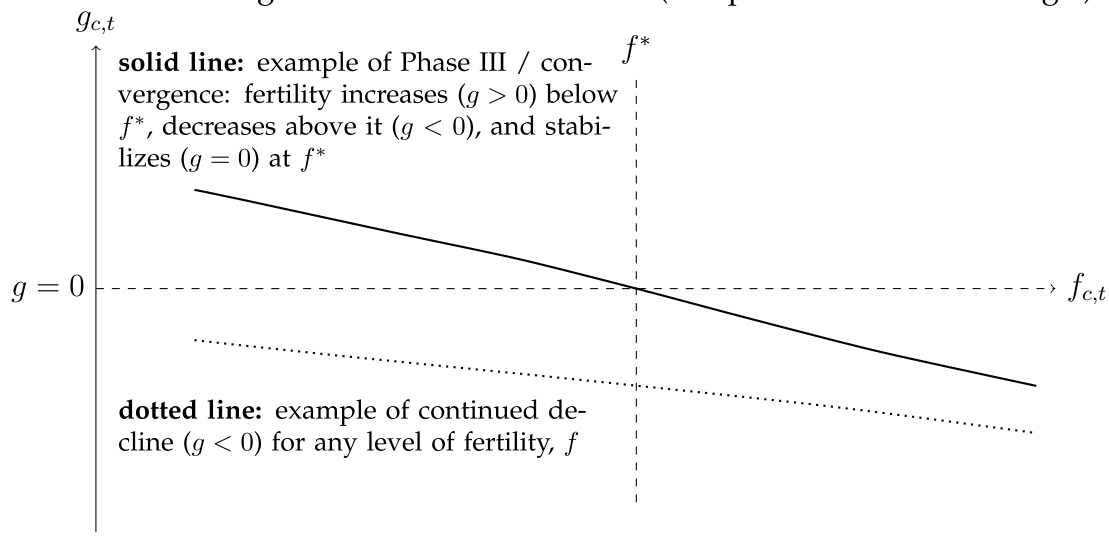
Figure 1: Motivation: Northern European fertility temporarily rebounded in the 2000s, but that rebound was not sustained, and downward trajectories continued in other high-income countries



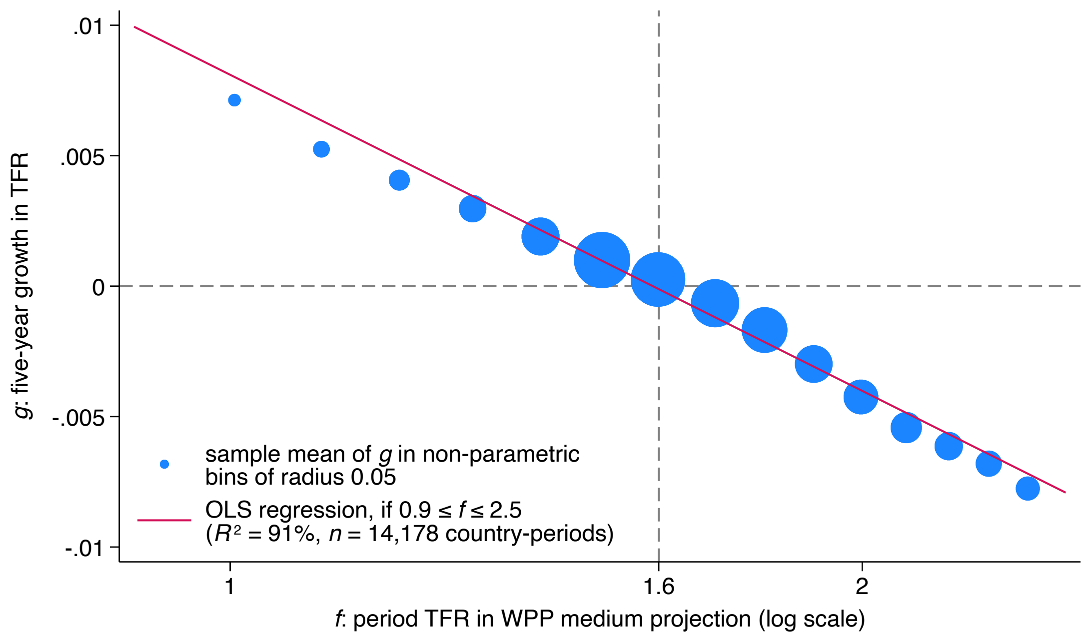
Note: The vertical axis plots total fertility rates, averaged into year-specific means at the country level. Data are from the 2024 United Nations World Population Prospects (UN WPP) (UN DESA, 2024). "Northern Europe" and "High Income" are as categorized by the UN WPP.

Figure 2: Fertility phase diagrams ($\mathbb{E}[g|f]$), in theory and as implemented in the 2024 WPP population projection

(a) illustration of TFR growth as a function of TFR (compare Alkema et al.'s Fig 6)

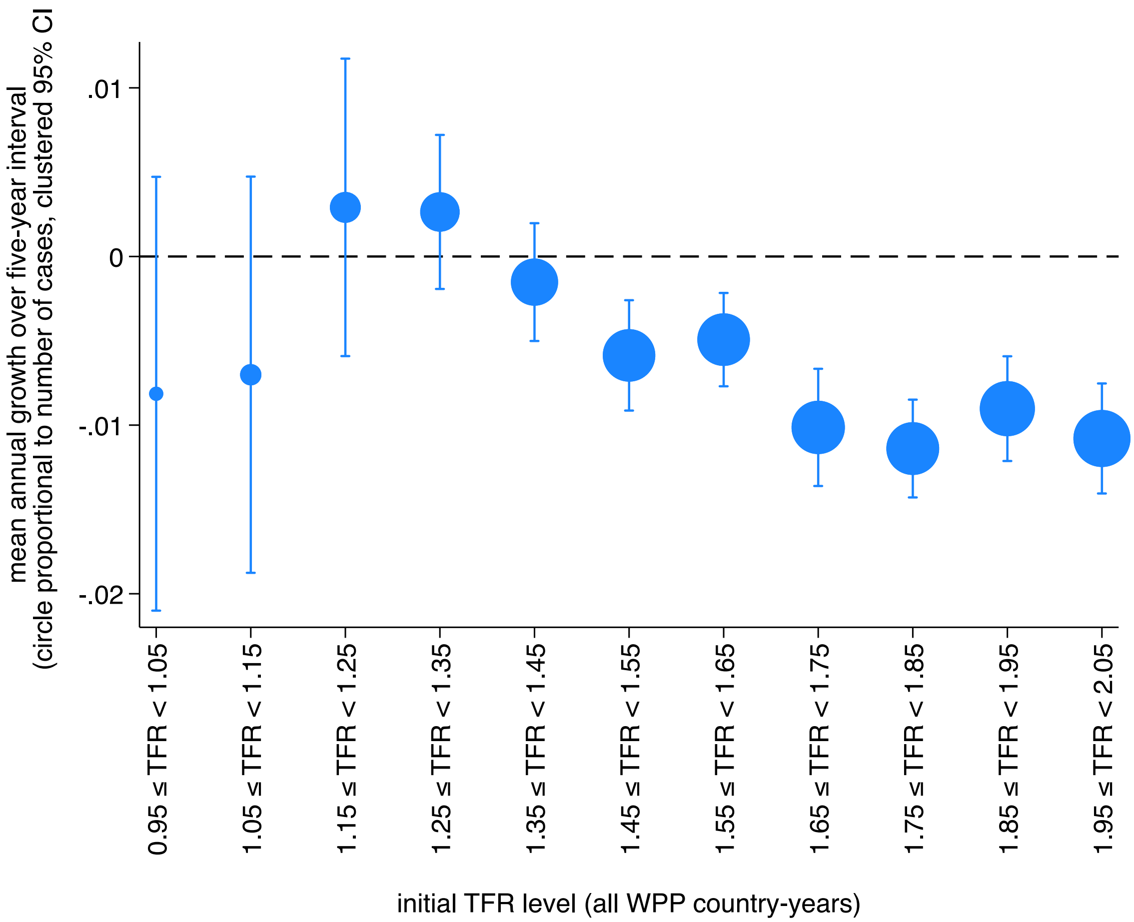


(b) $\mathbb{E}[g|f]$ in WPP 2024's Medium projection, 2027–2100, as a pooled, log-linear best fit



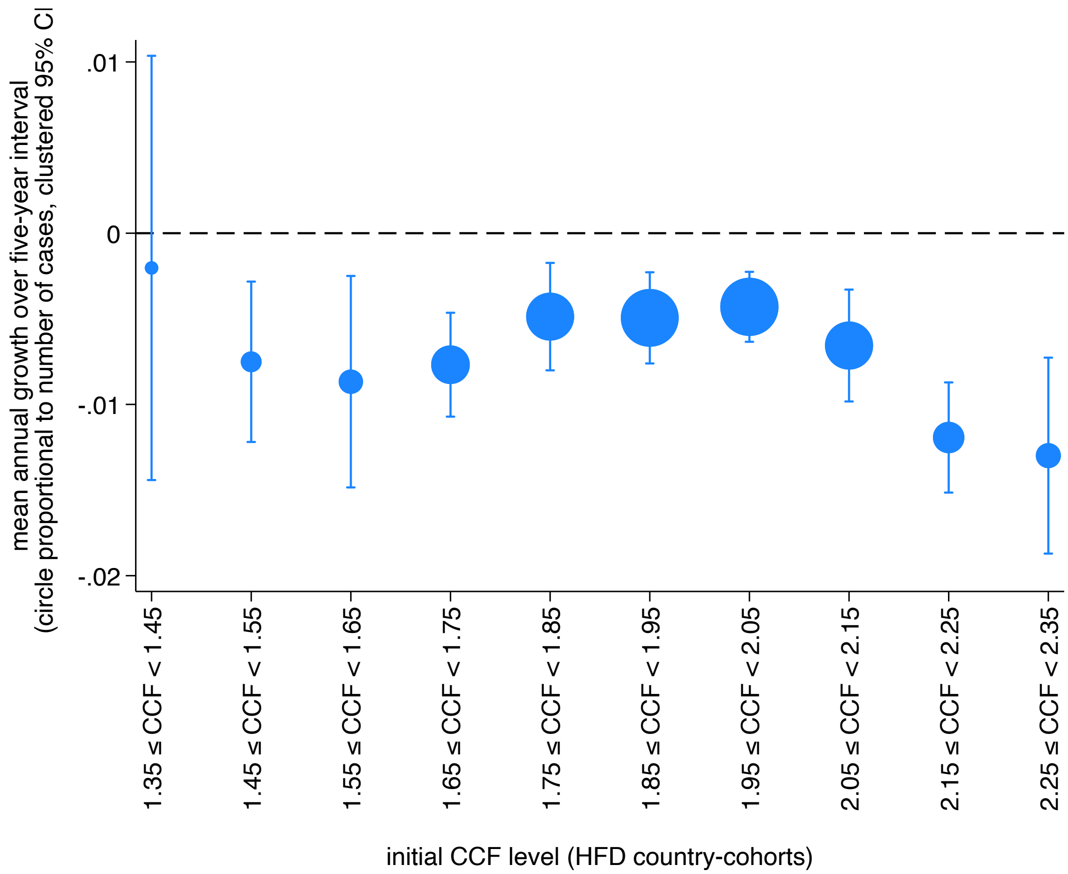
Note: Panel (a) is not an empirical estimate. It illustrates how the empirical figures in the paper should be read: A "Phase III model" of future fertility predicts a zero-growth point or region, but a law of motion with $g < 0$ throughout would predict negative TFR change, even at low fertility. Panel (b) plots binned means of future annualized five-year fertility growth rates, computed using Eq. 2, by fertility level in the index year. Marker sizes are proportional to the number of country-years composing each binned mean. Data are the country-level Medium projections from the 2024 UN WPP and span 2027 to 2100 (UN DESA, 2024). All countries are included except for the Vatican. The OLS regression line ($g_{c,t} = \beta_0 + \beta_1 \cdot lnTFR_{c,t} + e_{c,t}$) is fit to the underlying growth rates and fertility levels, not the binned means. The horizontal intercept of the log-linear regression identifies a pooled $f^* = 1.6$ (in fact the UN WPP projection model permits moderately differing f_c^* intercepts for different countries c). Figure A1 in the Appendix plots the underlying data points instead of binned means.

Figure 3: Average TFR change, conditional on TFR level, using past WPP country-years ($n = 3,372$)



Note: The figure plots binned means of historical annualized five-year fertility growth rates, computed using Eq. 2, by fertility level in the index year. The figure plots estimates for years 1950 to 2018. 95% confidence intervals are constructed using standard errors clustered by country. Marker sizes are proportional to the number of country-years composing each binned mean. Data are the historical estimates from the 2024 UN WPP, spanning 1950 to 2023 (UN DESA, 2024). Means, standard errors, and sample sizes are reported in Appendix Table A1.

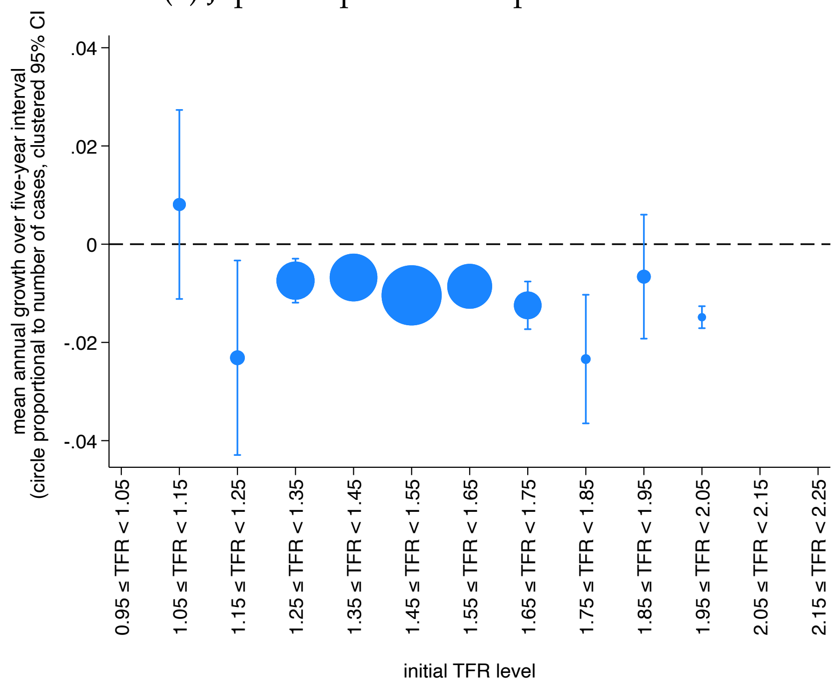
Figure 4: Average CCF change, conditional on CCF levels, using all past HFD country-cohorts ($n = 890$)



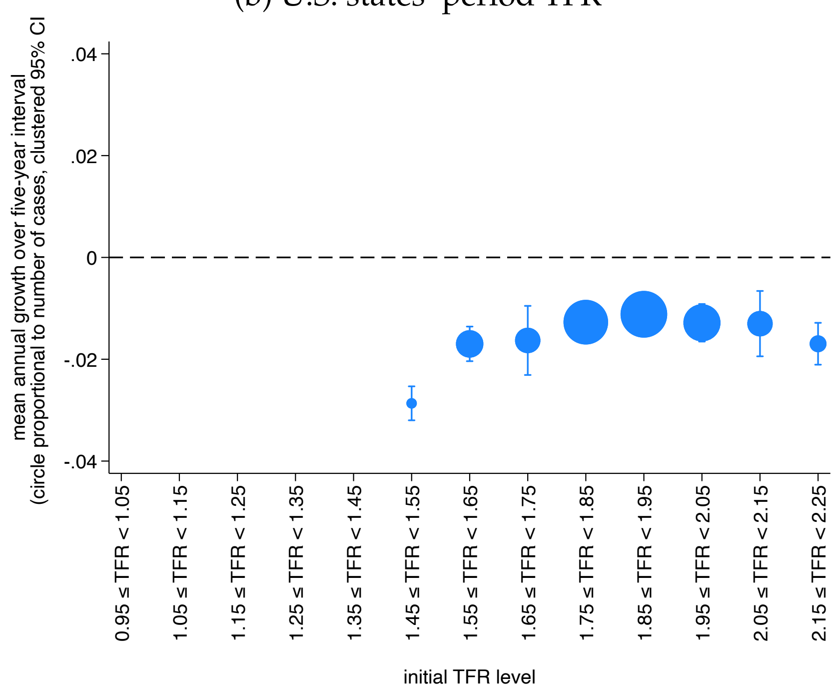
Note: The figure plots binned means of historical annualized five-year growth rates in completed cohort fertility, computed using Eq. 2, by completed fertility of the index cohort. The growth rate measures average change over successive cohorts separated by one birth year, rather than successive one-year periods as in Figure 3. The figure plots estimates for cohorts born between 1935 and 1975. 95% confidence intervals are constructed using standard errors clustered by country. Marker sizes are proportional to the number of country-cohorts composing each binned mean. Data are from the Human Fertility Database (Human Fertility Database, 2026). Means, standard errors, and sample sizes are reported in Appendix Table A1.

Figure 5: Evidence from subnational units since 2010 in Japan and the U.S.

(a) Japanese prefectures' period TFR

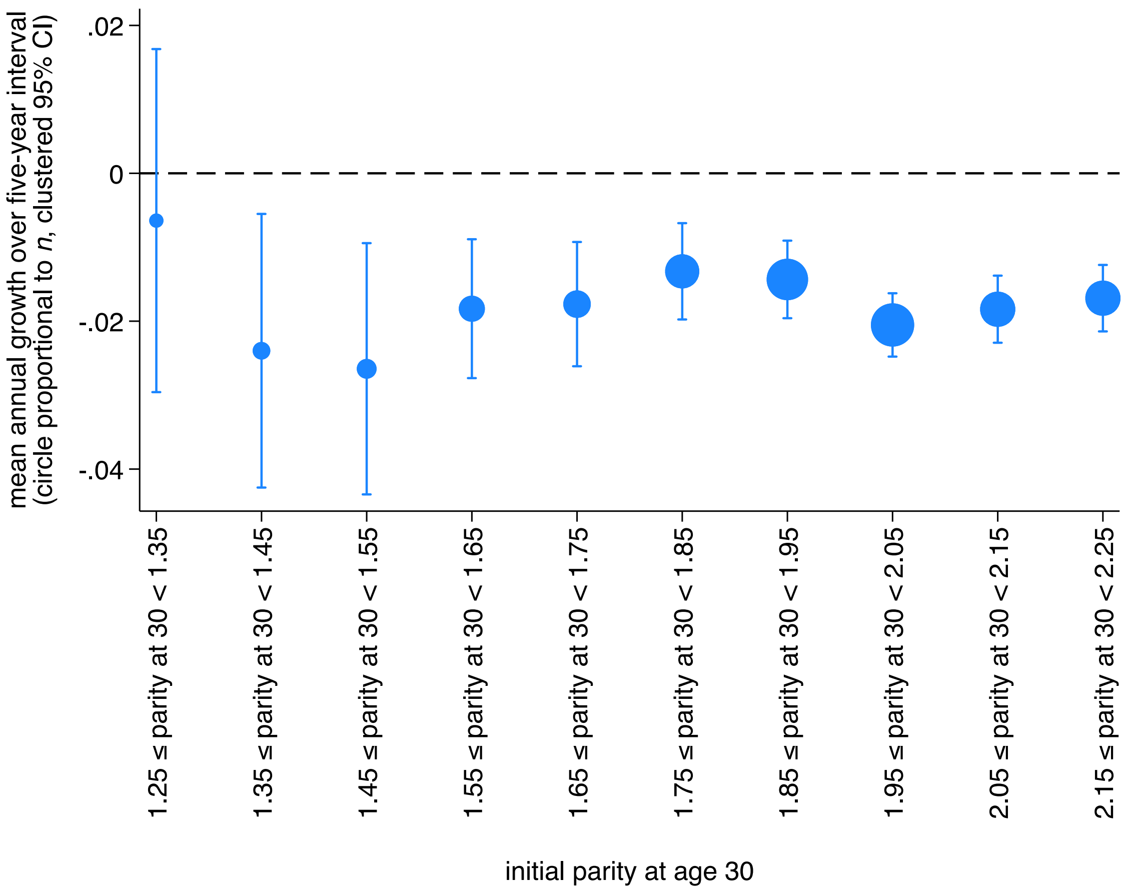


(b) U.S. states' period TFR



Note: Panel (a) plots binned means of historical annualized five-year fertility growth rates, computed using Eq. 2, by fertility level in the index year for Japanese prefectures from 2010 onwards. Data are from the 2024 Vital Statistics of Japan. Panel (b) plots binned means of historical annualized five-year fertility growth rates, computed using Eq. 2, by fertility level in the index year for US states from 2010 onwards. Data are from Galofré-Vilà and Gómez-León (2025). In both panels, 95% confidence intervals are constructed using standard errors that are the larger of heteroskedasticity robust standard errors and standard errors clustered by prefecture for Japan or state for the US. Marker sizes are proportional to the number of prefecture-years or state-years composing each binned mean.

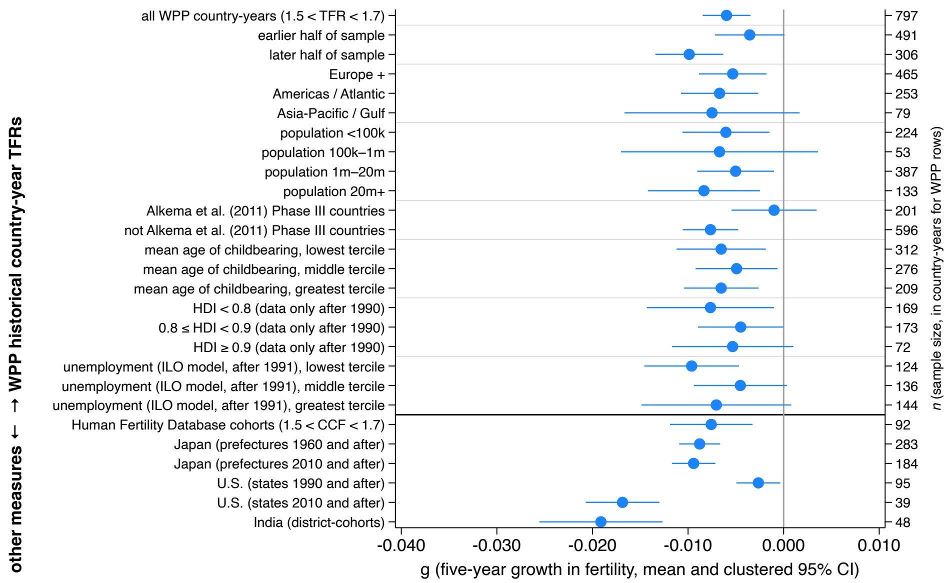
Figure 6: Evidence from Indian districts with low cohort fertility



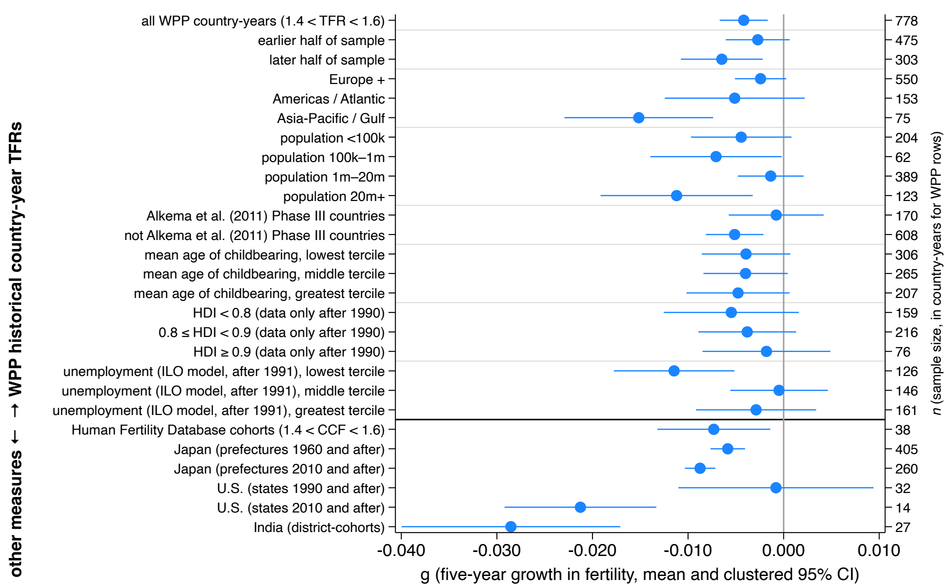
Note: The figure plots on the vertical axis binned means of the growth rate of mean parity at age 30, comparing those in the 1981–1985 birth cohort to those in the 1986–1990 birth cohort, for districts in India. The horizontal axis categorizes by cohort parity at age 30 for the 1981-1985 cohort. The growth rate measures average change in parity at age 30 over successive five-year birth cohorts. 95% confidence intervals are constructed using the greater of state-clustered or heteroskedasticity-robust standard errors. Marker sizes are proportional to the number of district-cohorts composing each binned mean. Data are from the 5th Indian Demographic and Health Survey, which collected data between 2019 and 2021.

Figure 7: Average fertility change is negative near 1.6 and near 1.5 in each of many subsamples

(a) Five-year change in fertility after total fertility is near 1.6

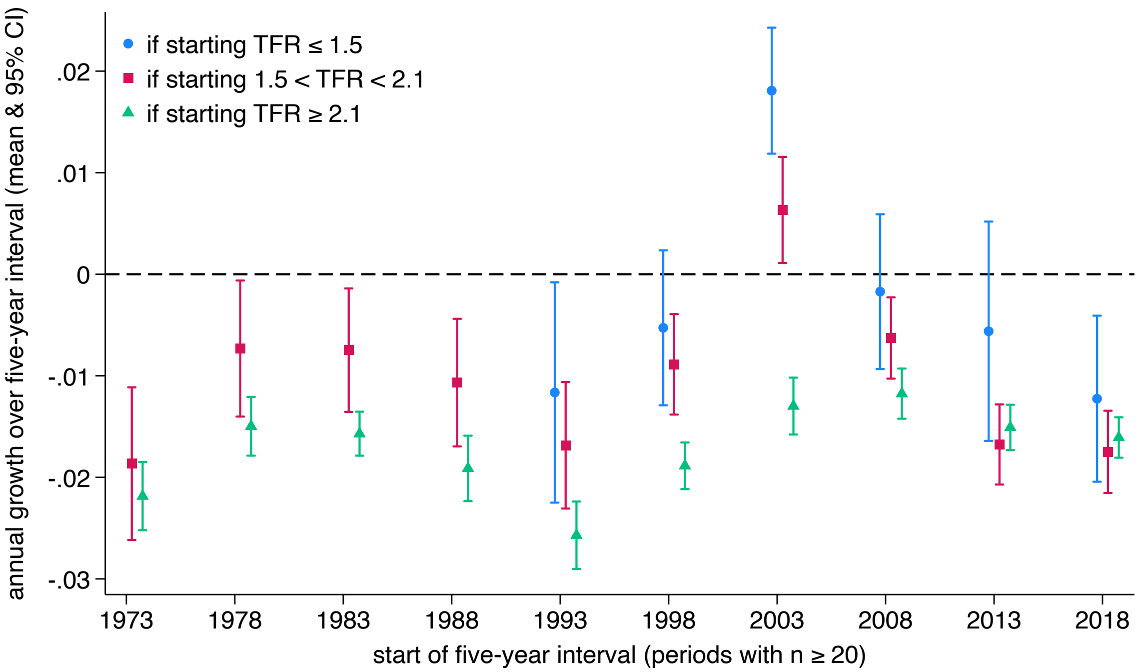


(b) Five-year change in fertility after fertility is near 1.5



Note: The figure plots mean historical annualized five-year fertility growth rates in different subsamples. In each panel, estimates above the horizontal black line split the 2024 UN WPP country-year TFR data into different subsamples. The estimates below the horizontal black line summarize findings from other data sources described in Section 3.
95% confidence intervals are constructed using clustered standard errors when countries, cohorts, Japanese prefectures, or US states appear more than once per estimate; otherwise heteroskedasticity-robust standard errors are used. t-statistics are adjusted for degrees of freedom.
In order to create the sub-samples of the 2024 UN WPP TFR data, we use the regional and population size classifications from the 2024 UN WPP (UN DESA, 2024). The mean age at childbearing is also from the 2024 UN WPP (UN DESA, 2024). HDI data are from UNDP Human Development Report (UNDP, 2025) and unemployment rates are from the World Development Indicators (World Bank, 2026, which sources from the International Labour Organization).

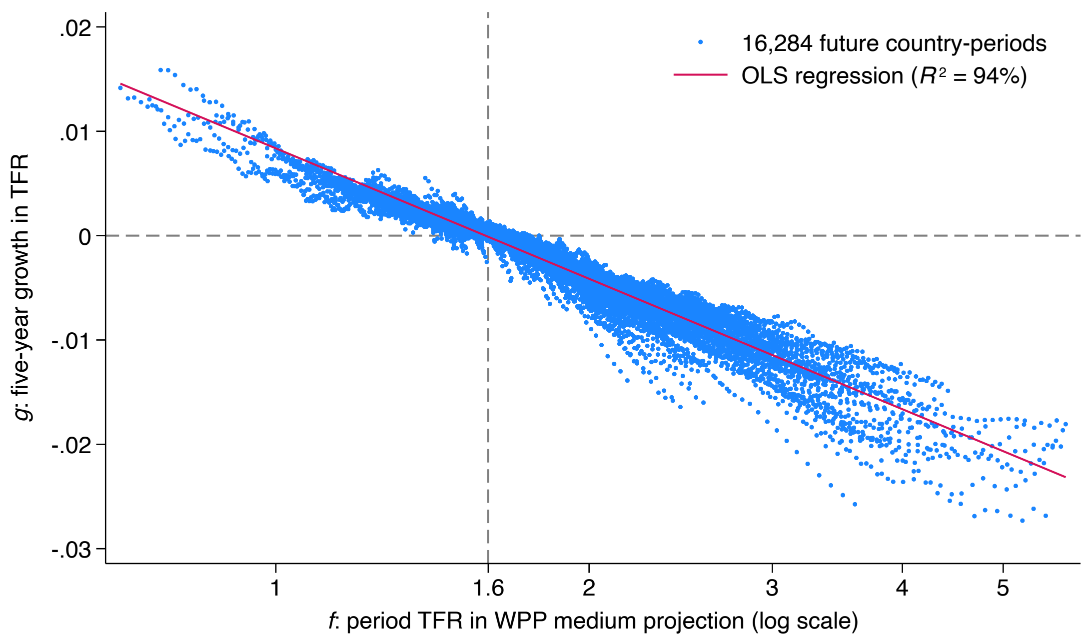
Figure 8: Average TFR change was different in the 2000s compared to other periods: All past WPP country-years, split by period



Note: The figure plots the mean historical annualized five-year fertility growth rate for each five-year period spanning the past four-and-a-half decades separately. Here, the horizontal axis indicates period rather than initial TFR ranges. Initial TFR ranges are represented by marker shapes: the blue circle represents an initial TFR ≤ 1.5, the red square represents an initial TFR such that $1.5 < TFR < 2.1$, and the green triangle represents an initial TFR ≥ 2.1. 95% confidence intervals are constructed using heteroskedasticity-robust standard errors because each country appears only once per estimate. Data are the historical estimates from the 2024 UN WPP, spanning 1950 to 2023.

A Appendix exhibits

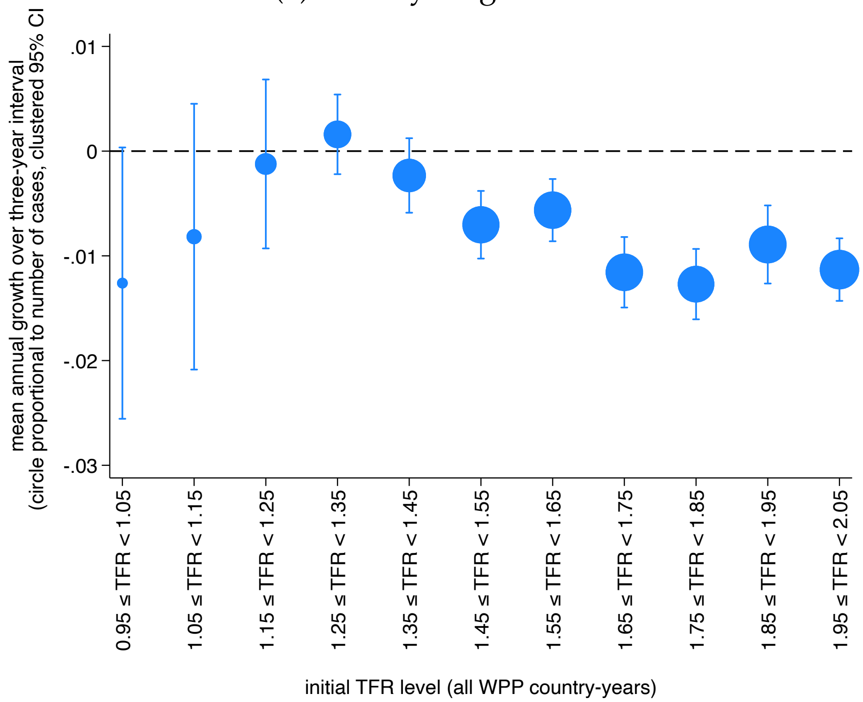
Figure A1: Fertility phase diagrams ($\mathbb{E}[g|f]$), in WPP 2024's Medium projection (disaggregating binned means from Figure 2)



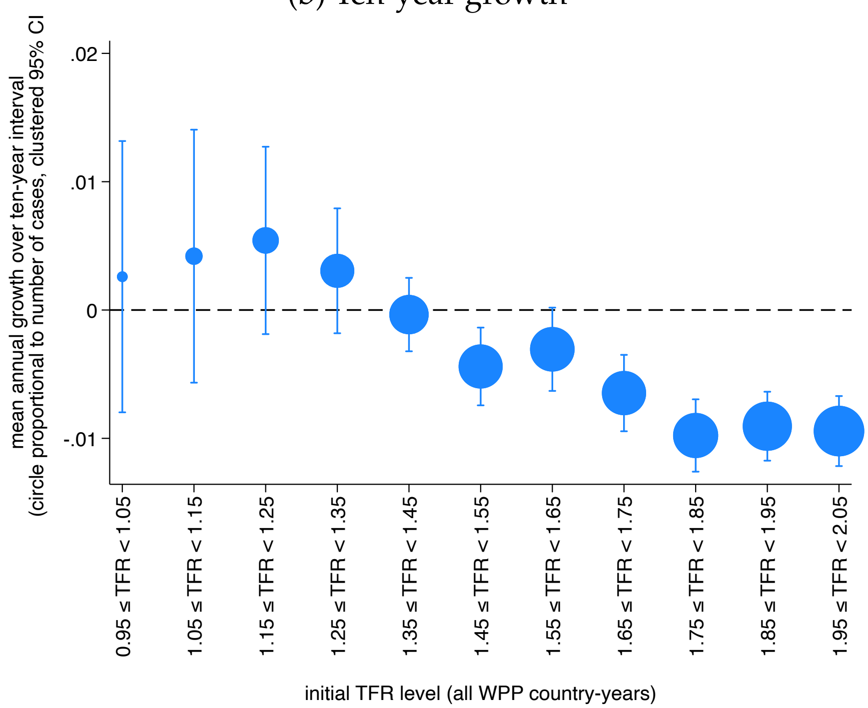
Note: The figure plots future annualized five-year fertility growth rates, computed using Eq. 1, by fertility level in the index year. This can be compared to Figure 2, which plots binned means. Data are the country-level Medium projections from the 2024 UN WPP and span 2027 to 2100. All countries are included except for the Vatican. The horizontal intercept of the log-linear OLS regression line ($g_{c,t} = \beta_0 + \beta_1 \cdot lnTFR_{c,t} + e_{c,t}$) identifies a pooled $f^* = 1.6$ (in fact the UN WPP projection model permits moderately differing f_c^* intercepts for some countries c, and models a first-order autoregressive time series, rather than a growth rate in logs).

Figure A2: Figure 3's main result is robust to using 3-year or 10-year growth instead of 5-year growth

(a) Three-year growth



(b) Ten-year growth



Note: The figure plots binned means of historical annualized three-year (in panel (a)) and ten-year (in panel (b)) fertility growth rates, by fertility level in the index year. This figure can be compared to Figure 3, which plots annualized five-year growth rates. Data are the historical estimates from the 2024 UN WPP, which includes estimates from 1950 to 2023, so the domain of panel (a) stops in 2020 and the domain of panel (b) stops in 2013. 95% confidence intervals are constructed using standard errors clustered by country. Marker sizes are proportional to the number of country-years composing each binned mean.

Figure A3: Log-linear approximations to the TFR phase diagram, by five-year period

Alkema et al. (2011), in their Figure 6, fit a two parameter model to 52 observations of the TFR law of motion in Phase III, with an intercept (which determines an equilibrium fertility point) and a slope (determining the speed of convergence). Motivated by this, we here fit a log-linear model to *all countries'* TFR changes over each five year period using UN WPP 2024 historical data. In particular, we fit two in each panel: One using the whole distribution of TFR, and another restricted to TFR < 2.1. (Our log-linear functional form, expressed in growth rates, is similar to the WPP's AR(1) functional form, expressed in levels, because the objects of interest are small percent changes.) Because each panel plots TFR change over one five-year interval, each country appears once in each panel.

To facilitate the comparison of these parametric results with the hypothesis that 1.6 is an average stable point for TFR (that is, that $f^* \approx 1.6$), each panel includes a confidence interval on the fit model's prediction for g at $f = 1.6$. These confidence intervals are all entirely negative except for growth from 2003-2008, which coheres with the evidence of heterogeneity by period in Figure 8. The final period starts in 2018 because the WPP 2024 historical data end in 2023.

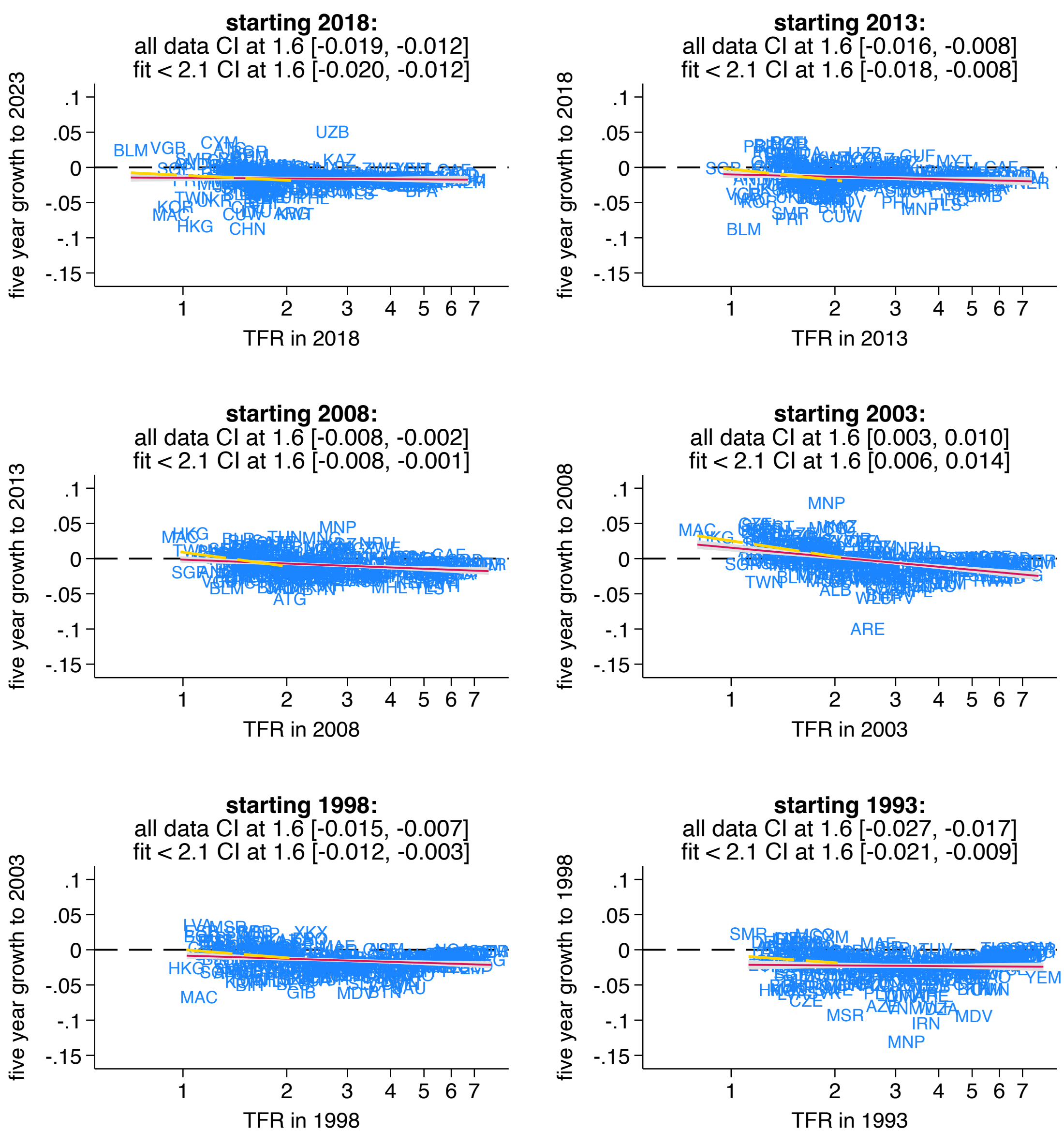
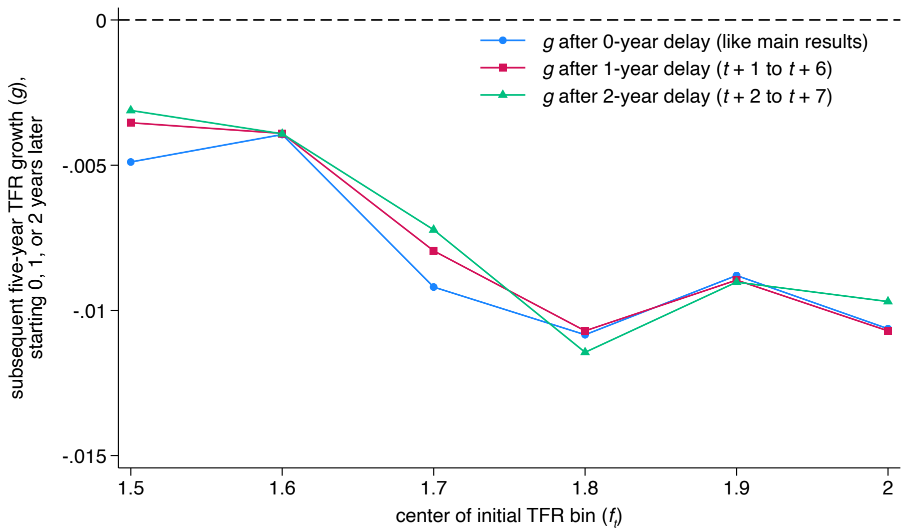
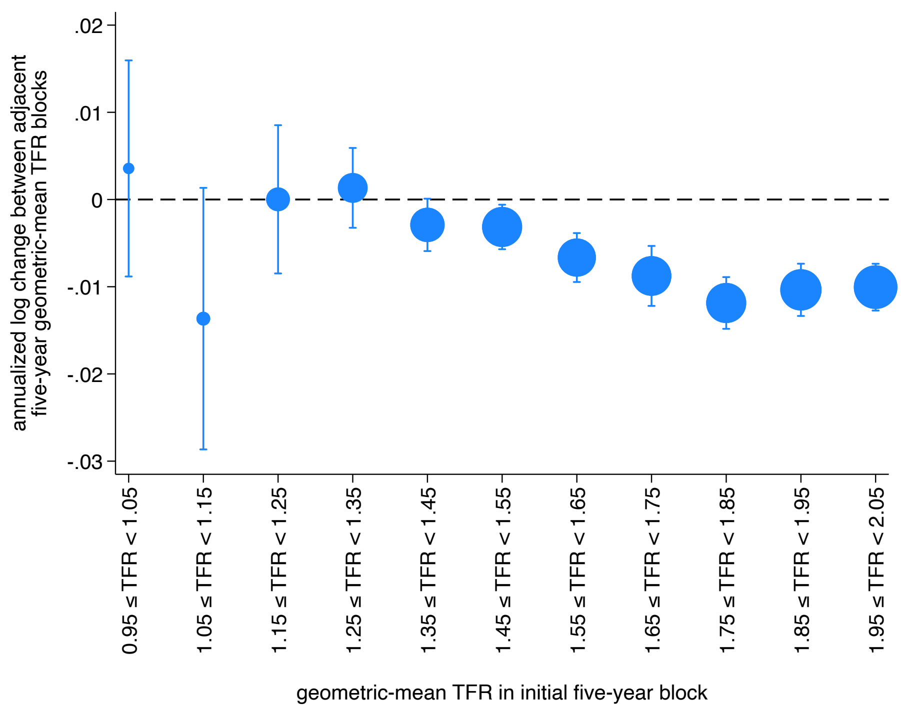


Figure A4: Figure 3's main result is robust to two alternative treatments of transitory TFR variation

(a) delaying the start of the five-year subsequent growth



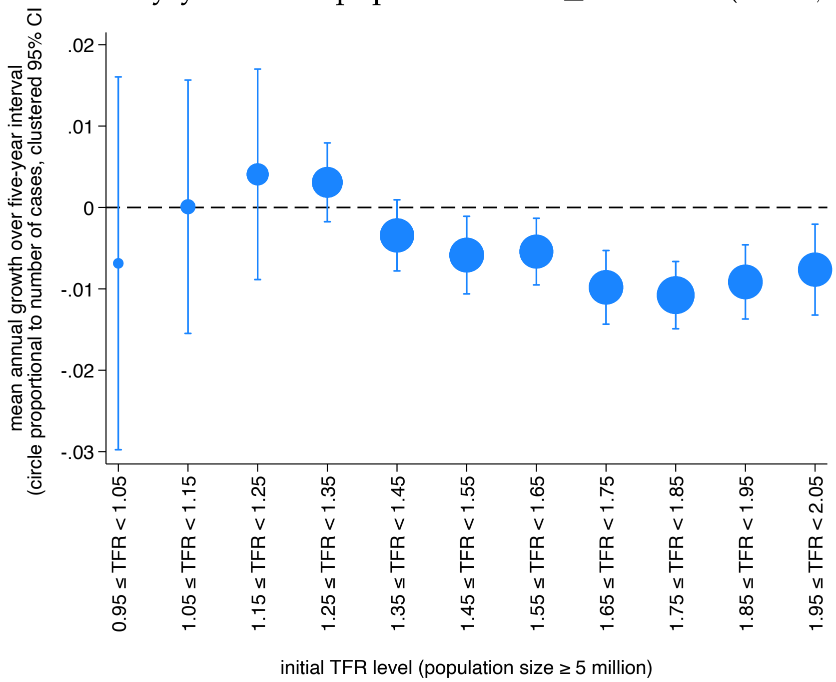
(b) measuring fertility with geometric means in adjacent five-year blocks



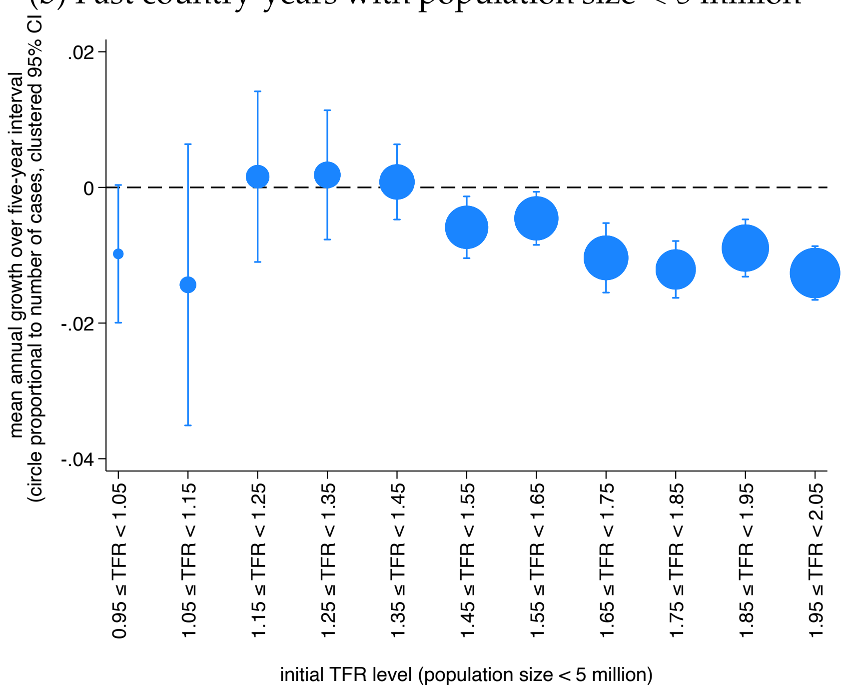
Note: The purpose of this robustness check is to reduce the influence of transitory variation in any single initial or final year (for example, growth g may appear negative because f in a starting year is unusually high). In panel (a), five year TFR growth g is computed after a 1-year or 2-year delay. Because this reduces the sample size, for this panel only the sample is held constant at the set of country-years with 2-year-delayed 5-year growth observed. In panel (b), for each country and index year t, initial fertility is the geometric mean of annual TFR over years $t - 4$ through t, and subsequent fertility is the geometric mean over years $t + 1$ through $t + 5$. Growth is the annualized log change between these adjacent five-year block means. The two blocks within each observation do not overlap; observations for adjacent index years do overlap.

Figure A5: Figure 3's main result is robust within country-years with larger and smaller populations

(a) Past country-years with population size ≥ 5 million ($n = 1,568$)

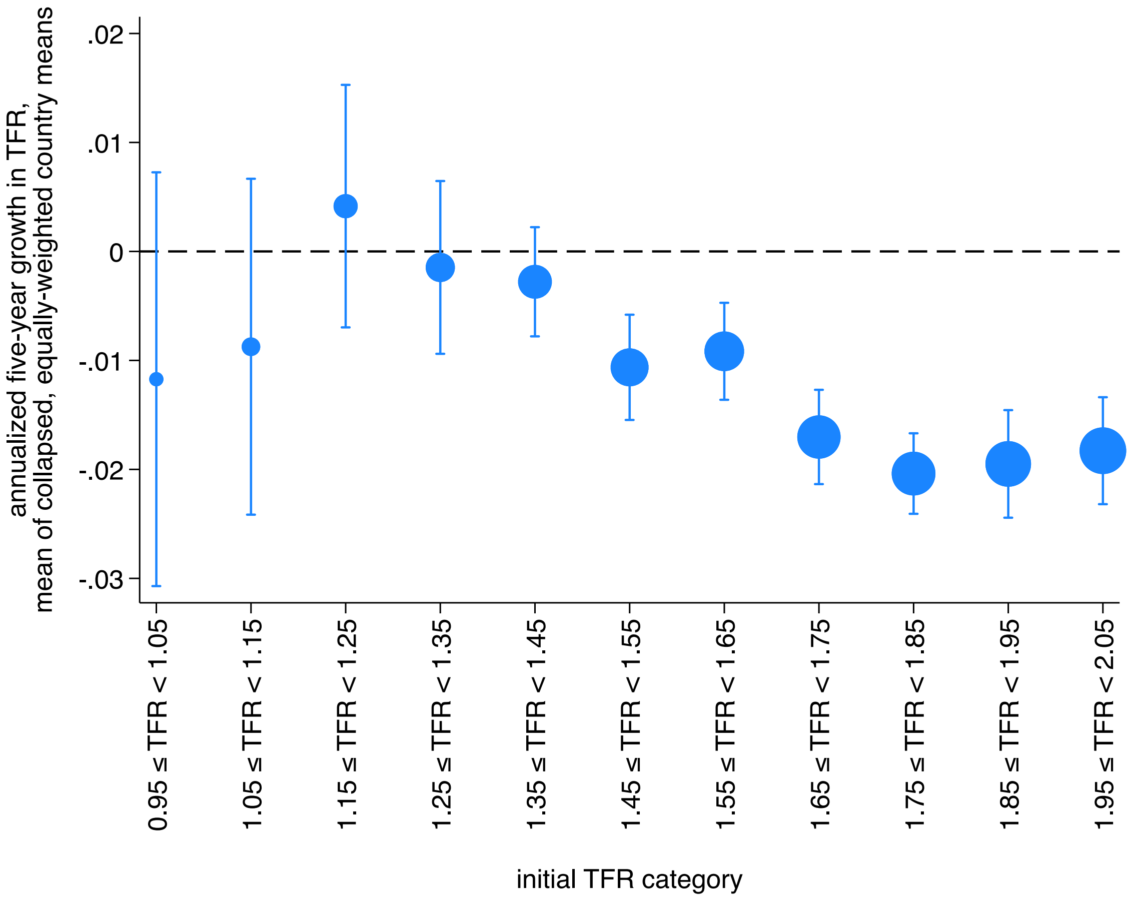


(b) Past country-years with population size < 5 million



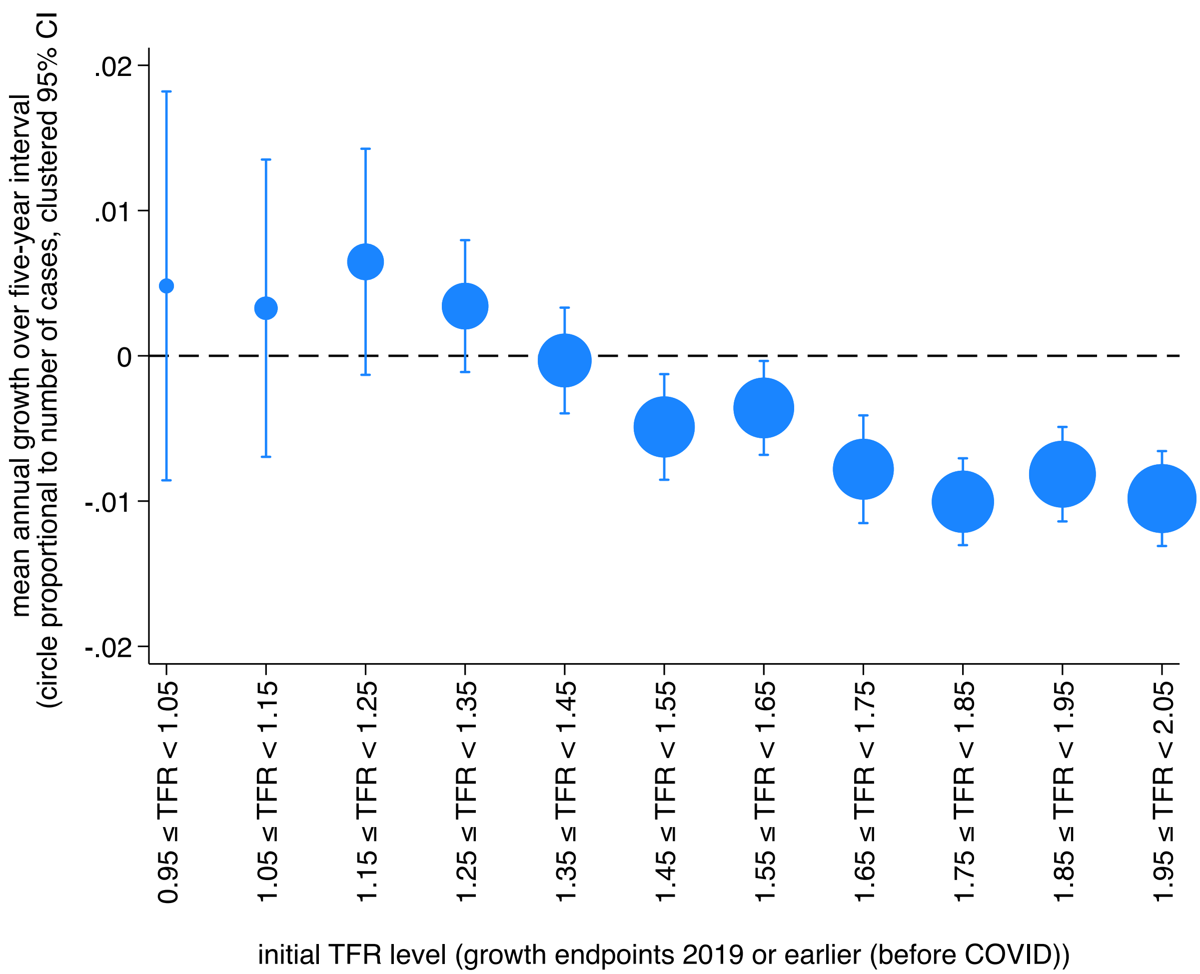
Note: The figure plots binned means of historical annualized five-year fertility growth rates, computed using Eq. 2, by fertility level in the index year. We split the sample roughly in half by population size because many country-years in the 2024 WPP historical estimates have very small population sizes. Panel (a) shows results from country-years with population size ≥ 5 million, and panel (b) shows results from country-years with population size < 5 million. The figure plots estimates for years 1950 to 2018. 95% confidence intervals are constructed using standard errors clustered by country. Marker sizes are proportional to the number of country-years composing each binned mean. Data are the historical estimates from the 2024 UN WPP, spanning 1950 to 2023.

Figure A6: Figure 3's main result is robust to first collapsing the sample within each bin into country-specific means, then weighting those means equally



Note: The main result in Figure 3 takes an average of (f, g) observations, and therefore weights a country more heavily within a bin if it has had more country-years at that TFR. An alternative is to give each country (with at least one observation) equal weight within a bin, by first collapsing into country-specific within-TFR-bin means of g. The figure plots binned means of country means. 95% confidence intervals are constructed using heteroskedasticity-robust standard errors. Marker sizes are proportional to the number of observations composing each binned mean. Data are the historical estimates from the 2024 UN WPP, spanning 1950 to 2023.

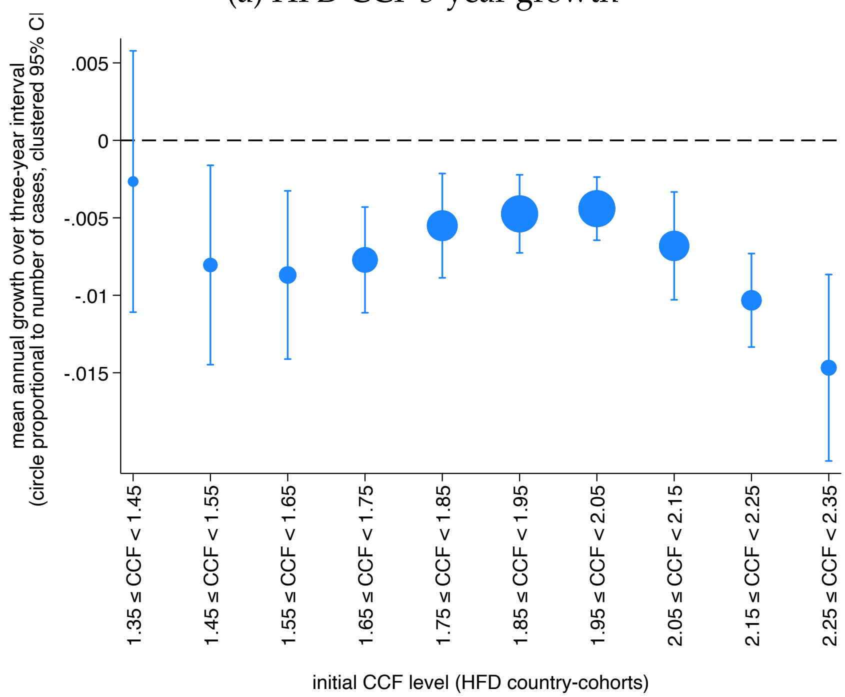
Figure A7: Figure 3's main result is robust to excluding years affected by COVID



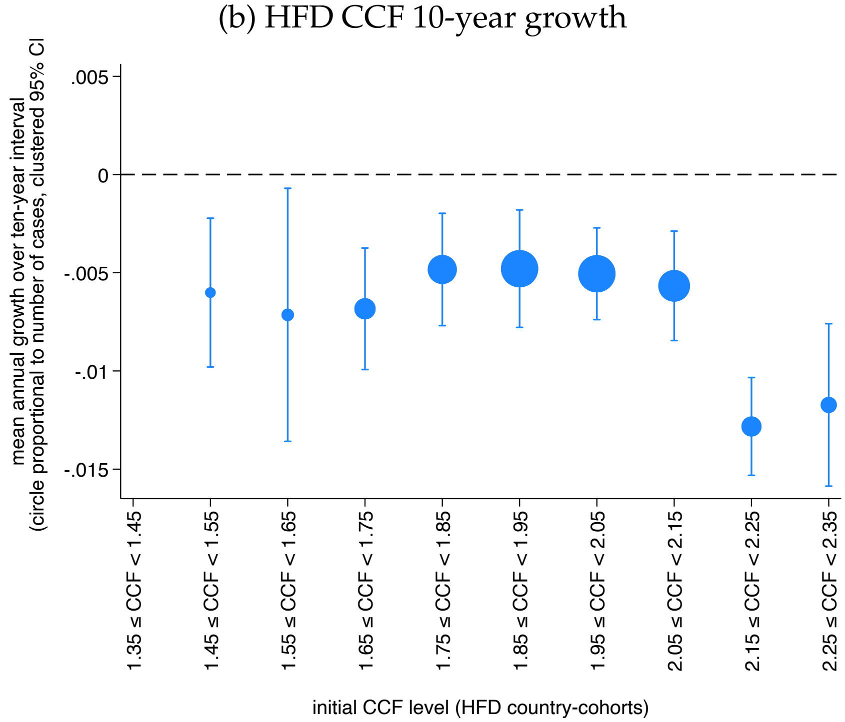
Note: The figure plots binned means of historical annualized five-year fertility growth rates, computed using Eq. 2, by fertility level in the index year. The figure plots estimates for years 1950 to 2014, so that subsequent annualized five-year growth rates do not include years affected by Covid-19. 95% confidence intervals are constructed using standard errors clustered by country. Marker sizes are proportional to the number of country-years composing each binned mean. Data are the historical estimates from the 2024 UN WPP, spanning 1950 to 2019.

Figure A8: Figure 4's cohort result is robust to using 3-year or 10-year growth instead of 5-year growth

(a) HFD CCF 3-year growth



(b) HFD CCF 10-year growth



Note: The figure plots binned means of historical annualized three-year in panel (a) and ten-year in panel (b) growth rates in completed cohort fertility, by completed fertility of the index cohort. This figure can be compared to Figure 4, which plots annualized five-year growth rates. The growth rate measures average change over successive cohorts separated by one birth year. The figure plots estimates for cohorts born between 1935 and 1975. 95% confidence intervals are constructed using standard errors clustered by country. Marker sizes are proportional to the number of country-cohorts composing each binned mean. Data are from the Human Fertility Database.

Table A1: Means, standard errors, and sample sizes for main results

	WPP			HFD			U.S.			Japan			India		
f	$\bar{g}$	s.e.	n	$\bar{g}$	s.e.	n	$\bar{g}$	s.e.	n	$\bar{g}$	s.e.	n	$\bar{g}$	s.e.	n
0.95–1.05	-0.0081	0.0066	23												
1.05–1.15	-0.0070	0.0060	59							0.0081	0.0098	6			
1.15–1.25	0.0029	0.0045	133							-0.0231	0.0101	8			
1.25–1.35	0.0026	0.0023	225							-0.0074	0.0023	63	-0.0064	0.0137	6
1.35–1.45	-0.0015	0.0018	328	-0.0020	0.0063	8				-0.0068	0.0021	100	-0.0240	0.0094	10
1.45–1.55	-0.0059	0.0017	411	-0.0075	0.0024	22	-0.0287	0.0017	3	-0.0104	0.0012	161	-0.0264	0.0087	13
1.55–1.65	-0.0049	0.0014	412	-0.0087	0.0032	31	-0.0170	0.0017	26	-0.0086	0.0019	88	-0.0183	0.0048	24
1.65–1.75	-0.0101	0.0018	424	-0.0077	0.0015	84	-0.0163	0.0035	22	-0.0125	0.0025	32	-0.0177	0.0043	27
1.75–1.85	-0.0114	0.0015	415	-0.0049	0.0016	132	-0.0127	0.0015	72	-0.0234	0.0067	3	-0.0133	0.0039	43
1.85–1.95	-0.0090	0.0016	454	-0.0049	0.0014	195	-0.0112	0.0011	79	-0.0066	0.0064	7	-0.0144	0.0039	65
1.95–2.05	-0.0108	0.0017	488	-0.0043	0.0010	198	-0.0128	0.0019	49	-0.0149	0.0011	2	-0.0205	0.0026	71
2.05–2.15				-0.0066	0.0017	133	-0.0130	0.0033	22				-0.0184	0.0032	46
2.15–2.25				-0.0119	0.0016	54	-0.0170	0.0021	9				-0.0169	0.0023	46
2.25–2.35				-0.0130	0.0029	33									

- WPP results correspond to Figure 3. Standard errors are clustered by country.
- HFD results correspond to Figure 4. Standard errors are clustered by country.
- U.S. and Japan results correspond to Figure 5. Because some estimates have few clusters, reported standard errors are the greater of either clustered (by state for the U.S. and prefecture for Japan) or heteroskedasticity robust.
- India results correspond to Figure 6. Because some estimates have few clusters, reported standard errors are the greater of either clustered (by Indian state, which is one division coarser than these district-level observations) or heteroskedasticity robust.

B The Phase III subsample

In this Appendix Section B, we step away from our main research question to consider the assumptions and methodological implications of Alkema et al. (2011)'s approach to projecting future fertility rates. Alkema et al. (2011) assumed that countries where fertility had fallen below replacement would converge towards an equilibrium at replacement fertility. The empirical task of Alkema et al. (2011)'s research was to estimate the speed of convergence to replacement fertility. Later studies have recognized that replacement-level fertility is an unlikely future, but have maintained a framework that assumes some equilibrium point (or a distribution of country-specific points) that fertility tends towards. For example, in the 2024 WPP projection, this assumption was changed to the assumption that there exists some fertility equilibrium, f^*, such that $\mathbb{E}[g|f \approx f^*] = 0$ and for lower values of fertility, a rebound should be expected: $\mathbb{E}[g|f < f^*] > 0$. This framework has endured even as fertility rates have continued to decline (Sobotka, 2026).

Our paper asks essentially the same question as Alkema et al. (2011) asked about low-fertility country-periods: How do fertility rates tend to change, once fertility rates have been low? Yet, we do not find evidence for the existence of a robust, stable f^*, particularly not at the levels conventionally assumed, such as 1.5 or 1.6. The main results of our paper are different from Alkema et al. (2011)'s in large part because they do not start from the assumption of convergence – so the sample is not conditioned on prior experience of fertility rebound – and in part because they incorporate more recent data.

Here, we investigate the impacts of conditioning on prior experience of fertility rebound. We find that Phase III classification does not identify a persistent regime (in the sense that it does not predict a country's future TFR stability), and that conclusions drawn from the Phase III subsample are sensitive to small perturbations of the sample.

B.1 Description of the "Phase III" subsample

In our main results, we estimate $\mathbb{E}[g|f]$ using all country-years near f. Alkema et al. (2011) investigated the low-fertility stage of the demographic transition using a selected group of 20 countries or areas, which they interpreted to have already started a rebound from a fertility floor. Eighteen of the twenty were in Europe. The other two were Singapore and the United States. None were in Latin America, Africa, or Oceania. This subset excluded Australia, Canada, Cuba, Japan, Mauritius, and South Korea, for example, even though each had experienced a TFR below 2 well before 2010.

Focusing on such a subset will be informative only if its experiences are likely to be generalizable. In this case, the set of 20—which Alkema et al. (2011) call "Phase III" countries—had experienced a sustained fertility increase starting from a level below replacement. Specifically, the Phase III subsample was defined as the countries: "in which two subsequent five-year increases below a TFR of 2 children have been observed." In this way, when estimating the probability and magnitude of future rebound, they exclude countries that had not experienced recent increases in fertility rates from some low.[32] As we

[32] We reprint the sample definition from Alkema et al. (2011, p. 820) for the reader's convenience: "we use a deterministic rule to identify the start of Phase III during the observation period for countries in which the

have discussed in the main text of the paper, the UN WPP adopted the Alkema et al. (2011) method for identifying Phase III countries starting with its 2010 population projections, and continuing up to the 2024 projections. The 2024 UN WPP has classified 39 countries as having already entered "Phase III."

To be classified as "Phase III," a country must demonstrate a fertility increase across the means, all under two, of three consecutive five-year periods spanning a 15-year period. Panel (a) of Figure B1 shows the mean TFR among the 39 "Phase III" countries identified by the 2024 WPP plotted against time, relative to when they entered "Phase III." The 39 countries entered "Phase III" at different times, so the horizontal axis centers country-periods on when they entered "Phase III." On average, the increasing-fertility 15-year span that qualified a country for "Phase III" has been followed by renewed fertility decline.

Panel (b) of Figure B1 shows, by fertility rate, the count of all country-year observations (solid blue line, the sample we use), as well as country-year observations in the Phase III subsamples reconstructed according to the methodology of the 2024 WPP (dashed blue) and Alkema et al. (2011) (dashed red). By restricting the sample to countries which have demonstrated a sustained fertility increase, the 2024 WPP and Alkema et al. (2011) leave out many country-year observations at similarly low fertility rates. For example, comparing the solid blue line to the dashed blue line, the WPP Phase III subsample only includes about one-fourth to one-third of the total number of country-year observations at the same fertility rate.

If the goal is to project fertility rates in a way that extrapolates from the historical record, constructing a subsample in this way amounts to selecting observations on the time-path of the dependent variable. Phase III entry is defined by a prior sequence of fertility increases, so the resulting estimation sample is deliberately enriched for upward-moving low-fertility episodes. Because Phase III estimation begins during the qualifying sequence, some changes used to establish entry also appear in the estimating sample. This is not automatically problematic if Phase III in fact constitutes a distinct and persistent regime. It does, however, make persistence, external validity, and sensitivity empirical questions. Figures B2–B4 examine whether the qualifying rebound continues and how sensitive fitted dynamics are to the inclusion or exclusion of the qualifying country-year observations.

B.2 Phase III countries do not avoid later fertility decline

In Figure B2 we compare the path over time of average TFR among Phase III countries (or UN areas) against the path among other countries that had low fertility but were not selected for Phase III analysis. Panel (a) compares the 20 Alkema et al. (2011) Phase III countries against the 75 countries that had similarly low TFR in 2010—similarly low in the sense of TFR no greater than France's 2010 TFR of 2.017, which was the greatest 2010 TFR among the 20 countries selected as Phase III. Panel (b) compares the 39 countries that

---

fertility transition has ended. The countries that have entered Phase III are defined as the countries in which two subsequent five-year increases below a TFR of 2 children have been observed. By this definition, in the period 2005–2010 the start of Phase III was observed in 20 countries (Belgium, Bulgaria, Channel Islands, Czech Republic, Denmark, Estonia, Finland, France, Germany, Italy, Latvia, Luxembourg, Netherlands, Norway, Russian Federation, Singapore, Spain, Sweden, United Kingdom, and the United States)."

were classified as Phase III in the 2024 WPP revision against 57 comparison countries with similarly low 2010 TFR.[33]

If the Phase III classification had enduring predictive power and robust external validity, we would expect this subsample to show a continuing fertility rebound towards an equilibrium. Figure B2 shows that it does not. In both panels, average fertility in the selected countries eventually falls in parallel to average fertility in the countries that were not selected.

Designation as a Phase III country "predicts" the initial fertility increase that was used to create the subsample, but does not predict a recent path that avoids continued fertility decline. Among the 96 countries that had a TFR no greater than 2.017 in 2010, both those selected as Phase III and those that were not have an average TFR near 1.4 in 2023. This suggests that the increasing pattern of pre-2010 TFR of countries that were classified as Phase III did not give robust reason to expect continued fertility rebound or convergence.

In Figure B3, we deepen Figure B2's finding of mean fertility decline after Phase III designation by adding a non-parametric statistical significance test, and by plotting the distribution of outcomes by countries. The theoretical expectation behind modeling a Phase III is that countries' TFRs will either "increase" or "continue" at a similar level (United Nations, Department of Economic and Social Affairs, Population Division, 2024, p. 30), rather than systematically fall. Here we evaluate this expectation by comparing countries' TFRs in 2023 (which is the latest year in the data) with the country's TFRs at the end (panel (a)) and at the midpoint (panel (b)) of the period that first qualified the country to be classified as Phase III. Here the question is whether TFRs have indeed increased or continued unchangingly after a Phase III qualification, or if they have instead fallen.

The data show fertility decline, even after a country has qualified for Phase III. We test the statistical significance of this pattern with a one-sided, exact binomial test that considers only the sign of the relationship. This test is low-powered and conservative because it dichotomizes the information into simply whether 2023 TFR is lower than the qualifying TFR. The test results show evidence of a systematic tendency for 2023 TFR to be lower than the Phase III qualifying TFR—that is to say, for TFR to fall during Phase III.

Panels (a) and (b) include 14 countries (in green hollow circles) beyond the 39 countries (blue solid circles) identified by the 2024 WPP as Phase III. These 14 have a 15-year run of three increasing five-year means (see Table B1), but they are not recognized by the WPP as Phase III. Our reconstruction suggests that this is generally because the WPP only classifies a country as Phase III *if the qualifying run of 15 years ends in a year divisible by five*. In other words, each of the 14 countries has a qualifying sequence of three consecutively increasing five-year means under two, but perhaps for the years 1992-2006 in the case of Ireland or 1997-2011 in the case of Azerbaijan, instead of, say, 1991-2005 or 1996-2010. We detail all 14 such countries in Table B1.

Notwithstanding this classification issue, the conclusion of systematic subsequent TFR decline is found among the 39 countries listed as Phase III by the 2024 WPP, among the 14 countries in Table B1, and among all 53 countries.

[33] These total counts of 95 and 96 do not match across panels because panel (a) averages Guernsey and Jersey to create the Channel Islands.

B.3 Subsequent fertility change, and implied fertility equilibria, are unstable as the Phase III subsample changes in small ways

Figure B4 documents that the way fertility behaves in Phase III country-years is sensitive to the inclusion of the initial period of increasing TFR that qualifies a country for Phase III. More precisely, the Phase III data for each country consists of two parts: first 8 country-years that are part of the 15-year run of increasing average TFR that generates the Phase III designation,[34] and second all of the subsequent country-years in the historical record. Figure B4 reports the results of omitting those 8 initial country years, starting with our reconstruction of the full Phase III sample and omitting each country's first Phase III year, first and second Phase III years, and so on. After omitting the first 8 years, only data *after* a country's qualifying run of increasing TFR is used in the estimation. The horizontal axis in each panel of Figure B4 is the number of qualifying years omitted.

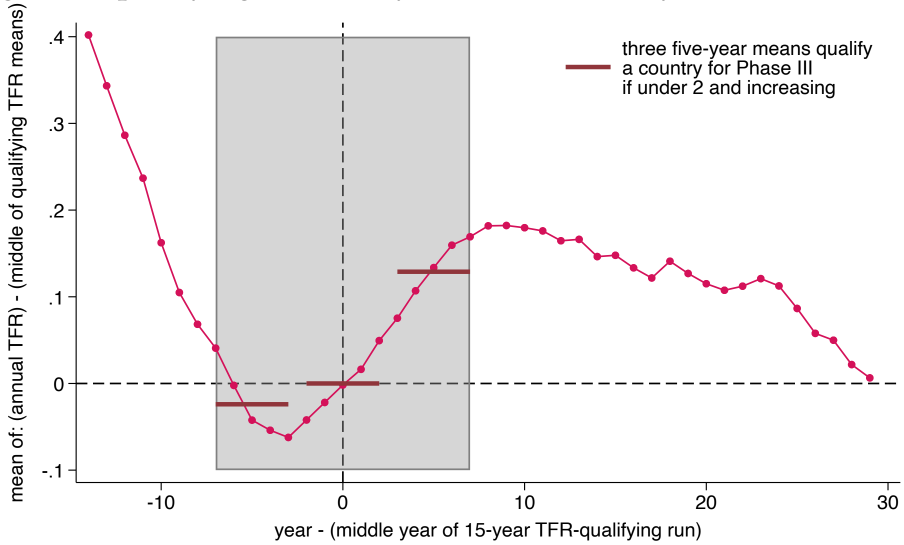
Panel (a) of Figure B4 shows that using both OLS and (the shared global part of) hierarchical models, the conclusion of zero TFR change near 1.6 is sensitive to the inclusion of the qualifying period of TFR increase. As these data years are omitted, the subsequent change in TFR at 1.6 becomes robustly negative. Panel (b) shows that the horizontal intercepts,[35] which are the implied zero-change TFR crossing points, of the same models as in panel (a) are highly sensitive to these changes in the sample. Estimates span far beyond the plausible range of low TFR values, and in many cases the denominator (the estimated slope) is not statistically significantly negative, so no attracting root is identified. Panel (c) verifies that these sensitivities occur even though mean TFR is only very slightly changed by excluding the increasing, qualifying observations. This coheres with the increasing TFR pattern, rather than the TFR level, as the source of the instability.

[34]This is 8 years and not 15 because the methodology starts the Phase III subsample from the midpoint of the 15-year run.

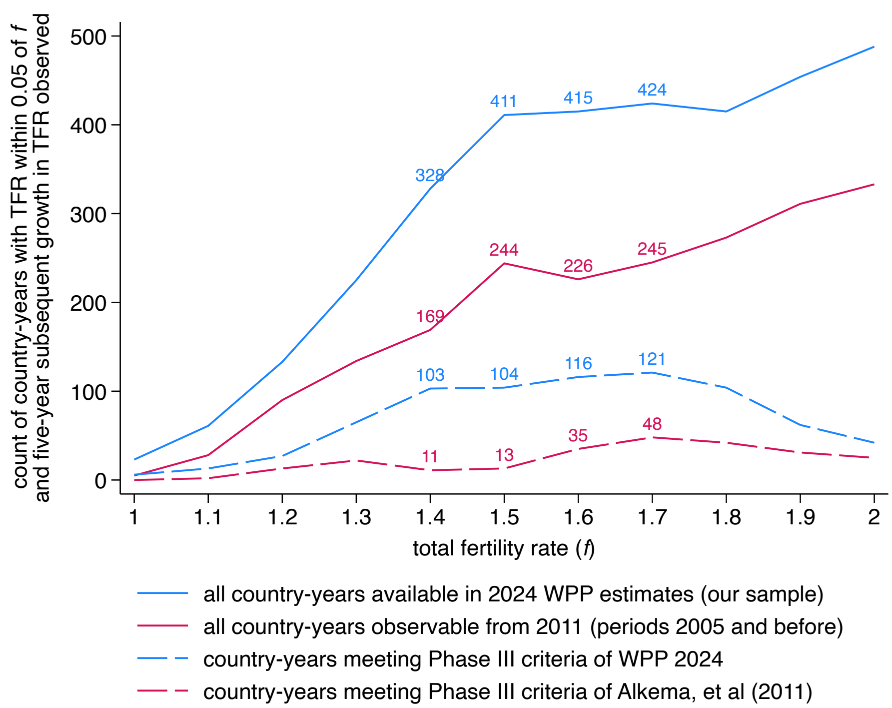
[35]Computed as $(-1 \times \frac{\text{estimated constant}}{\text{estimated slope}})$.

Figure B1: The construction and size of the WPP's Phase III sample

(a) The 2024 WPP constructs its Phase III sample of 39 countries by selecting country-years starting at the midpoint of three successively-increasing five-year TFR means, all under 2. On average, this qualifying run of 15 years is followed by renewed fertility decline.



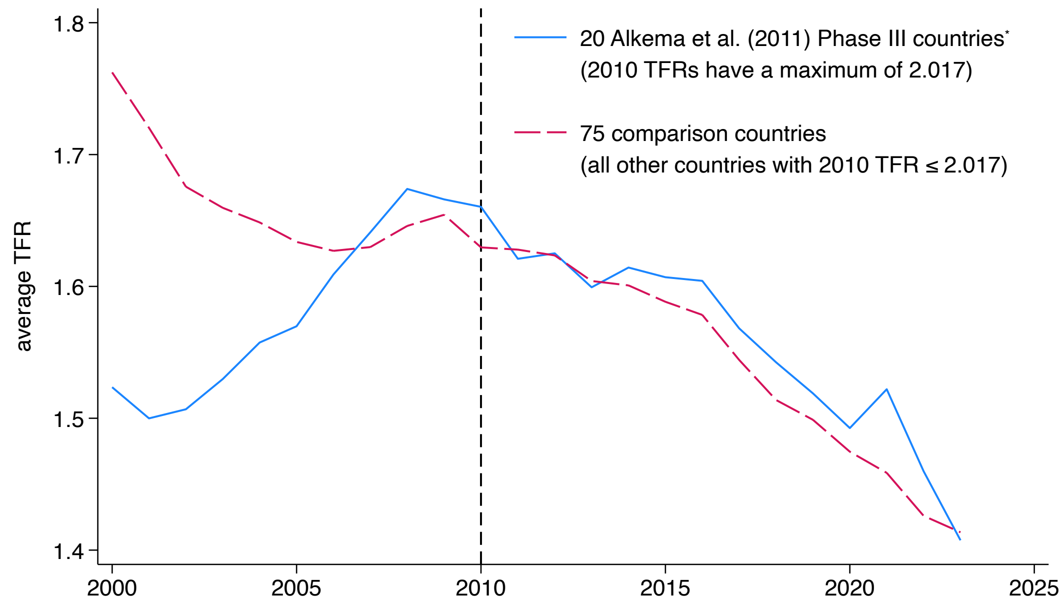
(b) Retrospective reconstruction of n counts corresponding to full and Phase III samples



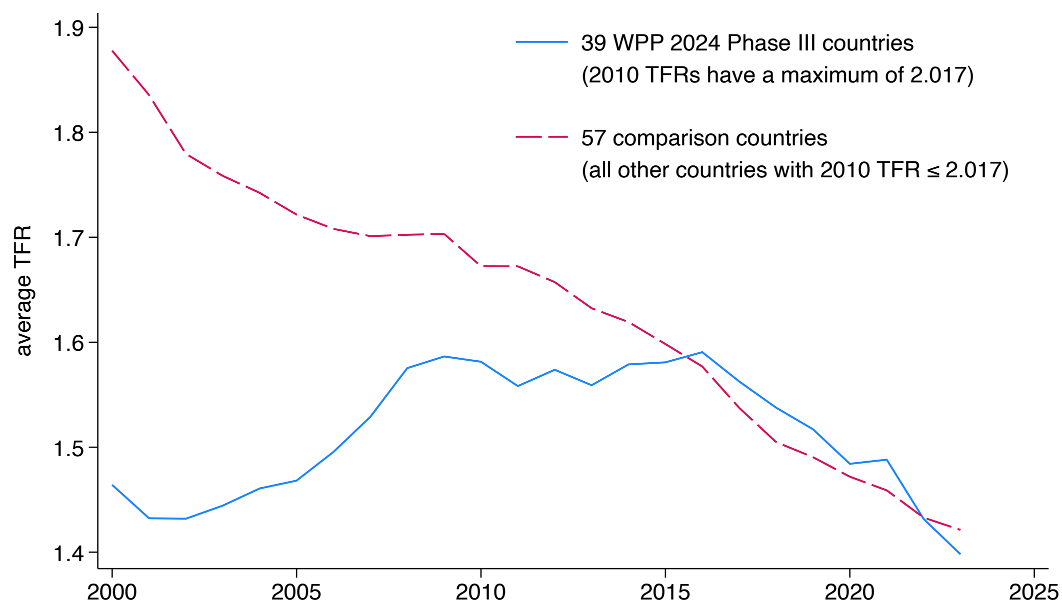
Note: Panel (a) shows the mean TFR among the 39 Phase III countries identified by the 2024 WPP just before and after having entered Phase III. The 39 countries entered Phase III at different times, and the horizontal axis centers country-periods on when they entered Phase III. Each event-time mean uses countries observed at that horizon, so country coverage declines at longer horizons. Panel (b) is a retrospective reproduction of sample inclusion criteria, not a mechanical replication of prior research. It shows, by fertility rate, all country-year observations in the 2024 UN WPP (solid blue), all country-year observations observable from 2011 (solid red), and country-year observations in the Phase III subsamples of the 2024 WPP (dashed blue) and Alkema et al. (2011) (dashed red). Data are from the 2024 UN WPP.

Figure B2: The countries identified for Phase III by Alkema et al. (2011) (panel (a)) and the 2024 WPP (panel (b)) later had downward fertility trajectories, on average

(a) As classified by Alkema et al. (2011)

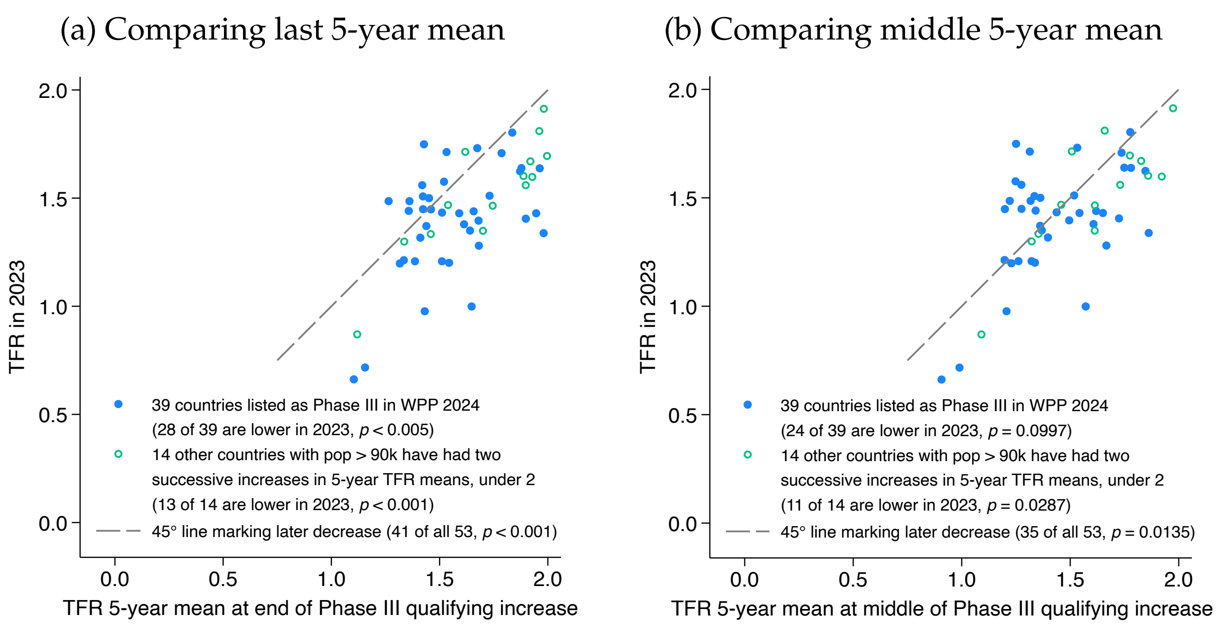


(b) As classified by WPP 2024



Note: Panel (a) compares the 20 Alkema et al. (2011) Phase III countries against the 75 countries that had similarly low TFR in 2010—similarly low in the sense of TFR no greater than France's 2010 TFR of 2.017, which was the greatest 2010 TFR among the 20 countries selected as Phase III. Panel (b) compares the 39 countries that were classified as Phase III in the 2024 WPP revision against 57 comparison countries with similarly low 2010 TFR. These total counts of 95 and 96 do not match across panels because Alkema et al. (2011) include the Channel Islands as one of the 20 Phase III countries, but the 2024 WPP includes Guernsey and Jersey separately, rather than the Channel Islands as a whole. For panel (a) only, and not elsewhere in our paper, we average the annual TFR in Guernsey and Jersey to create the Channel Islands.

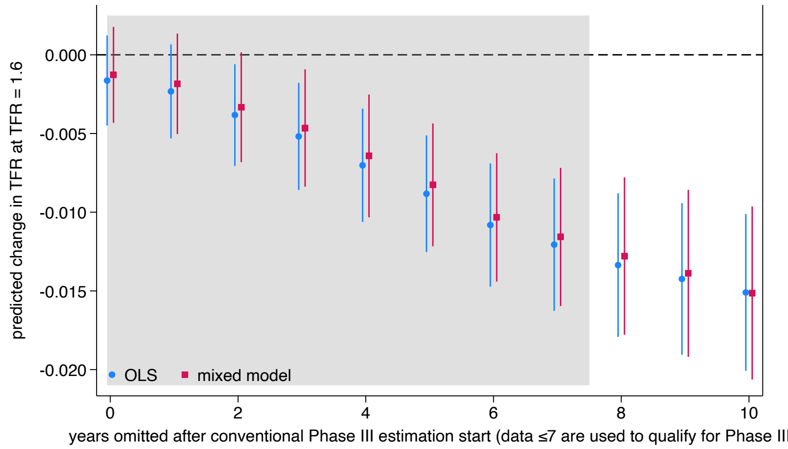
Figure B3: Phase III countries' (and unofficial Phase-III-like countries') TFRs systematically tend to be lower in 2023 than when the country met the criteria for Phase III classification



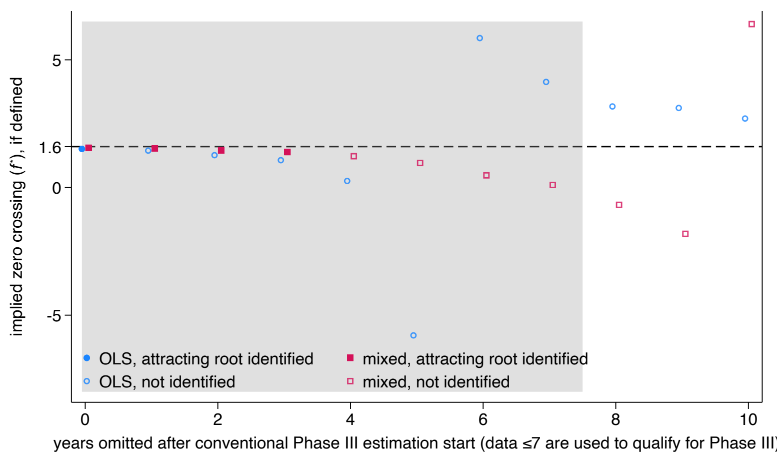
Note: The figure compares countries' TFRs in 2023 (which is the latest year in the data) with the country's TFRs in the period that first qualified the country to be classified as Phase III. Panel (a) compares to the country's TFR at the end of the qualifying period and panel (b) at the midpoint. The solid blue circles are the 39 countries identified by the 2024 WPP as Phase III. The hollow green circles are 14 countries that have a 15-year run of three increasing five-year means, thereby meeting the criteria for Phase III, but are not recognized by the WPP as Phase III. Our reconstruction indicates that this is generally because the WPP only classifies a country as Phase III if the qualifying run of 15 years ends in a year divisible by five.

Figure B4: Phase III results are sensitive to omitting from the estimating sample the initial years of increasing TFR that qualify a country for Phase III classification

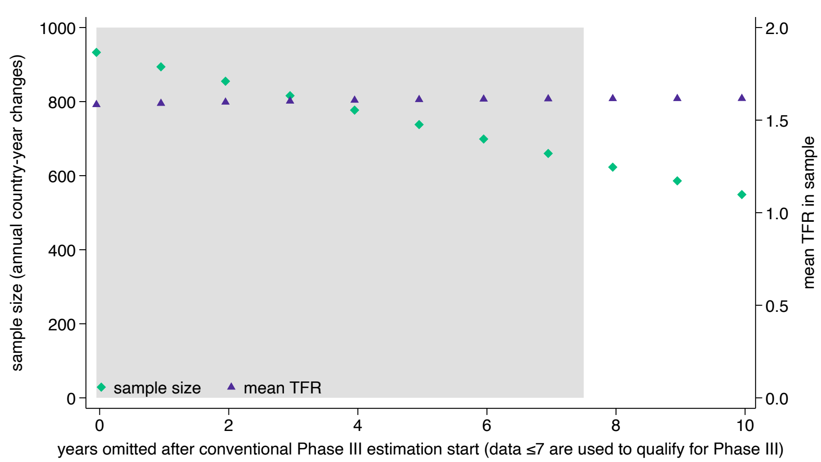
(a) Predicted change at TFR of 1.6 becomes negative without initial qualifying years



(b) Any implied crossing f^* is unstable as the sample changes and often unidentified



(c) About one-third of the Phase III sample is in the initial, qualifying years, but mean TFR is similar in non-qualifying years



Note: This figure is a diagnostic analogue, not a re-estimation of the full WPP model: Following the 2024 WPP and its designation of 39 Phase III countries, we fit linear AR(1) models for change in TFR as a function of TFR—both pooled OLS models and a hierarchical random-coefficients (or mixed-effects) model, from which we report the global component. Moving along the horizontal axis reflects progressively restricting the sample to omit no years, the first year, etc. that a country is classified as Phase III. Up to the 7th, these data years' increasing TFR profile is used to qualify a country for the Phase III sample. Panel (a) shows the model-predicted subsequent change in TFR at starting TFR of 1.6. Panel (b) shows the horizontal intercept from each model, which gives the implied f^*. Hollow markers indicate that the attracting root is not identified, which means that the slope of the linear AR(1) model is not statistically significantly negative. Panel (c) shows mean TFR and the sample size for each estimation sample.

Table B1: Countries appearing to meet Phase III criteria but not listed among WPP 2024 list of 39

country	qualifying run	1st mean	2nd mean	3rd mean	2023 TFR	2023 lower?
CAN	1981 - 1995	1.612	1.613	1.700	1.349	yes
IRL	1992 - 2006	1.894	1.922	1.928	1.598	yes
AZE	1997 - 2011	1.774	1.827	1.919	1.670	yes
GBR	1997 - 2011	1.676	1.730	1.898	1.560	yes
GRC	1997 - 2011	1.336	1.354	1.459	1.334	yes
HRV	1998 - 2012	1.449	1.459	1.539	1.468	yes
ABW	1999 - 2013	1.833	1.859	1.888	1.602	yes
ARM	1999 - 2013	1.457	1.507	1.619	1.714	no
GEO	1999 - 2013	1.591	1.659	1.961	1.810	yes
POL	2000 - 2014	1.277	1.322	1.337	1.299	yes
MKD	2002 - 2016	1.608	1.613	1.745	1.465	yes
IRN	2003 - 2017	1.768	1.775	1.996	1.695	yes
TWN	2005 - 2019	1.084	1.091	1.119	0.870	yes
VNM	2005 - 2019	1.905	1.974	1.982	1.913	yes

Note: To understand this table, consider the example of Ireland. Over the 15 years from 1992 to 2006, Ireland's period TFR showed three successive five-year means that increased and that were each under two. However, Ireland is not classified as a Phase III country in the WPP 2024. Our reconstructions suggest that this is because the WPP 2024 classifies a country as in Phase III only if the qualifying fifteen-year run ends in a year divisible by five (so a qualifying run ending in 2005 would count, but 2006 would not).[a] Such a restriction would be substantive and may not reflect fundamental demographic processes.

[a]Based on the published methodology, our investigations do not explain why the 2024 WPP does not classify Canada as a Phase III country (United Nations, Department of Economic and Social Affairs, Population Division, 2024, p. 30, footnote 48). We therefore reconstructed our computation using independently sourced Government of Canada TFR data:

5-year span	year 1	year 2	year 3	year 4	year 5	mean TFR
1981-1985	1.65	1.63	1.62	1.62	1.61	1.626
1986-1990	1.59	1.58	1.60	1.66	1.71	1.628
1991-1995	1.72	1.71	1.69	1.69	1.67	1.696

These published values yield two successive increases in five-year means below two. However, the first increase is only 0.002, so its sign may be sensitive to rounding. According to this national data series, every Canadian TFR after 2015 has been less than 1.60.

Source: Provencher and Galbraith (2024) `https://www150.statcan.gc.ca/n1/pub/91f0015m/91f0015m2024001-eng.htm`