

PNAS

1

² Supporting Information for

³ Masculinization of populations reverses sex difference in fertility

⁴ Henrik-Alexander Schubert, Thomas Spoorenberg, Christian Dudel and Vegard Fykse Skirbekk

⁵ Corresponding Author: Henrik-Alexander Schubert.

⁶ E-mail: schubert@demogr.mpg.de

⁷ This PDF file includes:

⁸ Supporting text

⁹ Figs. S1 to S12

¹⁰ Tables S1 to S6

¹¹ Legends for Dataset S1 to S9

¹² SI References

¹³ Other supporting materials for this manuscript include the following:

¹⁴ Datasets S1 to S9

15 **Supporting Information Text**

16 **1. Regression-based approach**

17 Male fertility is a topic in demography that is widely sidelined for some substantial and methodological reasons, despite
18 prominent calls to bring men back in (1). Among those reasons are data quality, data availability and also the utility of the
19 estimates, because most population projections take a female-centered approach. The data quality of paternal information is
20 usually inferior to female fertility, because men tend to under-report children when they do not live with them in the same
21 household (2) and vital statistics data contains usually a large fraction of missing values on paternal age or paternal age is only
22 measured in married couples (3). Recent methodological advancements attempted to overcome these problems (3, 4). This
23 resulted in a wider availability of data on male fertility (5–7), uncovering new insights.

24 We obtain male fertility data through an indirect demographic approach proposed by Keilman et al. (8), which relates the
25 TFR for men to the TFR for women and the sex ratio at reproductive age, see Equation 1. We have adjusted the model to be
26 better suited for contemporary estimations. The first adjustment (model 2) accounts for the fertility postponement, shifting
27 the age window for estimating the reproductive sex ratio from 20–39 to 25–44, which captures better the fertility intensities
28 after global fertility postponement (9). The second alternative (model 3) accounts also for the age gap between fathers and
29 mothers through estimating the sex ratio between 25–44 years old men to 20–39 years old women.

30 The model is trained on male fertility data described in Table S1 and following equation 1. The data sources contained
31 individual-level birth records with information on paternal age and maternal age (see top panel of Table S1), or readily available
32 data at the country level (see bottom panel of Table S1). If the paternal age was missing, the information was imputed using
33 the conditional approach proposed in Dudel and Klüsener (3). More information on the data curation of the subnational data
34 can be found in Schubert and Dudel (5), information on the data from the Human Fertility Collection can be obtained from
35 Dudel and Klüsener (6), and the estimation method for the DHS data is described in Schoumaker (4). The data from the
36 Human Fertility Collection, which is arguably the gold standard of male-fertility data, has been excluded from the model
37 training and is only used for the model validation, see Section 2. The model looks as follows,

$$\log(TFR_m) = \alpha + \beta_1 \log(TFR_w) + \beta_2 \log(SR) + \epsilon \quad [1]$$

38 where $\log(TFR_m)$ is the logarithm of the TFR for men, $\log(TFR_w)$ is the logarithm of the TFR for women, and $\log(SR)$ is
39 the logarithm of the sex ratio at reproductive age.

40 Using an indirect demographic approach makes several assumptions. First, the approach assumes that the model captures a
41 large fraction of variation in the male TFR. This is confirmed by a high R^2 . Second, it assumes that the relationship is the
42 same in the training as in the prediction data. The training data contains almost all countries and areas around the world and
43 several time episodes, which makes it a good representation of different demographic contexts. The out-of-sample prediction
44 confirms that the model is able to predict the TFR for men outside the sample, see Section 2.

45 **A. Prediction uncertainty.** We perform a robustness check accounting for the prediction uncertainty in the regression model to
46 ensure the robustness of the main results, in particular of the declining trend in TFR_m relative to TFR_w and the fertility
47 crossovers, and to understand the uncertainty in the regression model. The prediction interval contains very likely (with the
48 probability of $1 - \alpha$) the random future observation y_0 (10). We set $\alpha = 0.05$ and the prediction intervals are estimated the
49 following way:

$$x'_0 \hat{\beta} \pm t_{n-p}(1 - \alpha/2)\hat{\sigma}(1 + x'_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)^{1/2} \quad [2]$$

50 where x'_0 is the value for the predictor variables, $\hat{\beta}$ is the regression coefficients from model 1, t_{n-p} is the t-value from
51 the regression coefficients, $\hat{\sigma}$ is the standard deviation for the regression coefficients. Figures S3 and S4 show the uncertainty
52 around the predicted values of the TFR of men, largely confirming our main results.

55 **2. Out-of-sample Validation**

56 The model from the indirect approach to estimate TFR_m was evaluated using out-of-sample validation on data from the
57 Human Fertility Collection (6), which was withheld from the model training. The predicted values and the 90%-prediction
58 intervals are obtained from regression models using data from the WPP2024 on adult sex ratios and TFR_w . We are thus able
59 to compare the observed to the predicted value and also test the calibration of the uncertainty, see Figure S1. The figure shows
60 the observed TFR for men on the x-axis and the predicted TFR for men on the y-axis. Large deviations from the diagonal line
61 indicate larger prediction error. The vertical bars indicate the 90% prediction intervals. The model 3, in the bottom panel
62 labeled age-gap model, has the best prediction performance (smallest gaps between predicted and observed value), and the
63 best-calibrated prediction intervals. Furthermore, Figure S2 reveals that the average prediction error is the smallest.

64 **3. Demographic scenarios**

65 Furthermore, we account for different demographic scenarios of the future developments of population structures and fertility
66 rates implemented in the WPP2024 (for details on the estimation see 11). We include the following scenarios:

- 67 • Accelerated decline of adolescent birth rate (ABR) with recovery: Accelerated ABR decline with recovery of half of
68 reduced fertility once cohorts have aged 10 years

- Accelerated ABR decline: Age-specific fertility below age 20 declines by 20 per cent per year until ABR is below 10 births per 1000 women aged 15-19
- Constant mortality
- Fertility scenarios: Low fertility, High fertility, Instant replacement fertility, Constant fertility, No fertility below age 18
- Instant replacement zero migration
- 80 per cent prediction intervals: Lower 80 PI, Upper 80 PI
- 95 per cent prediction intervals: Lower 95 PI, Upper 95 PI
- Medium: the mean probabilistic projections for fertility and mortality, and median probabilistic projections for net migration
- Momentum: Instant-replacement fertility as of 2024, Constant mortality as of 2024, Zero migration from 2024
- Freeze rate: This approach assumes no change in the demographic parameters
- Zero migration

The results are displayed in Figures S5, S6 and S6, showing that the trends are largely parallel, but the scenarios differ in the level of the difference in the TFR for men to TFR for women. Scenarios with slower fertility decline and higher population growth rates, for instance the High scenario and the upper uncertainty boundaries, show more muted declines of the TFR for men relative to women. The instant replacement fertility scenario differs in the impact dependent on the current fertility level. If instant fertility would lead to a sudden increase in fertility, for instance among low fertility countries, it actually leads to higher TFR_m to TFR_w , see high-income panel in Figure S5. However, in regions with above replacement fertility, for instance in Sub-Saharan Africa, the relative difference drops more strongly, see bottom left panel in Figure S5. The low population growth scenarios and also lower prediction intervals are generally showing lower TFR_m relative to TFR_w .

To illustrate the impact of a WPP2024-scenario relative to the Medium-scenario used in the main results, we estimate the absolute difference between the scenario at hand and the Medium-scenario in the Figures S9 and S10. Thus, values higher than zero indicate that this specific scenario may lead to a higher relative difference than in the Medium-scenario. We can further distinguish between statistical scenarios (incorporating projection uncertainty) and demographic scenarios (modeling different deterministic scenarios for the evolution of demographic behavior), see Figures S7 and S8.

4. Standardization

We use standardization as a tool to demonstrate the impact of sex difference in the same age-groups on the difference in the Total Fertility Rate. The interpretation goes as follows, the estimate TFR for men would be the result if the mother and the father were always exactly the same age. While this assumption is not realistic (for evidence on parental age gaps see 12), it illustrates the impact of sex differences in the population structures. The results are presented in Figure S11 and corroborate the results from the indirect estimation model showing similar trends towards lower TFR_m relative to TFR_w over time, but the differences are a bit more muted, because those pattern are more reinforced by the actual age differences between parents.

5. Age-gap approach

The standardization simplifies the estimation by assuming that the mother is exactly the same age as the father, but one could also use empirical data on age differences between parents to obtain a more realistic estimation of the TFR_m that relaxes the assumption regarding the age differences. We do this in the age-gap approach by drawing on data on the conditional distribution of births by father's age y for each maternal age-group x from Dudel and Klüsener (6) and present the results in Figure S12. The conditional distribution $P(y | x)$ can be defined as the probability that a birth to a mother aged x is happening to a father aged y . The sum of the conditional distribution for each maternal age is 1: $1 = \sum_{i=15}^{55} P(y | x)$. Using such conditional distributions, we distribute the births using the following estimation: $B(y) = \sum_{i=15}^{55} B(x) \cdot P(y | x)$. The standardization approach is a special case of the age-gap approach in which for the case $x = y \Rightarrow P(y | x) = 1$ and for all other ages y is $P(y | x) = 0$, so that births are not shifted to other age groups.

The results from the standardization, age-gap, and regression-based approach show similar trends overall, see Figure S12. The age-gap and the regression-based approach are particularly close, which highlights that our regression-based approach, which is used for the main results, is capturing both sex imbalances and age gaps well. Moreover, the 90% Prediction Intervals contain the results from all three approaches, which indicates that the model uncertainty is well calibrated. The TFR ratio from the standardization approach is usually slightly lower than from the other approaches, which shows that age gaps partially offset the impact of population imbalances in age groups.

Table S1. The table summarizes the male fertility data used in the study to train the regression models, providing information on the country, the observation period, the spatial unit, the number of spatial units, and a link to the source.

Country	Period	Level	Units	Source
Australia	1990-2020	States, Territories	8	explore.data.abs.gov.au
Colombia	1998-2020	Departments	32	https://microdatos.dane.gov.co
Finland	1990-2020	Regions	19	https://www.stat.fi/
France	1989-2013	Departments	81	insee.fr/fr/statistiques
Germany	1990-2018	States	16	https://www.destatis.de
Mexico ^a	1990-2021	Regions	32	inegi.org.mx/programas/natalidad
USA	1969-2004	States	51	https://data.nber.org/nativity/
Spain	1998-2020	Provinces	32	https://www.ine.es/
Human Fertility Collection	1968-2016	Countries	17	https://www.fertilitydata.org/Data/DataAvailability#MTOTTable
Schoumaker's fertility data	2010	Countries	163	https://perso.uclouvain.be/bruno.schoumaker/data/
Schoen's fertility data	1963-1974	Countries	22	https://www.sciencedirect.com/science/article/pii/0049089X85900043

^a The time-series is not complete for all states. There is no information for Aguascalientes, Baja California, Baja California Sur, Campeche, Chiapas, Chihuahua, Ciudad de México, Coahuila de Zaragoza and Colima after 2015.

Table S2. This table summarizes the TFR for women from the WPP2024 over the entire time period by decade. The first column provides a row index, the second column lists the different variables, the third column indicates the decade, the third to seventh columns provide the decade-mean, standard deviation, minimum, and maximum.

No	Indicator	Decade	μ	SD	Minimum	Maximum
TFRw	1950-1960	5.352	1.621	1.836	8.253	
TFRw	1960-1970	5.237	1.736	1.632	8.249	
TFRw	1970-1980	4.566	1.915	1.337	8.597	
TFRw	1980-1990	3.976	1.885	1.133	8.770	
TFRw	1990-2000	3.357	1.702	0.932	8.377	
TFRw	2000-2010	2.911	1.506	0.798	7.700	
TFRw	2010-2020	2.638	1.295	0.705	7.531	
TFRw	2020-2030	2.299	1.085	0.661	6.325	
TFRw	2030-2040	2.108	0.830	0.753	5.339	
TFRw	2040-2050	1.981	0.632	0.839	4.461	
TFRw	2050-2060	1.895	0.495	0.914	3.780	
TFRw	2060-2070	1.834	0.395	0.976	3.268	
TFRw	2070-2080	1.789	0.321	1.034	2.882	
TFRw	2080-2090	1.756	0.265	1.085	2.612	
TFRw	2090-2100	1.731	0.221	1.125	2.403	

Table S3. This table summarizes the data from the WPP2024 and the estimated TFR for men over the entire time period. The first column provides a row index, the second column lists the different variables, the third column indicates the decade, the third to seventh columns provide the decade-mean, standard deviation, minimum, and maximum.

Indicator	Decade	μ	SD	Minimum	Maximum
TFRm	1950-1960	6.499	2.224	1.885	11.620
TFRm	1960-1970	6.306	2.405	1.696	11.317
TFRm	1970-1980	5.525	2.654	1.061	11.137
TFRm	1980-1990	4.812	2.682	1.138	11.683
TFRm	1990-2000	3.912	2.425	0.909	11.322
TFRm	2000-2010	3.299	2.124	0.759	9.931
TFRm	2010-2020	2.905	1.805	0.560	9.351
TFRm	2020-2030	2.446	1.501	0.574	9.520
TFRm	2030-2040	2.189	1.136	0.659	7.246
TFRm	2040-2050	1.981	0.840	0.726	5.279
TFRm	2050-2060	1.830	0.648	0.744	4.409
TFRm	2060-2070	1.770	0.505	0.759	3.648
TFRm	2070-2080	1.708	0.409	0.839	3.103
TFRm	2080-2090	1.652	0.341	0.894	2.761
TFRm	2090-2100	1.627	0.283	0.894	2.478

Table S4. This table summarizes the data from the WPP2024 and the estimated adult sex ratio (men aged 24-45 to women aged 20-39) entire time period. The first column provides variable name, the second column indicates the decade, the third to sixth columns provide the decade-mean, standard deviation, minimum, and maximum.

Indicator	Decade	μ	SD	Minimum	Maximum
Adult Sex Ratio	1950-1960	-0.139	0.160	-0.838	0.893
Adult Sex Ratio	1960-1970	-0.119	0.161	-0.713	1.361
Adult Sex Ratio	1970-1980	-0.134	0.173	-0.750	1.380
Adult Sex Ratio	1980-1990	-0.139	0.172	-0.670	1.346
Adult Sex Ratio	1990-2000	-0.090	0.167	-0.505	1.097
Adult Sex Ratio	2000-2010	-0.065	0.174	-0.619	1.456
Adult Sex Ratio	2010-2020	-0.039	0.187	-0.490	1.383
Adult Sex Ratio	2020-2030	-0.004	0.192	-0.556	1.324
Adult Sex Ratio	2030-2040	0.001	0.169	-0.539	1.141
Adult Sex Ratio	2040-2050	0.031	0.140	-0.304	0.960
Adult Sex Ratio	2050-2060	0.065	0.128	-0.180	0.915
Adult Sex Ratio	2060-2070	0.049	0.109	-0.162	0.797
Adult Sex Ratio	2070-2080	0.056	0.100	-0.165	0.706
Adult Sex Ratio	2080-2090	0.073	0.094	-0.109	0.640
Adult Sex Ratio	2090-2100	0.068	0.085	-0.098	0.596

Table S5. This table summarizes the difference between the TFR_m and the TFR_w in the period between 1950 and 2100 by decade. The first column provides variable name, the second column indicates the decade, the third to sixth columns provide the decade-mean, standard deviation, minimum, and maximum.

Indicator	Decade	μ	SD	Minimum	Maximum
TFR Difference	1950-1960	0.201	0.132	-0.387	0.921
TFR Difference	1960-1970	0.183	0.133	-0.560	0.716
TFR Difference	1970-1980	0.176	0.144	-0.553	0.753
TFR Difference	1980-1990	0.162	0.151	-0.546	0.731
TFR Difference	1990-2000	0.106	0.149	-0.498	0.492
TFR Difference	2000-2010	0.073	0.148	-0.616	0.586
TFR Difference	2010-2020	0.047	0.146	-0.605	0.520
TFR Difference	2020-2030	0.010	0.147	-0.593	0.599
TFR Difference	2030-2040	-0.001	0.129	-0.542	0.469
TFR Difference	2040-2050	-0.028	0.104	-0.485	0.261
TFR Difference	2050-2060	-0.054	0.091	-0.469	0.185
TFR Difference	2060-2070	-0.047	0.075	-0.427	0.127
TFR Difference	2070-2080	-0.054	0.067	-0.391	0.094
TFR Difference	2080-2090	-0.066	0.062	-0.364	0.072
TFR Difference	2090-2100	-0.064	0.055	-0.341	0.047

Table S6. This table summarizes the frequency of crossovers over the entire time period. The first column provides variable name, the second column indicates the decade, the third to sixth columns provide the decade-mean, standard deviation, minimum, and maximum.

Indicator	Decade	μ	SD	Minimum	Maximum
Crossover	1950-1960	0.004	0.062	0	1
Crossover	1960-1970	0.010	0.100	0	1
Crossover	1970-1980	0.007	0.084	0	1
Crossover	1980-1990	0.016	0.127	0	1
Crossover	1990-2000	0.017	0.128	0	1
Crossover	2000-2010	0.005	0.071	0	1
Crossover	2010-2020	0.014	0.119	0	1
Crossover	2020-2030	0.011	0.103	0	1
Crossover	2030-2040	0.007	0.082	0	1
Crossover	2040-2050	0.013	0.111	0	1
Crossover	2050-2060	0.008	0.088	0	1
Crossover	2060-2070	0.005	0.071	0	1
Crossover	2070-2080	0.005	0.071	0	1
Crossover	2080-2090	0.005	0.068	0	1
Crossover	2090-2100	0.004	0.065	0	1

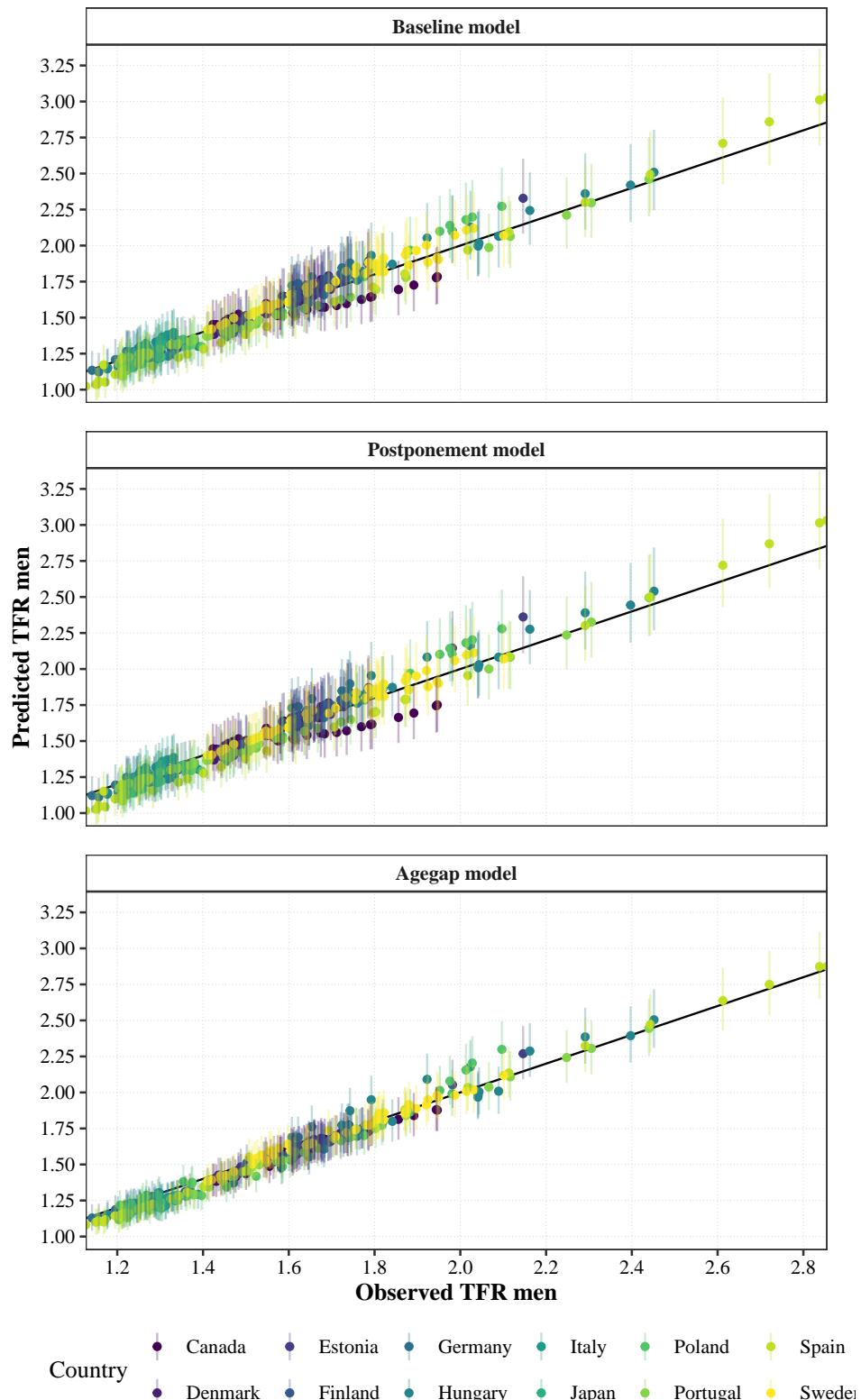


Fig. S1. This figure illustrates the out-of-sample validation of the baseline model, the postponement model and the age-gap model on data from the Human Fertility Collection (6). The x-axis shows the observed TFR_m from the Human Fertility Collection, while the y-axis shows the predicted TFR_m and the 90%-prediction intervals from the different regression models using data on adult sex ratios and TFR_w from WPP2024.

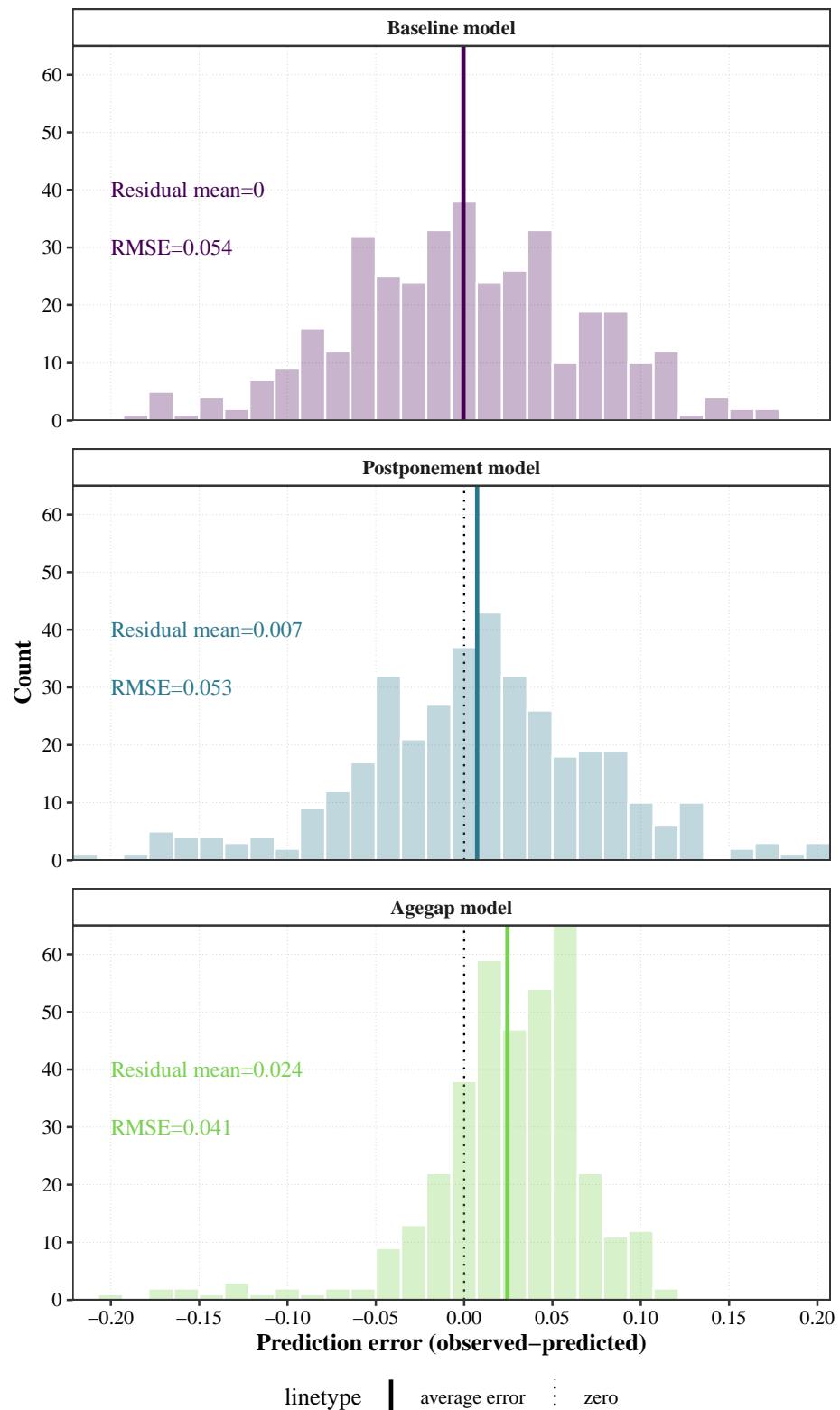


Fig. S2. Out-of sample validation error of the regression-based approach using data from WPP2024 on the data in Dudel and Klüsener (6). The x-axis provides e_i error between the observed and predicted value, and the y-axis shows the count of observations.

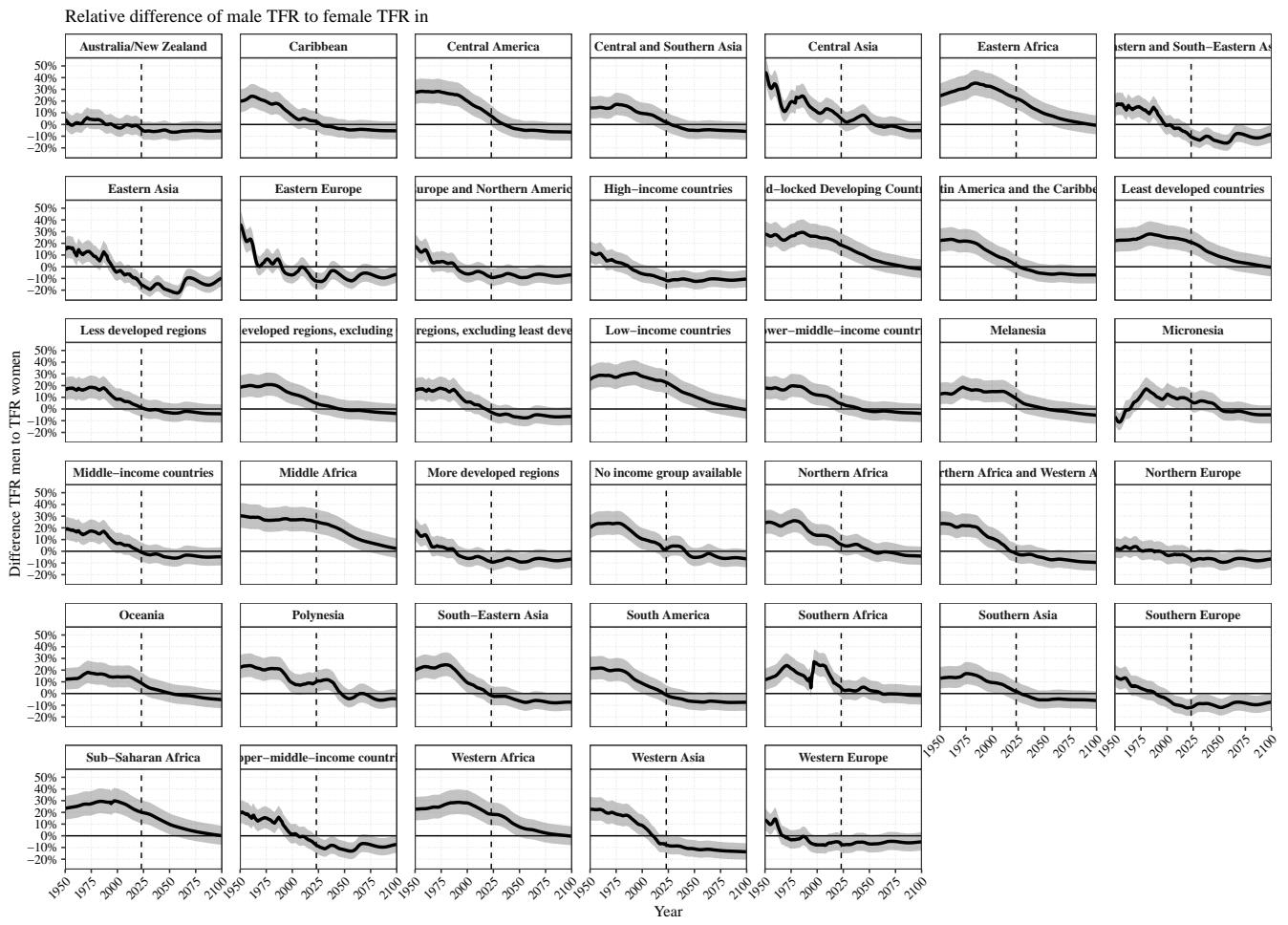


Fig. S3. This figure illustrates the relative difference between the TFR_m to TFR_w (y-axis) in the period between 1950 and 2100 (x-axis) accounting for the prediction uncertainty (grey shading) for the different regional classifications.

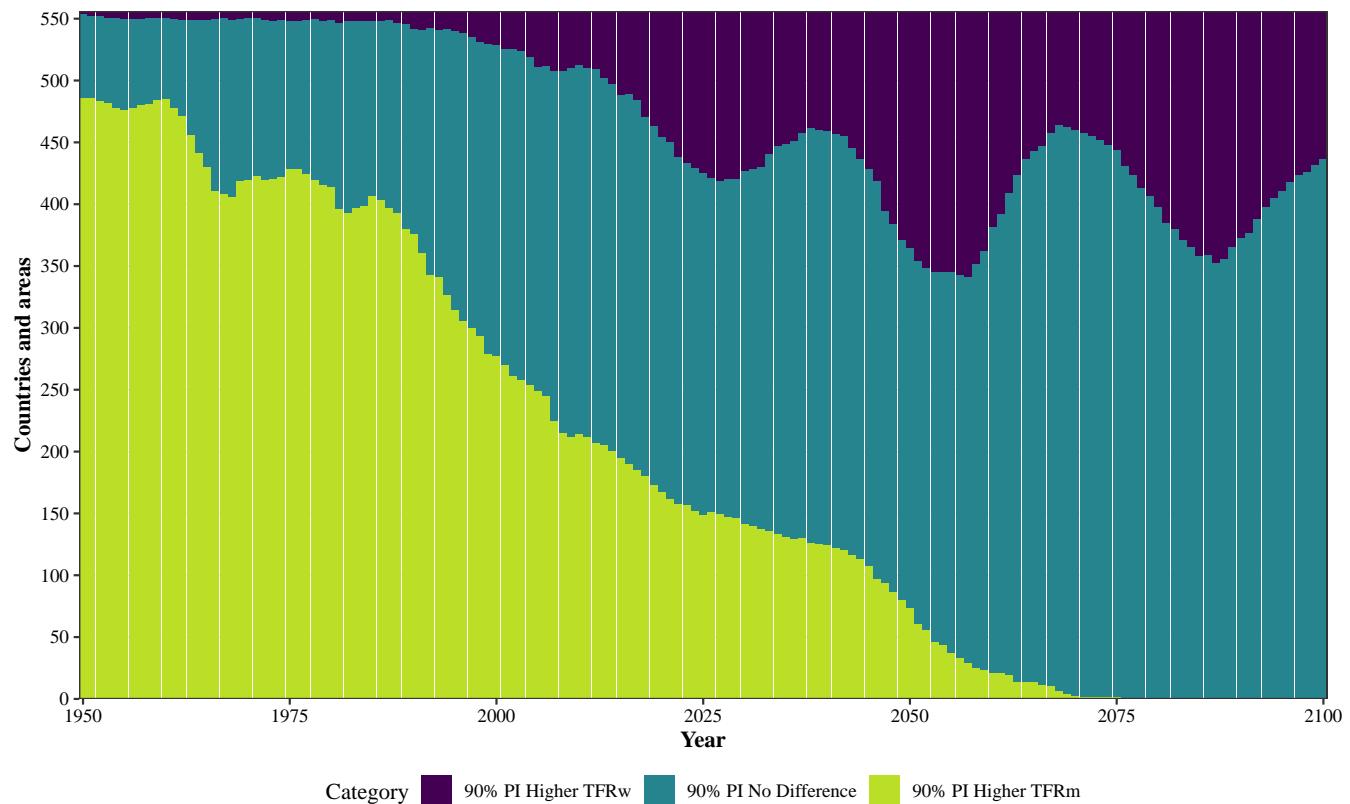


Fig. S4. This figure illustrates on the y-axis the number of countries and areas with higher TFRm than TFRw (green), higher TFRw than TFRm (purple) and countries and areas where the difference is not statistically significant (blue) in the period between 1950 and 2100 (x-axis) accounting for the prediction uncertainty.

Relative difference of male TFR to female TFR in

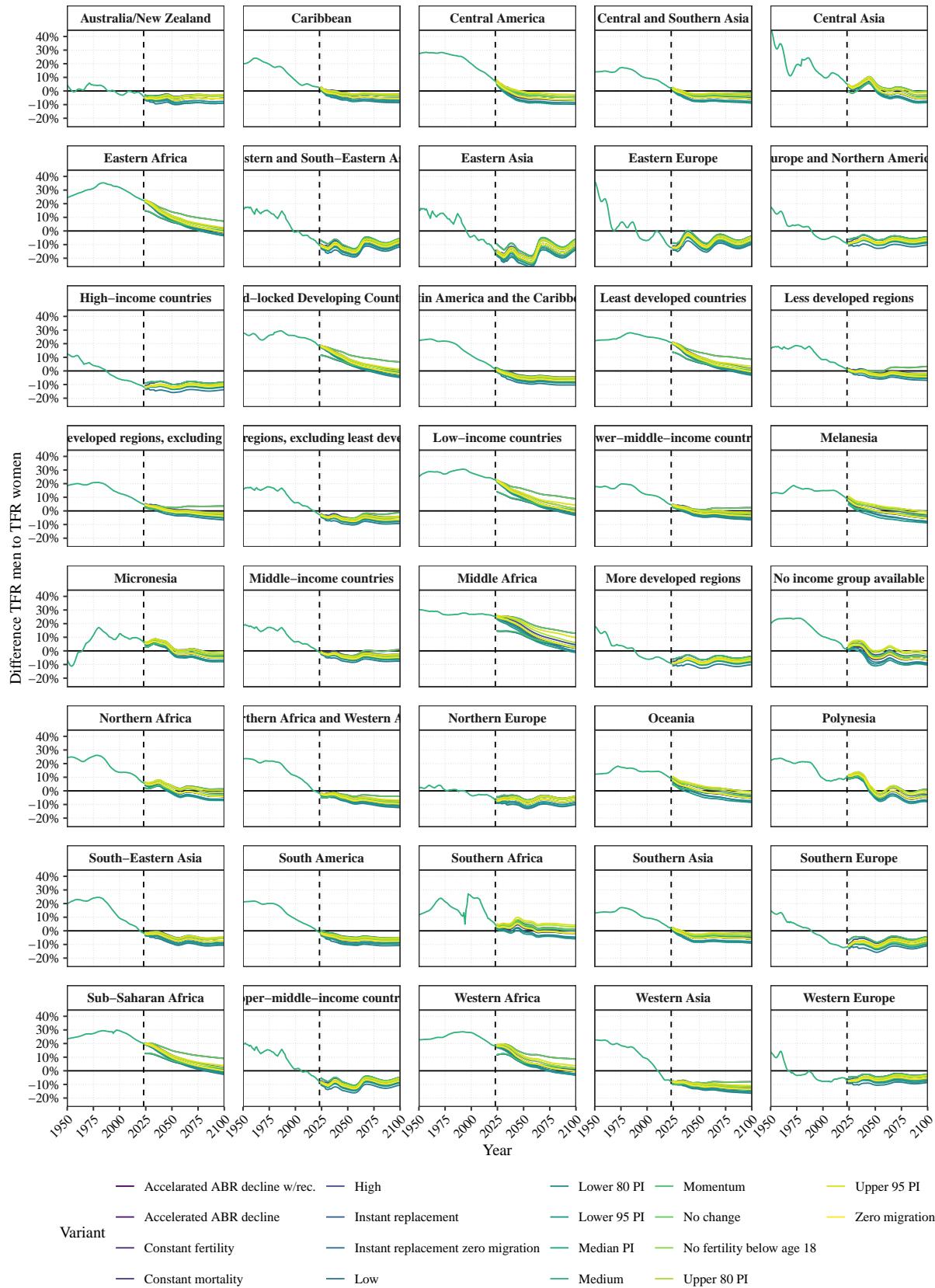


Fig. S5. This figure illustrates the impact of different WPP2024-scenarios (colours) on the relative difference between the TFR_m to the TFR_w in per cent (y-axis) across country groups (different panels) between 1950 and 2100 (x-axis). Values higher than 0 indicate higher TFR_m relative to the TFR_w , and values below 1 indicate higher TFR_w relative to the TFR_m .

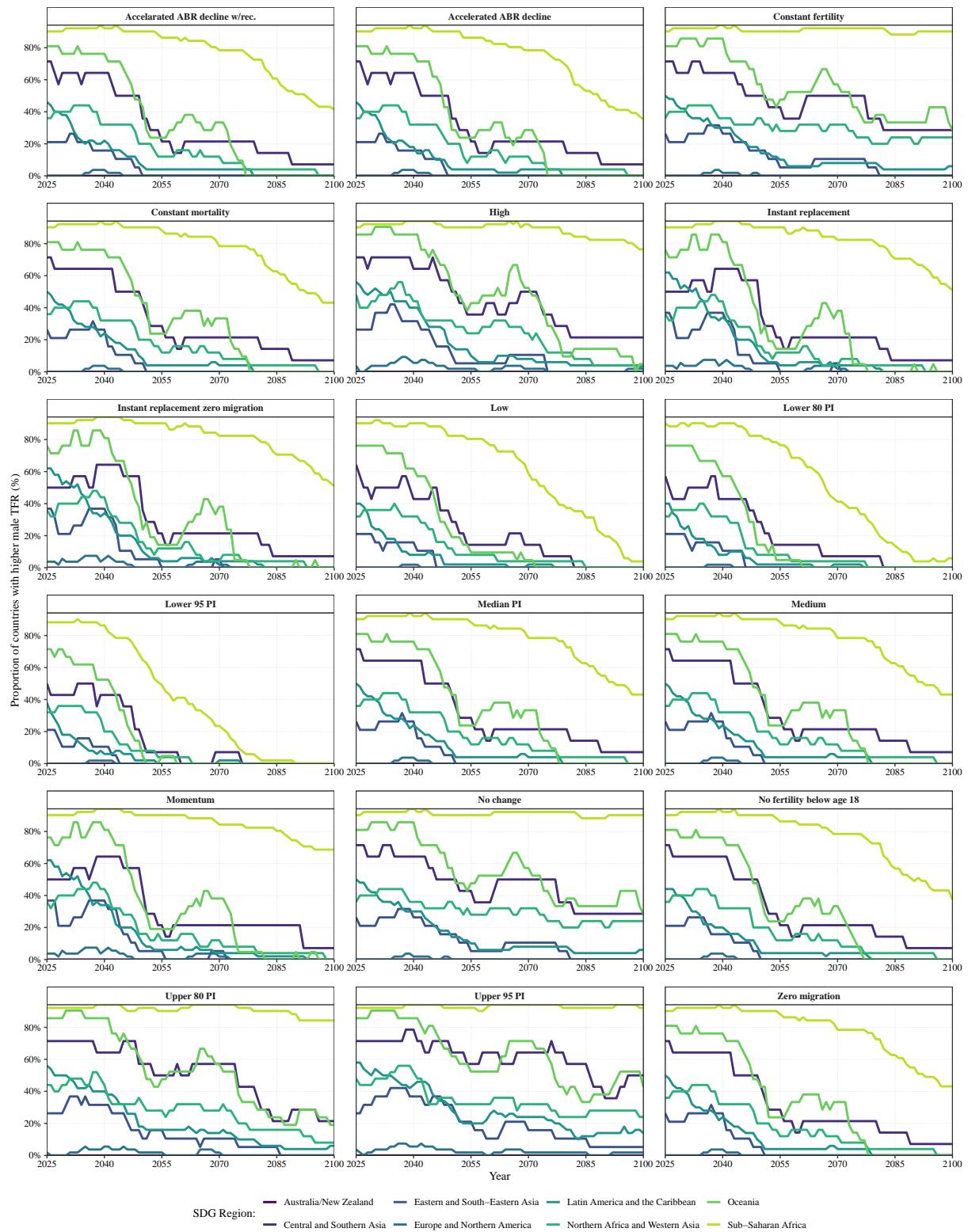


Fig. S6. This figure illustrates the share of countries and areas with higher TFR_m relative to the TFR_w in per cent (y-axis) across different WPP2024-scenarios (different panels) and geographic regions (colours) between 2025 and 2100 (x-axis).

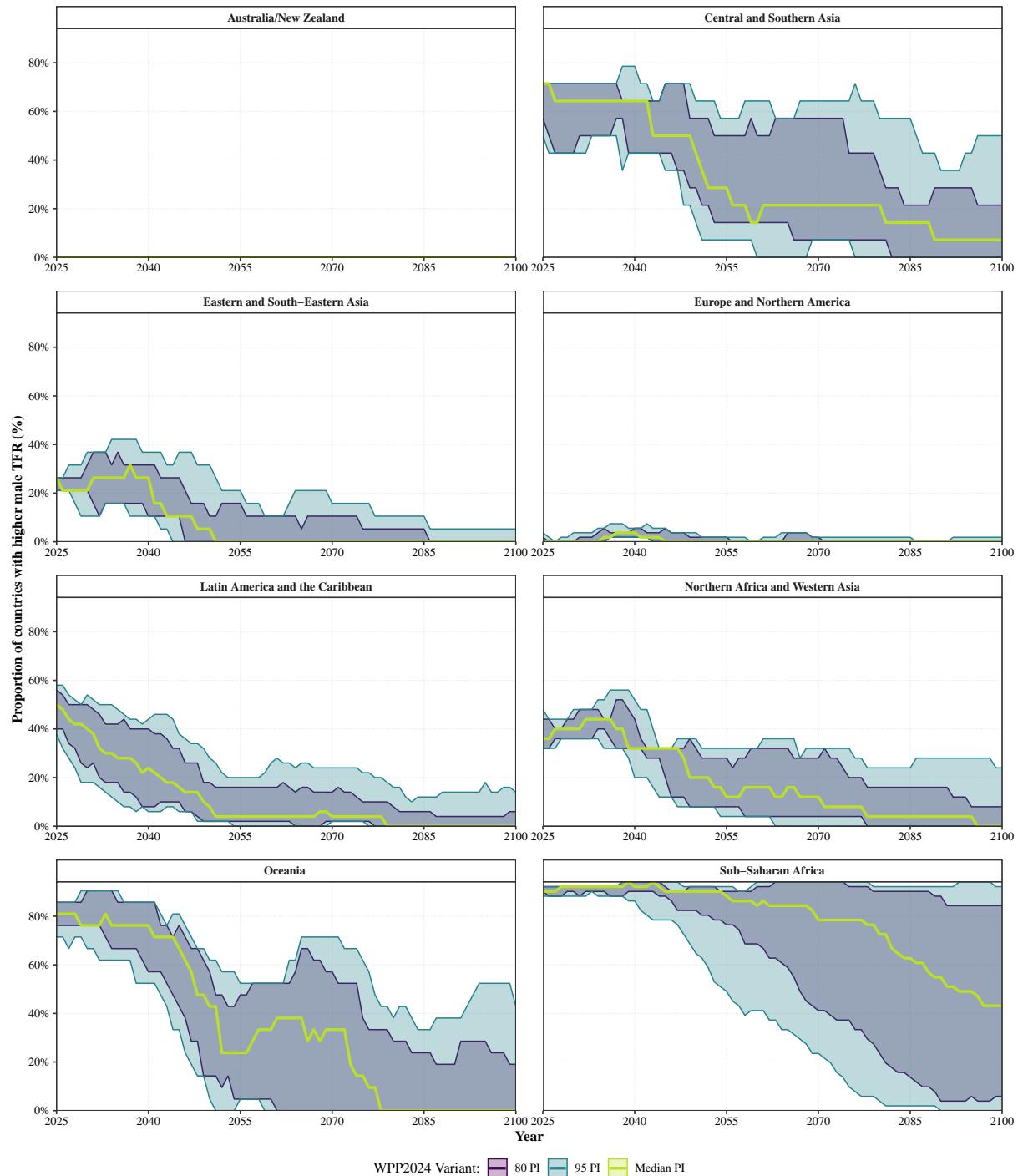


Fig. S7. This figure illustrates the share of countries and areas with higher TFR_m relative to the TFR_w in per cent (y-axis) incorporating the uncertainty in the estimation from the WPP2024 probabilistic model (different colours) and geographic regions (panels) between 2025 and 2100 (x-axis).

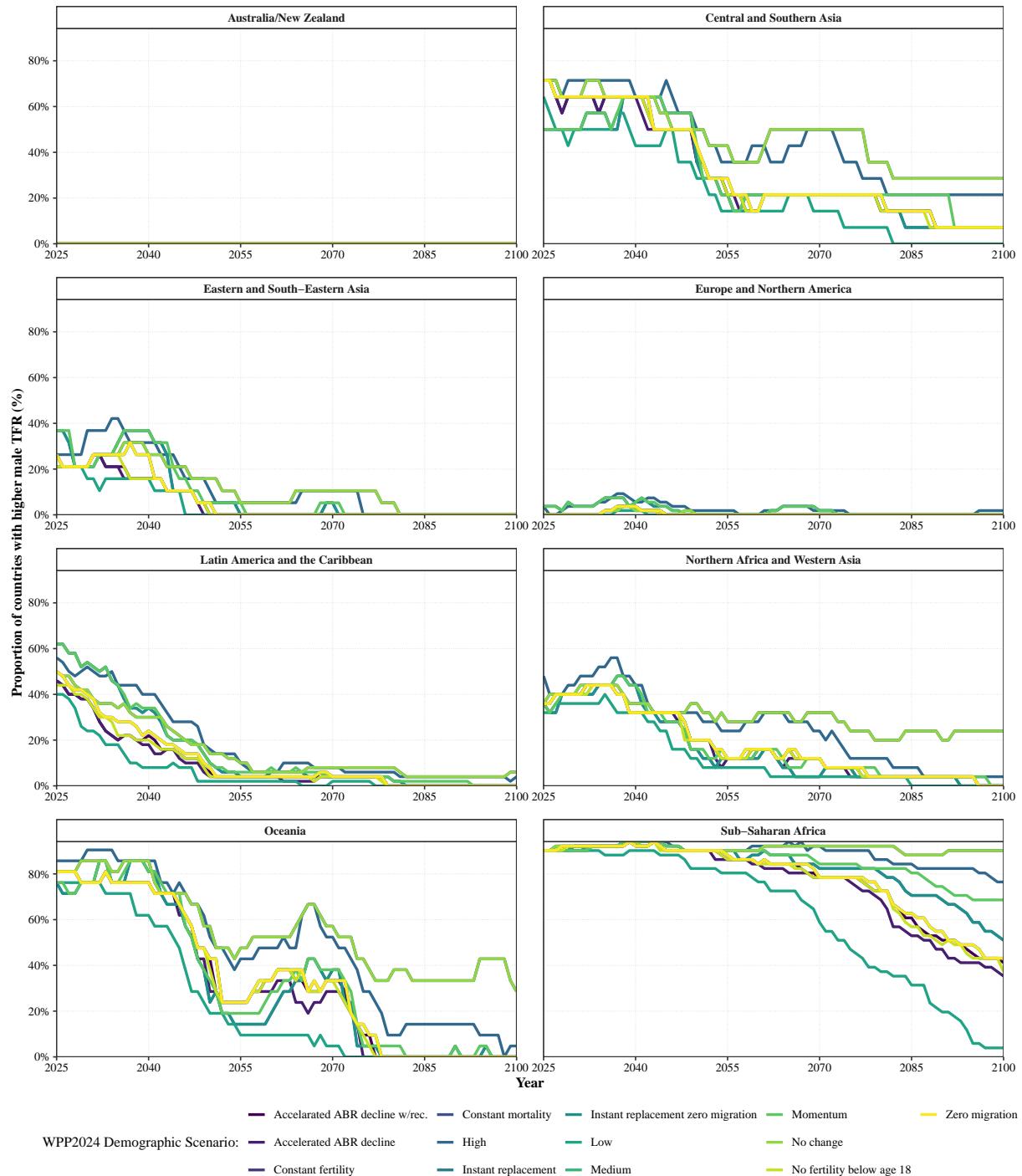


Fig. S8. This figure illustrates the share of countries and areas with higher TFR_m relative to the TFR_w in per cent (y-axis) across different WPP2024-scenarios (different colours) and geographic regions (panels) between 2025 and 2100 (x-axis).

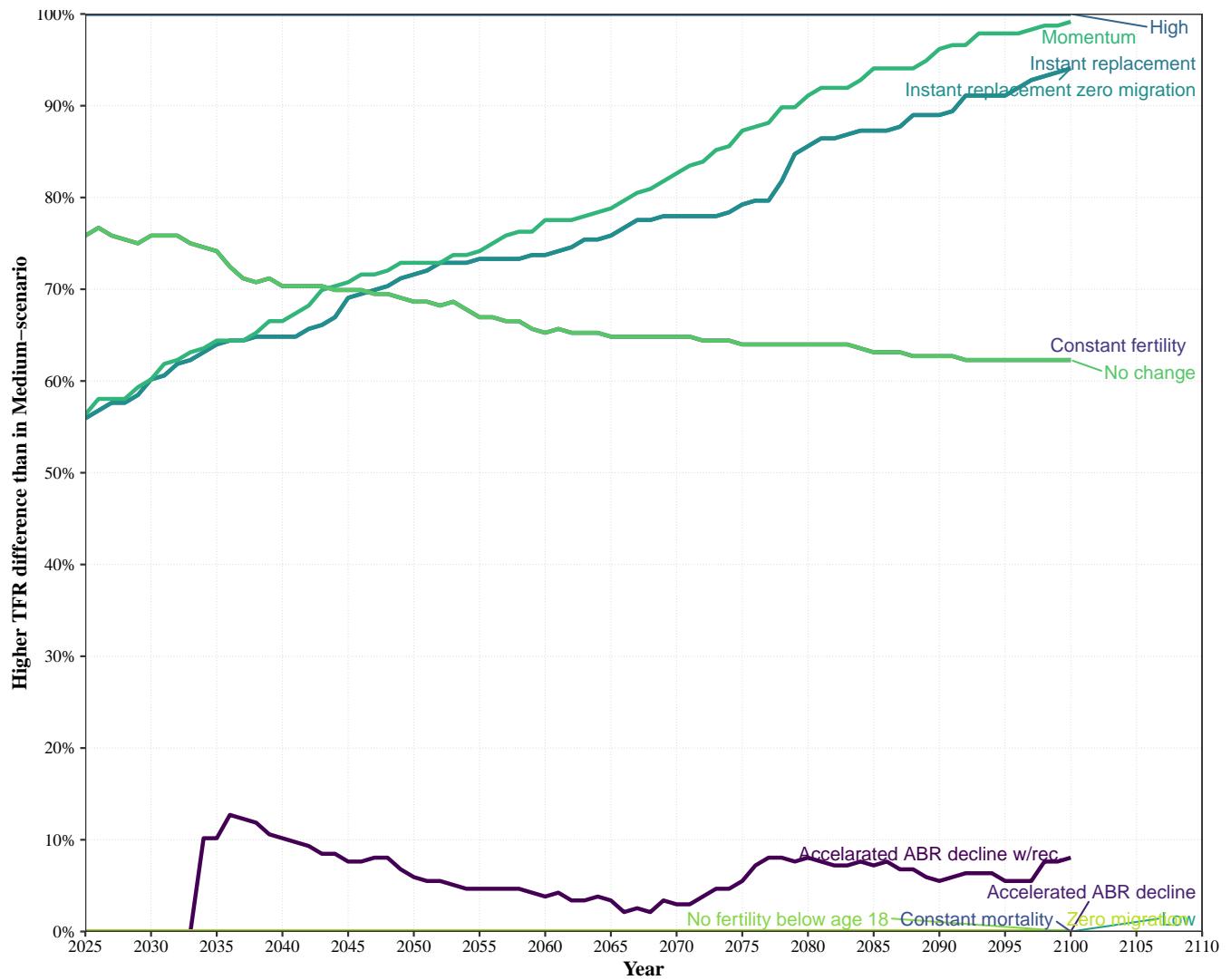


Fig. S9. This figure displays the percentage of countries and areas where the specific scenario has a higher male-female TFR difference than the medium scenario. Interpretation: the higher the share the more often the assumptions of the scenario contribute to higher TFR_m relative to TFR_w

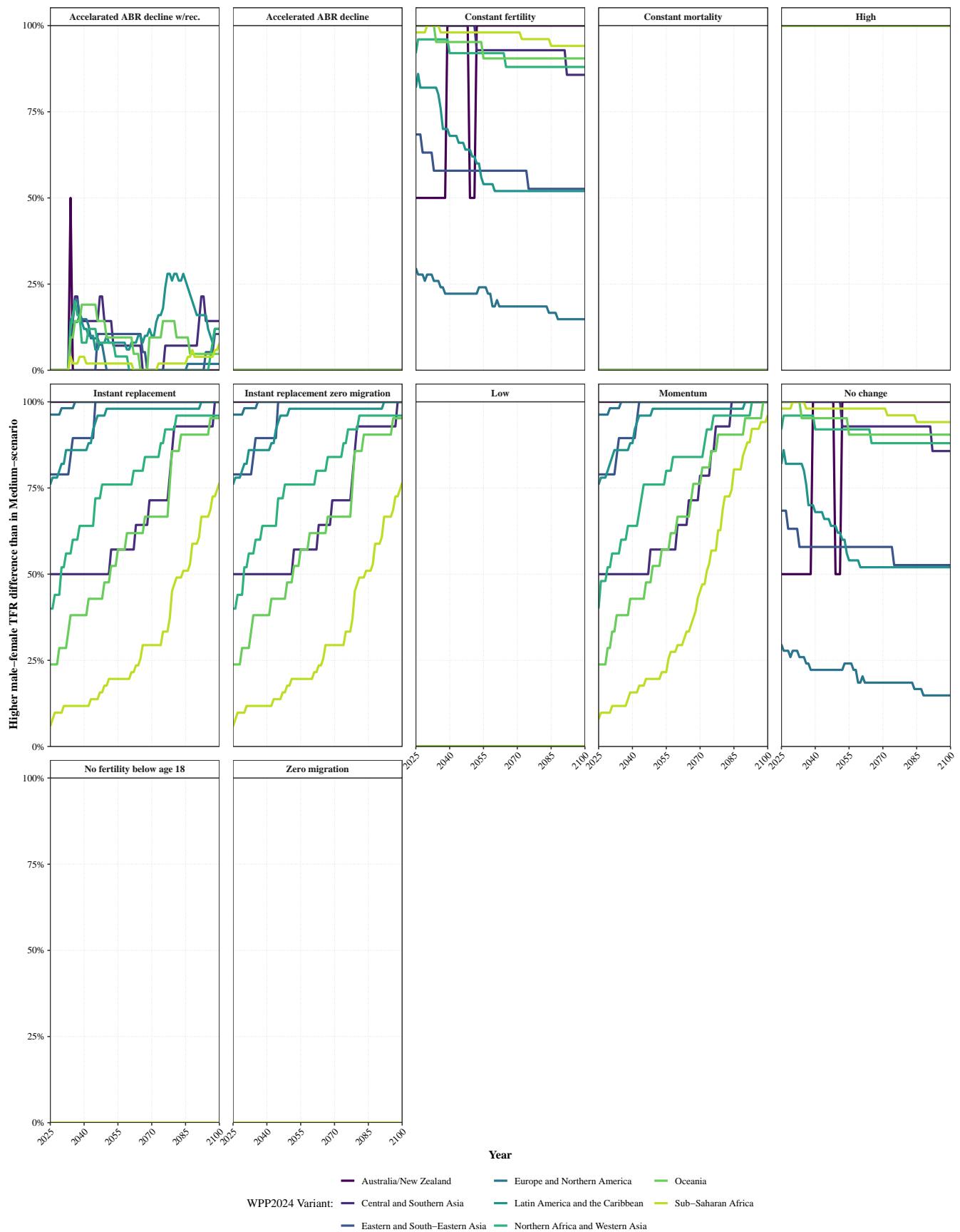


Fig. S10. This figure displays the per cent of countries and areas where the specific scenario has a higher male-female TFR difference than the medium scenario by SDG region. Interpretation: the higher the share the more often the assumptions of the scenario contribute to higher TFR_m relative to TFR_w

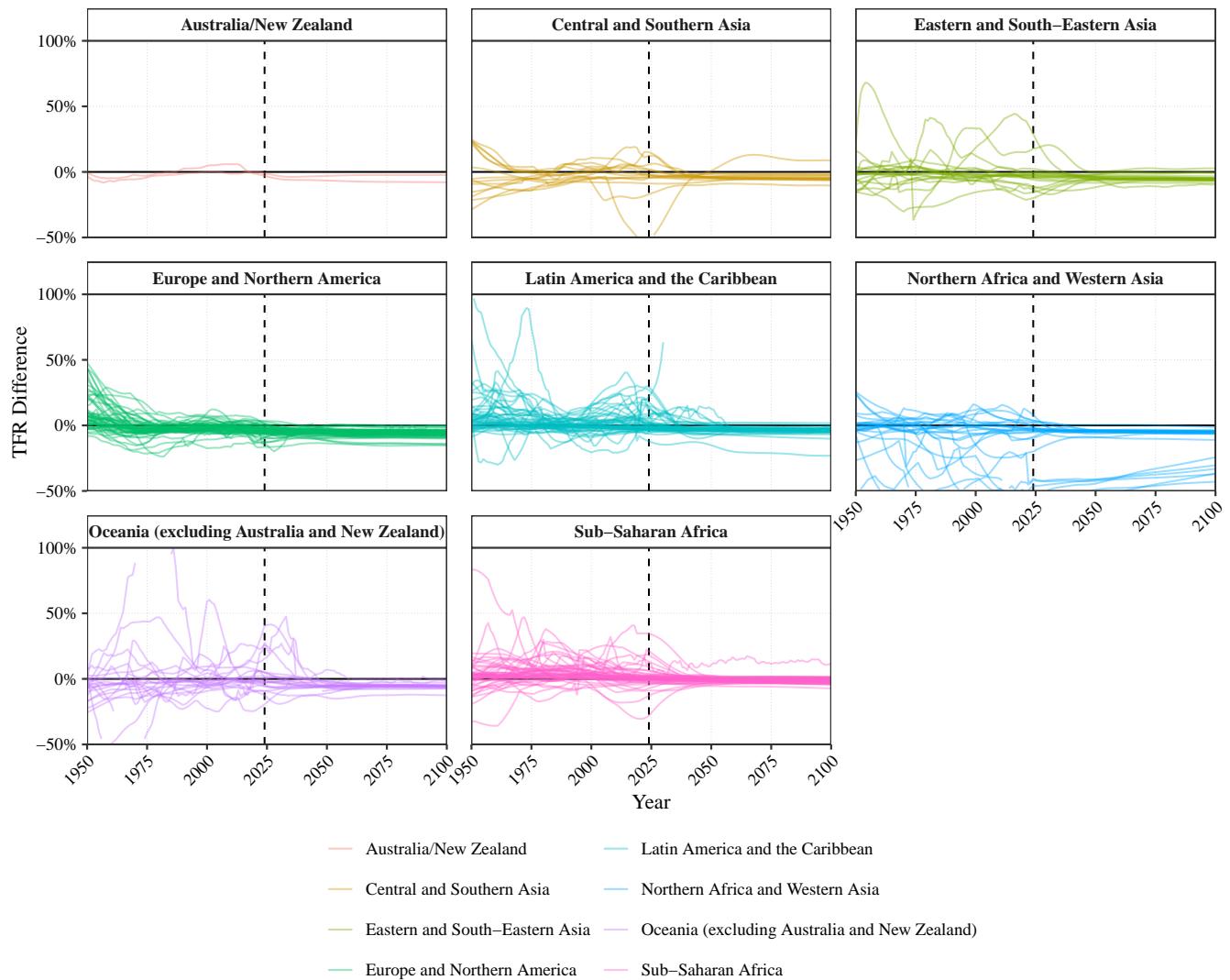


Fig. S11. Percentage difference in male to female TFR (y-axis) in the period between 1950 to 2100 (x-axis) using the standardization approach. Positive values indicate a higher TFR among men and negative values indicate a lower TFR among men. The vertical lines indicate the crossover from higher male TFR to higher female TFR.

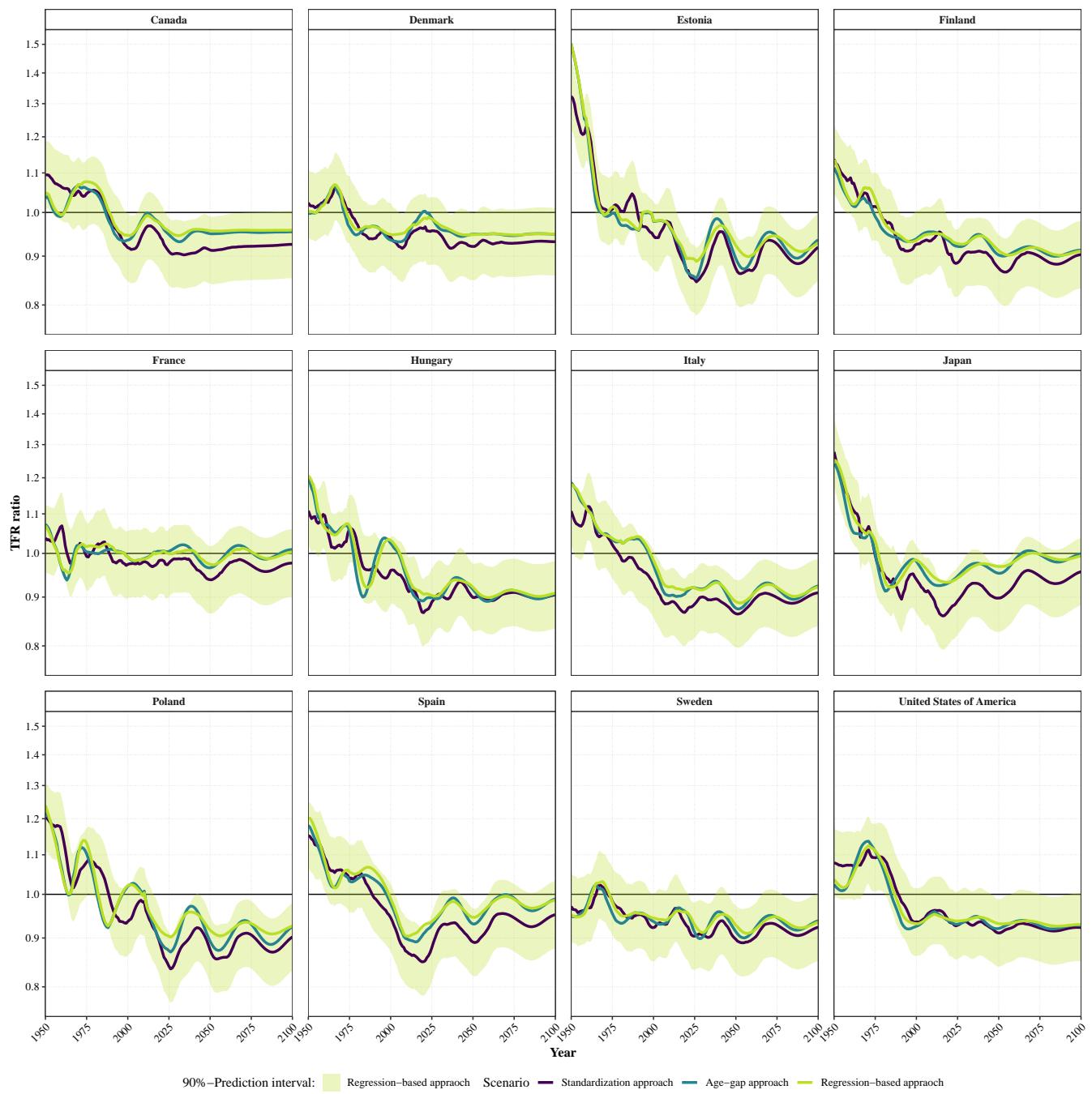


Fig. S12. This figure illustrates the impact of the estimation method on the TFR ratio ($\frac{TFR_{ml}}{TFR_{w}}$) for selective countries over the period between 1950 to 2100. The purple line displays the results from the standardization method, the dark green line shows the results from the age-gap approach, and the light-green line illustrates the results from the regression-based approach including the prediction errors.

118 SI Dataset S1 ([https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20\(Standard\)/CSV_FILES/WPP2024_Life_Table_Complete_Medium_Female_1950-2023.csv.gz](https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20(Standard)/CSV_FILES/WPP2024_Life_Table_Complete_Medium_Female_1950-2023.csv.gz))
119
120 This file contains a compressed-version of the WPP2024 Female life tables for the period between 1950 and 2023.

121 SI Dataset S2 ([https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20\(Standard\)/CSV_FILES/WPP2024_Life_Table_Complete_Medium_Male_1950-2023.csv.gz](https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20(Standard)/CSV_FILES/WPP2024_Life_Table_Complete_Medium_Male_1950-2023.csv.gz))
122
123 This file contains a compressed-version of the WPP2024 Male life tables for the period between 1950 and 2023.

124 SI Dataset S3 ([https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20\(Standard\)/CSV_FILES/WPP2024_Life_Table_Complete_Medium_Female_2024-2100.csv.gz](https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20(Standard)/CSV_FILES/WPP2024_Life_Table_Complete_Medium_Female_2024-2100.csv.gz))
125
126 This file contains a compressed-version of the WPP2024 projected Female life tables for the period between 2024 and 2100.

127 SI Dataset S4 ([https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20\(Standard\)/CSV_FILES/WPP2024_Life_Table_Complete_Medium_Male_2024-2100.csv.gz](https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20(Standard)/CSV_FILES/WPP2024_Life_Table_Complete_Medium_Male_2024-2100.csv.gz))
128
129 This file contains a compressed-version of the WPP2024 projected Female life tables for the period between 2024 and 2100.

130 SI Dataset S5 ([https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20\(Standard\)/CSV_FILES/WPP2024_Fertility_by_Age1.csv.gz](https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20(Standard)/CSV_FILES/WPP2024_Fertility_by_Age1.csv.gz))
131
132 This file contains the fertility information by single-ages of the mother for the period between 1950 and 2100.

133 SI Dataset S6 ([https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20\(Standard\)/CSV_FILES/WPP2024_Population1JanuaryBy_Medium_1950-2023.csv.gz](https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20(Standard)/CSV_FILES/WPP2024_Population1JanuaryBy_Medium_1950-2023.csv.gz))
134
135 This file contains the estimated population structures single-ages for men and women for the period between 1950 and 2023.

136 SI Dataset S7 ([https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20\(Standard\)/CSV_FILES/WPP2024_Population1JanuaryBy_Medium_2024-2100.csv.gz](https://population.un.org/wpp/assets/Excel%20Files/1_Indicator%20(Standard)/CSV_FILES/WPP2024_Population1JanuaryBy_Medium_2024-2100.csv.gz))
137
138 This file contains the projected population structures single-ages for men and women for the period between 1950 and 2023.

139 SI Dataset S8 (https://www.fertilitydata.org/File/GetFile/Zip/m_HFC_ASFRstand_TOT.zip)
140 This website contains the country-level male fertility data in the Human Fertility Collection by Dudel and Klüsener (6).
141 SI Dataset S9 (<https://perso.uclouvain.be/bruno.schoumaker/data/A.%20Estimates%20of%20male%20and%20female%20fertility.xlsx>)
142
143 This file contains the country-level male fertility data reported in Schoumaker (7).

144 References

- 145 1. Frances K. Goldscheider and Gayle Kaufman. Fertility and Commitment: Bringing Men Back In. *Population and Development Review*, 22:87, 1996. ISSN 00987921. .
- 146 2. Kara Joyner, H. Elizabeth Peters, Kathryn Hynes, Asia Sikora, Jamie Rubenstein Taber, and Michael S. Rendall. The Quality of Male Fertility Data in Major U.S. Surveys. *Demography*, 49(1):101–124, February 2012. ISSN 0070-3370, 1533-7790. .
- 147 3. Christian Dudel and Sebastian Klüsener. Estimating men's fertility from vital registration data with missing values. *Population Studies*, 73(3):439–449, September 2019. ISSN 0032-4728, 1477-4747. .
- 148 4. Bruno Schoumaker. Measuring male fertility rates in developing countries with Demographic and Health Surveys: An assessment of three methods. *Demographic Research*, 36:803–850, March 2017. ISSN 1435-9871. .
- 149 5. Henrik-Alexander Schubert and Christian Dudel. Subnational Birth Squeezes? Male-Female TFR Differences across Eight High- and Middle Income Countries over Time, 2025.
- 150 6. Christian Dudel and Sebastian Klüsener. Male–Female Fertility Differentials Across 17 High-Income Countries: Insights From A New Data Resource. *European Journal of Population*, 37(2):417–441, April 2021. ISSN 0168-6577, 1572-9885. .
- 151 7. Bruno Schoumaker. Male Fertility Around the World and Over Time: How Different is it from Female Fertility? *Population and Development Review*, 45(3):459–487, September 2019. ISSN 0098-7921, 1728-4457. .
- 152 8. Nico Keilman, Krzysztof Tymicki, and Vegard Skirbekk. Measures for Human Reproduction Should Be Linked to Both Men and Women. *International Journal of Population Research*, 2014(1):908385, 2014. ISSN 2090-4037. .
- 153 9. Éva Beaujouan and Tomáš Sobotka. Late Motherhood in Low-Fertility Countries: Reproductive Intentions, Trends and Consequences. 2017.
- 154 10. Ludwig Fahrmeir, Thomas Kneib, Stefan Lang, and Brian D. Marx. *Regression: Models, Methods and Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2021. ISBN 978-3-662-63881-1 978-3-662-63882-8. .
- 155 11. United Nations Department of Economic and Social Affairs. World Population Prospects 2024: Methodology of the United Nations population estimates and projections. 2024.
- 156 12. Christian Dudel, Yen-hsin Alice Cheng, and Sebastian Klüsener. Shifting Parental Age Differences in High-Income Countries: Insights and Implications. *Population and Development Review*, 49(4):879–908, 2023. ISSN 1728-4457. .