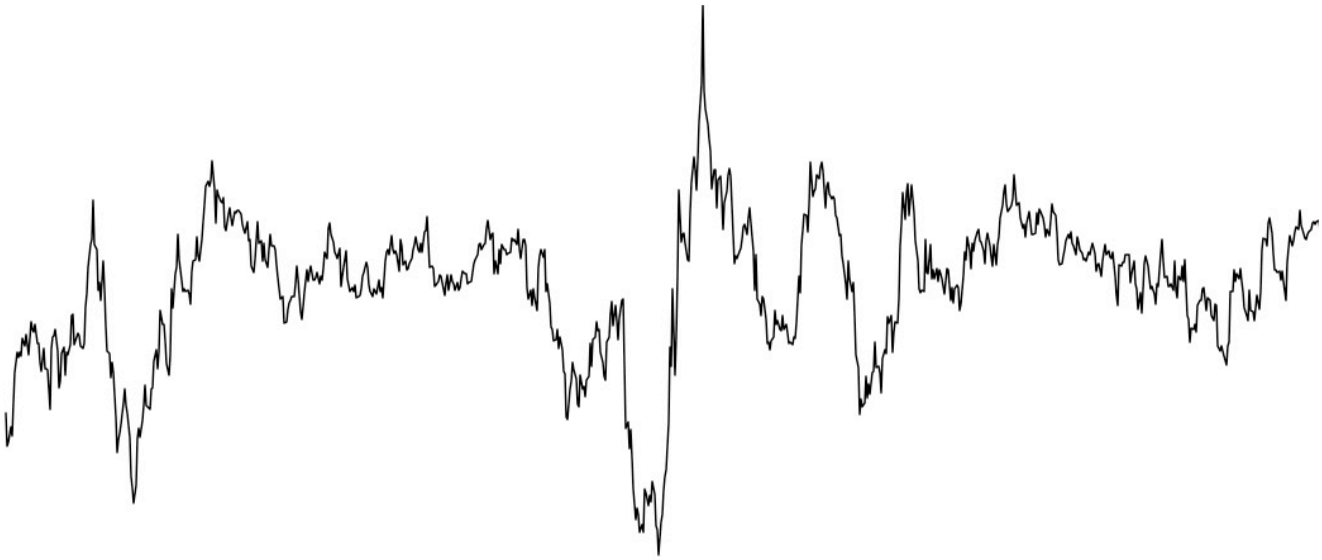
ALPHA SOURCES

MARCH, 2024



DELAYED GRATIFICATION

Why are global birth rates falling, and does it matter?

Global birth rates are falling, and the decline appears to be accelerating in one country after the other. Why is this happening? The previous chapter ended with the observation—by Kaplan et al (2002)—that increasing returns to investment in embodied capital for women, proxied by longer education and a higher rate of full-time labour force participation, could plausibly explain the emergence of sustained sub-replacement level fertility. In doing so, the paper connects evolution and anthropology—predominantly focused on the quantum effect of fertility—and the social sciences.

The social sciences have been pondering birth postponement and sub-replacement fertility since the 1980s under the moniker of the second demographic transition, SDT, a term coined by Lesthaeghe and van de Kaa (1986) and van de Kaa (1987). At its core, the SDT is concerned with the same question that has been confounding evolutionary anthropologists. Why does fertility seem to be falling

well below replacement level in one country after the other, and should we expect this to be a permanent state of affairs.

The SDT attempts to describe and explain the second major onset of falling fertility in a post-Malthusian context, specifically the trend after the Second World War, and in particular since the 1970s. The analysis of tempo effects of fertility that takes place within the broad socioeconomic field of life course research Hunt (2005), which is concerned with the sociological analysis of the timing of key events such as the age of leaving home, first job, marriage, first birth, retirement etc. The life course is sometimes mentioned interchangwably with the life cycle, but I will keep a sharp distinction here. I will refer to the life cycle in its narrow form from economics where it describes the flow of income, consumption and savings over an individual's life time.

The main distinguishing feature between a first and a second demographic transition is the empirical observation of a sustained rise in women's mean age of first birth. This phenomenon drives a marked fall in period fertility, TFR, and potentially as a result, total cohort fertility as women end up having fewer children than they want, or because desired fertility itself falls over time. The distinction between period and total cohort fertility is a crucial one in the SDT. The likelihood that so-called missing births today, due to postponement, will arrive later must be held against the obvious near-term effects on population ageing from falling fertility today, and the potential for low TFR to feed through to total cohort fertility over time. Population ageing is a key socioeconomic phenomenon worthy of analysis in its own right, which I will treat in subsequent chapters. More generally, for fertility research, the issue for scholars is that cohort fertility is the variable of interest, but that it is only observed with a significant lag.

Marked and sustained birth postponement is the primary empirical manifestation of the SDT for an analysis focused on fertility, but it is part of a broader palette of social, cultural and technological changes that characterise the idea of an SDT. These include a decline in the rate of marriage, an increase in cohabitation outside marriage, a rise in the share of births outside marriage, the emancipation of women—higher education and labour force participation rates—and an increase in the prevalence and use of contraceptive technology. Lesthaeghe (2010) describes three revolutions; a gender revolution, a contraceptive revolution and a sexual revolution, all tied into a revolt and re-consideration of key institutions such as the state, family and religion.

The following juxtaposition between a first and a second demographic transition, alongside three dimensions, is inspired by Table 1 in Lesthaeghe (2014).

Marriage - A transition from a high rate of marriage at low age, low divorce rates, and a high rate of re-marriage after divorce of widowhood to falling marriage rates, an increase in age of first marriage, a rise in cohabitation without marriage, rising divorce rates and a low incidence of re-marriage after divorce or widowhood.

Fertility - A transition from overall falling birth rates due reduced fertility at older age—but still above replacement fertility—falling mean age at first birth and deficient contraception to birth postponement, rising mean age at first birth, sub-replacement fertility, efficient contraception, rising non-marital fertility, and rising childlessness for married women.

Socioeconomic factors - A transition from basic lower order needs—income, food and housing etc—to higher order needs such as self-realisation, individual autonomy, and recognition. A trans-

ition from strong ties and memberships to community, political and religious networks to a more individual and atomised structure. A transition from a first secularisation—organisation around high order political and social structures—to a second secularisation characterised by rejection of traditional authorities, a sexual/gender revolution and the like. A transition from fixed to fluid gender roles, and a shift in equality between genders.

Lesthaeghe (2010) and (2014) identify three intellectual sources for the SDT. French historian Philippe Ariès, Ariès (1962) and Ariès (1980), suggest that the fall in fertility during the demographic transition is driven by two distinct motivations, which follows closely the framework described up until now. The early stages of the DT is characterised by Ariès as the "child-king era", in which falling fertility is "unleashed by an enormous sentimental and financial investment in the child". This is just another way to frame the quantum effect of fertility discussed above via Becker and Kaplan in which parental investment is increasingly devoted to offspring quality, instead of quantity. Granted, Lesthaeghe (2014) invokes the idea of a "an altruistic investment in child quality", which denotes a somewhat different process than the incentive-driven shifts in trade-offs described by Kapland and Becker, but it produces the same result in the end. In the second part of DT, however, Ariès describes a shift towards adult self-realisation, and presumably resource accumulation. This, in turn, moves resources away from reproduction or at least shifts the timing of births.

The second driving force of the SDT is linked directly to the observation of the rising prevalence of sub-replacement level fertility, and more specifically, to the failure of the mechanisms to ensure stable fertility at replacement levels in the final stages of the transition. Lesthaeghe (2014) specifically takes aim at Easterlin's cyclical fertility theory which predicts stabilisation in underlying fertility

over time as the kids of low birth cohorts will have tendency to have higher fertility than their parents, and vice versa.

The third driving force is less defined, but is broadly related to cultural and, according to Lesthaeghe (2014), "ideational" shifts which have worked alongside shifts in economic structures and incentives to produce a new demographic regime, clearly distinct from the theoretical end-point predicted by original transition theory. This sounds profound, but it is also a broad catch-all argument that is difficult to verify, let alone refute. It gets more concrete with the reference to a shift along the lines of Maslov's theory of changing needs as a function of income and overall security. As society has become wealthier, needs and desires have shifted from basic needs such as security, survival and solidarity to self-realisation, social status, recognition etc. This, again, sounds intuitively comfortable, especially in the context of the idea from LHT that an increasing share of resources devoted to somatic growth—investment in oneself—can be a drain on reproductive effort. In a sociological context, however, this explanation also sets up a rather crude dichotomy between nominally "old" values such as family, community and "new" values such as individualism and self-realisation. It is not clear to me that such a distinction is sophisticated enough to describe the underlying cultural and shifts associated with a fall fertility in the latter part of the 20th century.

Cultural evolution theory offers a framework to understand how such shifts can permeate over time via the idea of socially-transmitted ideas and values Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985). CE theory operates with two broad categories of diffusion of ideas; vertical diffusion which denotes behaviour and values passed on between direct descendants; from mother to daughter, father to son and so on. And horizontal diffusion, which covers the permeation of ideas and values in society as a whole

through media, social interaction, work relationships and everything in between. Collran and Lu (2023) provide evidence, Murphy (1999) Reher et al. (2008) and Jennings et al. (2012) for an increase in generationally transmitted behaviour of key variables such as age at first birth, age at marriage, ideal family size, contraceptive use and childrearing practices. The key question, however, is whether such transmission reflects the permeation of cultural values, according to Murphy (1999) Reher et al. (2008) and Jennings et al. (2012), or whether they're grounded in genetics, or both.

In an evolutionary context, the idea of horizontal diffusion of ideas and values is one way to introduce the idea of maladaptive fertility in a modern context. This is because the structures that drive horizontal value diffusion are non-kin, (Newson et al., 2005), and as such have no direct stake in the reproductive success of the women and men they influence (Hamilton, 1964). Specifically, following the idea of memetic evolution, in which memes and ideas battle for supremacy in an evolutionary contest, a contradiction could emerge between non pro-natal values for reproduction and the reproductive behaviour that would be optimal for the individual. This, in turn, could be one explanation for the evolutionary conundrum in which modern fertility behaviour seem to contradict basic evolutionary fitness tenets, and why such behaviour is growing in prevalence as countries move up the economic and informational value chain.

Even with the aid of cultural evolution, the SDT is a difficult framework to grapple with, mainly because it allows for so many different interpretations of the post-1970s demographic landscape that it is difficult to see how a unifying explanation is possible. The SDT literature, and more generally the socioeconomic literature on fertility in a modern context spans everything from empirical studies to social constructionist research, which relies on interpretation and sometimes normative value judgement. The question is whether

▸▸

the SDT as set out above is able to adequately capture and contain the evolution of fertility and birth rates, primarily in the developed world, while allowing for the existence of significant idiosyncratic factors across countries and regions. I think it can, just about, if we allow ourselves to pick a number of the key ideas from the SDT

SHOW ME THE NUMBERS

The SDT hypothesis makes a number of postulates and predictions, many of which can be easily verified by empirical analysis. The first is a decline in marriage rates, and by extension a fall in the kind of union formation that would normally be associated with relatively high and stable fertility. It is assumed in the SDT framework that an rise in the incidence of out-of-wedlock births is positively correlated with overall falling fertility.

Data from *Our World in Data* show that marriage rates have been in broad decline since the beginning of the 1970s in developed economies. In the US and the EU, the marriage rate per 1000 people stood at 10.6 and 7.9 in 1970, respectively, and by 2016, it had declined to 7.00 and 4.40, respectively. The trend is similar in countries such as Japan, South Korea and Australia,

The main feature of the SDT, meanwhile, is an increase in the incidence of sub-replacement fertility due to birth postponement, or tempo effects. This, apart from being associated with a decline in marriage rates, also is linked to a number of socioeconomic shifts and the advent of more efficient contraception. In the first chapter, I showed that a rising number of countries now have sub-replacement fertility, but let’s do a quick re-cap.

By the end of the 2010s, global fertility was still above the replacement level of just over 2 children per woman. Based on complete

data for 2022, the UN recently estimated that global period fertility, TFR, will fall gently over the the next 40 years, dipping below the replacement level in 2065. This prediction flies in the face of the speed with which fertility is falling based on the most recent data, and as I show below, global fertility will fall below the replacement level much sooner than that, based on the current trend.

In the developed world, fertility in Europe has long since fallen below the replacement level, and more recently, in the 2010s, it has declined below two in North America too. Elsewhere, fertility was on the verge on falling below two in both Latin America and Asia by the end of the 2010s, and these regions can now likely, for all intent and purposes, be characterised as having sub-replacement level fertility. In Africa, fertility is falling but is still well above 2.

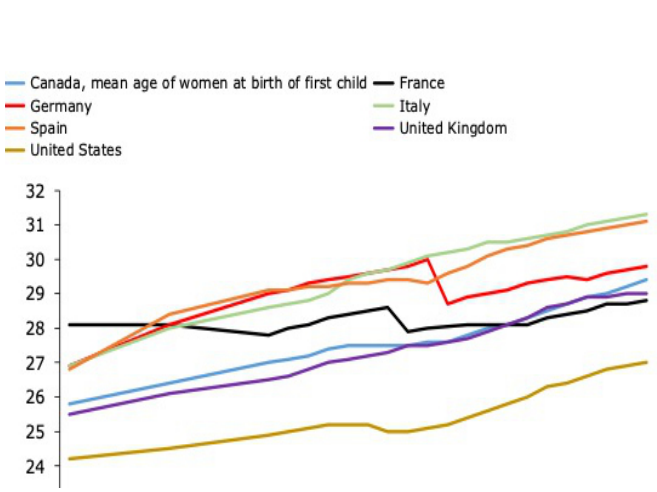
At of the end of the 2010s, the number of countries in the world with sub-replacement fertility was still rising, but the number of countries with very low fertility, defined here as a TFR below 1.5, seems to have peaked in the beginning of the 2000s, hinting at some catch-up from tempo effects Goldstein at al. (2009). It is important here to recall the distinction between period fertility—TFR—which is a snapshot of fertility in a given year, and cohort fertility, which measures completed fertility for women in a given cohort. A negative quantum effect—the tendency of women to have fewer children through their total reproductive career—will impact period and cohort fertility equally. By contrast, tempo effects will reduce period fertility, but not necessarily cohort fertility as births are pushed forward in time. The key question, however, is the extent to which strong tempo effects exert a downward effect on cohort fertility as the missing births are not fully recuperated over time.

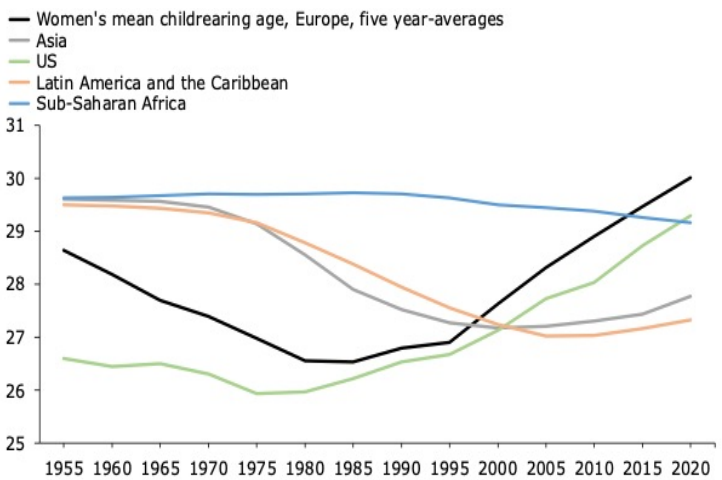
The decline in period fertility to the sub-replacement level in one country after the other is clearly linked to tempo effects. We have two pieces of data to show this. First, numbers from UNECE showing the mean age of women at the birth of their first child. Secondly, the UN publishes age-specific fertility data, which is "the annual number of births to women of a specified age or age group per 1,000 women in that age group" UN (2012).

The analysis below relies on a 2021 panel from the UN, with estimates through 2050 in line with the UN's general population projections. This dataset is used to compute the third indicator, the so-called mean age at childrearing indicator—MAC—which is defined as the "the mean age of mothers at the birth of their children if women were subject throughout their lives to the age-specific fertility rates observed in a given year" UN (2012).

The UNECE data display odd kinks in the data for Germany and the UK, but broadly show that that women's mean age at first birth has increased in the developed world since the beginning of the 1990s, with the exception of France. These data also reveal significant differences in level across countries. In the US, for example, the mean

CH 01 / Births are happening later - **CH 02 /** The tempo effect in action





Source: UN and UNECE

age at first birth had increased to around 26 years in 2019, from 24 in 1990, while the corresponding numbers in Italy and Spain were around 30, up from 27 in 1990. In other words, the UNECE dataset suggests that the mean age of women at first birth in the US today is what it was in Italy and Spain in 1990. This is consistent with overall fertility in the US being substantially higher during this period than in southern Europe.

The UN's mean childrearing age data paint a more complete picture, stretching back to the 1950s, but remember that these data need to be compared to overall fertility to provide a perspective on true tempo effects. This is because the MAC measures the average age for mothers at all births, rather than the average age at first birth. This means that in countries with high fertility, the MAC can be high even though many women give first births relatively early.

The second chart above shows that the mean childrearing ages in the US and Europe fell from 1950 to the end of the end of the 1970s before rising steadily since. According to these numbers the MAC in the US troughed at some 26 in the middle of the 1970s, rising to 29 by the end of the 2020s. In Europe, the MAC bottomed a little later than in the US, at the start of the 1980s, just under 27. It then rose slowly from 1980 to 1995 after which the increase has since accelerated. By 2020, the UN estimates that it had risen to around 30. Given that the rise in MAC corresponds to a period in which total fertility was also falling, we can say that the increase is related to tempo effects.

In Asia and Latin America, meanwhile, the numbers suggest that these regions are on the cusp of a significant shift. The MAC fell steadily from the 1960s to the beginning of the 2010s, but it now seems to be rising in both regions, alongside what increasingly appears to be a move towards sub-replacement fertility. The increase

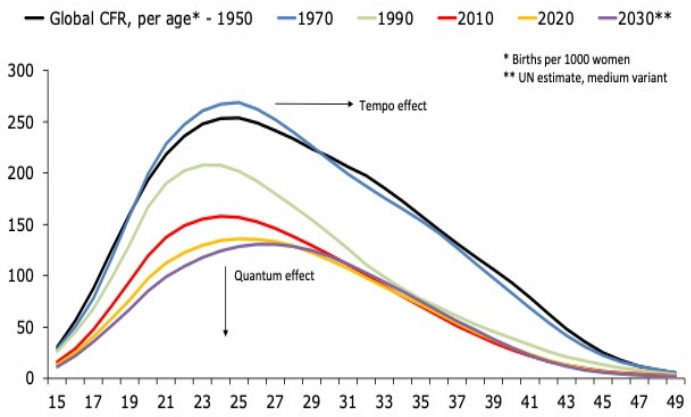
is particularly pronounced in Asia, where the UN data signal an accelerated rise in the latter part of the 2010s. Finally, in sub-Saharan Africa, the MAC was stable from 1950 to 1990, between 29 and 30, and has only recently started to decline towards 29. This is because women in this region tend to have many children throughout their fertility career.

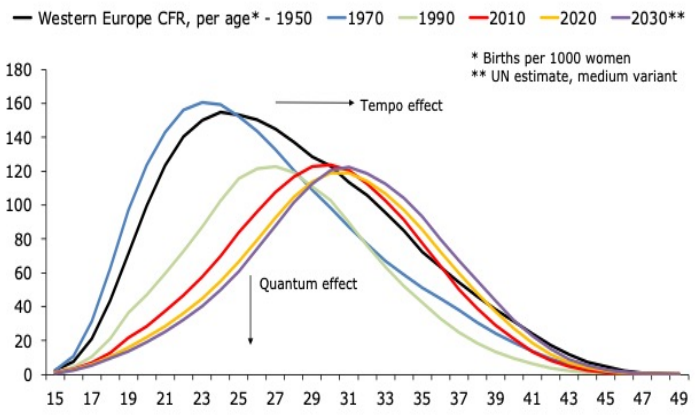
THE SDT, A THEORY BASED ON EUROPEAN DATA

The easiest way to make sense of the age-specific fertility data is to look at the change in the raw data over time. The two charts below plot these data for the world and western Europe, with UN medium estimates for 2030. The arrows show the difference between quantum effects—a decline in births irrespective of timing—and tempo effects, which are linked to birth postponement.

The first chart reveals a clear negative quantum effect in global fertility, with a break between 1970 and 1990. Indeed, from 1950 to 1970, the number of births to women in the 20s rose slightly, a picture that has since changed substantially. From 1970 to 1990 the CFR for women in their 20s dropped from around 250, to just under

Ch 03 / Quantum on a global level… - **Ch 04 /** …But tempo in Europe





Source: UN, authors' calculations

200. It fell further between 1990 and 2010 to 150. For 2020 and 2030, the UN estimates the CFR for women in their 20s at 125 and 117, respectively. For women in their 30s, we see an equally strong fall in the number of births from the peak in the 1950s and 1970s to 1990, but the negative quantum effect then wears off. Indeed, for women aged 33 and over, the CFR has been virtually stable since 1990. Put differently, the main driver of lower global birth rates between 1950 and now is due to fewer children born to woman in the 20s.

By contrast, we see little evidence of tempo effects on a global level. This is to say, women's age at their peak fertility is little changed over time. From 1950 to 2020, the age of peak fertility for women was virtually stable at 23-to-25 years old, with the UN predicting only a slight shift to 27 by 2030. We can see this too if we measure the ratio of births for mothers aged 35-to-25, and those aged 45-to-35. Birth postponement would show up in these number via a rising ratio as births are pushed forward into older age, but on a global level, these ratios are lower today than they were in the 1950s and 1970s.

The picture in Western Europe is different. Granted, we see a negative quantum effect over time too with the CFR falling in the 1990s from relative stability in the 1950s and 1970s. This is in line with the global data. But after that, the tempo effect takes over. The peak CFR was little changed from 1990 to 2020, but its timing across women's age changes over that period from 25 to 30. By the 2000s, the age of peak fertility had shifted to around 30 from the early 20s between the 1950s and 1970s. More tellingly, since the 2000s, the crude birth rate for women aged 30+ have been higher than in the 1950s, despite the fact that overall fertility fell significantly between these two periods.

The difference between the chart on global and European birth rates above highlights an important point about the SDT; it is a theory generated primarily in a European context. As such, it makes sense to first look at the evidence of the SDT in Europe, before trying to apply the theory to the rest of the world. The Special Collections 7 by Demographic research—see here—featuring work by Thomas Frejka, Tomáš Sobotka and others is a good place to start.

This work, published in 2008, produces case studies on several European countries, including Germany, France, and the UK. Research on European countries in the first decade of the 2000s, when a substantial amount of work was done in this area finds strong evidence for the SDT in north and western Europe starting in the 1970s, and in central and eastern Europe at beginning of the 1990s. The literature—including Frejka et al (2008)— broadly points to delayed union formation and fertility, and a change in the values attached to marriage, as predicted by the SDT. This, in particular, is linked to increasing acceptance of cohabitation out-of-wedlock, non-marital fertility and outright childlessness.

Frejka et al (2008) identify two routes through which the SDT has evolved in Europe, separating the relatively rich north and western parts of Europe and poorer central and eastern areas. In the former, a rise in economic affluence drove a shift in values and culture towards self-realisation and individualism—in particular among women—driving a shift in the pace and nature of family formation, and ultimately birth postponement. In the latter, meanwhile, the fall in birth rates, driven by postponement, starts in the poorer parts of society, in response to the economic and societal convulsions after the fall of the Berlin Wall and the dissolution of the Soviet Union. This paints an altogether bleaker picture of birth postponement in a post WWII context, driven by falling living standards and rising economic uncertainty. Interestingly, the evolutionary framework de-

scribed in a previous chapter offer support for how both these pathways can drive down birth rates over time. In the first case, the return on investment in somatic capital, non-reproductive effort, increases, incentivising birth postponement. In the second case, the tempo effect occurs as a result of more difficult external, or in this case social, conditions increasing the cost of reproduction and successful child rearing.

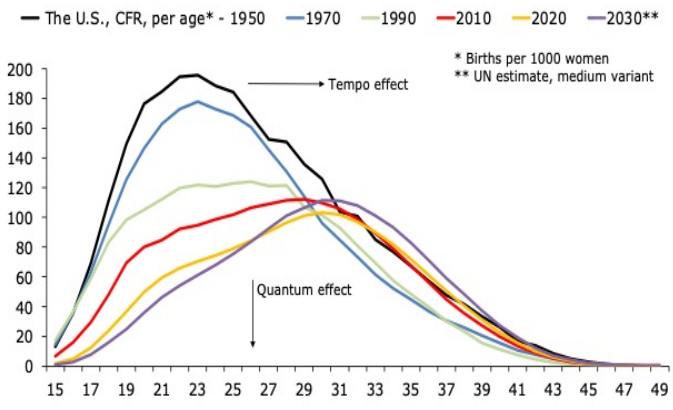
THE END OF US EXCEPTIONALISM?

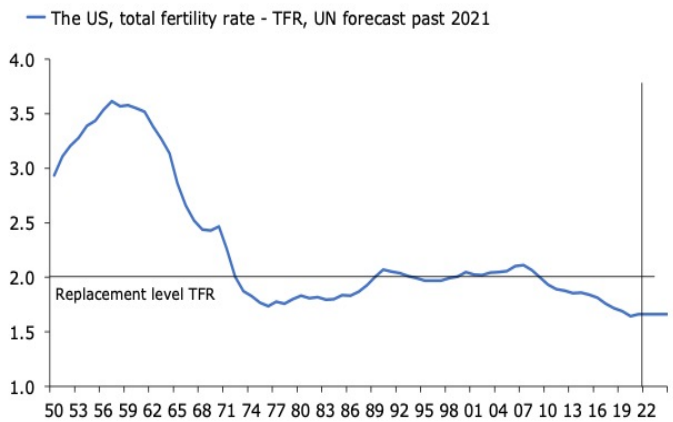
Lesthaeghe and Neidert (2006), an analysis of fertility in the US, is confident in the broad application of the SDT:

"(…) every characteristic of the second demographic transition has spread to the majority of industrialized Western populations, including Mediterranean and Central European countries."

This is a bold statement for the US in the middle of the 2000s. Fertility in the US *rose* at the start of the 1980s, from 1.8 to just over replacement levels at the beginning of the 2000s. Higher fertility of non-Whites was a key contributor to this increase, especially the fact that, according to Lesthaeghe and Neidert (2006), total fertility

CH 05 / Tempo effects in the US - **CH 06 /** No longer exceptional





Source: UN, authors' calculations

rates of the Hispanic US populations was consistently above replacement levels through the 1990s, and into the 2000s. Indeed, the paper asserts that an ethnicity-based analysis lends itself to the conclusion that;

"(…) the US is a textbook example of the SDT where immigration and higher immigrant fertility compensate for subreplacement fertility of much of the native population."

This, however, is only a part of the story. Relatively high fertility in the US has given rise the idea of US exceptionalism, linked, according to Lesthaeghe and Neidert (2006) and Carlson (2005), to a high and stable share of people in the US self-identifying as religious. It follows from this, controversially, that it is conservatives in the US who are doing the heavy lifting on fertility, in contrast to their less-family oriented liberal counterparts. Is this a true characterisation of US fertility from 1980 into the early 2000s? It is is a difficult question to answer clearly. It makes sense to the extent such a trend is associated with stability of religious and family-oriented norms supporting relatively early family formation and childrearing. It also chimes with the idea that it was the fall from grace of such values in Europe, which the SDT pinpoints as a driver of sub-replacement level fertility in that region.

Lesthaeghe and Neidert (2006)'s evidence is somewhat cherry-picked. The paper suggests that votes for the George W. Bush presidency in 2004, which was ostensibly a Christian/evangelical conservative political project, is highly correlated with higher fertility and a lower score of the authors' composite SDT index.

Lesthaeghe and Neidert (2006) concludes that;

"Yes, there is an "American exceptionalism" among a non-negligible section of the population. That section is mainly located in the Mid-west, the Great Plains, and the South. It is on average much more rural than metropolitan, less well educated, adheres more to Evangelical Christianity or Mormonism."

The question is whether such exceptionalism is unique to the US? It could be, but more generally, the idea that a group or sub-population in a country has higher fertility than the country as a whole, and that this is linked to religious values and culture, is too general a proposition to claim as a unique US phenomenon. Lesthaeghe and Neidert (2006) admits as much when they say:

"As indicated before, the French, Swiss, Belgian, German, Portuguese, and Italian historical first demographic transitions all exhibited clear connections with the political maps and this has continued to be the case for the regional patterns of the second demographic transition as well."

Unbeknownst to Lesthaeghe and Neidert (2006), their analysis was published just as fertility in America hit an inflection point. By 2007, the US TFR remained above the replacement level at 2.1. As of 2021, it had declined to just below 1.7, according to UN data, closely mirroring the slide in many European countries in the latter part of the demographic transition.

What happened?

Chart 05 above shows evidence of both quantum and tempo effects of US fertility over time. The quantum effect was particularly strong from the 1950s to 1990 with a significant fall in the number of

births for women in their 20s and early 30s as the baby-boom petered out. The chart also offers a glimpse into the drivers of the so-called American exceptionalism; namely, the green line which shows a relatively constant number of births for women in their 20s and into the early 30s throughout the 1990s. This is to say, we see no tempo effect of US fertility in the 1990s. But when we skip forward to 2010, evidence of birth postponement becomes clearer, hand-in-hand with a continued quantum effect. The number of births fall for women in the 20s, partially offset by a rise in births for women in the 30s. This effect continues in the 2010s.

The timing of the fall in US birth rates coincide almost perfectly with the onset of the Great Recession, which has given rise to the idea that the surge in economic uncertainty pushed birth rates lower. It probably did, and not just in the US, but by how much?

Sobotka et al (2011) argue that that fertility rates does indeed exhibit pro-cyclical behaviour in the developed world—leading to the expectation that fertility falls during recessions—but also that the effect is quite small, by "a few percentage points" and over "short durations." This implies that while recessions can impact the timing of births for women in their child-rearing age, it is unlikely to drive a shift in total cohort fertility, a conclusion based on the observed fertility declines during the Great Depression in the 1930s and Oil Crisis in the 1970s. This makes sense, but it is important to consider the interaction between quantum and tempo effects.

Specifically, it is likely that a recession-induced rise in birth postponement will have a lagged effect on total quantum, for some women. This is especially the case if the rise in economic uncertainty happens in a context of an already-pronounced tempo effect or if the recession is particularly onerous and prolonged. It is difficult to verify the first of these conditions in the US, though it seems clear

that some tempo effects were present by the onset of the GFC, as evidenced by chart 05 above. The GFC, as it turns out, was a particularly severe crisis, especially in relation to its epicentre in the housing market and the associated hit to household formation, Lee and Painter (2013). In this way, it is reasonable I think to believe that the timing of the big recession in 2008 and 2009 had an outsize cyclical effect on birth rates in the US, and elsewhere too.

More specifically in the US, the fall in fertility to below replacement levels suggests that the two factors linked to the idea US exceptionalism—high fertility in the immigrant population and a religious/conservative trait for high birth rates—reversed by the end of the 2010s. Kearney et al. (2022) find evidence of one of these. Their data confirm a significant tempo effect of fertility over the past two decades, and a sharp fall in fertility among Hispanic women, and to a lesser extent black women. From 1990 to 2005, the number of births to Hispanic women was almost twice the number of women born to white non-Hispanic. By 2020, the difference between the has all but vanished. Granted, the shift in crude birth rates across ethnicities does not correct for the relative size in these population groups. But the decomposition done by Kearney el al (2022) shows that falling teen births by Hispanics is the single largest contributor to the decline in U.S. Births from 2007 to 2019, by 14.0%. In total, declining birth rates for Hispanic women account for almost 20% of the decline over that period. The remaining decline is explained by falling births for white and black teens, and whites in their early 20s, pointing to significant tempo effects.

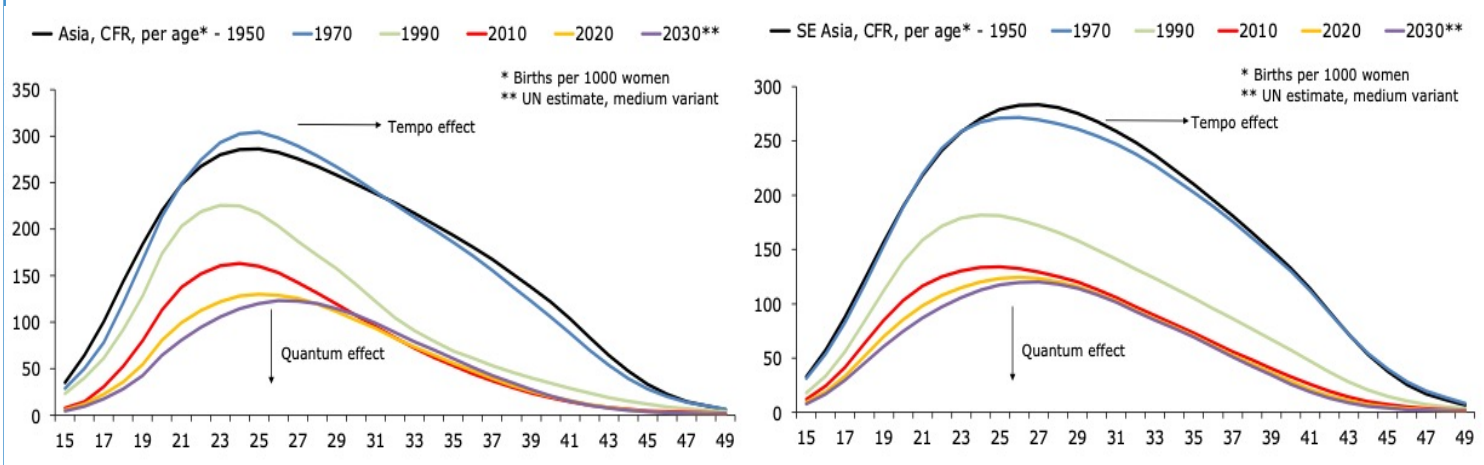
THE DEARTH OF BABIES IN ASIA

Lesthaeghe (2010) suggests that the SDT is spreading to non-western countries. The paper finds that the very low fertility rates recorded in south east Asia by the beginning of the 2000s are driven by

significant tempo effects, mirroring the experience in Europe. Using data from Frejka and Sardon (2009), Lesthaeghe (2010) shows that birth postponement—births through the year 27—for cohorts of women born 1965 to 1980 show remarkably similar profiles across Europe and Asia. The similarity continues as we track these cohorts of women into their final reproductive years, with some evidence of recuperation, but less so than in Europe. The broad conclusion is that we see a similarity between Europe and select south eastern Asian countries, primarily Hong Kong, Japan and South Korea, and more recently, China too.

One of the issues in generalising the analysis is that the sample under scrutiny changes significantly across studies and the questions asked. This is especially the case when attempting to generalise across a broad region such as south east Asia. With that in mind, Lesthaeghe (2010) remains confident that the underlying socioeconomic markers of the second demographic transition—delayed marriage, births out of wedlock and cohabitation—all are present in selected south east Asian economies.

CH 07 / Mainly Quantum in Asia - **CH 08 /** And not much different in SE Asia

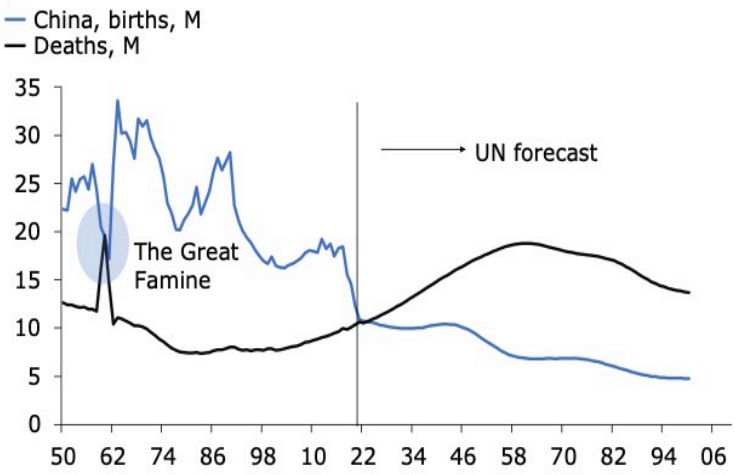


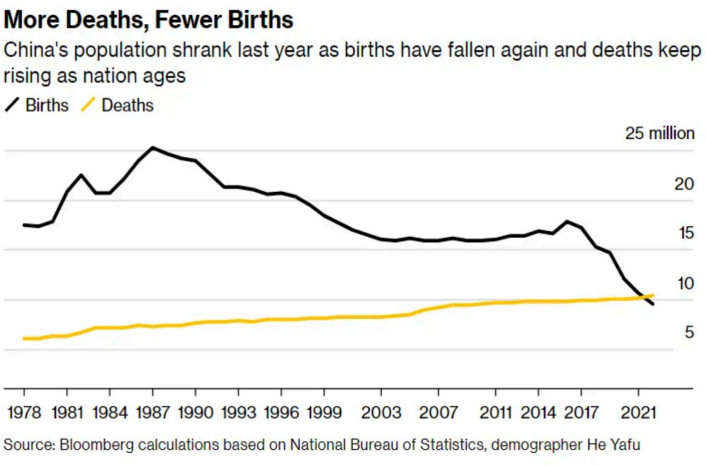
Source: UN, authors' calculations

The two charts above seem to contradict the findings in Lesthaeghe (2010) somewhat, indicating that falling Asian fertility is mainly propelled by quantum effects, and only very recently tempo effects. Specifically, tempo effects are visible in the 2010s, but the charts also point to a sharp drop in the peak number of children born to women in their early-to-mid 20s. The data also point to a significant fall in births at older age groups, especially in the four decades from 1950 to 1990. This coincided with the one-child policy in China, and it is impossible understand the demographics of Asia without understanding Chinese experience.

China's population shrunk in 2022, for the first time in 60 years, by 850K, the net result of 9.6M live births, and 10.4M deaths. It is worth taking these numbers with a pinch of salt. Accurately accounting for some 1.4B people is difficult, especially down to a sub-1M difference between deaths and births. It's possible that future revisions will show that China's population has been shrinking since the beginning of the 2020s, or that it won't start shrinking until 2025 or beyond. Whatever the precise numbers are, however, outright and sustained population decline is coming. China's fertility rate has long since declined below the replacement level, and mortality is now rising as the population ages.

CH 09 / The great shift in China - **CH 10 /** A turning point in China's demographics?



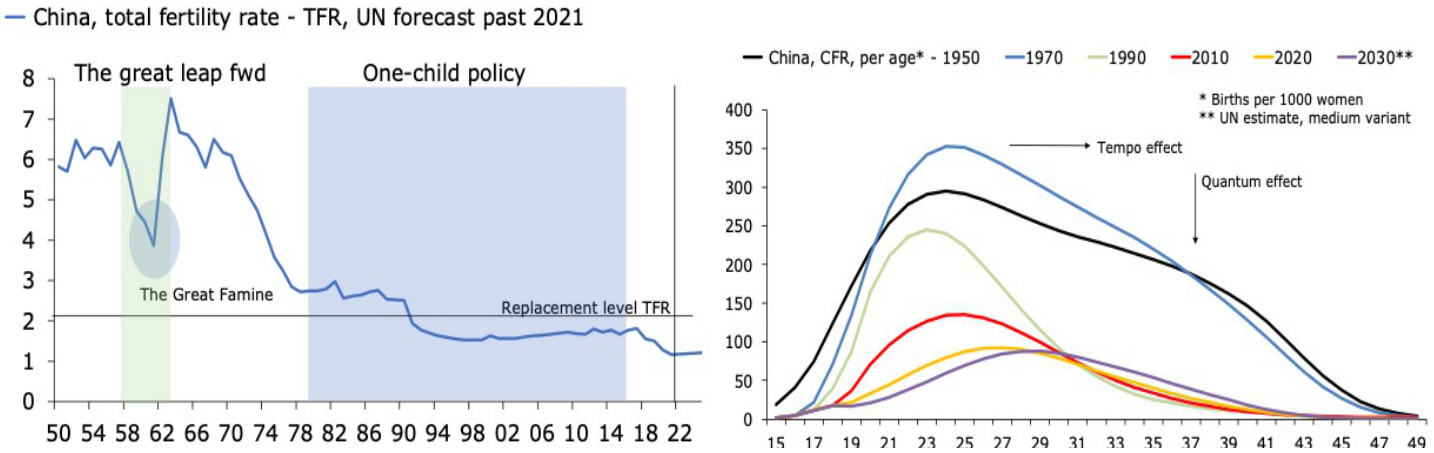


Source: UN, authors' calculations and Bloomberg

Chart 09 plots UN estimates for births and deaths as of 2021, while chart 10 plots a Bloomberg chart with the latest estimates from the domestic statistical office. According to this picture, China's shrinking population is the result of a steady decline in all-age mortality since the end of the 1990s, and a sharp fall in live births since 2017. The increase in mortality is at this point a function of population aging. Lift expectancy at birth is still rising, but as the probability of individual mortality rises with age, this is now dominating the aggregate picture. The sharp decline in the number of births since 2017 is more difficult to explain, though it chimes with an overall falling fertility rate over time. Some economists have speculated that the drop in births between 2017 and the beginning of the 2020s is the result of statisticians adding in a decline they suspect, or know, happened earlier. It is certainly odd that the drop in births, which is the principal reason for why China's population is now shrinking, occurred just as the country relaxed its one-child policy.

Charts 11 and 12 plot China's TFR, and the shift in birth rates over time, across age groups. China's one-child-policy was introduced in 1980, but family-planning policies, trying to limit population

CH 11 / A fertility trap in China? - **CH 12 /** quantum and tempo in China



Source: UN, authors' calculations

growth, started in the 1970s. Broadly speaking, and in line with the experience in other Asian countries, the evidence suggests that the country's demographic transition was already well underway by the beginning of the 1980s. This, in turn, makes it difficult to separate policy-effects from more general transition dynamics.

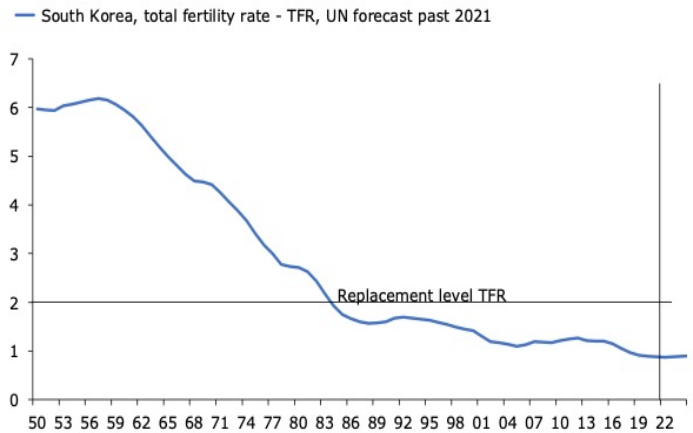
China's fertility rate dropped below the replacement level in 1991, continuing its decline to just over 1.5 by 1998. It then rebounded to 1.8 by 2017, before collapsing to just over 1 by the beginning of the 2020s. This most recent plunge reverses a period during which Chinese fertility had been following a pattern of a rebound in fertility, which looks like reversing tempo effects, after an initial fall below replacement levels. But the devil is in the detail. As I argue below, it now seems more likely that the most recent fall in fertility is driven by lagged tempo effects.

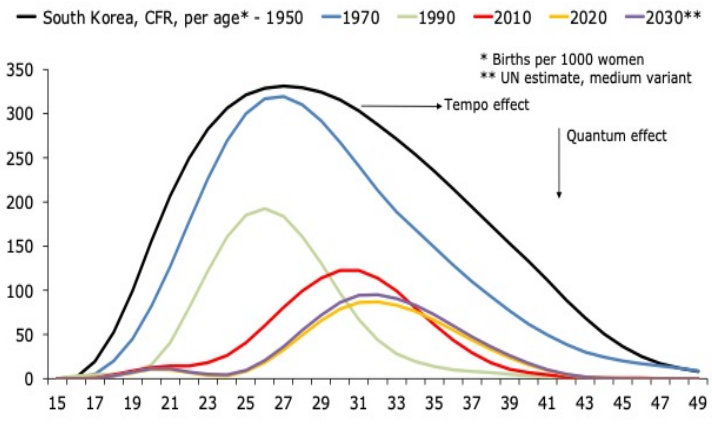
Specifically, UN data suggest that quantum effects have been the dominant driver of falling fertility over time, making the most recent plunge in birth rates look somewhat odd. The charts above clearly show the significant shift in birth rates in response to the country's family planning policies. The profiles of age-specific fertility in 1950 and 1970 are extraordinary. Birth rates remained high for women through their 30s, and even into the 40s, contributing to the country's high fertility during that period. By 1990, however, the profile of birth rates take on a more traditional function with birth rates peaking in the early 20s, before falling significantly to negligible levels past the mid 30s. We see another lurch lower in birth rates due to the quantum effect between 1990 and 2010, before the beginning of more meaningful tempo effects.

UN data show that the peak age of fertility in China was broadly stable from 1950 to 2010, at 23-to-25 years, before shifting towards 27 by the 2020s. The UN predicts a further shift to 29 by 2030, due effectively to extrapolation of the most recent trend. This could be a crucial data-point for understanding the slide in period fertility since 2017. It is likely that the plunge in birth rates since that period is due to tempo effects. The UN age-specific data support this hypothesis, pointing to accelerating birth postponement since 2010, amid a further reduction in quantum effects. It is difficult to know for sure until we see numbers through 2025, but it makes theoretical sense. The negative quantum effect of fertility will tend to peter out as period fertility falls below 1.5, unless you assume a significant increase in childlessness among women.

In South Korea, the recent collapse in fertility indicates that the country could be one of the most extreme examples of the second demographic transition. Recently, we learned that South Korea's total fertility rate fell to an astonishing 0.78 in 2022, from 0.81 in 2021, the lowest period fertility rate on the planet. The next two charts paint a clear picture. The first, chart 13, shows the sustained decline in fertility rates, which began in the 1960s. In 1960, South

CH 13 / How low can it go? - **CH 14 /** Both tempo and quantum effects in South Korea





Source: UN, authors' calculations

Korean women were having about six children per women, a number which had declined to just over four by 1970 and just over two by 1980. By the middle of the 1980s, fertility fell below the replacement level, and the decline has continued since, despite temporary rebounds at the start of the 1990s and again at the beginning of the 2000s. Period fertility resumed its decline around 2015, and South Korea now has the lowest total fertility rate in the world.

The second, chart 14, plots birth rates across age groups. It paints a clear picture. The fall in birth rates from 1950 to 1990 was driven exclusively by quantum effects. Specifically, women's age at their peak fecundity was little changed in that period, around 26-to-27, but the number of births to women in that age group almost halved over that period, from some 330 per 1000 women in 1950 to just under 200 in 1990. From 1990 onwards, the tempo effect accelerates. In the twenty years ending 2010, the peak birth rate age rose to 30, and further to 32 in 2020. This shift coincided with an accelerating quantum effect, to just under 100 births per 1000 women.

These are remarkable data, but do they overstate the decline in cohort fertility? Probably. Yoo and Sobotka (2018) estimate a tempo-adjusted fertility rate at 1.5 in 2014, compared to a recorded TFR of 1.2 in that year. This is a significant difference, but not one that changes the picture of a sustained and significant fall in birth rates over time. Based on evidence from the UN data that tempo effects recently have accelerated beyond the numbers analysed in Yoo and Sobotka (2018), we can use the ratio between tempo-adjusted fertility and TFR in 2014, at 1.25, to approximate the present tempo-adjusted fertility rate at around 1.0; that's still very low. Importantly, Yoo and Sobotka also present evidence indicating that the decline in fertility to "ultra-low" levels has been driven mainly by quantum effects, specifically by a fall in first and second births. This makes the decline in South Korean fertility even more aston-

ishing. It means that an increasing share of women in the country aren't having any babies at all, by choice, never mind only one.

Many of the generic drivers of the SDT are present in South Korea. The marriage rate, for instance, has plunged; it stood at 3.7 in 2022, about half the rate in 2010, and down from 10 at the beginning of the 1990s. And because of conservative values limiting out-of-wedlock births, the decline in marriage rates is a strong driver of falling birth rates , compared to in other countries.

Then there are gender roles, also a function of South Korea's relatively conservative culture. Many women in South Korea are consciously choosing not to have children, or to postpone having their first child, to object to what they perceive as a culturally-driven expectation that they sacrifice their careers to take care of children and take up traditional role as non-working mohter. In a piece for the Atlantic—*The Real Reason South Koreans aren't Having Babies*, March 2023—Anna Louie Sussman investigates the case of plunging fertility rates in South Korea, tracing it to fundamental distrust between young men and women. For the latter, the decision not to have children seem in some cases to be the ultimate action through which to explicitly reject South Korean culture itself, and more specifically, its conservative male-dominated foundations. One of the female protagonists in Sussman's story says; "I try to have faith in guys and not to be like, 'Kill all men,'" she says. "But I'm sorry, I am a little bit on that side—that is, on the extreme side."

South Korean men, for their part, harbour resentment for women, primarily for being too picky, and for not realising that the country's conservative culture also comes with a price for them. The man and his family are expected to shoulder the up-front cost of starting a new family, mainly via the purchase of a new home, which is prohibitively expensive in many areas of the country.

Sussman's article paints a picture of an extreme version of new feminism and an extreme counter-response by some men. This shift has upended the culture of family formation in South Korea that used to produce births in a conservative version of the general model where the man is the breadwinner, and the women gives up most of her career to take care of the home and children.

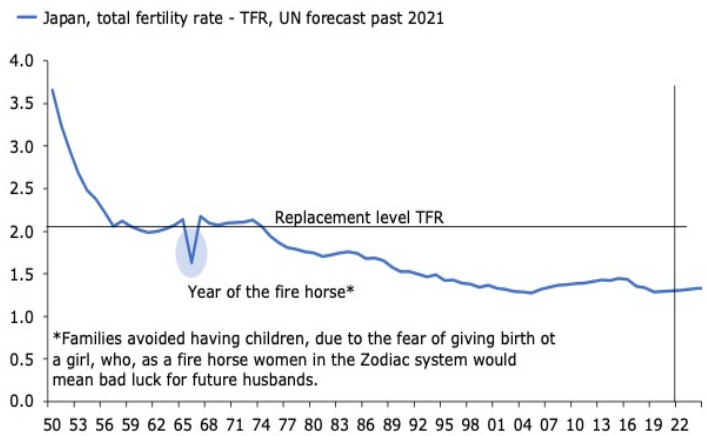
For the government in South Korea, falling fertility presents a number of economic challenges in the context of the effects of population ageing on government finances and economic growth. Estimates suggest that government has spent $200B in the past 16 years to combat falling birth rates in part via direct subsidies to parents. It has little to show for it. More generally, in the context of a debate about gender roles, low fertility is a double-edged sword. It is a problem if falling fertility is driven by financial and economic barriers to family formation. If, however, low fertility is seen through a lens of a conscious choice by part of women, who refuse to conform to conservative values of child-rearing and stay-at-home roles, it suddenly becomes more difficult to articulate as a problem.

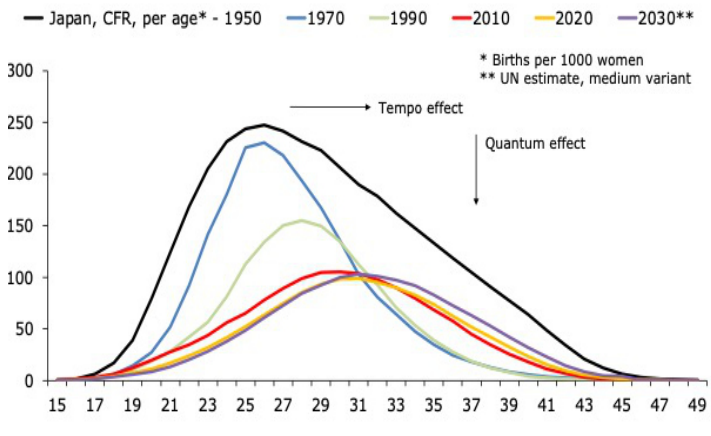
This is to say, it is difficult to articulate this as a problem without running headfirst into the accusation of being a misogynist, or anti-feminist. This is because it is the argument that women in South Korea should take one for the proverbial team and allocate more resources to having babies, whatever the cost to their careers. Because the answer invariably is that it is a combination of these general drivers of lower fertility—exogenous socioeconomic factors and endogenous cultural factors—it is difficult to offer an objective answer to the question of whether South Korea's ultra-low fertility is a genuine problem, as opposed to a logic outcome of the confluence of trends mentioned above.

The third case study in Asia takes us to Japan, which, measured by median age, is the oldest country on earth, excluding the greying millionaires of Monaco and the some-5,000 people on British St. Helena. At the end of 2021, Japan had a median age of 48.4, well ahead of the second major country on the list, Italy, with a median age of 46.8. Japan is about to get older still. According to preliminary estimates, the country's fertility rate fell further in 2022, while the gap between births and deaths remained wide as ever. The number of live births fell by 5.0% in 2022, to 770.774, while deaths rose by 9.0%, to 1.57 million. Japan's rapidly ageing population is the result of a quicker and more sustained post-1945 fertility transition than in other developed economies.

Japan in effect has undergone two separate fertility transitions since the end of the Second World War Tsuya (2015). The first was driven by a broad-based decline in fertility across all ages, while the fall in fertility to below replacement levels has been driven mainly by tempo effects. Tsuya (2015) emphasises that the initial post-war transition, which took the fertility rate from almost four in 1950 to around 2 in 1960, was driven to a large extent by falling birth rates among women aged 30 and higher. The second chart below con-

CH 15 / Fertility has plunged in Japan… - **CH 16 /** …Due to quantum and tempo effects





Source: UN, authors' calculations

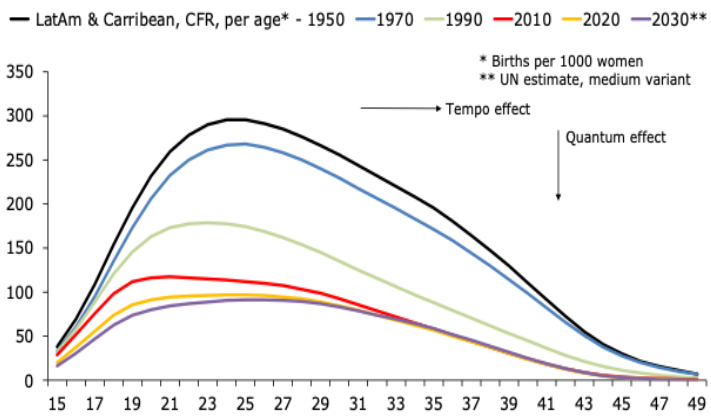
firms this. It shows that the number of children born to women at the peak of their fertility, at the age of 25-to-27, fell only slightly from 1950 to 1970, in contrast to a collapse in births to women aged 30 or more. Put differently, the initial and quick fertility transition in Japan was primarily due to women reducing births dramatically in the latter part of their reproductive life. From 1960 to the middle of the 1970s, fertility rates stabilised, with the exception of the bizarre drop in 1966, due to superstition that women born in that year—the year of the fire horse—would bring bad luck to their future husbands. Fun fact; 2026 is the year of fire horse again, so keep your eyes peeled for another drop in births in a few years.

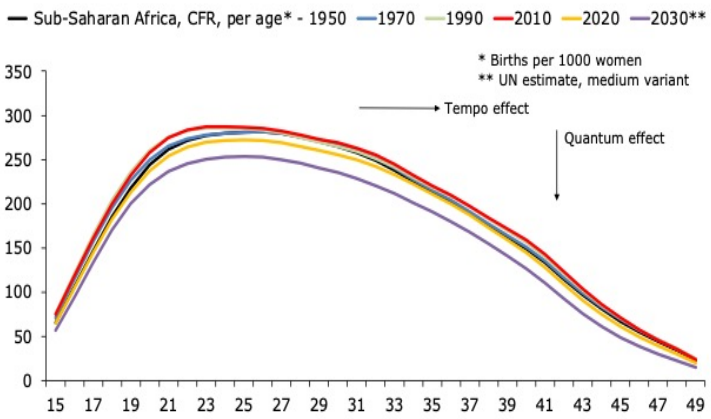
By 1990, fertility had fallen further, driven by both quantum and tempo effects. The number of births to women at peak fertility fell by almost a third, and the mean age of first childbirth rose. This development continued through the 1990s and into the 2000s. The fertility rate rebounded from a low of 1.27 in 2005 to 1.43 in 2016, but it has since dropped back, and if the early 2022 estimates are correct, it is now at a new low. The UN's most recent forecasts predicts little change in the decade from 2020 to 2030, except from an increase in births to women in the latter part of their fertility career. I am not sure where they get that prediction from.

The drivers of the decline in Japanese fertility broadly follow the tenets of the second demographic transition. Marriage rates have declined dramatically in the past 50 years, and because the country's relatively conservative values all but exclude out-of-wedlock child births—a mere 1-2% of live births have taken place in non-married unions since 1960—the relationship between falling marriage rates and lower fertility has been particularly strong. Japan's experience, in other words, mirrors that of South Korea.

The numbers cited by Tsuya (2005) are remarkable. Among women aged 25-to-29, the unmarried proportion in 1975 stood at 18%. By 2010, this number had soared to 60%. For women aged 30-to-34, just over a third of women remained unmarried by 2010, up from 8% in 1970. For women aged 35-to-39, the proportion of unmarried was 23% in 2023, rising from 5% in 1970. The shift for men is even more dramatic, supporting the culturally charged meme that young men either offer too little in the way of companionship to women. Women simply have better thing to do in a modern era than to get married and have children with what they perceive as hapless and subpar men. In 2010, the share of unmarried men aged 25-to-29 stood at 72%, up from 48% in 1975, while it had increased to 47% for men aged 30-to-34, from 14% in 1960. For men aged 35-to-39, the share of unmarried had increased to 36% in 2010, from 6% in 1975. In other words, in the middle of the 1970s, almost all men were married by their late 30s, but by 2010, this number had declined to around 60%. In 1975, the celibacy rate for men was just 2%. By 2010, it had soared to 20%.

CH 17 / No tempo in LatAm, yet - **CH 18 /** Sub-saharan Africa is stuck





Source: UN, authors' calculations

BI-MODAL LATIN AMERICA AND FROZEN-IN-TIME AFRICA

The final two regions in our whistle-stop tour of modern global fertility trends is Latin America and Sub-Saharan Africa. The first chart below plots age-specific birth rates for women in Latin America, and it tells a clear picture of a rapid demographic transition since the 1970s, due to mainly to quantum effects. Fertility in many Latin American countries fell below the replacement level in the early 2000s, indicating a shift along the lines observed in other key regions. But there is a twist to the story in Latin America. The decline is due almost exclusively to quantum effects.

The number of births to women in their peak reproductive age stood at 270 per 1000 women in 1970, only down slightly from 1950. By 1990, it had declined to just under 180, and it fell further in the five years ending 2020, to just under 100. This is consistent with evidence in other regions, but LatAm stands out in terms of tempo effects. Indeed, the age of the peak crude birth *fell* since 1950, from 25 to 22-to-24 in the early 2000s. It was 24 in 2020, according to the UN. In other words, the decline in the region's total fertility rate from around 6 in the early 1960s to just over 2 in 2010 was due exclusively to quantum effects. Lima et al. (2018) analyse the relative lack of birth postponement in Latin America, invoking the idea of a bi-modal fertility regime, split along the axis of women's educational attainment. Young, and relatively un-educated, women are having their first child relatively early, and maintain relatively high levels of fertility through their early reproductive years. By contrast, women with tertiary education are now engaged in significant postponement, in line with evidence in developed economies. Lima et al. (2018) provide convincing evidence of this hypothesis via early 2000s consensus data from Chile and Brazil. In Chile, the age of first birth for women with primary and secondary education cluster in the early 20s, while it jumps to the early 30s.

This gap is even larger in Brazil, with the age of first birth for women with tertiary education is in the mid-30s.

These data also offer an important clue to the very recent shifts in Latin American fertility. From 2010 to 2020, the total fertility rate in Latin America fell from just over the replacement level, to slightly below and initial data from the start of the 2020s suggest that it is now well below the replacement level. This accelerated decline in fertility coincides with a market rise in the number of women proceeding to tertiary education. According to data from the World Bank, the number women aged more than 25 completing at least a short-cycle tertiary education rose sharply in all key Latin American countries from around 2005 to 2020. Put differently, countries in Latin America is now undergoing their version of the second demographic transition, and in line with the experience in developed economies, we should expect period fertility to fall much further.

In sub-Saharan Africa, the second chart above suggests that we can skip this region altogether. Specifically, the charts points to little in the way of either quantum of tempo effects of fertility overall time, indicating that countries in this region are yet to start their demographic transitions, at least not in the traditional sense.

THE LOW FERTILITY TRAP

The onset of rapidly falling fertility to sub-replacement levels in one country after the other prompted researchers in the middle of the 2000s to ask whether some countries are at risk of falling into a fertility trap Lutz et al (2006) and Lutz and Skirbekk (2005). And if there is, should governments attempt to do something about it?

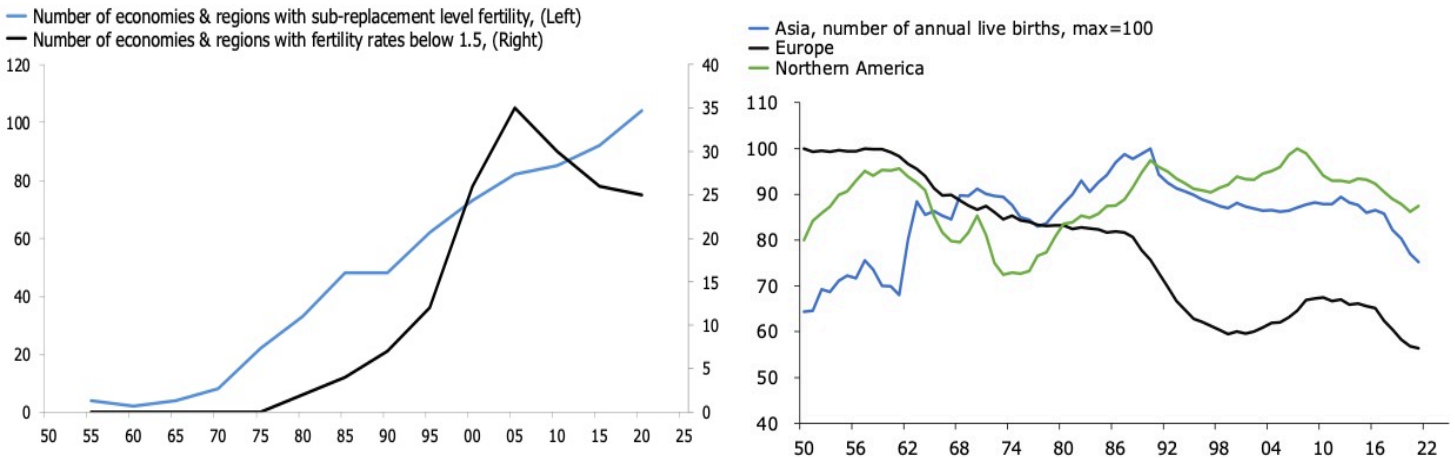
Lutz et al (2006) identifies three self-reinforcing pathways through which a fertility trap can emerge. The first is negative population

momentum. A sustained period of fertility significantly below the replacement level will shrink the potential pool of mothers over time. In absolute terms, this means that a rapid decline in births today means fewer births tomorrow, and if this process is sustained over a sufficiently long period of time, it can be virtually impossible to get the number of births back to the initial condition. This is true even if we assume catch-up in period fertility due to tempo effects. In other words, this first self-reinforcing mechanism operates on the absolute number of births.

The second one operates on period and cohort fertility via sociological and behavioural pathways. The idea is that life course traits such as ideal family size, the timing of marriage and first births, and overall childlessness perpetuate over time, mainly among women, to lock-in low period fertility, which spill over into cohort fertility over time. Cultural evolution, which explains how behaviour and ideas influece behaviour over time and across populations, is one framework through which to understand this.

The third self-reinforcing mechanism that operates to reduce fertility is based on Easterlin's relative income hypothesis. If economic

CH 19 / Breaking the fertility trap? - **CH 2O /** Never coming back?



Source: UN, authors' calculations

aspirations of young cohorts are increasing over time, but they find that such aspirations are persistently unfulfilled, the tendency to postpone family formation and childrearing, or perhaps even forgo it entirely, will accelerate, driving down fertility. The link between the two, acting as a self-reinforcing mechanism, is that the population ageing precipitated by falling fertility itself acts to drive down economic growth and disposable income of working-age cohorts.

The first of these reinforcing mechanisms is simple in the sense that it is based on the straightforward idea of negative momentum. Falling fertility over a long period of time will reduce the number of childrearing women in future generations, driving down the overall number of births over time. This is especially the case as period fertility falls below the replacement level for a sustained period as this will also, in time, push down total cohort fertility. The second and third reinforcing mechanisms are closely related to the theoretical and conceptual frameworks used to explore the idea of a second demographic transition.

The second reinforcing mechanism, as stated, relates principally to the hypothesis that ideal family size is a lagged function of actual fertility, and that ideas about fertility in society as a whole can influence household formation and fertility decisions by individuals and couples. But more broadly, it attempts to capture idea that the factors contained in the SDT can push countries into a fertility trap.

This idea expands on the more traditional framework for understanding the gaps between ideal and realised family sizes, across developing and developed economies. In the former, completed cohort fertility often exceeds the ideal family size giving rise to public policies aimed at offering ways for couples to reduce the number of children, predominantly via more easy access to contraception.

By contrast, in developed economies, the SDT framework is based on the idea of a fundamental shift in the relative allocation of resources to reproduction and somatic investment or self-actualisation, to use a socioeconomic parlance, mainly for women. This trend in birth postponement, as described, will then lead to a situation in which women end up having fewer children than they want Demeney (2003), a trend which can become entrenched over time. In this context, survey-based indicators of desired fertility is often considered an upper bound for realised fertility.

The second reinforcing mechanism in the fertility trap goes a step further than the idea of a shift in the resource allocation trade-off, by proposing a decline in ideal family size, or desired fertility, as a lagged decline of actual fertility. This, in turn, is based on early evidence in the latter part of the 1990s and into the 2000s, Goldstein (2003), of exactly such an effect, and a model for how this effect might perpetuate over time, Testa and Grilli (2006). The problem is that accurate data on ideal family size and desired fertility is patchy. The World Database does compile data on desired fertility, but the availability of data across time and countries is too sparse to do a proper empirical analysis, especially since a true statistical test of Lutz et al. (2006)'s hypothesis requires a panel data analysis with a big cross-section and time span. At this point such a panel is not available as far as I can tell. Often, data on actual and wanted fertility tend to be relatively rich in developing countries, which drives an imperative in these countries policymakers to offer ways for women countries to have fewer children.

More generally, it is difficult to separate the weak form of the second reinforcing mechanism from the strong form. The former is simply the observation that low fertility today, and the behaviours and trade-offs associated with it, begets low fertility tomorrow, without a specific link to a fall in desired fertility. The latter is based

on the explicit assumption that desired fertility is falling over time, and perhaps that the share of women wanting no children at all, is rising, in the developed world. Anecdotal evidence suggests that it is, but it is difficult to verify.

Skirbekk (2022)—*Decline and Prosper - Changing Global Birth Rates and the Advantages of Fewer Children*—devotes a chapter to the measurement and shifts in ideal family size and desired fertility. Measuring these variables is not an exact science. Ideal family tends to be captured by qualitative survey data, which invariably will contain biased information and measurement error. In addition, these survey data are difficult to compile, and tricky to compare across countries and time. Still, Skirbekk offers evidence to suggest that the ideal family size has shrunk in a post WWII context, converging on a two-children ideal family. This, according to Skirbekk, is the result of pressures on ideal family size from both the bottom and the top. In the former, most women believe that having one child, or none at all, is suboptimal, while in the latter, the number of women who want three or more children has declined. Generally, ideal fertility and family size tend to converge on two as countries develop. Ideal family size, in this context, is still large in Africa.

Within countries, ideal family size and fertility tend to fluctuate across gender, age and socioeconomic groups. The question we're interested in here, however, is whether the low fertility ideal can overshoot, driving down fertility below two, perpetuating a fertility trap. It is difficult to tell for certain. Skirbekk draws loosely on the ideas of cultural evolution in the discussion of how "low fertility role models" in politics, sports and popular media can influence fertility decisions and perceptions of ideal family size. This makes sense, but it is difficult to accurately quantify. Another key question is whether families are having fewer children than they would like. Evidence suggest that they do, in the US, UK, Norway, Singapore

and other parts of East Asia. This, in turn, raises the question of whether this is reflects mismeasurement of ideal family size, or whether institutional or socioeconomic factors are preventing women from having the children they want. Seen from the point of view of fertility trap hypothesis, the possibility that men and women are nt able to have the children they want is the key issue.

Outright childless is an extreme result of the second reinforcing mechanism in the fertility trap. Skirbekk (2022) shows that it is on the rise, especially in the developed world. Again, socioeconomic shifts associated with the SDT are highlighted as the key drivers of rising childlessness; the rise in opportunity cost of reproduction due to better education and economic options for women, better contraception—allowing those who don't want children to meet their objective—the delay and decline in marriage, the increase in the cost of starting a family—Easterlin—and perhaps too an increase in the social acceptance of childlessness. Data presented in Skirbekk (2022) suggest that anywhere from one in five to one in four women in a sample of developed economies were childless towards the end of their reproductive career.

It is important in this respect to understand that childlessness is an ultra-lagging indicator. This is because it cannot be measured at the cohort level until the very end of women's reproductive career. So, childlessness measured today is based on data for women born 40-to-45 years ago. Indeed, with assisted reproductive technology—ART—now prolonging women's reproductive span, current measures of childlessness for a given cohort could be revised up significantly in the latter part of many cohort's reproductive career. More generally, it is tempting to correlate the rapidly falling period fertility with a rising share of childless, but we have to be careful making that link. Skirbekk (2022) shows that the variance of childlessness for women born in 1972, observed in 2013-to-2014, across low-fertility

countries is significant. A high propensity of outright childlessness is one way to get to low fertility, but it is not the only one.

The third reinforcing mechanism of the fertility trap, based on Easterlin's cohort analysis, relies on the hypothesis that fertility is positively correlated to the ratio between income expectations and economic aspirations, and that this ratio has been declining for young cohorts since the the 1970s. Lutz et al. (2006) offer empirical evidence to suggest that this ratio is indeed falling over time in Japan, Sweden, Italy and the United Kingdom. Whether this trend is conclusively related to falling fertility in these countries is more difficult to verify. Indeed, there is an inherent contradiction between the second and third reinforcing mechanisms. This is because the former is driven, in part, by improving conditions for women in the workforce driving birth postponement, while the latter is driven by the opposite in the form of worsening relative economic prospects, prompting a delay in family formation. It is possible, in theory, for both mechanisms to operate at the same time across different countries, cohorts and income groups, but to verify them empirically, let alone disentangle them, is difficult.

IS THE FERTILITY TRAP REAL?

The literature since Lutz et al. (2006) has come to think about the fertility trap as the idea, inspired by Macdonald (2005), that once period fertility falls below a certain level, it can be very difficult to raise it back above this level again. This is the idea that fertility falling below this level is strong indicative evidence that one or more of the three reinforcing mechanisms are in place. Lutz et al. (2006) is unwilling to be pinned down on where this threshold is, though the literature has, by now, formally associated the idea of a fertility trap with a TFR level below 1.5. The idea of a fertility threshold below which lock-in mechanisms for low birth rates set in is, in the

early fertility trap literature, linked to a concrete policy advice. This is the idea that the returns for pro-natal policies are high if used to prevent fertility from falling below 1.5 in the first place.

To perform a simple test of the validity of the fertility trap today, it is useful to think about the hypothesis in its strong and weak form. The strong from stipulates that once period fertility falls below a certain level—I'll be using a TFR of 1.5 despite Lutz et al. (2006)'s objections—it will not recover. The weak form rests on the simpler heuristic of negative population momentum. The two charts above marks the fertility trap to market with recent data.

To the extent that the strongest form version of the fertility trap states that a fall in TFR below 1.5 is un-recoverable data since the middle of the 2000s seem to falsify it. Chart 19 is a repeat of a chart shown in a previous chapter. It shows that the number of countries and regions with a TFR below 1.5 peaked by around the time that Lutz et al. (2006) was published. This shows that a fall in period fertility to below 1.5 isn't as conclusive as assumed in the strong form of the fertility trap hypothesis. Specifically, this challenges the idea of the second reinforcing mechanism in favour of the idea that prolonged birth postponement will, in some cases, give way to a catch-up, allowing period fertility to recover.

In its weak form, however, the fertility trap hypothesis is standing tall. This is especially the case in the context of the most recent decline in period fertility in Anglo-Saxon economies, which now seems to be entrenching a TFR below the replacement level in these countries. The number of countries and regions with below-replacement fertility was still rising by 2020, and the second chart above shows that the number of annual live births in Europe, Asia and North America, the major regions where fertility is lowest overall, is now well below their post-WWII peak. Judging by recent trends in period

fertility in the major countries in these regions, the number of births will continue to fall for the forseeable future.

WHAT HAPPENS NEXT?

The combination of persistently low fertility in the developed world, and rapidly falling fertility in many parts of the developing world means the global period fertility rate will soon fall below the replacement level, defined as just over two children per women.

In 2019, the UN estimated global fertility at 2.47 in the five years ending 2020, with a projected decline to 2.42 by 2025. These numbers are well out of date with the latest UN population update in 2022. In the 2022 data, and forecasts, global fertility is set at 2.31 in 2021, with an estimated decline to 2.30 by 2025. In both datasets, the UN continues to assume that global fertility will remain above the replacement level for several decades. In the 2019 dataset, the UN's medium variant forecast assumes that global fertility will be above replacement levels until 2065, like the assumption in the 2022 dataset.

In other words, UN continues to assume that the decline in global fertility rates will soon peter out. This flies in the face of even the simplest empirical analysis. For instance, fertility fell more quickly the two years ending 2021 than expected by the UN in 2019. In addition, from 2016 to 2021, the UN data shows that the global TFR fell by an average of 0.04 per year, quicker than the five-year averages in the preceding five years. Put differently, the decline in global fertility is accelerating. If we extrapolate the trend, the TFR will hit the replacement level, 2.1, by 2026, a cool 49 years before schedule. I wonder whether the UN will realise it before then.

The prospect of global fertility falling below replacement level in the next few years adds to the evidence of the weak-form fertility trap hypothesis, which relies on the idea that a sustained fall in fertility leads to negative population momentum. But readers shouldn't just take my word for it.

Pointing fingers at official long-term forecasts for fertility is a time-honoured practice in demographics research. Lutz et al. (2006) ponder why UN and Eurostat forecasts, from 1999 and 2005 respectively, assume long-term stabilisation in fertility rates, despite evidence to the contrary. Lutz et al. (2006) says:

"This deviation from the conventional rules of trend analysis must have to do with strong beliefs that somehow there is a powerful force that will stop and even reverse the trend, i.e., that at the individual level, people will always want children, and that at the aggregate level, human populations would not voluntarily shrink and age to an extent that would be socially disruptive or in the very long run might even mean shrinking to insignificance. From an evolutionary perspective, these are seemingly reasonable assumptions because a species without a drive to reproduce would not have survived to this day."

The idea of some baseline rate of reproduction below which the rate of fertility becomes evolutionarily suboptimal has long been dogging demographic research, as I have discussed in earlier chapters. Mulder (1998), for example, raises the question of whether modern fertility trends are maladaptive, starting an inquiry that continues to this day, and which remains unresolved.

Some researchers have long since come to the conclusion that replacement level fertility is, in the famous words of Demeney

(1997), an "implausible end-point of the demographic transition". Lutz et al. (2006) counter the idea of some unbreakable evolutionary lower bound of fertility by noting that the advent of modern contraception has broken the link between sex and procreation, rendering the latter a simple function of changeable, and fickle, “individual preferences and culturally determined norms”. To me, Kaplan (1996) and (2002) are the best attempts to reconcile this debate, but this is a discussion that won’t go away, primarily because of the lower bound problem.

It is one thing to have a hypothesis, consistent with evolutionary theory, of why birth rates will decline to very low levels and stay there. It is an altogether more complicated proposition to come up with an evolutionary hypothesis for declines in period and cohort fertility to one or lower. Specifically, it is difficult to explain why a growing share of couples would either choose to have no children, or postpone having their first child for so long that they risk having significantly fewer children than they want, or end in up in outright childlessness. Looking beyond the evolutionary debate about whether modern fertility behaviour is maladaptive, the persistent fall in global birth rates raises broader question across the social sciences about the idea of an inherently optimal level of fertility. In short, this is the debate about whether the rapid fall in global birth rates is a good thing or not? This is a question which can be answered in many ways, depending on your perspective.

To the extent that the accelerating decline in fertility is driven by countries moving from high-to-low fertility regimes—sub-Saharan Africa, and many countries in LatAm and Asia in the past two decades—almost everyone would argue that it is a good thing. This is because falling fertility in most of those cases are strongly correl-

ated with rising economic wealth and GDP per capita. The hypothesis is that falling fertility is a necessary, but not sufficient, conditions for moving up the economic value chain. This idea that a transition from high to low fertility is associated with, if not directly related to, an increase in economic development, rising living standards and falling mortality is a cornerstone of development economics. And, as I explained in a previous chapter, it is the basis of most modern accounts of the drivers of the demographic transition, Galor and Weil (1996, 1999 and 2000).

On the flip side, falling fertility in countries with already low fertility—which implies a sustained decline in period and cohort fertility below the replacement level—raises a number of complex questions in the socioeconomic sciences. If this is happening as a result of women's entry into the labour market, inferring a conscious choice by women to allocate resources to their own development rather than reproduction—it is difficult to argue that it is a negative trend. This is to say, you can, if you're willing to argue that women ought to spend more time reproducing than working; good luck with that!

More traditionally, a common argument is that modern society, mainly via political intervention, ought to make it attractive for women to combine giving birth, child-rearing, and labour market participatiion. This sounds like an attractive solution, rooted in politically correct ideas such as the promotion of gender equality by opportunity, if not by outcome. The problem is that empirical evidence suggests that it is difficult for public policy to lift birth rates, and iron out traditional gender roles. This is linked, I suspect, to two factors. First, men and women in couples will tend to opt for traditional gender roles after childbirth—less work for the woman and the man as the breadwinner—to a much higher extent than those

seeking gender equality would prefer, or like to admit. Secondly, evening out the burden of reproduction between the sexes, as difficult as that is, isn't the main factor. Within competition among women for labour market success and wealth accumulation is fierce too. Put simply, in a fundamental biological framework in which women bear a relatively high cost of reproduction, it is self-evident that women who forgo reproduction, either entirely or early on, will have a significant advantage over those that don't. This suggests that the potential relative returns for women foregoing reproduction in a modern economy with rapidly improving labour market opportunities for women are high.

A central question remains whether sustained sub-replacement fertility—especially birth postponement—is because women end up having less children than they really want. Or is the fall in fertility to below replacement levels a structural result of shifts in family-formation and cohabitation trends, marriage rates and women's control over their reproduction. The fertility trap literature allows for both explanations to co-exist, but from the perspective of a policy-maker, the distinction is key.

They intersect exactly in the debate and discussion about how much public policy can affect birth rates in developed economies, and whether they should. Finally, population ageing is the most widely agreed negative (economic) impact from a sustained decline in fertility rates, mainly via a strain on the financing of public services, low economic growth, and the negative externality from too much savings and too little investment, in an open economy.

IS LOW FERTILITY A PROBLEM?

Most of the literature on the rapid decline in fertility during the second demographic transition, as well as the founding literature on the fertility trap, sees low fertility in developing economies as a problem, or more specifically, something which can and should be adjusted by public policy. A theme in the literature, for example, is that if falling birth rates is because women, and couples, face obstacles in having the babies they want, governments should attempt to mitigate this.

A more general theme is that falling birth rates are a problem for socioeconomic reasons. This is mainly linked to the challenge of population ageing for economic growth and the sustainability of public finances due to rising costs of pensions and healthcare. This is best understood by starting with the observation that initial conditions matter, and that path dependency is an important phenomenon in modern market economies. Most developed capitalist market economies operate with some form of social contract by which tax revenue paid by the working-age population is used to pay for the health, consumption and leisure of those that do not work. A sustained decline in birth rates can become a problem in such a system because it risks making this social contract non-viable, leaving a number of unpalatable choices.

You can tax a shrinking working age population harder to pay for the provision of health and care for the elderly, or you can reduce the quality of such health and care. In a democracy where the elderly hold the majority vote, the former is the most likely outcome, which creates its own problems. Excessively high taxation on entrepreneurs in the working-age population can adversely impact international competitiveness and it could also, following Easterlin's rel-

atively income hypothesis, reduce living standards relative to expectations, entrenching falling fertility and population ageing, starting a vicious circle. ***In a nutshell, it is possible to make a strong economic case for the position that modern mixed capitalist economies work best over time if fertility is close to the replacement level.*** We see a contradiction here between forces that operate on the individual and social level to reduce fertility to below replacement levels and the economic viability of key economic institutions and structures as fertility falls.

More recently, however, the position on falling fertility has become fragmented and more polarised with different groups and discourses now entrenched at opposite extreme of the question of whether sustained below replacement-fertility is a good thing.

The literature is now openly asking the question of whether low fertility could in fact be desirable. Striesness and Lutz (2014) rejects the original analysis that sustained below-replacement level fertility is detrimental, due mainly to adverse socioeconomic factors, and instead introduce the idea of an education-weighted dependency ratio, EWDR and climate change to argue that;

"The first very tentative results seem to suggest that perhaps longer-term fertility levels somewhere between 1.5 and 1.8 are the best for our planet and will, at the same time, result in future higher welfare as long as we invest more in the education of our slowly declining number of children."

The EWDR relies on the intuition from Becker and Kaplan to posit that the optimal level of fertility is one that produces relatively highly educated populations—less quantity, more quality—and low

public entitlements—highly educated people don't need them—which, according to the authors' simulations happen at fertility rates well below the replacement level. The introduction of climate change into the argument for lower fertility is controversial, mainly because it is a position that lends itself to extreme views. Striesness and Lutz (2014) states the obvious:

"If we were to care only about this environmental dimension, there would be little doubt that fewer people would be better and the resulting OLF [optimal level of fertility] would be zero."

Striesness and Lutz (2014) settles on a weight of 20% for the environmental dimension in the end, but this is an arbitrary number. Lutz (2017) discusses population ethics and climate change, and arrives at the same conclusion; a smaller and more highly educated global population is most desirable, and that this only happens with fertility levels well below replacement levels for an extended period.

The shift in the demographic literature towards a broader and more nuanced perspective on the effect of sub-replacement level fertility has coincided with a more radicalised and polarised discourse about falling fertility in the volatile cross-section between groups arguing in favour of mitigating climate change and environmental degradation, the survival of humanity and gender roles. It's possible to identify, at least, two extreme positions in this cross-section.

1) Reproduction is needed for the survival of the human race, groups, and lineages - The proponents of this view argue in favour of the most uncompromising version of the argument that sustained below-replacement fertility is maladaptive. According to this position, the rapid fall in fertility across many developed eco-

nomies is a result of fundamental cultural and socioeconomic flaws in modernity, which manifests themselves in weakness of the individual and the group. In this view, societies with very low, and falling fertility, are unhealthy for two reasons.

The first is linked to the mathematical reality that a population with below-replacement fertility will breed itself out of existence over time. Given the timeframe with which evolution works it is impossible to evaluate the optimality of birth rates in real time, but in countries with extremely low, and still-falling, fertility, it seems relevant to ask whether maladaptive forces are now at work.

The second, however, is more emotionally charged. It is associated with the fear that low fertility increases the risk that groups and societies are taken over and outcompeted by other groups with higher fertility. A classic version of this argument arises in the context of the fear—primarily on the political right—that western cultures are at risk of being out-bred by non-western cultures. This argument is often made in the context of immigration, where overpopulation in one region drives immigration to another, and the idea that immigrant population have higher fertility rates than the incumbent, often so-called, native population.

2) Low fertility is good, even necessary, because it helps mitigate the link between destructive climate change and environmental degradation from overpopulation - This argument is most often presented as the left-wing anti-thesis of the position stated above, though it comes in an extreme right-wing version too. According to the left-wing version of this view, humanity is a burden on planet earth, and to the extent that modernity limits reproduction it is a good thing. In the extreme version of this argu-

ment, having children is seen as a crime against the planet and eco-system, and is often linked to overall anxiety over the future of the planet. It is, in effect, a neo-Malthusian position. The most left-wing version of this argument shows itself in the so-called "de-growth" narrative in developed economies, where some interest groups argue that the modern economy itself, and by derivative the people in it, are engines for environmental degradation.

This argument comes in a right-wing version too, which follows from the argument above that low fertility in some cases is detrimental to the insider-group. In other words, this position takes the form of an insider-outsider view of the ability and right to existence and enjoyment of the wider environment for one group relative to another. National Socialism's view of the superiority of the Aryan ancestry and the need for lebensraum for the German people, is an extreme version of this story, grounded in the idea of a superior nation, and race. But this insider-outsider argument is not confined to the nation as the unit of analysis. It can exist on multiple levels of group and kin analysis. It is also sometimes linked to seemingly objective reproduction programs such as the idea to limit reproduction to the smartest people with the "best genes", or so-called eugenics. In the most "benevolent" version, it is a position that breeding should be managed for the greater good.

In both the left-wing and right-wing version of the neo-Malthusian argument, the exercise of power to achieve the desired results is paramount. This is because the stakes are high. After all, if we're at risk of making the planet uninhabitable due to over-population, doesn't it make sense to coerce people to have fewer children. Similarly, the argument in favour of brute force to suppress the existence of one group in favour of another is also easily justified with

an emergency narrative. All it needs is a story of a imminent danger for one group or nation from another, which legitimises the use of power to limit the reproduction, or existence, of the opposing group. The conflict in the Middle East between Hamas and Israel has clear undertones of these themes.

The extreme right-wing and left-wing neo-Malthusian positions are kindred spirits. If we are indeed in a climate emergency, it isn't a huge stretch to assume that power and coercion to limit population growth will be exercised across national groups and identities by those who are able to wield the biggest stick. This is easily the same argument levied in favour of an insider-group using force to drive out an outsider-group in a zero sum game over resources, territory or overall influence.

These extreme positions are canonical in the sense that almost everyone can identify with elements in them, if rarely their fully fledged versions. They are anchors for a debate about the optimal size of the population and the appropriate rate of reproduction, which has been a feature of public and scientific debate since the early days of civilisation, Izazola and Howett (2010). The success of a narrative at any given point in time is just as much a consequence of the political economy of the time as it is about the objective scientific position stated by any given researcher.

Confucius, Plato and Aristotle all treated the question of an optimal population size, mainly from the point of view of offering a foundation for sound government, and generally saw an expanding population as a good thing Izazola and Howett (2010). The first concrete proposal for an optimal global population size came in 1679 from Dutch scientist Antoni van Leeuwenhoek, who proposed that the

earth's land surface could support 13.4B people. This estimate was based on his calculation of a 1:13400 ratio between the landmass of the Netherlands and the earth's total landmass, from which it follows that the 1M people in the Netherlands at the time could be extrapolated to an optimal 13.4B for earth as a whole. Such estimates of earth's "carrying capacity" vary dramatically, by anywhere from 1 billion to 1 trillion, according to McGuigan (2022).

The discussion about an optimal population size evolved in the 18th century, mainly through a debate about the threat of earth reaching its carrying capacity, due to fears that the supply of food would not be able to keep up with the rise in population. This debate culminated with the treatise by Robert Malthus in 1798, famously linking the idea of an arithmetic expansion in the supply of food, which was inconsistent with a geometric increase in the size of the population. The timing of Malthus' intervention was as important for the popularity of the narrative as was the validity of the underlying argument Izazola and Howett (2010). This is ironic because the narrative underpinning the story told by Malthus peaked just as economic conditions were changing to render their predictions and assumptions mute. Izazola and Howett (2010) notes that;

"Malthus' theory and implications thrived at an academic level during a period which, in retrospect, would falsify them; the technological changes and demographic transition that took place in Europe during the end of the 19th century and beginning of the 20th have no doubt proved the specific implications and corollaries of Malthus' theory wrong."

After WWII, the population-optimist work of Danish economist Ester Boserup in 1965, and the pessimistic story propelled by the Club of

Rome—in particular Denish Meadow's book *The Limits to Growth*, and Paul R. Erchlic's 1960 book *The Population Bomb*—are important milestones in the battle between these two positions. Boserup turns Malthus' thesis on its head, arguing that it is population and population density which forces agricultural output to increase through technological progress rather than the former acting as a binding constraint on the latter.

By contrast, Erchlic and Meadows gloomily predicted the collapse of the global eco-system by 1980 and 2025, respectively. Meadows still has a few years to go as I type this, but I suspect he will be proven wrong, eventually. It is difficult to refute the idea that earth could potentially reach carrying capacity such that Malthusian forces kick in to limit population growth. Indeed, in some countries, characterised by the absence of economic development, you could probably argue that such mechanisms are in place today. But it is also difficult to deny the potential for Boserup's tenet to be true; namely that causality works from population growth to technological progress, which will allow earth to sustain more people as the pressure to innovate rises.

The state of play in this debate at the start of the 2020s play on all the themes discussed above, and the positions are getting increasingly extreme. We see at least two clear fronts; first between those who argue that rapidly falling birth rates represent a fundamental malfunction in the social fabric and political economy, and those who argue that falling fertility is the natural result of conscious choice, especially by women, to forgo family formation for other goals. The former position is supported by the general query within evolutionary grounded sciences of whether modern reproductive trends are in fact maladaptive. In practice, however, the argument

that low birth rates are a problem is often elevated to an argument how women and couples ought to behave. This then rouses the latter position, which is quintessentially feminist, namely that a pro-natal argument is chauvinist because it, in a modern context, forces women to spend more resources on reproduction than they would like. Of course, the feminist position is not clear cut here. In a recent *The Nation* article, Donegan (2023), describes the dichotomy between two feminist positions, one, the natural birth movement, which emphasises the role of women as mothers and their unique reproductive importance and a modern feminist position which sees this as a narrow and constricting interpretation of women today.

The key point is that the contradiction between the positions above is even more about value rather than a discussion about the evolutionary viability of sustained below replacement level fertility.
The second front is underpinned by the encroachment of the climate change/emergency discourse on the discussion about global population growth and birth rates. In the extreme version of this story, having children is seen as a crime against humanity, and women who decide not to have children are elevated for exercising their free choice. The counterargument runs along two axes; one argues that climate change is much less a threat than assumed by the consensus and the other, correctly in my view, highlights the thorny question of which people it is that we should get rid off, or prevent from procreation, to keep the size of the global population down to protect the climate.

CONCLUSION

This chapter has covered an awful lot of ground, but mercifully, its main conclusion is reasonably simple. The fall in fertility to below replacement in one country after the other since the 1970s and 1980s is driven to a large extent by accelerated tempo effects, or birth postponement. The speed and breadth of this process varies across countries, but it is now clear in all countries and regions, save Sub-Saharan Africa. Judging by the shifts towards pronounced postponement since the 2010s in Anglo-Saxon economies, Asia and Latin America, it is clear that global period fertility will soon fall below the replacement level, and that this will be primarily a result of tempo effects. As of 2022, the UN's population projections do not yet take this shift into account, but even the slow-moving population forecasts by the UN will soon have to acknowledge this. Postponement implies catch-up as women recuperate births later in their fertility career, but it is also increasingly clear that sustained tempo effects have a lagged impact on quantum effects.

Why is this happening?

In a recent investigation by *the Financial Times* on the continued decline in global birth rates, the Finnish demographer Anna Rotkirch throws her hands up in air;

"The strange thing with fertility is nobody really knows what's going on. The policy responses are untried because it's a new situation. It's not primarily driven by economics or family policies. It's something cultural, psychological, biological, cognitive."

I sympathise with Ms. Rotkirch, but hopefully, the analysis above offers some progress in terms of an explanation.

The hypothesis of a second demographic transition links falling fertility via birth postponement to two broad shifts; the first is a shift in values towards individualism and self-realisation driving a decline in marriage rates and a fall in the traditional cohabitation and union formation habits which have been instrumental in keeping fertility close to replacement levels. This mechanism diffuses through cultural evolution pathways that act horizontally across different groups in society, and vertically across shifts in behaviour passed down from parents to children over generations.

The second shift is socioeconomic and comes in a positive and negative version. In the former, modernity has increased the opportunity cost of early reproduction and cohabitation. This is especially true for women with the increase in their educational attainment and higher relative labour force participation rates since the 1970s, a process which is still underway today. Coupled with more widely available contraception and assisted reproductive techologies, foregoing or delayging reproduction has become both easier, more convenient and more lucrative in monetary terms.

In the latter negative version, a decline in *relative* living standards for younger generations—linked to Easterlin's relative income hypothesis—is driving a delay in household formation and childrearing. This is particularly related to the rising cost of housing and education, which forces young people to delay family formation.

Crucially, these two drivers are not mutually exclusive. It is perfectly reasonable to imagine both working at the same time, on different cohorts and generations of women, to drive down fertility.

Is falling fertility a problem?

In the late 1990s and early 2000s, the trend towards sustained sub-replacement fertility was understood and described as a challenge for public policy and the economy more generally. The notion of a fertility trap describing the risk of an entrenched decline in birth rates once period fertility falls below a certain level is the clearest example of this position. It emphasises the risk that rapidly falling birth rates is the result of economic conditions pushing fertility below the desired level and increasing unwanted childlessness. It also highlights the issues arising from population ageing in the stability of modern market economies with welfare institutions.

More recently, however, the debate on the continuing fertility decline has become nuanced, but also increasingly polarised. Climate change has entered the proverbial chat via the focus on falling fertility as potentially advantageous if it reduces the ecological and environmental foot print of humans on the planet. A pro-low fertility feminist view also has emerged, emphasizing the right of women to choose to have fewer children later, or none at all.

This is stands in contrast to a the view that a rise in the incidence of sub-replacement fertility, and increasingly very low fertility in many large countries, reflects something fundamentally wrong, and indeed maladaptive. This position does not just latch on to the obvious mathematical fact that a population with sub-replacement fertility will cease to exist over time. It is a broader criticism of the shift in values described above towards self-realisation and individuralism away from the family as the central unit of importance. It is also, increasingly in Conservative groups, seen as a threat to the dominance of West and liberal democracies, who from this vant-

age point, are in an existential contest with fundamentally opposing cultures, where the ability to match the fecundity of one's opponent is a key prerequisite for victory.

What happens next?

Judging by the persistence of below-replacement fertility in countries who are furthest along in their demographic transitions and accelerating tempo effects in those next in line, global period fertility will fall sharply in the next decade. This, in turn, means that the global TFR will soon fall below the replacement level. This undoubtedly will further polarise the debate on the trend in global birth rates. If this is true, all evidence suggests that it will have a lagged effect on cohort fertility. It is possible that we are entering a period with a prolonged period of sub-replacement global fertility. Is this bad? The answer, however unsatisfactory is; it depends.