# Exploring age-specific and cumulative cohort rates using composite fertility lattice plots: an international comparison of Human Fertility Database and Human Fertility Collection data

# Abstract

## BACKGROUND

The Human Fertility Database (HFD), and the related Human Fertility Collection (HFC), provide highly disaggregated data on age-specific fertility rates for 45 countries. These sources provide a wealth of opportunity for learning about the development of different pathways of transition to low fertility both within and between countries and geographical regions.

## OBJECTIVE The aim of this paper is to use composite fertility lattice plots, which combine information from different visualisation methods of the Lexis surface, such as heat maps and shaded contour plots, to explore changes in age-specific fertility rates and ensued cumulative cohort fertility rates across countries and geographical regions.

## METHODS

Standard shaded contour maps use both shade and contour to represent the same variable. In our plots we instead use colour/shade to indicate age-specific fertility rates, and a series of distinct contour lines to indicate the cumulative fertility rates reached by different cohorts at different ages. These figures are then ranked by fertility in the last commonly observed period, and colour coded according to geographic region.

## RESULTS

By looking first at the thick solid contour lines from left to right in each population figure, we can see at which age different cohorts either reached or last reached replacement fertility levels. Other contours help understand the degree of shortfall below replacement levels for different cohorts. It appears that, once countries have lost a replacement fertility level, they tend to not regain this fertility level. Exceptions are Norway and the USA, which saw rising fertility rates for cohorts born after 1950s and late 1960s, respectively. The age-specific fertility trends, as well as broader political and socioeconomic conditions in these countries, are very different, suggesting different paths by which replacement fertility rates can be achieved.

## CONTRIBUTION

Complex data visualisations show, in an intuitive way, how age-specific fertility rates are related to cumulative cohort fertility rates. Combining this information enables us to explore differences between countries and geographical regions and can make an important contribution to comparative demographic research on fertility and other trends.

# Introduction

One of the key drivers underpinning population change is the continuing transition to low fertility (Billari and Kohler 2004; Lutz and Skirbekk 2005). The fertility transition, as a process of fertility decline from high to below replacement levels, is one of the cornerstones of the demographic transition, both in its classic proposition (Davis 1945; Notestein 1945, 1953) and in the proposition of the ‘second demographic transition’ (Lesthaeghe 1995, 2010; Lesthaeghe and van de Kaa 1987). The transition to low fertility interrelates with improvements in life expectancy (Oepen and Vaupel 2002; Shkolnikov et al. 2011; Sánchez-Barricarte 2017) and international migration patterns (Adsera and Ferrer 2014; Andersson, 2004; Dubuc, 2012; Kulu, 2005; Sobotka, 2008).

Since the late 1960s, most countries from the Western world have undergone a rapid fertility decline which have been accompanied by an increasing postponement of childbearing to older ages. Within the European context, there is an emerging debate about the persistence of low fertility, with some studies reporting recuperation in cohort fertility trends (see Caltabiano, Castiglioni and Rosina (2009) for Italy; Andersson et al. (2009) for Nordic countries and Sobotka et al. (2011) for German-speaking countries and the United States) and others revealing a reversal in period fertility trends among European countries (Goldstein, Sobotka, and Jasilioniene 2009; Myrsylä, Goldstein, and Cheng 2013). Yet the transition to low and later fertility has become a characteristic feature which is not solely restricted to the European and, more generally, Western regions, but encompasses an increasing number of non-Western countries, notably those in East and South-Asian regions (Basten, Sobotka, and Zeman 2014; Frejka 2017; Frejka, Jones, and Sardon 2010; Jones 2009).

The transition to low fertility has been shaped by a complex and dynamic combination of technological, institutional, socio-economic, political, cultural and normative factors. The driving forces behind the transition to low fertility include the diffusion of modern contraception (Montgomery and Casterline 1996; Kohler, Behrman, and Watkins 2001), educational expansion (Blossfeld and Huinink 1991; Skirbekk 2008; Lutz, Butz, and Samir 2014), the large-scale entry of women into the labour force (Adsera 2011; Oppenheimer 1994; Sundstrom 1992), family policies (Thevenon 2011; Rindfuss and Choe 2015), globalisation and economic uncertainty (Mills and Blossfeld 2005), ideational changes (Lesthaeghe and Surkyn 1988), gender roles and gender equality (Anderson and Kohler 2015; Esping-Andersen and Billari 2015; Frejka 2017;.Goldscheider et al. 2015, McDonald 2002, 2006).

The aim of this paper is to provide a description and an illustration of different fertility trends across 45 countries and regions from major geographical clusters or macro-regions within and outside Europe (for full details see the online appendix). The latter reflect different institutional, societal, economic and cultural conditions which have shaped different pathways of low fertility transition. Changes in age-specific and cumulative cohort fertility rates are explored by using composite fertility lattice plots, which combine information derived from different visualisation methods of the Lexis surface, namely heat maps and shaded contour plots (see also Schöley and Willekens, 2015).

In particular, we present and describe the processes involved in producing a lattice plot which shows both age-specific fertility rates (ASFRs) available by year, and implied cumulative cohort fertility rates (CCFRs) for dozens of countries whose data is available either within the HFD or HFC. The end result is a lattice of Lexis surface visualisations, arranged by cohort on the horizontal axis and age on the vertical axis. Within these visualisations the ASFRs are represented graphically by cell shade, CCFRs by a series of easily distinguishable contour lines, and the strips which label each of the small multiples within the visualisation are coloured according to geographic region. We refer to this visualisation as a Composite Fertility Plot (CFP), as both colour/shade, and contour lines, are used to represent different variables.

We suggest that the full visualisation produced is best viewed by being printed out in colour as an A3 or A2 poster, which we have found to greatly facilitate ‘at-a-glance’ comparison between countries; (Vaupel, Gambill, and Yashin 1987) to this end we make a high resolution version of the final visualisation available online, along with the R code used to produce this final visualisation and additional figures for individual countries. Within this paper we present this final visualisation, as well as smaller visualisations comprising a small subset of countries to illustrate the value of this approach for both within-country and cross-country comparison of fertility trends.

# Methods

## Data

To provide broad geographic and temporal coverage, data from the Human Fertility Database (HFD) and Human Fertility Collection (HFC) were combined(Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria) 2015; Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria) 2016). The Human Fertility Database (HFD) includes age-specific fertility rates (ASFRs) for 28 countries over a range of periods, and is entirely based on official vital statistics from each country; data from each country are rigorously checked and standardised according to a detailed methods protocol. In addition to overall ASFRs birth-order specific fertility rates are also included in the HFD, alongside summary statistics such as crude birth rates, mean age at birth, and total fertility rates;(Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria) 2016) however for the present visualisations only overall ASFRs are used. The HFC is a complementary database to the HFD, created by the same institutions, which aims to supplement the high-quality, official vital statistics contained in the HFD with data from additional sources whose data quality may not be of the same standard. (Grigorieva et al. 2015) Data disaggregated by age in single year was used from both the HFD and HFC, and Lexis squares (one year by one year) rather than Lexis triangles or Lexis parallelograms were used, as in practice the use of squares only has limited effect on precision of cohort estimates. (Caselli and Vallin 2005) In addition to the 28 countries from the HFD, data from 48 countries were used, but not all were included in the final lattice visualisation of 45 countries because of limited periods of observation. Further details are provided in the online appendix.

Data from the HFD and HFC are overlapping for some years and countries. Where more than one country- and year-specific ASFR was available, the value from the HFD was used in first preference; otherwise, records from the HFC were used in the following order of preference according to the ‘collection’ field of the HFC dataset: 1) STAT (Official Statistical Data); 2) ODE (Data from the European Demographic Observatory, L’Observatoire Démographique Européen); 3) RE (Research estimates). For almost all countries, this approach produced a dataset comprising ASFRs for many contiguous years; linear interpolations of ASFRs were used for the few countries and years where this was not the case. The code for both combining data across sources and interpolating values where necessary are available in the online appendix.

## Lexis surface mappings

For each country, ASFRs were arranged onto a Lexis surface with birth cohort year on the horizontal axis and age in years on the vertical axis; as mentioned previously Lexis squares rather than triangles or parallelograms were used throughout, so these are not true cohort estimates, but are sufficient to illustrate the principles of the visualisation, and the R code is freely available for researchers to iterate the approach further. ASFRs for each country and year were mapped to colours and shades using the Spectral colour palate from the RColorBrewer R package. (Neuwirth 2014) The R packages Lattice and LatticeExtra were used to produce the visualisations. (Sarkar 2008)

For birth cohorts where ASFRs were available from age 15 years, CCFRs were produced for each age and cohort year. If refers to the ASFR for country , in year , and at age , then the CCFR for age can be defined as , where is the simple index of birth cohort (. Within the Lexis surfaces, contour lines were added at positions across the cohort-age surface where reached specific values. More specifically, a thin dashed contour indicates of 1.30 babies per woman, a thin solid contour indicates of 1.50, a thick dashed contour indicates of 1.80, and a thick solid contour indicates a replacement fertility level of 2.05. Because is a cumulative quantity, the contour lines will always have the same monotonic ordering - thin dashed, thin solid, thick dashed then thick solid - from bottom to top. The position, presence or absence, and trajectories of these contour lines across different birth cohorts, and between countries, are all useful indicators of how fertility trends have changed over time and place, and how far short of replacement fertility levels most affluent world nations fall. All countries are arranged in order of highest CCFR in 2007, the last common year of observation for all populations being compared.

# Results

This section will first provide the visualisation for two or more countries at a time, in order to both introduce the visualisation and illustrate its utility for both cross-cohort and cross-country comparisons. The complete visualisation will then be presented as the final figure, though as discussed previously we recommend this last visualisation be downloaded and printed in colour as a large landscape poster.

## Low fertility European countries: East Germany, West Germany, Italy and Spain

Figure 1 shows the CFP for Spain, Italy, East and West Germany. Each of these countries currently have fairly low fertility levels, and by looking at both the colour of the cells and the trajectories of the contours we can learn more about other similarities and differences between these countries. Firstly, if we look at the trajectory of the thick solid contour line we can identify the last birth cohorts that reached replacement fertility levels; this is around 1938-9 for Germany, 1945 for Italy, and 1954 for Spain. The thick dashed lines, indicating 1.80 babies/woman, are more dissimilar within Germany, and were first not achieved in West Germany for the 1944 birth cohort, and for East Germany, for the 1949 cohort; in East Germany this fertility level was then re-established briefly, for cohorts born around 1952-1960, before being lost again; this fertility level was last for Italian cohorts born in 1955, and Spanish cohorts born in 1960. The thin solid line, 1.50 babies/woman, looks like it may last be achieved for cohorts born between 1970-75 in all four countries; and in all four populations it appears levels of 1.30 will continue to be met for future cohorts, though by around the age of 35 years in East/West Germany, and by age 40 in Spain and Italy; Spain and Italy therefore appear at greater risk of future cohorts not reaching even the 1.30 level, and so for total fertility in these countries to fall below those in Germany, though recent research suggests a fertility rebound may have occurred since the 2000s in Italy [REF for Italy] , and may also be increasing in Germany [REF for Germany – Sobotka]. In all four populations, a shift from lower to higher ages of peak fertility is observed by noting that cell colours between around age 20 and 25 years change from yellow to green over time.

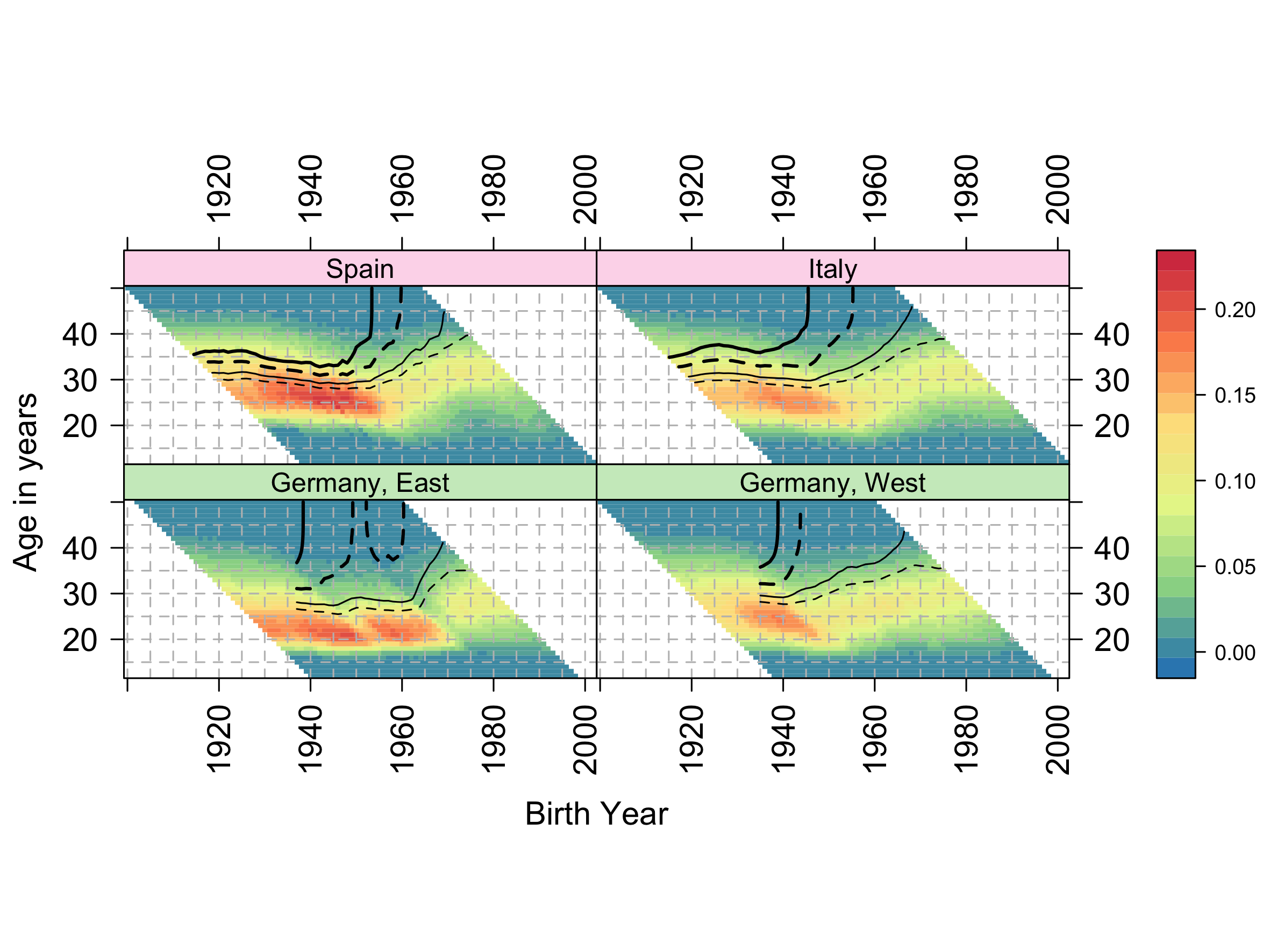


Figure 1 Composite Fertility Lattice Plot for Spain, Italy, East and West Germany. Data source: HFD/HFC

## Norway and non-European English-speaking countries

Figure 2 compares fertility trends of Norway with two pairs of English-speaking countries - New Zealand and Australia; and the USA and Canada. Each of these countries have relatively high fertility for affluent world nations. Compared with Australia, New Zealand has somewhat higher fertility with total cohort fertility levels (thick solid contour) being achieved by age 40 years in the 1975 cohort, and with this contour appearing to level off at around this age; this is in contrast to Australia, where replacement levels were last met for the 1966 cohort.

Norway and the USA, though different in many ways, are similar in their fertility trends, in that they are the only populations within the dataset where replacement fertility (RF) levels by cohort were ‘recovered’ by latter cohorts after being ‘lost’ by earlier cohorts; this can be seen by noting the parallel pairs of thick black vertical contour lines in both countries. In Norway, replacement fertility was first ‘lost’ after around the 1953 cohort then ‘recovered’ after the 1956 cohort; and in the USA RF was ‘lost’ by the 1950 cohort then ‘recovered’ after the 1964 cohort. Within Norway RF levels look set to disappear with latter cohorts, as the RF line hovers around age 43 in the latest period, whereas in the USA it appears to have stabilised around age 36. Of these populations, Canada has the lowest fertility, and appears to have stabilised at levels between 1.50 and 1.80 babies/woman.

Looking at the colour and shading of cells, the USA is distinct amongst affluent world nations in having a wider range of ages of high fertility, with less evidence of fertility postponement within the population overall than in most other comparable nations.

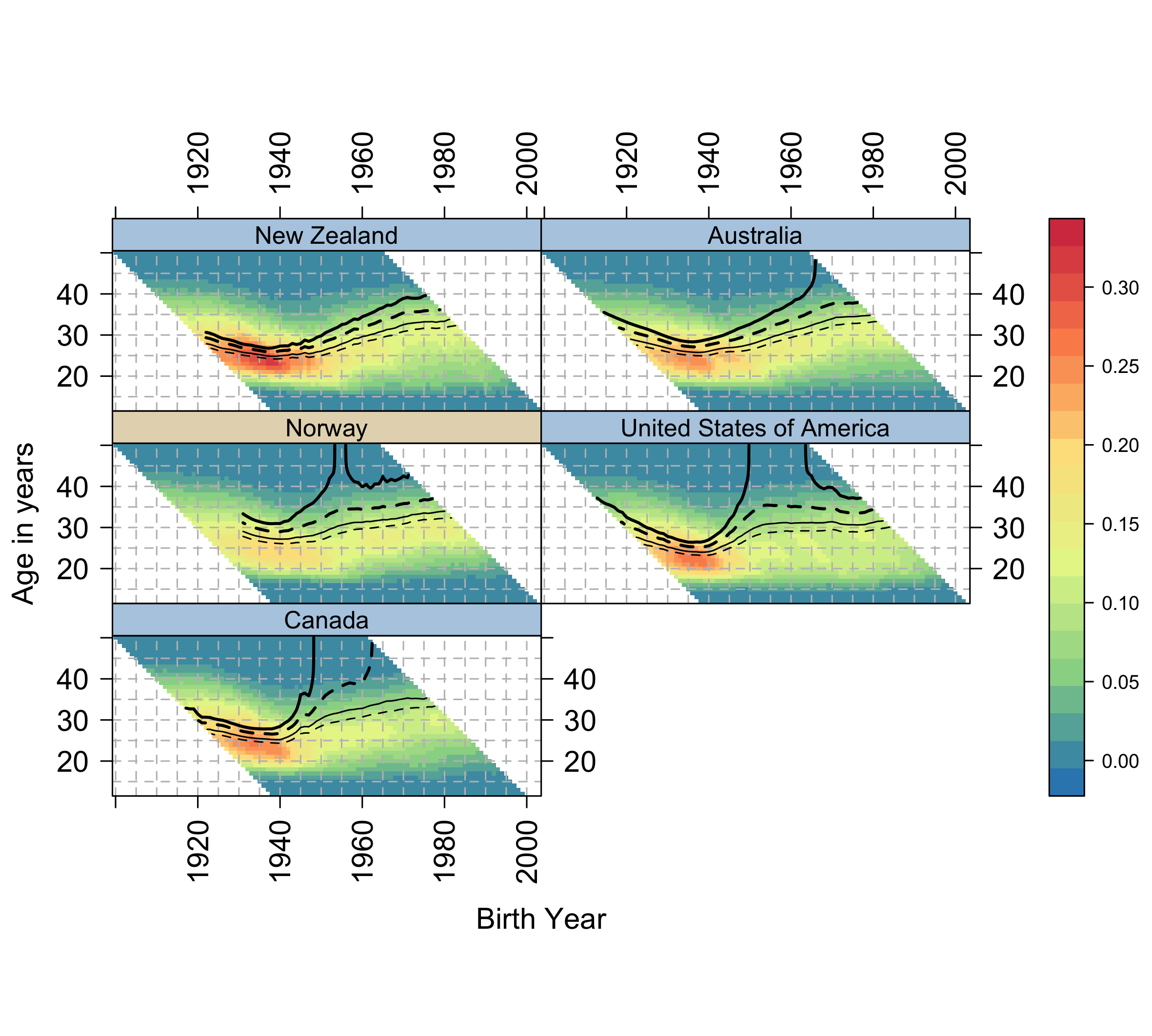


Figure 2 Composite Fertility Lattice Plots for Norway and non-European English-speaking countries. Data source: HFD/HFC

## Ireland and Great Britain

As Figure 4 indicates, the British Isles comprise populations with very different fertility trends. Both the Republic of Ireland (RoI) and Northern Ireland are relatively small populations, withfairly high fertility. RoI appears likely to stabilise above RF, and Northern Ireland to either do so as well or only fall slightly below RF. England & Wales – treated as one population – lost RF after the 1951 birth cohort, but appear to have stabilised slightly above 1.80 babies/woman. Scotland lost RF by the same birth cohort, but then also lost fertility levels of 1.80 babies/woman after the 1960 birth cohort.

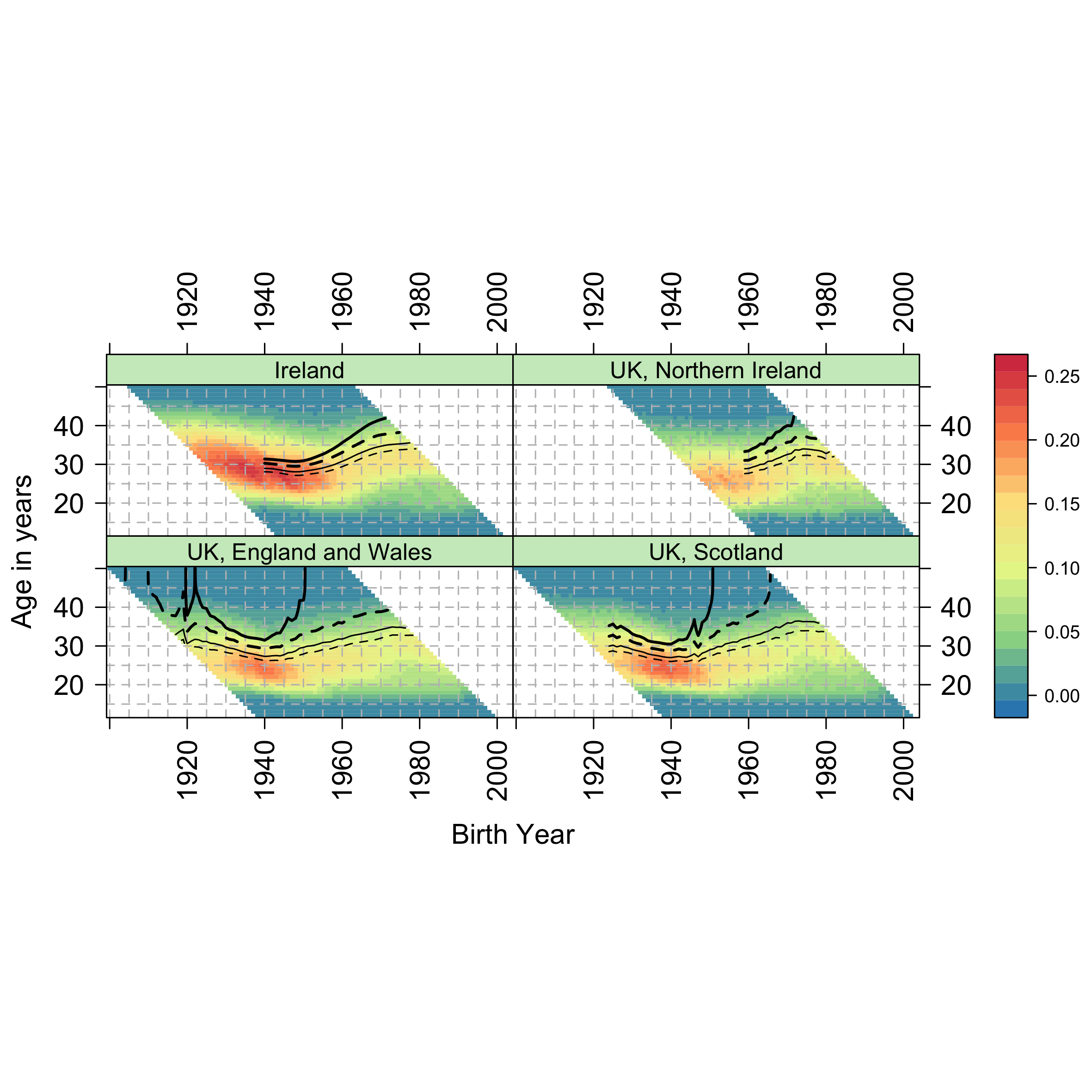


Figure 4 Composite Fertility Lattice Plots for Ireland, Northern Ireland, England and Wales, and Scotland. Data source: HFD/HFC

## East and South Asian Countries

Figure 3 compares Taiwan, Japan and the Republic of Korea. All countries industrialised very rapidly, leading to very different experiences by cohort. This is reflected in very rapid changes in both total fertility by cohort, and the ages of peak fertility. Japan is characterised by long-standing low fertility, and fertility levels by cohort were very close to 2.05 for cohorts born between around 1933 and 1944, with some cohorts reaching this level but not others of very similar birth year this variability between cohorts as to whether RF levels were met results in a series of vertical lines after around the age of 43 for these cohorts. After the 1944 cohort, fertility then fell briefly below 1.80 for around the 1945 cohort, then again after the 1960 cohort; cohorts born around 1970 look unlikely to meet 1.50, and the age at which the latter cohorts are reaching 1.30 babies/woman is now approaching 40 years, so fertility may decline yet further.

Taiwan perhaps saw the fastest decline of the three populations in fertility, moving from a higher fertility for cohorts born up to around 1955, to very low fertility for cohorts born after around 1970. Of the three populations, the Republic of Korea appears more likely to sustain fertility levels of between 1.50 and 1.80 babies/woman, and so fall less far behind RF levels than the other countries.

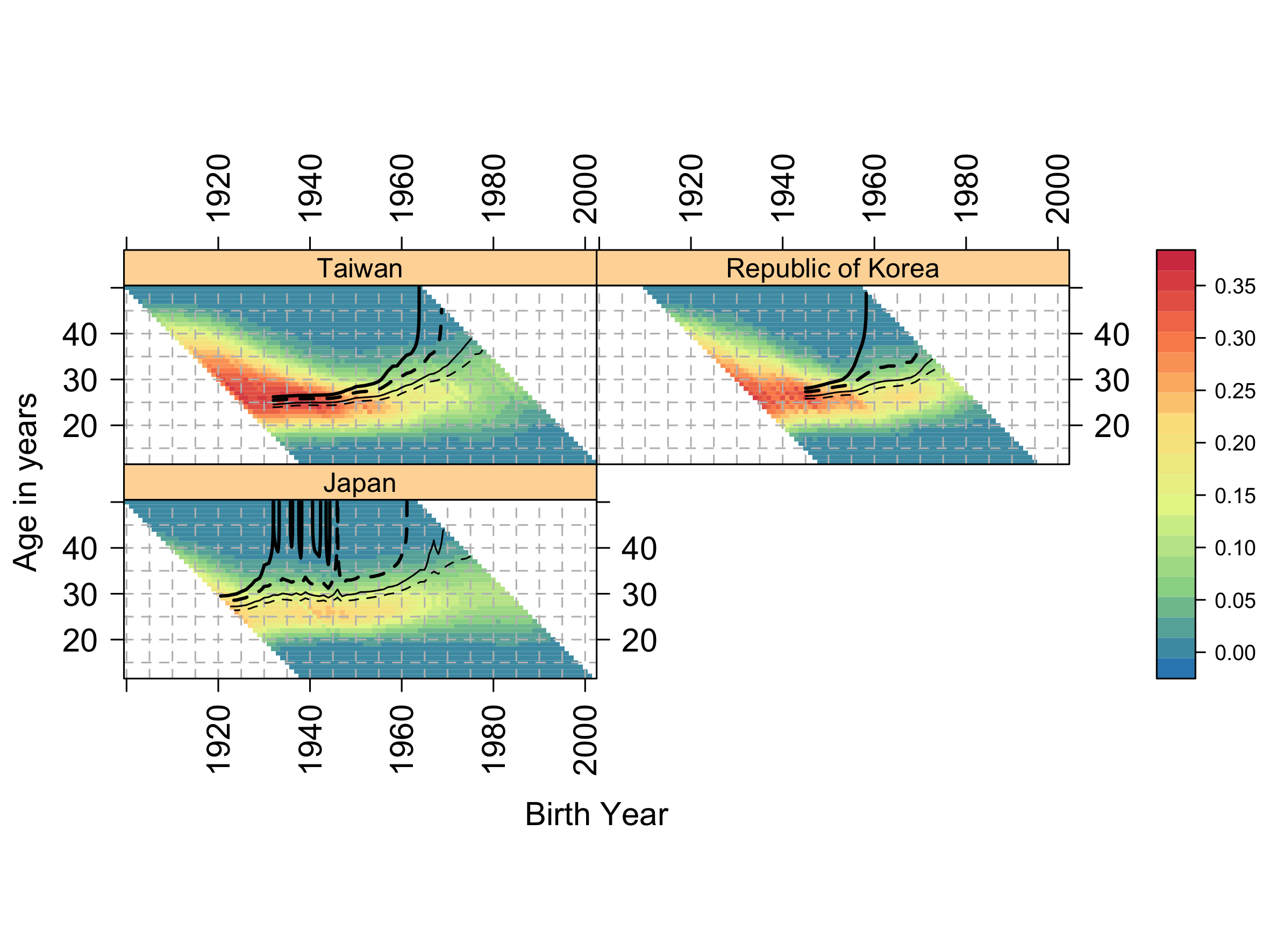


Figure 3 Composite Fertility Lattice Plots for Taiwan, Republic of Korea, and Japan. Data source: HFD/HFC

## All countries

Figure 5 shows the composite fertility lattice plots for 45 countries, and so represents a visual summary of hundreds of thousands of separate ASFRs. Because of the number of populations being compared, HFC/HFD population codes are used and no gridlines are added to avoid overplotting. A separate version of the figure is available at a higher resolution, with full country names and gridlines, in the online appendix. A common scale is used to for ASFRs for all countries. As mentioned previously, the colour of the strips which label each country indicates the broad geographic macro-regional grouping of countries ,and fertility trends can be further explored by comparing the fertility trends of countries by these groups.

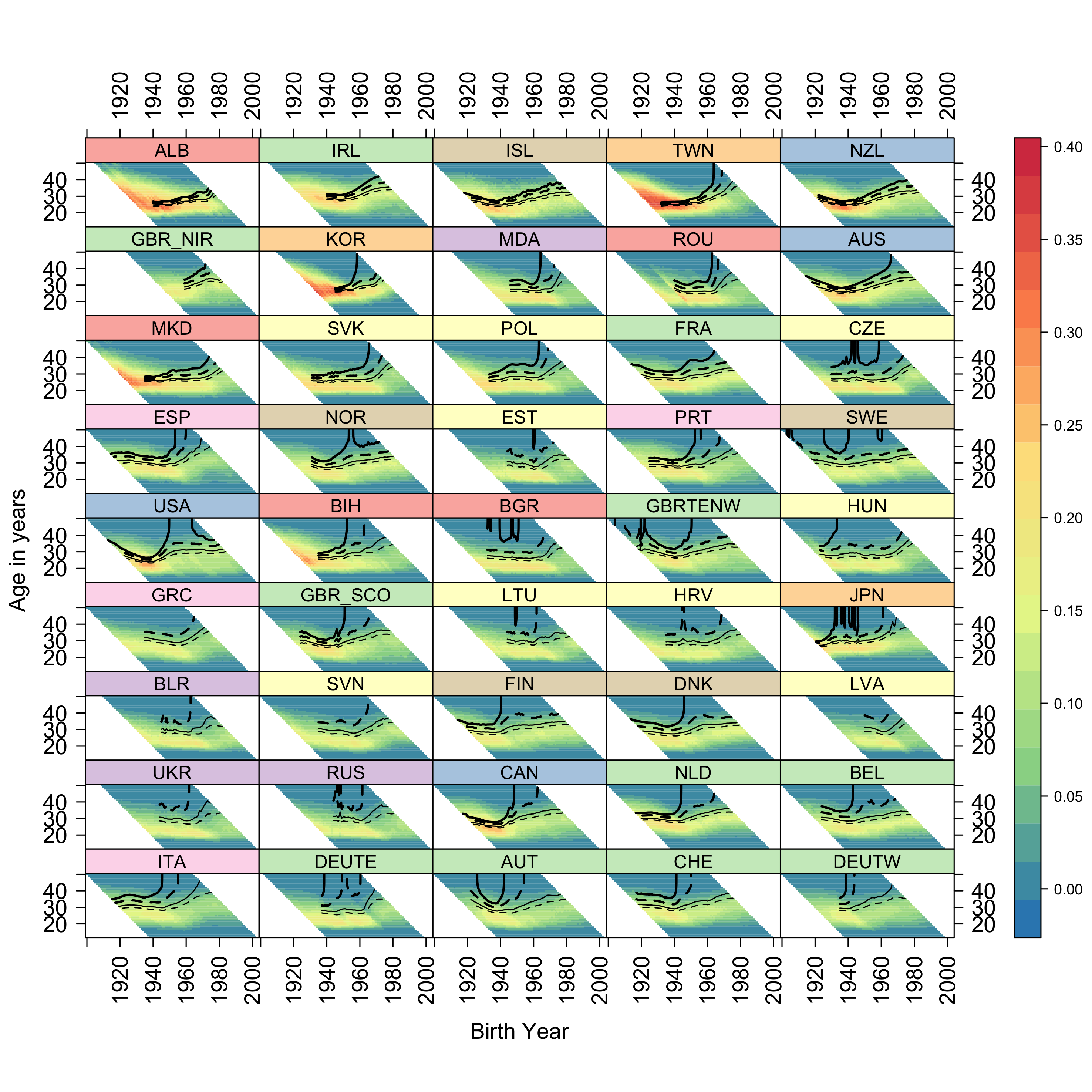


Figure 5 Composite Fertility Lattice Plot for 45 countries. Data source: HFD/HFC

# Summary and conclusions

Complex data visualisations through composite fertility lattice plots reveal a considerable degree of heterogeneity in the development of fertility trends over time both within and across geographical macro-regions. Fertility patterns reconstructed by combining data visualisations of age-specific and implied cumulative cohort fertility rates appear to reflect differences in the institutional, socio-economic, political, cultural and normative contexts, suggesting different pathways of transition to low fertility.

Among low fertility countries, we illustrated the cases of Italy and Spain versus East and West Germany. For Germany, it appears that replacement fertility levels were lost for cohorts born in the period preceding World War II, compared to Italy and Spain, where replacement fertility was loss for cohorts born in subsequent periods. In line with other studies (Frejka 2017; Sobotka et al 2011), in East and West Germany fertility rates continued to steadily decrease across successive cohorts and they appear to have stabilised at levels between 1.30 and 1.50 babies/woman. By contrast, Spain and Italy display signs for future cohorts of not being able even to achieve very low levels of cohort fertility (1.30 babies/woman).

Notable exceptions in the HFD/HFC dataset are the cases of Norway and the USA which appear to have ‘lost’ and subsequently ‘recovered’ replacement fertility over time, by moving from earlier to more recent cohorts. This is in line with previous studies (Burkimsher 2017; Shkolnikov et al. 2007). Recuperation patterns may be due, among other factors, to the influence of immigrant fertility (for USA, see Choi 2014; Fernandez and Fogli 2006, 2009; for Norway, see Østby 2002 and for alternative views, see Kravdal 2016 and Tønnessen 2015).

When comparing European English-speaking countries, the Republic of Ireland and Northern Ireland seem to have stabilised around replacement fertility. By contrast, England and Wales appears to sustain fertility levels of around 1.80 babies/woman, whereas Scotland displays lower cohort fertility levels. Recent studies suggest that immigration may play a significant role in current UK fertility levels (Dubuc 2012; Kulu et al. 2017; Robards and Berrington 2016, Tromans, Natamba, and Jefferies 2009).

Among Asian countries, Japan features long-standing low fertility patterns, with earlier cohorts showing a high level of variability in high fertility and recent cohorts stabilising at very low fertility levels. Taiwan stands out for having witnessed the fastest transition to low fertility, moving from high to very low cohort fertility in the space of 15 years, whereas the Republic of Korea appears to have stabilised at higher cohort fertility levels. The fertility patterns among Asian countries may reflect different socio-economic, policy and cultural contexts (Fukuda, 2016; Gauthier 2016; Lee and Lin 2016; Straughan, Chan, and Jones 2009; Suzuki 2013; Westley, Choe, and Retherford 2010).

The composite data visualisations which were presented as illustrative cases in this paper provide an initial basis intended to encourage further investigation in the development of different pathways of the transition to low fertility across ages, cohorts and countries. We hope that at-a-glance comparisons of these different analytical dimensions may contribute to help to formulate new research questions and hypotheses on the possible mechanisms underlying different fertility pathways.