Probabilistic Models for Human Migration Forecasting and Residency Imputation

Nathan G Welch

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:
Adrian E Raftery, Chair
Adrian Dobra
Emilio Zagheni

Program Authorized to Offer Degree:
Statistics

University of Washington

Abstract

Probabilistic Models for Human Migration Forecasting and Residency Imputation

Nathan G Welch

Chair of the Supervisory Committee:
Adrian E Raftery
Department of Statistics

I develop probabilistic models to enhance the estimation and forecasting of human migration flows and residency. Using a Bayesian hierarchical approach, I first propose a model for forecasting global bilateral migration flows among the 200 most populous countries, producing well-calibrated projections that reduce error rates compared to existing methods. This model is integrated into a population projection framework to forecast migration flows by age and sex, providing the first probabilistic forecasts of international bilateral migration flows through 2045. Next, I address the influence of age structure on much longer-term migration forecasts by introducing the Migration Age Structure Index (MASI) that adjusts net migration rates, offering narrower prediction intervals and more accurate projections of population change, especially for aging populations. Finally, I improve the Person-Place Model (PPM), a key tool used by countries without population registers for census and intercensal population estimation, by developing the Bayes PPM—a Bayesian hierarchical model that refines residency estimates from administrative records. This model eliminates crude approximations to a well-defined statistical model currently in use, enhancing the accuracy of uncertainty intervals in demographic estimates. Collectively, these contributions offer more capable tools for forecasting migration and demographic changes, supporting policymakers in navigating complex global migration dynamics.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number Page

LIST OF TABLES

GLOSSARY

ABC: Applied Bayesian Computational Statistics Working Group

APC: Administrative Population Census

AR: Administrative Records

BFM: Bayesian Flow Model

ERP: Estimated Resident Population

EU: European Union

GCC: Gulf Cooperation Council

MASI: Migration Age Structure Index

NRFU: Nonresponse Follow-Up

PPM: Person-Place Model

UN: United Nations

UNHCR: United Nations High Commission for Refugees

UW: University of Washington

WPP: World Population Prospects

ACKNOWLEDGMENTS

I would like to express my immense gratitude to my Ph.D. advisor, Adrian E. Raftery. Your guidance and encouragement have been instrumental to me. Learning from you has been the opportunity of a lifetime, and I am profoundly grateful for your mentorship.

I am also deeply appreciative of Hana Ševčíková for your invaluable guidance throughout the development of this thesis. I will always be thankful for the chance to learn from you.

I thank the Department of Statistics faculty and staff for their support and guidance through the years. Thomas Richardson and Ellen Reynolds were especially invaluable to my time in the department. I will miss Friday mornings in the Applied Bayesian Computational Statistics (ABC) Working Group, especially the thoughtful suggestions from fellow ABC members Daphne Liu, Michael Pearce, and Nicholas Irons.

I am grateful to my committee members for their guidance in developing this thesis. Adrian Dobra's mentorship leading up to my General Exam is among the most formative experiences of my time at the University of Washington. Emilio Zagheni's demography courses are foundational to the work outlined in this thesis. Mark Ellis's research with administrative records data helped inspire my interest in statistical methods tailored to such datasets.

Lastly, but most importantly, I want to express my heartfelt appreciation to my wife, Amanda. Your love and patience made it possible to navigate the highs and lows throughout this journey. Thank you for the countless sacrifices you made, the encouragement you provided, and for always standing by my side.

DEDICATION

To my family

Chapter 1

INTRODUCTION

1.1 Motivation

Understanding global human migration flows and their impact on population dynamics has become increasingly important in recent decades as policymakers seek deeper insights into the most volatile component of demographic change—migration. Traditional deterministic models for migration forecasting often fail to capture the inherent uncertainty in future migration patterns, limiting their utility in long-term planning. This dissertation aims to develop probabilistic models that produce more accurate migration forecasts and assess the associated uncertainty.

Recent methodological advances, such as those by Abel (2013a) and Azose & Raftery (2019a), have made it possible to generate plausible estimates of international migration flows at a global scale. However, flow forecasting method development has lagged behind progress in flow estimation (Bijak, 2006). This dissertation addresses this gap with a Bayesian hierarchical model that uses recently developed bilateral migration estimates from 1990–2020 to generate probabilistic bilateral migration flow forecasts among the 200 largest countries through 2045.

Migration is highly correlated with age, with younger populations generally having a higher propensity to migrate. Despite this, existing models often overlook the impact of a country's population age structure on migration patterns. This dissertation addresses this gap with the Migration Age Structure Index (MASI), a novel approach to overcoming limited age-specific migration rate estimates. The MASI provides a theoretically sound approach to population age-standardized historic and forecast net migration rates for very long term global population forecasting.

As administrative records become more integrated into demographic estimation processes, particularly in countries lacking population registers, the need for accurate residency imputation model has become more pressing. This dissertation develops the Bayesian Person-Place Model (Bayes PPM), which introduces a well-defined statistical foundation for estimating residency probabilities, addressing critical limitations associated with the current Person-Place Model (PPM) (National Academies of Sciences, Engineering, and Medicine, 2023; Brown et al., 2023; Rastogi & O'Hara, 2012; Morris, 2014; Keller et al., 2018). The Bayes PPM addresses a number of specific recommendations outlined in (National Academies of Sciences, Engineering, and Medicine, 2023) in particular.

1.2 Background

This dissertation includes three distinct but interconnected contributions to the field of human migration forecasting and demographic estimation using administrative records.

Abel (2013a) pioneered a method to estimate the minimum number of people who must have changed their country of residence to explain the change in migrant stocks among all countries of the world. Azose & Raftery (2019a) extended these estimates to produce well-calibrated pseudo-Bayes estimates of bilateral migration flows, which have been crucial for crafting effective migration policies (Willekens, 2018). These developments enable a deeper understanding of historic trends and were used to demonstrate critical deficiencies in existing migration flow forecasting methods. However, this dissertation includes the first probabilistic global migration flow forecasting method resolving these deficiencies.

The association between migration and age is long established. Rogers & Castro (1981a) identified a consistent pattern in the age profile of out-migrants and proposed a model migration age schedule. Kupiszewski (2002) discussed differences in migration forecasts when the sending population age-structure is ignored, while Fertig & Schmidt (2005) argued that neglecting population age in calculating migration rates obscures the underlying population dynamics. Kolk (2019) reviewed age-specific migration and the analogy to age-specific and total fertility rate definitions. (Director of National Intelligence, 2021; Münz, 2013) high-

light the need for quantitative methods accounting for changes in population age structure to replace qualitative conjectures of what could occur. This dissertation addresses this gap.

Since World War II, demographic trends in the United States have predominantly been derived from censuses and carefully designed surveys (F. C. Billari & Zagheni, 2017). While censuses offer wide coverage, they are infrequent. Surveys are more timely but often suffer from low response rates and small sample sizes. F. C. Billari & Zagheni (2017) highlight how *Big Data* technologies and massive stores of nontraditional and administrative data could be combined to close the gap. The National Academies Report on the US 2020 Census highlighted the successful use of administrative records (AR) to supplement census operations and recommended further research. One significant area of such research at the US Census Bureau is the Person-Place Model (PPM). Rastogi & O'Hara (2012) provided a foundational assessment of AR-based enumeration in the United States, proposing a logistic model to select the optimal person-address pair for individuals with multiple possible addresses. Morris (2014) offered one of the first comprehensive descriptions of the PPM, while Keller et al. (2018) further refined the model. This dissertation proposes the Bayesian Person-Place Model (Bayes PPM), which introduces a well-defined statistical foundation for estimating these probabilities, addressing the shortcomings of the current model. The Bayes PPM is shown to outperform extant PPM specifications in terms of accuracy and uncertainty estimation, offering a more robust approach capable of supporting higher fidelity demographic research, including intercensal population estimation and decennial census operations (National Academies of Sciences, Engineering, and Medicine, 2023; Brown et al., 2023).

1.3 Key Contributions

This dissertation makes several significant contributions to statistical demography, particularly in the areas of migration forecasting and demographic estimation.

In Chapter 2, a Bayesian hierarchical model is introduced for forecasting global bilateral migration flows among the 200 most populous countries. This model leverages historical data to produce well-calibrated probabilistic projections of migration flows, accounting for

uncertainty. The model's ability to pool information across time periods and individual flows enhances its predictive accuracy, reducing the mean absolute error by 61% compared to existing models. This advancement provides policymakers and researchers with a more reliable method for understanding and forecasting international migration patterns.

Chapter 3 addresses the critical gap in migration modeling by incorporating the role of population age structure in forecasting migration flows. The Migration Age Structure Index (MASI) is introduced as a novel method for adjusting net migration rates based on the age distribution of the population. By accounting for the varying migration propensity across different age groups, MASI improves the accuracy of migration forecasts, particularly for countries with rapidly aging populations. This contribution not only refines migration predictions but also enhances long-term demographic projections by integrating age-adjusted net migration flows into population forecasts.

Chapter 4 develops the Bayesian Person-Place Model (Bayes PPM) to improve residency estimation using administrative records. The Bayes PPM addresses the limitations of the existing Person-Place Model (PPM) by providing a well-defined statistical foundation for estimating the probability that an individual resides at a given address based on their administrative records data. This model enhances the accuracy of demographic estimates and the reliability of uncertainty intervals, making it particularly valuable for population estimation in countries without population registers.

Collectively, these contributions advance the state of the art in migration forecasting and demographic estimation, providing new methods to improve the accuracy and reliability of population estimates and projections.

1.4 Organization of Dissertation

The remainder of this dissertation is organized as follows. Chapter 2 presents the Bayesian Flow Model for forecasting global bilateral migration flows. Chapter 3 introduces the Migration Age Structure Index (MASI), a novel approach for summarizing population age dynamics and method for adjusting net migration rates based on population age structure. Chapter 4

develops the Bayesian Person-Place Model (Bayes PPM) for improving residency estimation using administrative records, addressing the limitations of existing models and providing a robust statistical foundation for demographic analyses. Finally, Chapter 5 concludes the dissertation by summarizing the contributions of the research and suggesting directions for future work.

Chapter 2

PROBABILISTIC FORECASTS OF BILATERAL MIGRATION FLOWS FOR 200 COUNTRIES

2.1 Introduction

We propose a method for forecasting global human migration flows. A Bayesian hierarchical model is used to make probabilistic projections of the 39,800 bilateral migration flows among the 200 largest countries from 2020 through 2045. The model is fit to estimates of quinquennial bilateral migration flows from 1990 through 2020. We find that the model produces well calibrated out-of-sample forecasts of bilateral flows, as well as total country-level inflows, outflows, and net flows. The mean absolute error decreased by 61% using our method compared to a leading model of international migration. Out-of-sample analysis indicates that simple models of migration flows offer accurate projections of bilateral migration flows in the near term. Our method matches or improves on the out-of-sample performance using these simple deterministic alternatives, while also accurately assessing uncertainty. Finally, we adapt a fully probabilistic population projection model to generate bilateral migration flow forecasts by age and sex for the 200 most populous countries.

This chapter is closely based on Welch & Raftery (2022). Projected migration flow summaries for all countries are available as supplemental material to the published article.

Recent methodological advances have made it possible to generate plausible estimates of international migration flows at a global scale. Abel (2013a) pioneered a method to estimate the minimum number of people who must have changed their country of residence to explain the change in migrant stocks among all countries of the world. Azose & Raftery (2019a) extended the minimum migration estimates to produce well calibrated pseudo-Bayes estimates of bilateral migration flows. This was found to perform best among six methods

for estimating international migration, in the sense of having the highest correlation with several common measures of migration (Abel & Cohen, 2019).

Accurate estimates of historic migration trends and forecasts of future trends are essential to crafting effective migration policies (Willekens, 2018), but flow forecasting method development lags progress in flow estimation (Bijak, 2006). Gravity models use *push* factors to help explain the magnitude of out-migration from a country along with *pull* factors to help explain the magnitude of country-level in-migration. Recent work with these models use estimates of migration flows from most of the world to a few wealthy countries and vice versa to quantify the influence of push and pull factors on the magnitudes of migration flows (Cohen et al., 2008; K. Kim & Cohen, 2010). Alternatives to the gravity model approach are concerned with migration flow forecasting for a subset of countries or regions (Wiśniowski et al., 2015a; Raymer & Wiśniowski, 2018).

We address the problem of probabilistic forecasting of international migration flows between all pairs of the most populous 200 countries. Our approach uses a Bayesian hierarchical model. A Bayesian hierarchical model pools information across time periods and individual flows (Gelman & Hill, 2006). This property helps compensate for the small number of periods where migration flow estimates are typically available. The Bayesian approach also makes it possible to encode outside information in the model, helping to reign in implausibly large forecast variability.

This paper describes a Bayesian hierarchical model that builds on a key idea from (Willekens & Baydar, 1986): *once the overall level of migration is controlled for, spatial distribution patterns are found to be remarkably stable.* Rogers, Raymer, and colleagues (Rogers et al., 2002; Raymer et al., 2020) used this idea to model regional or subnational migration flows. Our approach estimates spatial interactions conditional on the country of origin and total origin country outflows in each period. This conditional origin framework makes it possible to capture the variability in spatial interactions across flows from the same origin. It also yields forecasts of net migration by country that sum to zero by construction.

Out-of-sample evaluation results indicate that our model generates plausible, well-calibrated

forecasts of bilateral migration flows over a short time horizon. These accommodate the possibility of major migration shocks in the future. We use our fitted model to generate probabilistic forecasts of global bilateral migration flows for quinquennial periods from 2020-2025 period to 2040-2045.

2.2 Data

BFM out-of-sample results and forecasts are based on the pseudo-Bayes estimates of flows during five-year periods starting in 1990 and running through 2020 (Abel & Cohen, 2019; Azose & Raftery, 2019a). The flow matrix for C countries in the period starting at time t is

$$M_t := \begin{pmatrix} 0 & m_{1,2,t} & \dots & m_{1,C,t} \\ m_{2,1,t} & 0 & \dots & m_{2,C,t} \\ \vdots & \vdots & \ddots & \vdots \\ m_{C,1,t} & m_{C,2,t} & \dots & 0 \end{pmatrix}, \tag{2.1}$$

where $m_{i,j,t}$ is the flow of individuals from country i to country j during the period starting at time t. The off-diagonal entries show the total number of movers whose place of residence at the end of the period was different than the one they had at the beginning of the period. The diagonal entries are set to zero as we are only interested in modeling the magnitude of migration flows.

The sum of the entries in row i, $m_{i,+,t} = \sum_j m_{i,j,t}$, is the total number of people whose residence was in country i at the start of period t but was some other place at the end of period t. This sum approximates the outflows from origin i during the period. Similarly, the sum of the entries in column j, $m_{+,j,t} = \sum_i m_{i,j,t}$, is the total number of people whose residence was in country j at the end of the period but was somewhere else at the beginning of the period. This is the total of the inflows to destination j during period t. The net flow for country c is $r_{c,t} = m_{+,c,t} - m_{c,+,t}$.

The flow matrix can underestimate the total number of people who migrated during each

five-year period. Some people will start and finish the period with residence i even though they established multiple residences other than i throughout the period. A person might also start the period with residence i, establish several residences throughout the period, and reside in country j at the end of the period. The flow estimate would only capture the change from i to j.

Flow estimates are available for $C = 200$ countries during the six five-year periods starting in years $t \in \{1990, 1995, 2000, 2005, 2010, 2015\}$. This translates to 39,800 bilateral migration flows observed during six periods starting in 1990 and ending in 2020. We do not address uncertainty in the underlying estimates to generate our forecasts; however, future work might account for such uncertainties in the underlying data.

2.3 Methods

2.3.1 Bayesian Flow Model

We fit a Bayesian hierarchical model to all available flow estimates. The model takes advantage of three properties of the data generating process: 1) every individual must start from one origin, 2) every individual can choose just one of (C-1) possible destinations, and 3) the spatial distribution of migration flows remains relatively constant over time. We exploit these three properties in the specification of the hierarchical model which is as follows:

$$
\begin{aligned}
&\textbf{Observations} \begin{cases} m_{i,\cdot,t} \mid \pi_{i,\cdot,t}, \delta_{i,t} \overset{\text{ind}}{\sim} \text{Multinomial}\left(N_{i,t}, \pi_{i,\cdot,t}\right) \end{cases} \\
&\textbf{Outflow} \begin{cases} N_{i,t} = \lfloor \delta_{i,t} P_{i,t} + 1/2 \rfloor \\ \log \delta_{i,t} \overset{\text{ind}}{\sim} \text{Normal}\left((1-\phi)\,\mu_i + \phi \log \delta_{i,t-1}, \sigma_i^2\right) \\ \phi \sim \text{Uniform}\left(0,1\right) \\ \mu_i \overset{\text{iid}}{\sim} \text{Normal}\left(\nu, \tau_0^2\right) \\ \nu \sim \text{Normal}\left(\mu_0, 100^2\right) \\ \sigma_i \overset{\text{iid}}{\sim} \text{Beta}\left(a_0, b_0\right) \end{cases} \\
&\textbf{Inflow} \begin{cases} \pi_{i,j,t} = \exp \eta_{i,j,t} / \sum_{j=1}^{J} \exp \eta_{i,j,t} \\ \eta_{i,j,t} \overset{\text{ind}}{\sim} \text{Normal}\left(\kappa_{i,j}, \psi_{i,j}^2\right) \\ \kappa_{i,j} \overset{\text{iid}}{\sim} \text{Normal}\left(0, 10^2\right) \\ \psi_{i,j} \overset{\text{iid}}{\sim} \text{Beta}\left(p_0, q_0\right) \end{cases}
\end{aligned}
$$

where "ind" means independently distributed, "iid" means identically and independently distributed, and $\lfloor x \rfloor$ denotes the floor of x, i.e. the largest integer smaller than x. Table 2.1 defines each term of the model.

Under this specification, the expected flow from origin i to destination j during period t is

$$\mathbb{E}\left[m_{i,j,t} \mid \pi_{i,j,t}, \delta_{i,t}\right] = \pi_{i,j,t} N_{i,t} \approx \pi_{i,j,t} \delta_{i,t} P_{i,t}. \tag{2.2}$$

This expected value encodes the three defining features of the generative process. Each flow is composed of individuals who belonged to the population of the origin country at the start of the period. Only a subset of the country's population migrates, and that fraction for origin country i in period t is $\delta_{i,t}$. Every individual leaving origin i in period t migrates to

Table 2.1: Bayesian Flow Model

$m_{i,j,t}$	Integer-valued flow (data) from origin i to destination j during period starting in year t
$N_{i,t}$	Number of migrants departing origin i for period starting in year t after rounding to the nearest integer value
$P_{i,t}$	Period t number of person-years in origin i and period t (000s)
$\delta_{i,t}$	Out-migration rate in migrants per 1,000 person-years for origin i for period starting in year t
$d_{i,t}$	Observed out-migration rate in migrants per 1,000 person-years for origin i for period starting in year t
$\pi_{i,\cdot,t}$	Vector of destination weights among migrants with origin i in period t such that $\sum_j \pi_{i,j,t} = 1$ for all i and t
ϕ	Global weight on mean departure rate function
μ_i	Long-term mean out-migration rate for origin i
σ_i^2	Temporal variation around departure rate from origin i
$\eta_{i,j,t}$	Destination weight at time t for destination j among migrants from origin i
$\kappa_{i,j}$	Mean destination weight for j among migrants from origin i
$\psi_{i,j}^2$	Variance of destination weights j from origin i
ν	Grand mean of long-term departure rates
μ_0	User-specified prior for long-term departure rate grand mean
τ_0^2	User-specified variance about the long-term departure rate grand mean
a_0, b_0	User-specified parameters for temporal variation around departure rates
p_0, q_0	User-specified parameters for destination weight variation for origin i
C	Number of countries (200)

only one destination and the distribution of destinations is stable over time; this is what the vector $\pi_{i,\cdot,t}$ encodes.

For each origin, the out-migration rate is modeled on the logarithmic scale as a weighted average of two quantities: the logarithm of the last departure rate, $\log \delta_{i,t-1}$, and the long-term mean rate for country i, μ_i. Stochastic variation around log out-migration rates is captured by σ_i^2. This model ensures that departure rates never drop below zero.

The mean of $\pi_{i,j,\cdot}$ represents the long-term relative tendency of migrants from origin i to move to destination j. Variation about the long-term tendencies from period to period is represented by a random intercept mixed effects model. Spatial distribution tendencies

$(\pi_{i,C-1,t})$ are encoded by an over-parameterized model equivalent to the centered logratio (clr) transformation (Aitchison, 1986), namely

$$\text{clr}(\pi_{i,j,t}) = \log \frac{\pi_{i,j,t}}{\left(\prod_{j=1}^{C-1} \pi_{i,j,t}\right)^{1/(C-1)}}. \tag{2.3}$$

This model provides a flexible framework that could be extended to take origin- and/or destination-specific variables into account, e.g. colonial relationships, shared language, and economic differences creating push/pull forces among country pairs. A sum-to-zero constraint on $\kappa_{i,j}$ makes this model identifiable.

The model is implemented by Markov chain Monte Carlo (MCMC) using the R NIMBLE software package (de Valpine et al., 2017, 2021; R Core Team, 2021). NIMBLE is designed to sample from the posterior distribution of a Bayesian hierarchical model and is optimized for computational efficiency; however, sampling the model remained computationally slow. We overcame the computational challenges by splitting the model into an outflow component and an inflow component.

The outflow component includes all parts of the model involving $\delta_{i,t}$. We estimated μ_i, ϕ, and σ_i by approximating $\delta_{i,t}$ with the observed out-migration rate for origin i, $d_{i,t}$. The spatial distribution parameters, $\pi_{i,j,t}$ were estimated from the observations $clr\big(p_{i,j,t}\big) \sim \text{Normal}\big(\kappa_{i,j}, \psi^2_{i,j}\big)$, where $p_{i,j,t} = \frac{m_{i,j,t}+k}{\sum_j m_{i,j,t}+k}$ for $k \in (0, 1]$. These approximations make it possible to parallelize the model, reducing processing time by several orders of magnitude. Parallelized performance gains, however, make it necessary to specify a number of hyperparameters that might otherwise be estimated simultaneously with the rest of the parameters in the model.

Prior Specification

There are several user-specified parameters in the BFM. We set the prior parameters using empirical aggregate metrics across all flows. This preserves the information sharing benefits

of hierarchical modeling while improving computation time by several orders of magnitude.

- a_0, b_0: These define the prior distribution of the σ_i. More than 97% of the standard deviations of $\log d_{i,t}$ by origin are smaller than 1. Values of σ_i near 0 can lead to implausibly narrow intervals for some small countries. Values of σ_i larger than 1 permit a few highly variable outflows to obtain implausibly massive proportions of the population to out migrating. We therefore set a_0 and b_0 so that values of σ_i close to zero or above 1 are unlikely. We do this by minimizing the sum of the differences of the 2.5% and 97.5% quantiles from Beta(a_0, b_0) and (0.15, 0.99).

- p_0, q_0: We found p_0 and q_0 by minimizing the sum of the differences between the 2.5%, and 97.5% quantiles of the Beta(p_0, q_0) distribution and the quantiles of the empirical standard deviation of the means of $\text{clr}(p_{i,j,t})$ by origin for all positive flows.

- τ_0: This parameter represents the amount of variation around the long term outflow mean for each origin. We set this value equal to approximately three times the standard deviation of the mean log origin outflow rates.

- μ_0: This parameter represents the mean of long-term means of the log outflows for each origin. We set it equal to the mean of the log origin outflow rate means.

2.3.2 Forecasting

After drawing S samples from the posterior distribution of the parameters of the BFM, we generated probabilistic forecasts as follows:

1. Sample a departure rate for origin i during period t from the posterior predictive distribution for $\delta_{i,t}^{(s)}$ and convert it to an integer $N_{i,t}^{(s)} = \text{round}\big(\delta_{i,t}^{(s)} P_{i,t}\big)$.

2. Sample an allocation vector from the posterior predictive distribution of $\pi_i^{(s)}$.

3. Sample a flow vector from the density $m_{i,\cdot,t}^{(s)} \mid N_{i,t}^{(s)}, \pi_i^{(s)} \sim \text{Multinomial}\big(N_{i,t}^{(s)}, \pi_i^{(s)}\big)$.

4. Distribute outflow counts by age group according to a Rogers-Castro migration schedule (Rogers & Castro, 1981b) and sex according to the observed proportions present in the population.

5. Increment t and repeat until reaching the last period of the forecast.

For C countries, T periods, and S samples, this procedure leads to a $C \times C \times T \times S$ dimensional array. Each flow matrix sampled using this procedure is guaranteed to yield zero total net global migration, i.e. to ensure that the sum of inflows and the sum of outflows across countries are equal. Furthermore, country inflows, country net flows, regional flows, and global flows are all implicitly defined by the model.

Population projections were generated from a custom implementation of the `bayesPop` `R` package (Ševčíková & Raftery, 2016a). Our implementation makes it possible to trace inflows back to specific origins by age and sex for each period. Models of mortality and fertility implemented in `bayesPop` make forecasts generated from our implementation bona fide population projections based on bilateral migration flows rather than country-level net migration.

We put thresholds on the fraction of the population leaving in any one period in order to avoid unprecedented decreases in origin country populations. The outflow rates from countries with one million or more people were constrained so that no more than 16% of the population departs in any single period. Gulf Cooperation Council (GCC) member countries (United Arab Emirates, Bahrain, Kuwait, Qatar, Oman, Saudi Arabia) were constrained so that no more than 42% of the population departs in any single period. Finally, outflow rates from GCC labor-supplying countries (Bangladesh, Egypt, India, Indonesia, Pakistan, Philippines) were constrained so that no more than 3% of the population departs in any one period. All three constraints were calculated using the flow and population data. The GCC bounds were calculated using the maximum observed departure rate from these countries. The non-GCC bound is equal to the 99th percentile of historic mean departure rates among non-GCC and non-GCC labor countries.

2.3.3 Gravity Model

We evaluated the Bayesian hierarchical model by comparing out-of-sample performance to a gravity model of migration (K. Kim & Cohen, 2010). After removing flows with zero migrants, we used ordinary least squares to fit the gravity model

$$\begin{aligned}
\log m_{i,j,t} = \beta_0 &+ \beta_1 \log P_{i,t} + \beta_2 \log P_{j,t} + \beta_3 \log D_{i,j} \\
&+ \beta_4 \log PSR_{i,t} + \beta_5 \log PSR_{j,t} \\
&+ \beta_6 \log IMR_{i,t} + \beta_7 \log IMR_{j,t} \\
&+ \beta_8 \log urban_{i,t} + \beta_9 \log urban_{j,t} \\
&+ \beta_{10} \log LA_i + \beta_{11} \log LA_j \\
&+ \beta_{12} LL_i + \beta_{13} LL_j \\
&+ \beta_{14} LB_{i,j} + \beta_{15} OL_{ij} + \beta_{16} COL_{i,j} \\
&+ \beta_{17} (t - 2000) + \beta_{18} (t - 2000)^2 + \varepsilon_{i,j,t}.
\end{aligned}$$

We used UN estimates of population, distance between capitals, potential support ratio, and infant mortality ratios. All other variables were obtained from the CEPII database (Mayer & Zignago, 2011) or were manually coded. Table 2.2 gives the definition for each covariate in the gravity model.

2.3.4 Hurdle Model

More than half of the historic flows are zero, but the gravity model approach (Cohen et al., 2008; K. Kim & Cohen, 2010) does not account for these migration flows explicitly. To deal with this, we also fit a gravity-like hurdle model using a two-stage method that removes the need to censor the data.

A hurdle model is a two-stage alternative to a zero-inflated mixture model. Hurdle models explicitly model the generative process leading to counts equalling zero. If an observation is greater than zero, then the *hurdle* is crossed and the second stage of the model is fit to the

Table 2.2: Gravity Model Covariates

$P_{i,t}$	Population of country i at the start of period t
$PSR_{i,t}$	Potential support ratio for country i at the start of period t, i.e. the number of people aged 15-64 per person aged 65+
$IMR_{i,t}$	Infant mortality ratio for country i at the start of period t
$urban_{i,t}$	Percentage of the population living in urban settings for country i at the start of period t
$D_{i,j}$	Distance between capitals of country i and country j
LA_i	Land area of country i
LL_i	Land locked indicator for country i
$LB_{i,j}$	Indicator of shared land boarder between countries i and j
$OL_{i,j}$	Indicator of shared official language in countries i and j
$COL_{i,j}$	Indicator of colonial relationship between country i and j
t	First year of period
$\varepsilon_{i,j,t}$	Variation not explained by the model

positive counts (Jackman, 2009). The gravity hurdle model with a Poisson count component and a binomial zero component is as follows:

$$\begin{aligned} m_{i,j,t} \mid m_{i,j,t} > 0 &\overset{\text{ind}}{\sim} \text{Poisson}\big(\lambda_{i,j,t}\big), \\ \mathbb{1}_{m_{i,j,t}>0} &\overset{\text{ind}}{\sim} \text{Binomial}\big(1, \omega_{i,j,t}\big), \end{aligned}$$

with positive count covariate matrix $X^{[1+]}$ and zero covariate matrix $X^{[0]}$,

$$\begin{aligned} \log \lambda_{i,j,t} &= X^{[1+]}_{i,j,t}\beta, \\ \text{logit } \omega_{i,j,t} &= X^{[0]}_{i,j,t}\gamma. \end{aligned}$$

The mean for the positive component, $\lambda_{i,j,t}$, uses the same regressors as the gravity model, yielding estimates for the parameter vector $\beta = (\beta_0, \beta_1, \ldots, \beta_{18})$.

The mean for the zero component includes the populations $P_{i,t}$ and $P_{j,t}$, the distance between the capitals $D_{i,j}$, a shared land border indicator $LB_{i,j}$, and a shared official language indicator $OL_{i,j}$. Hence the parameter vector for the zero component is $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_5)$.

Table 2.3 shows that the out-of-sample coverage for the truncated Poisson hurdle model was quite good, but that the the prediction errors for both the gravity model and the hurdle model were much larger on average than for the other models considered.

2.3.5 Mean Absolute Percent Error

For a total of F flows in the flow matrix M and the flow matrix estimate $\tilde{M}$, the mean absolute percent error is defined as

$$\text{MAPE}\big(M, \tilde{M}\big) = \frac{100}{F} \sum_{i \neq j} \frac{|m_{i,j} - \tilde{m}_{i,j}|}{m_{i,j} + 1}. \tag{2.4}$$

Normalizing the errors by the magnitude of the observed flow puts the magnitude of errors into context. This means that the errors are measured relatively to the size of the underlying flows. We use $m_{i,j} + 1$ instead of $m_{i,j}$ in the denominator to avoid degeneracies that arise when flows are equal to zero.

2.4 Results

2.4.1 Out-of-Sample Validation

We fit five competing models to estimates of bilateral migration flows for the first five periods for which flow data were available, i.e. 1990-2015, using the data that would have been available at the start of the 2015-2020 period. We fit each model to the flow estimates calculated using the 2015 Revision of the World Population Prospects (WPP) (United Nations. Department of Economic and Social Affairs, 2015) and associated flow estimates (Azose & Raftery, 2019a). Those models were then used to predict the bilateral flows for the 2015-2020 period using the most updated flow data and estimates available in 2021, including the 2019 Revision of the WPP and most recent estimates of migration flows (Abel & Cohen, 2019). We calculated the mean absolute error (MAE) and the mean absolute percentage error (MAPE) for all models and the prediction interval coverage for the probabilistic methods (Bijak et

al., 2019; Czado et al., 2009).

Table 2.3: Out-of-sample mean absolute error (MAE) in thousands of migrants per period, mean absolute percentage error (MAPE) and 95% prediction interval coverage for models fit to 1990-2015 flow estimates and tested on 2015-2020 flow estimates.

			95% PI			
Method	MAE	MAPE	Flow	In	Out	Net
Historic Mean Flow	1.2	139	-	-	-	-
Persistence	**1.0**	79	-	-	-	-
Gravity Model	3.0	1,565	86	77	80	99
Poisson Hurdle Model	10.0	25,649	90	66	65	48
Bayesian Flow Model	1.2	**71**	**94**	**86**	**93**	**94**

The MAE measures the average absolute difference between each flow forecast and the observed flow for the 2015-2020 period. It summarizes prediction error on the same scale as the data for all flows. The MAPE normalizes the error by the magnitude of the observed flow, after adding 1 to both the observed and predicted flows to avoid infinite values. Normalizing errors by the magnitude of the observed flow puts the magnitude of errors for different pairs of countries on the same scale.

We also evaluated competing probabilistic methods by comparing their 95% prediction interval coverages. If a probabilistic flow model is well calibrated, then the 95% prediction intervals from a model fit to the 1990-2015 flows should include about 95% of the 2015-2020 flows. Prediction interval coverage estimates that differ from the nominal value (95% in this case) indicate that a model may be poorly calibrated, misspecified, or both.

Table 2.3 summarizes the performance of the Bayesian Flow Model (BFM) alongside two simple deterministic approaches, a standard gravity model, and the Poisson hurdle model, which is a more complex gravity model (Cohen et al., 2008; K. Kim & Cohen, 2010; Jackman, 2009). Among the probabilistic models, the BFM was the best calibrated for bilateral flows, total in-flows, total out-flows, and net migration flows. The BFM had the lowest MAPE among the methods considered, and clearly outperformed both gravity models in terms of

both MAE and MAPE. Interestingly, the simple deterministic models performed similarly to the BFM in terms of MAE and MAPE, but they fall short in that they do not produce prediction intervals.

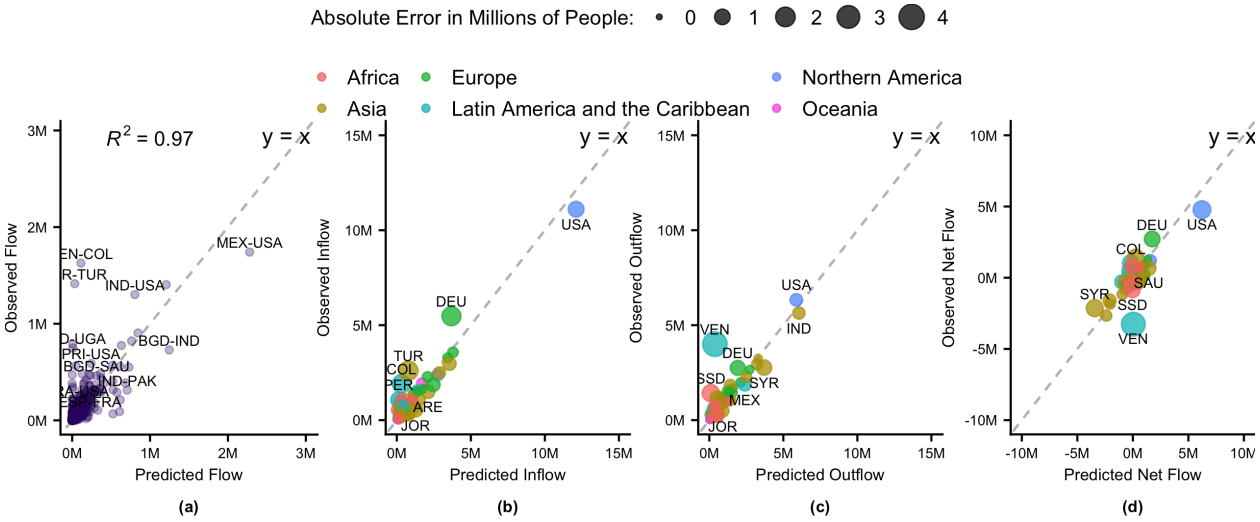


Figure 2.1: Observed 2015-2020 (a) bilateral flows, (b) total country inflows, (c) total country outflows, and (d) total country net flows compared to Bayesian hierarchical model median forecasts colored by United Nations Area and sized according to the absolute error in millions of people.

Figure 2.1(a) summarizes the distribution of observed and predicted flows for the BFM. Points that fall along the dashed line $y = x$ indicate perfect agreement between the single best forecast and the observed flow. The estimated R^2 values from the Poisson hurdle model, $R^2 = 0.93$, improves on the gravity model, $R^2 = 0.83$, and the BFM $R^2 = 0.97$ improves on both. The BFM is a good model for most flows, but there are several examples where the predicted flow leads to large errors. Examples include the flows from Venezuela to Columbia, from Syria to Turkey, from Mexico to the United States, and from South Sudan to Uganda. Large errors for flows originating in Venezuela and in Syria arise from major political crises in those countries (Bahar et al., 2018; Haddad, 2011). Departures from historic norms in the Mexico to United States flow could be partly explained by actual or perceived changes in immigration policy by the Trump administration between 2017 and 2020 (Hoekstra &

Orozco-Aleman, 2021). Large errors associated with South Sudan are reasonable since the country was founded only in 2011 (Martell, 2011). These cases show that major migration shocks can generate observations that fall far in the tail of the forecast distribution.

Figure 2.1(b)-(d) show country-level forecasts of 2015-2020 flows compared to the best available estimates of country-level flows. Median BFM inflow forecasts for Germany (DEU) and Turkey (TUR) were smaller than the estimated values for the 2015-2020 period. The median BFM outflow forecast for Venezuela (VEN) was smaller than the observed outflow while the median Syria outflow forecast was too large. Country-level inflow and outflow forecasts implicitly define net flow forecasts. The error in Venezuela's outflow forecast was carried over into the net flow forecast as shown in Figure 2.1(d): the net flow was smaller than the median forecast. This is consistent with the very large number of people that left Venezuela during the period.

2.4.2 Forecast Evaluation

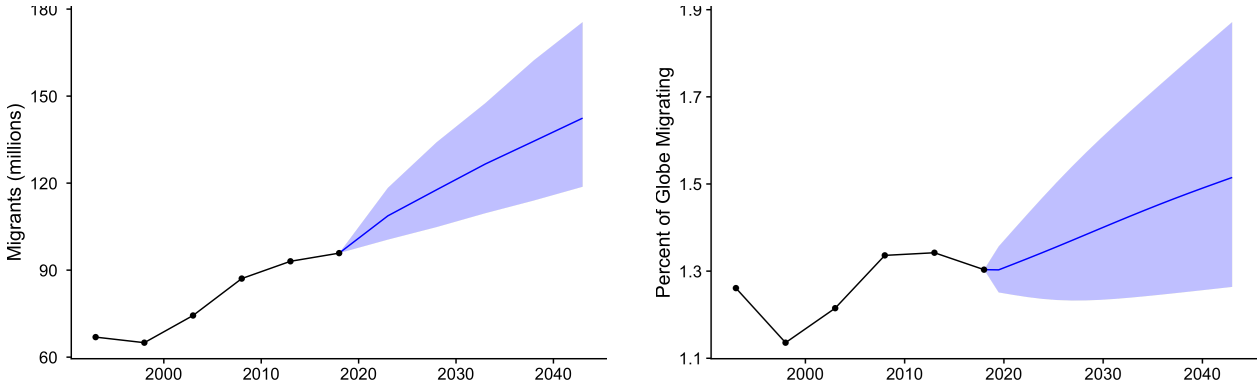
We produce forecasts for all 39,800 flows for 2020–2045. One appeal of well-calibrated probabilistic migration flow forecasts is that aggregate quantities or functions of flows lead to valid approximations of the statistical distribution associated with the function applied to individual flow forecast trajectories. We use this fact to evaluate the BFM forecasts of country-level inflows, outflows, net flows, and the percentage of the globe migrating. A summary of each bilateral flow forecast is included in the *Supplemental Index* of Welch & Raftery (2022).

Figure 2.2 shows the estimated and forecasted number of people migrating globally in each period. The median forecast is that the number of people migrating in 2040-2045 will be nearly 50% larger than the number of people migrating in the 2015-2020 period. However, much of this increase is due to the projected increase in world population. After accounting for global population growth, the percentage of world population migrating increases from about 1.3% in the 2015-2020 period to 1.5% in the 2040-2045 period–an increase of only about 16%.

Table 2.4 summarizes the estimated increase in the total number of people migrating

Table 2.4: Total global migration in millions of migrants per five year period and percentage of the population migrating. The columns correspond to the 5th, 50th (median) and 95th percentiles of the predictive distribution.

		2040-2045 Forecast		
	2015-2020	5%	50%	95%
Sum of Global Flows (millions)	96	119	142	176
Percent of Population Migrating	1.3	1.3	1.5	1.9



(a) (b)

Figure 2.2: (a) Global migration flows in millions of migrants and (b) in percentage of global population migrating during five year periods observed from 1990 through 2020 with median forecast (solid line) and 90% prediction interval for five year periods from 2020 through 2045

around the world for the last period where data are available (2015-2020), and the last period in the forecast (2040-2045). Growth in global migration will be driven first by the increasing global population and to a lesser extent by growth in the out-migration rates in a few large countries. The out-of-sample performance of the forecasts of individual flows and aggregate measures of migration indicates that the BFM is well calibrated in the short term.

Case Study: Germany Flow Forecast Figure 2.3 shows the BFM forecasts for Germany in millions of migrants per five year period from 2020 through 2045. Figure 2.3(a) shows

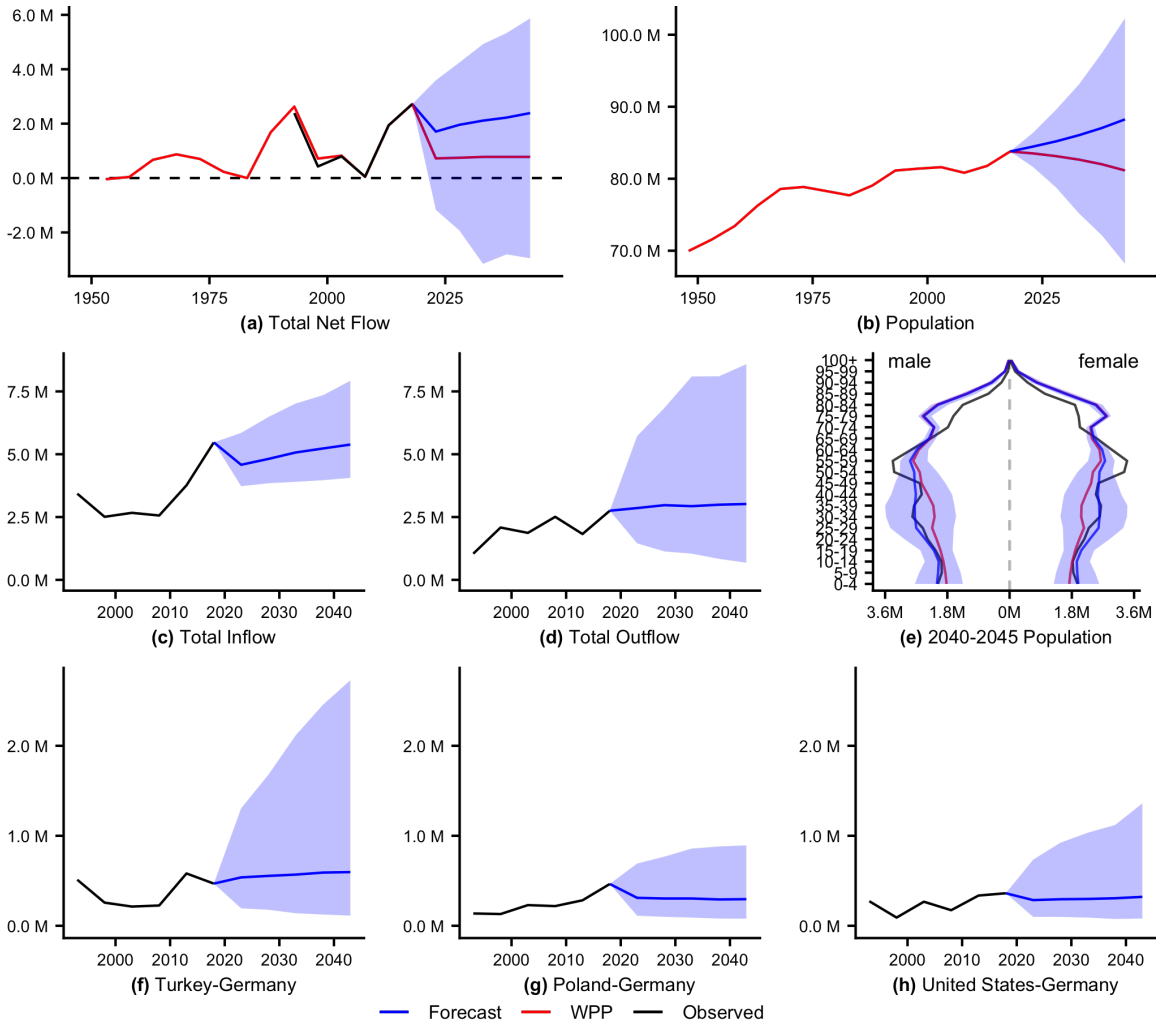


Figure 2.3: Observations and 90% prediction interval forecasts in millions of people per five year period for (a) total net flow, (b) total inflow, (c) total outflow, (d)-(f) bilateral flows with Germany as origin in descending order by historic magnitude, (g)-(i) bilateral flows with Germany as destination in descending order by historic magnitude.

the historic net migration rate in Germany from 1950 through 2020 along with the UN projection of net migration from the 2019 Revision of World Population Prospects (WPP) (United Nations. Department of Economic and Social Affairs, 2019). The UN net migration projection falls within the 90% prediction interval constructed from joint forecasts of bilateral flows into and out of Germany. Even though the median forecast and prediction interval

suggest more positive net flows into Germany using the BFM, our net migration forecast interval contains the UN's net migration projection for all forecast periods. The BFM forecast of net flows also appears plausible given the range of past net flows; however, the 90% prediction intervals from the BFM indicate that net outflows from Germany are possible over the coming years. This net outflow could occur if for example many of the migrants who fled humanitarian crises over the last decade return to their home countries once the humanitarian situation ends.

Figure 2.3(b) shows the range of uncertainty about the total population of Germany generated by the net migration flow forecast through 2045. The median population forecast using the BFM flow model is higher than the WPP 2019 projection and shows a wide range of uncertainty due to forecast uncertainties in total outflows from Germany and large inflows from a few countries.

Figure 2.3(c)-(d) shows the total estimated flows into and out of Germany every five years from 1990-2020 and forecasts of flows from 2020 through 2045 implied by the BFM. The outflow forecast effectively continues the historic pattern with relatively wide uncertainty. The median inflow forecast moderates from a historic peak in the 2015-2020 period before growing again through 2045. Inflow forecast patterns like this reflect a return to long-term historic destination preferences encoded in the destination component of the model and/or a fall in the total outflow from one or more countries from one period to the next. Inflow forecast counts tend to increase with the increasing populations in sending country populations.

The cumulative impacts of migration uncertainty on the age profile of the German population by the end of the forecast period in 2040-2045 are shown in Figure 2.3(e). The median forecast of the German age profile for people aged 0-44 remains similar to the current profile. As a result, economic and budgetary challenges that could arise from too few working age adults supporting people currently aged 45-65 may not be realized in the 2040-2045 period. However, the width of the 90% prediction interval shows that this future is not a sure bet.

Figure 2.3(f)-(h) shows estimates of the largest flows into Germany from 1990-2020 and

the BFM forecasts for those flows from 2020-2045. The median forecasts are approximately equal to the mean of the historic flows. Flow forecasts from Turkey to Germany contribute the largest number of migrants to the median flow forecast along with the largest degree of uncertainty in future inflows. See the *Supplemental Index* of Welch & Raftery (2022) for all country-level forecast summary plots.

2.5 Discussion

Bayesian probabilistic forecasting methods for demographic processes have become increasingly widespread in the last 15 years (Bijak & Wisniowski, 2010; Alkema et al., 2011; Raftery et al., 2013; Wiśniowski et al., 2015a; Azose et al., 2016a), and here we extend this approach to forecasting bilateral international migration flows. Our approach builds on the multiplicative components model (MCM) (Raymer et al., 2020), arguments for separating the outflow rate and spatial interaction process (Rogers et al., 2002), and the relative stability of the spatial distribution of global flows over the past three decades (Abel et al., 2021; Willekens & Baydar, 1986). The mean departure rate functional form is inspired by the net migration flow model of (Azose & Raftery, 2015b), and reflects the fact that the net migration rate can be decomposed into inflow and outflow components.

We model out-migration and spatial interaction conditionally on the origin and population at risk of out-migration. This approach separates the magnitude of a flow from the spatial interaction as in (Willekens & Baydar, 1986; Rogers et al., 2002; Raymer et al., 2020); however, conditioning on the origin and population at risk is a departure from these methods. We also model spatial interactions jointly, conditional on the origin and magnitude of outflows. This approach leads to implicit estimates of country-level inflows and net flows. It also ensures that the magnitude of global outflows equals that of global inflows.

Spatial interaction is modeled with a centered logratio model (Aitchison, 1986). Our spatial interaction model does not use covariates, but could accommodate covariates in estimating the spatial interactions. Others have suggested the logratio model for the multinomial likelihood model of migration flows (Raymer & Willekens, 2008). We evaluated that formula-

tion as well, but chose the centered logratio formulation to remove the influence of a baseline country on forecasts of destination preferences of all other countries.

One challenge to implicitly defined inflows lies in the implementation of country-level in-migration limits. If a country implements an in-migration cap or if unconstrained forecasts lead to implausibly dense populations at some destinations, then our forecasting approach requires additional steps to account for these limiting factors. They could be incorporated at the flow level using covariates in the spatial distribution component of the model, or by iteratively rerouting outflows from each origin to alternative destinations.

It is too early to quantify the impact of the COVID-19 pandemic (Fontanet et al., 2021) on global migration flows over the 2020-2025 period. Early indications from some parts of the world suggest that global migration may have fallen during the 2020-2021 period due to strict border controls (European Commission – Press release, 2021). In other places, flows may be on track to hit historic highs despite the pandemic (Hackman & Caldwell, 2021).

The flow estimates that we used to fit the BFM rely on migrant stock estimates that are compiled every five years (Abel & Cohen, 2019). Data generated from social media platforms, search engine inquiries, and other digital trace data might offer alternative and more timely sources of migrant stock estimates (Zagheni & Weber, 2012; Zagheni et al., 2014, 2017; Cesare et al., 2018; Fiorio et al., 2021; F. Billari & Zagheni, 2017; Lazer & Radford, 2017; Alexander et al., 2020). Bias-corrected migrant stock estimates derived from big data have the potential to improve the time resolution of migration flow estimates, especially for regional and subnational contexts where platform adaptation among the population is more widespread. However, early applications of digital trace data used to study demographic and public health trends suggests that additional work may be needed to resolve the signals present in big data (Cesare et al., 2018; Ruths & Pfeffer, 2014; Lazer et al., 2014).

Large outflows from India, China, Indonesia, and several African countries have little impact on the overall population of those countries through 2045. However, uncertainty about the magnitude of outflows from these countries generates substantial uncertainty in the population age distribution for people aged 0-30 in both sending and receiving countries

by the end of the forecasting period. Also, very large outflows from India, China, Indonesia, and populous countries in Africa are possible, and could generate very large inflows for a few popular destinations, unless destination countries constrain the flow of migrants into their countries.

A time invariant destination model means that the share of migrants leaving one region for another will be relatively stable over time. An origin's outflow rate and population size are the main factors influencing the magnitude of outflows in the BFM. Even if departure rates continue to follow historic patterns, the absolute number of migrants leaving high-fertility countries or ones with currently young populations.

Chapter 3

ACCOUNTING FOR POPULATION AGE DISTRIBUTION IN FORECASTING MIGRATION

3.1 Introduction

Preparations for the next great demographic era are underway around the world (C. J. Kim, 2019; Director of National Intelligence, 2021). Decades of declining and then sub-replacement fertility created a demographic dividend, where the share of the working-age population far outweighed the share of children and older adults no longer participating in the workforce (Bongaarts, 2009). Now the average ages of populations around the world are increasing at the fastest rate in history (Bloom & Zucker, 2023). International migration may blunt the impact of this dynamic, but migration alone will not fully mitigate the realities of the next demographic era (RAND, 2005; Coleman, 2008). Accounting for past and forecast population age structure reveals that migration may be an even less effective mitigation than currently understood (Münz, 2013; Lee, 2011). However, existing forecasting methods fail to account for this eventuality, leaving the role of population age structure in migration dynamics to imprecise qualitative conjecture.

We propose a probabilistic net migration forecasting approach that accounts for the impacts that population age structure had on historic migration rate estimates. We use these to generate new probabilistic population forecasts for all countries. The resulting forecasts quantify how different the global population distribution could be after accounting for past and projected shifts in the population age structure.

This chapter is closely related to a manuscript submitted to *Demography* by Welch, Ševčíková, and Raftery.

The chapter is organized as follows. Section 3.2 reviews the existing work that is foun-

dational to our study. Section 3.3 summarizes the data and methods used to generate age standardized migration forecasts. This is followed by an evaluation of the age-standardized migration method in Section 3.4. Then Section 3.5 summarizes differences between population forecasts with and without accounting for population age structure in the migration component of the forecasts. We conclude in Section 3.6 with a discussion of the contributions and limitations of our approach.

3.2 Background

Rogers & Castro (1981a) identified a consistent pattern in the age profile of out-migrants and proposed a model migration age schedule. Even though modern migration estimation and forecasting methods use migration age schedules to account for the impact of migration on the sending and receiving populations, population age structure is rarely an explicit consideration in estimation or forecasting. This is reasonable over short-term forecast horizons given the relatively slow pace of population age structure changes, but less so over longer horizons.

Kupiszewski (2002) discussed differences in migration forecasts when the sending population age-structure is ignored. Specifically, it was argued that ignoring the limited number of migration-age individuals in Poland and future population aging contributed to unrealistically large out-migration forecasts from Poland to other European Union (EU) countries before accession to the EU in 2004.

Fertig & Schmidt (2005) argued that neglecting population age in calculating migration rates obscures the underlying population dynamics. They proposed an adjusted net migration rate that is calculated as the number of migrants aged 0-39 divided by the sending country population aged 0-39. This approach reduces distortions created by the population of older people who tend to make up a relatively small proportion of migrants, but effectively sets older age migration to zero and ignores the age structure of the critical 0-39 age group.

Kolk (2019) reviewed age-specific migration and the analogy to age-specific and total fertility rate definitions. It was argued that population-level measures should be adopted for both subnational and international migration estimation. The authors outlined key chal-

lenges to extending such methods to an international context, specifically properly accounting for the population at risk of in-migration and differences in the definition of an international migrant from country to country.

Skjerpen & Tønnessen (2021) incorporate the age profiles of people from three groups of origin countries to project Norwegian migration inflows through 2100. They find that incorporating origin population age structure in their models equates to a reduction of 0.1 children per woman in the total fertility rate by 2100, compared to official Norwegian assumptions. The authors also highlight that existing migration forecasting methods overlook the relative stability of age structure forecasts, which should be leveraged to improve migration forecasts.

Fully probabilistic population forecasting is an active area of demographic research (e.g. Hyndman & Booth, 2008; Raftery et al., 2012, 2014; Wiśniowski et al., 2015b; Yu et al., 2023). Azose et al. (2016b) developed a probabilistic net migration model and used it to produce probabilistic population forecasts for all countries through 2100. Welch & Raftery (2022) proposed a fully probabilistic population forecasting method that uses a probabilistic model of bilateral migration flows for all countries through 2045. However, none of these population forecasting methods systematically account for population age structure changes in both the historic data used to fit these models and in forecasting.

Raftery & Ševčíková (2023) proposed a method for accounting for population age structure in probabilistic migration and population forecasts, and applied it to very long-term population forecasts, to 2300, motivated by the problem of estimating the social cost of carbon. Like us, they used net migration data, which has advantages of data availability and analytical simplicity over methods based on more complete migration data, such as between-country flows or in- and out-migration flows for countries. However, a problem with this is that it does not take account of the different age distributions of in- and out-migrants (Rogers, 1990). They addressed this issue by assuming that during periods of net in-migration, all movement was classified as in-migration with no out-migration, and similarly, net out-migration periods were treated as solely out-migration. Here we develop a new method that avoids this unrealistic assumption while still using net migration data.

3.3 Data & Methods

Data

The United Nations defines a long-term migrant as a person who moved to a different country and stayed for at least one year (McAuliffe & Oucho, 2024). Recent methodological innovations make it possible to estimate globally consistent, country-level total migration inflows and outflows from Census counts of the number of people by country of birth in all countries (Abel, 2013b; Azose & Raftery, 2019b). Abel & Cohen (2019) found that the pseudo-Bayes method of Azose & Raftery (2019b) to estimate bilateral migration flows led to the most accurate migration estimates among several extant methods. Unfortunately, pseudo-Bayes migrant flow estimates over 5-year periods are available for only six 5-year periods starting in 1990 and ending in 2020 (Abel & Cohen, 2019). Estimating and forecasting bilateral migration flows has a number of advantages (Welch & Raftery, 2022), but population age structure changes too slowly to solely rely on these estimates for long-term migration forecasting.

The United Nation's World Population Prospects (WPP) publishes globally consistent net migration estimates, defined as the difference between migration inflows and outflows divided by population size, back to the 1950-1955 period (United Nations, 2022). The pseudo-Bayes flow estimates contain more detailed information in the form of bilateral flows. We use strengths of both data sets to develop a new combined data series.

Migration flow estimates are calculated with respect to specific revisions of the WPP. We use WPP 2019 revision (United Nations. Department of Economic and Social Affairs, 2019) and the associated flow estimates for all analyses. The 2022 revision transitioned from 5-year period to 1-year period estimates and forecasts (United Nations, 2022). The 2024 revision was the first to use a fully probabilistic migration forecasting model (United Nations, 2024). We use the 2019 revision to show the differences in long-term forecasts using the last available deterministic migration projections for 5-year periods in the WPP 2019 compared to a probabilistic migration model that ignores population age structure and one that explicitly accounts for past and future population age structure changes.

We fit a model, described in Section 3.3, to the 1990–2020 period from which both in-migration and net migration rates can be estimated for all countries with populations of 100,000 or more. This model is then used to approximate the contribution that in-migration made to the net migration rate prior to 1990. Our model-based approach is an efficient solution that allows for the estimation of globally consistent age-standardized in-migration and out-migration rates for 5-year periods starting in 1950 and running through 2020. The successful decomposition of net migration rates into in-migration and out-migration rates is critical to calculating migration rates adjusted for the population age structure that gave rise to those rates.

Age Standardization of Net Migration Rates

For reasons of data availability across all countries, and analytic simplicity, the United Nations (UN) uses all-age net migration as part of the basis for its population projections for all countries (United Nations. Department of Economic and Social Affairs, 2019). The WPP 2024 (United Nations, 2024) replaced a deterministic projection net migration method with the probabilistic approach of Azose & Raftery (2015a). However, neither the UN's deterministic approach nor the probabilistic approach of Azose & Raftery (2015a) takes account of historic or future changes in population age structure.

We propose to modify the approach of Azose & Raftery (2015a) to account for population age structure. We do this by creating an age-standardized version of the net migration rate. This poses a challenge, as historical net migration data are not disaggregated by age, making it difficult to standardize net migration rates directly. This was one of the arguments of Rogers (1990) for not using net migration at all.

We propose a model-based approach to help address these concerns. We start by developing a method for decomposing historic net migration into in- and out-migration. This uses a subset of the data for which estimates of both in- and out-migration are available (i.e. 1990–2020), and uses this to estimate in- and out-migration for the entire data period (i.e. 1950–2020). We then develop a method for age-standardizing out-migration. This turns

out to be very simple, relying on a quantity we call the *migration age structure index* (MASI). We extend this to in-migration, and hence to net migration. Finally we apply the probabilistic forecasting method of Azose & Raftery (2015a) to the age-standardized net migration rates, and input the resulting forecasts to the overall probabilistic population projection method.

Table 3.1 summaries the notation required for our methods. By convention, we refer to each period using the first year (inclusive) and the last year (exclusive). For example, the period 2000–2005 represents the interval from January 1, 2000, to December 31, 2004, i.e., [2000, 2005). Also, a "+" subscript in place of an index value denotes the marginal sum over that index.

Decomposing Historic Net Migration into In- and Out-migration

Net migration age-standardization takes place on the inflow and outflow components of the net migration rate. We use a mixed-effects model to relate the 1990–2020 in-migration rate, $\text{IMR}_{i,t}$, to the net migration rate, $\text{NMR}_{i,t}$, for each country i over these six 5-year periods, t, from 1990 through 2020. Our mixed-effects model is defined as follows:

$$\begin{aligned} \text{IMR}_{i,t} &= \beta_{0,i} + \beta_1 \max\left(\text{NMR}_{i,t}, 0\right) + \varepsilon_{i,t}, \\ \beta_{0,i} &\sim \text{Normal}\left(\beta_0, \sigma^2_{between}\right), \\ \varepsilon_{i,t} &\sim \text{Normal}\left(0, \sigma^2_{within}\right). \end{aligned} \tag{3.1}$$

This model uses a random intercept term to account for the average country-specific in-migration rate associated with the net migration rate for each period from 1990–2020. Country-specific intercepts, $\beta_{0,i}$, are drawn from a distribution with global mean, β_0. The difference in a country's random intercept from the global mean, β_0, is determined by the average difference from the global mean for that country in all periods. Out-migration rates are implicitly defined by this model.

The random intercept model was fit using the linear mixed-effects model implementation

Table 3.1: Notation for migration age-standardized migration method.

$O_{i,t}$	Integer-valued outflow from origin i in period starting with year t
$O_{i,t,a}$	Integer-valued outflow from origin i in period starting with year t of age group a–$(a+4)$
$I_{i,t}$	Integer-valued inflow to destination i in period starting with year t
$N_{i,t}$	Integer-valued net migrant flow in country i in period starting with year t
$P_{i,t,a,s}$	Integer-valued population of origin i at the start of period t for age group a-$(a+4)$ and sex $s \in \{\text{male}, \text{female}\}$
$P_{i,t,+,+}$	Integer-valued population of origin i at the start of period t
$\tilde{P}_{i,t,+,+}$	Integer-valued population of origin i at risk of migration over period t to $t+5$, defined as $P_{i,t+5,+,+} - N_{i,t}$
$\text{OMR}_{i,t}$	$= O_{i,t}/\tilde{P}_{i,t,+,+}$. Out-migration rate for country i in the period starting in year t
$\text{OMR}_{i,t,a}$	$= O_{i,t,a}/\tilde{P}_{i,t,a,+}$. Out-migration rate for country i, period starting in year t and age group a
$G_{i,t}$	$= \sum_a \text{OMR}_{i,t,a}$. Gross Migraproduction Rate (GMR) for country i and period starting in year t
$\text{IMR}_{i,t}$	$= I_{i,t}/\tilde{P}_{i,t,+,+}$. In-migration rate for period starting in year t and country i
$\text{NMR}_{i,t}$	$= \text{IMR}_{i,t} - \text{OMR}_{i,t}$. Net migration rate for country i and period t
$\pi_{i,t,a}$	$= \tilde{P}_{i,t,a,+}/\sum_a \tilde{P}_{i,t,a,+}$. Share of population from age group a in period t
$\pi^*_{i,t,a}$	Reference population age distribution for country i in period t
R_a	Reference migration age schedule normalized so that $\sum_a R_a = 1$
$C_{i,t}$	$= \sum_a R_a \pi_{i,t,a}$. The migration age structure index (MASI), a scalar accounting for the population and migrant age structure in the migration rate at time t for country i
$\check{C}_t$	$= \sum_{i,a} R_a \pi_{i,t,a}$. MASI for the world at time t
$\text{OMR}^\star_{i,t}$	$= \text{OMR}_{i,t} \times C_{i,\text{baseline}}/C_{i,t}$. Age-standardized out-migration rate
$\text{IMR}^\star_{i,t}$	$= \text{IMR}_{i,t} \times \check{C}_{\text{baseline}}/\check{C}_t$. Age-standardized in-migration rate
$\text{NMR}^\star_{i,t}$	$= \text{IMR}^\star_{i,t} - \text{OMR}^\star_{i,t}$. Age-standardized net migration rate

in the *lme4* R package implementation (Bates et al., 2015). Starting with $\text{NMR}_{i,t}$, the estimated mean in-migration rate is $\text{IMR}_{i,t} = \beta_{0,i} + \beta_1 \max(\text{NMR}_{i,t}, 0)$. The net migration rate and estimated mean in-migration rate are then used to calculate the out-migration rate, $\text{OMR}_{i,t} = \text{IMR}_{i,t} - \text{NMR}_{i,t}$.

Figure 3.1 summarizes the 1990–2020 data, model fit, and association between estimated and actual country-level migration flow rates. Points in Figure 3.1(a) show the observed $\max(\text{NMR}_{i,t}, 0)$ on the horizontal axis and $\text{IMR}_{i,t}$ on the vertical axis. The discontinuity at $\text{NMR}_{i,t} = 0$ reflects the fact that the in-migration rate must be non-negative. In the

$\text{NMR}_{i,t} \leq 0$ portion of the domain, the country-specific intercept, $\beta_{0,i}$, of model 3.1 establishes the minimum average magnitude of in-migration corresponding to a negative net migration rate for that country. When $\text{NMR}_{i,t} > 0$, the model in-migration rate increases from this minimum average in-migration rate by an average of β_1 per unit increase in $\text{NMR}_{i,t}$. The global mean model in-migration rate is shown by the blue line in Figure 3.1(a).

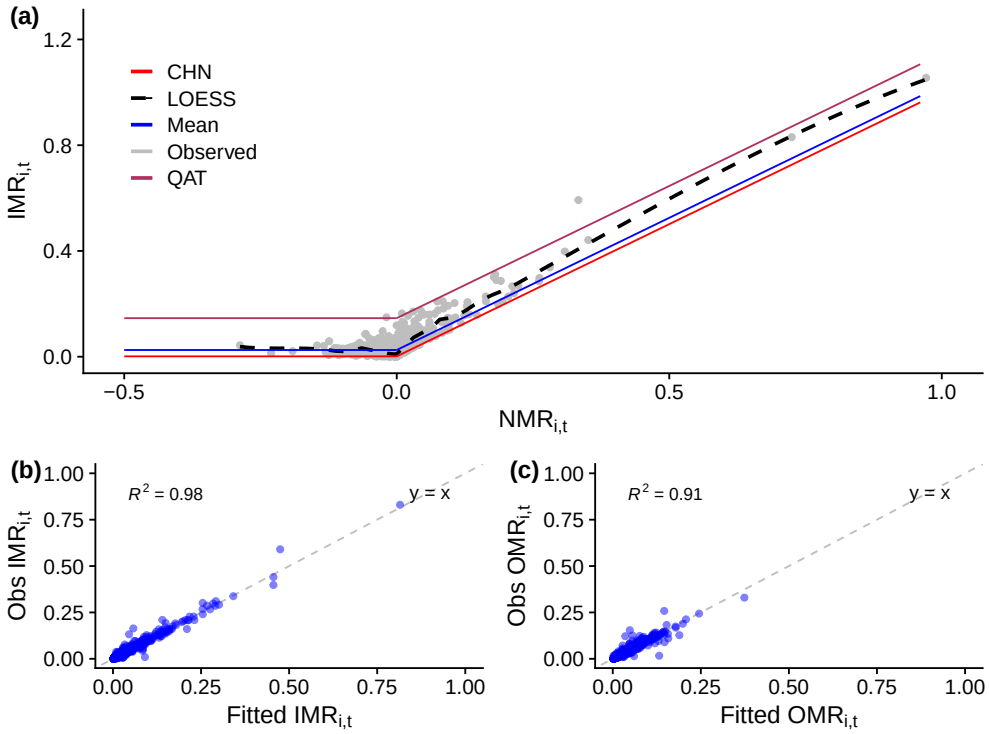


Figure 3.1: (a) Observed in-migration rates versus net migration rates for all countries (points) with mean model (blue), LOESS line (black dashed), and country-specific models for China (CHN, red) and Qatar (QAT, green); (b) Observed in-migration rates versus fitted in-migration rates compared to IMR Observed=Fitted (dashed line); (c) Observed out-migration rates versus fitted out-migration rates compared to OMR Observed=Fitted (dashed line) for 5-year-periods from 1990–2020.

Many observations in Figure 3.1(a) are highly concentrated around some values. Locally Estimated Scatterplot Smoothing (LOESS) is a non-parametric regression method used to reveal underlying scatterplot trends without assuming a specific global form for the relation-

ship (Cleveland, 1979). In this context, agreement between the LOESS line and the mean of model (3.1) indicates that the model recovers the general trend of the observed data. In the $\text{NMR}_{i,t} < 0$ portion of the domain, the similarity between the LOESS line and the $\beta_{0,i}$ values shows that the mean intercept term provides an accurate summary of the observed in-migration and net migration rates. Near $\text{NMR}_{i,t} = 0$, however, the disagreement between the model mean and LOESS line shows that many $\text{IMR}_{i,t}$ observations near $\text{NMR}_{i,t} = 0$ could be overstated with the model mean alone. The largest LOESS line departure occurs for large positive values of $\text{NMR}_{i,t}$. This is primarily due to the influence of large positive $\text{NMR}_{i,t}$ observations for a few countries.

Finally, Figure 3.1(a) shows two examples of country-specific average association between $\text{IMR}_{i,t}$ and $\text{NMR}_{i,t}$. The top line shows the model fit for Qatar (green), which experienced some of the highest migration rates in the world since 1990. The bottom line shows the model fit for China (red), a country with stable net out-migration over the same period. These two examples underscore the importance of a model capable of accommodating the significant variation observed between countries.

Figure 3.1(b) shows the fitted and observed $\text{IMR}_{i,t}$ for all countries from 1990 to 2020. The high level of agreement between fitted and observed in-migration rates with an $R^2 = 0.98$ is notable considering the simplicity of model (3.1). The agreement between fitted and observed out-migration rates in Figure 3.1(c) is lower with $R^2 = 0.91$, but is still strong. Taken together, Figures 3.1(b)-(c) show that our model-based decomposition of the net migration rate is reasonable for the periods where net migration, in-migration, and out-migration rates are all available. See Appendix A.2 for additional evaluations of model 3.1.

Figure 3.2 shows the observed net migration rate for 5-year periods starting in 1950 and running through 2020 on the original scales for the United States, Germany, Turkey, and Saudi Arabia. These countries were selected to demonstrate how our model-based approach to estimating in- and out-migration rates compares to direct estimates of these quantities when the net migration rate was large and positive for a sustained period (United States), was more dynamic (Germany), experienced an abrupt departure from a stable historic norm

(Turkey), or was generated in a Gulf Cooperation Council (GCC) member country (Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, United Arab Emirates). The in-migration and out-migration rates from 1990 to 2020 used to estimate our model are shown in the second and third columns (black dashed lines) along with the model-based in- and out-migration rate estimates for all 5-year periods (solid blue lines). Model-based estimates relating net migration rates to in- and out-migration rates are similar to the observed rates for 1990–2020. Estimated in-migration and outflow rates prior to the 1990–1995 period appear plausible for the periods where pseudo-Bayes estimates are available.

While the model-based decomposition is not perfect, Figure 3.2 shows that the mixed-effects net rate decomposition approach leads to plausible estimates of in-migration and out-migration rates. The correlation between the observed and predicted rates was over 0.97. Broad agreement between observed in-migration and out-migration rates for 1990–2020 periods suggests that net migration rates prior to 1990 can be similarly decomposed into in-migration and out-migration rates. See Appendix A.2 for all historic net rate decomposition plots.

Raymer et al. (2023) propose an alternative approach to decomposing net migration into in-migration and out-migration. This method uses a fixed constant to approximate the magnitude of net migration attributed to total in-migration and out-migration in proportion to a country's population. This method could also be used to decompose the net flow into total inflow and outflow in place of model (3.1). However, model (3.1) offers a systematic approach to specifying a quantity similar to the fixed constant used in Raymer et al. (2023), but one that is country-specific, which our analyses suggest is needed.

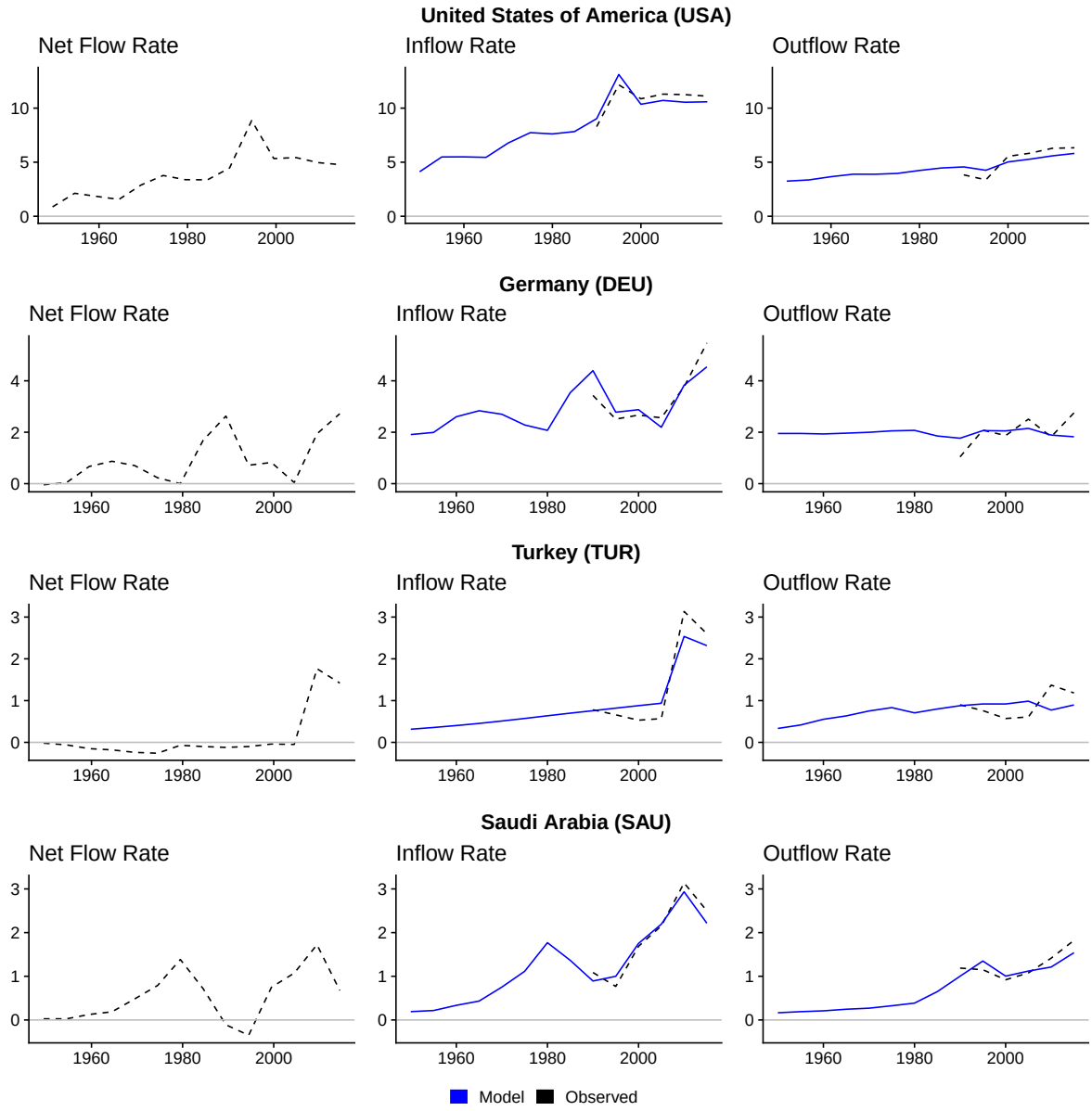


Figure 3.2: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people, compared to mixed-effects model estimates for the United States, Germany, Turkey, and Saudi Arabia. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

Age Standardization of In- and Out-migration Rates

The net migration rate, $\text{NMR}_{i,t}$, for country i over period t–$t+5$ is defined as the difference between the in-migration rate, $\text{IMR}_{i,t}$, and out-migration rate, $\text{OMR}_{i,t}$:

$$\text{NMR}_{i,t} = \text{IMR}_{i,t} - \text{OMR}_{i,t} = \frac{N_{i,t}}{\tilde{P}_{i,t,+,+}} = \frac{I_{i,t} - O_{i,t}}{\tilde{P}_{i,t,+,+}}. \tag{3.2}$$

Here, $I_{i,t}$ denotes the total inflow count, $O_{i,t}$ the total outflow count, and $N_{i,t} = I_{i,t} - O_{i,t}$ the total net flow count over the period starting in year t. We define the denominator of the net migration rate, $\tilde{P}_{i,t,+,+} = P_{i,t+5,+,+} - N_{i,t}$, as the population of country i at the end of the period starting in year t before factoring in the change due to net migration. This denominator specifies the population at risk of out-migration over the period.

The definition of the net migration rate shown in equation (3.2) departs from more conventional specifications, e.g., Preston et al. (2001) definition stated in terms of person-years. We use this net migration rate specification to estimate model parameters on the same scale as the *bayesPop* probabilistic population forecasting implementation (Ševčíková & Raftery, 2016b).

Historic net migration rate estimates are not disaggregated by age for most countries, but we aim to standardize migration rates to remove the effects of population age structure differences among countries in the same period and within countries across periods. Migration age patterns are known to be relatively consistent over time and country of origin (Rogers & Castro, 1981a), but, as Rogers (1990) discussed, the rate defined in equation (3.2) obscures the influence of the sending and receiving country population age structures. A population age-standardized net migration rate should account for the age structure of the sending populations in both components of the net rate as the migrant age distribution and totals are primarily linked to the origin population age structure.

Let $\pi_{i,t,a}$ denote the population in age group a as a proportion of the total population in country i at time t, namely

$$\pi_{i,t,a} = \frac{\tilde{P}_{i,t,a,+}}{\sum_a \tilde{P}_{i,t,a,+}}\,. \tag{3.3}$$

Since the origin population of inflows to country i consists of every country other than i in period t, we approximate the proportion of the global population in age group a by

$$\breve{\pi}_{t,a} \approx \frac{\tilde{P}_{+,t,a,+}}{\tilde{P}_{+,t,+,+}} \quad \text{where} \quad \tilde{P}_{+,t,a,+} = \sum_{i,s} \tilde{P}_{i,t,a,s} \quad \text{and} \quad \tilde{P}_{+,t,+,+} = \sum_{i,s,a} \tilde{P}_{i,t,a,s}\,. \tag{3.4}$$

Rogers & Castro (1981a) introduced the concept of the *Gross Migraproduction Rate* (GMR) for country i and period t, defined as $G_{i,t} = \sum_a \mathrm{OMR}_{i,t,a}$, where $\mathrm{OMR}_{i,t,a}$ represents the out-migration rate for country i over period t and age group a. This is a measure of overall migration that is not affected by population age distribution, similar to how the Total Fertility Rate (TFR) is a measure of overall fertility that is independent of the age distribution of women.

The values of $\mathrm{OMR}_{i,t,a}/G_{i,t}$, which give the age pattern of age-specific out-migration rates, tend to be stable over time and place, reflecting a tendency for international migration to be largely concentrated among people aged 15–35 and their dependent children, peaking in the twenties, as pointed out by Rogers & Castro (1981a). They also proposed a parametric model for this pattern, the famous Rogers-Castro curve. As an approximation, we thus consider the situation where this ratio is constant over time and space, so that $\mathrm{OMR}_{i,t,a}/G_{i,t} = R_a$ for all i, t, where $\sum_a R_a = 1$. Under this assumption, the quantity R_a remains constant across all countries and time periods, but it could be modeled using a Rogers-Castro curve or estimated empirically.

We then have an exact result for the out-migration rate for a reference population with a given age distribution:

Theorem 1. *Consider a population that has the same age-specific out-migration rates as country i in period t, but a different population age distribution given by $\pi^*_{i,t,a}$. Then this population has out-migration rate*

$$OMR^*_{i,t} = OMR_{i,t} \frac{\sum_a \pi^*_{i,t,a} OMR_{i,t,a}}{\sum_a \pi_{i,t,a} OMR_{i,t,a}}. \tag{3.5}$$

If both populations have the same age-specific pattern of migration rates, R_a, then

$$OMR^*_{i,t} = OMR_{i,t} \frac{C^*_{i,t}}{C_{i,t}}, \tag{3.6}$$

where $C_{i,t} = \sum_a \pi_{i,t,a} R_a$ is the migration age structure index (MASI).

The proof of this theorem can be found in Appendix A.1.

This indicates that one can age-standardize the out-migration rate to a given reference population using equation (3.6). This is a remarkably simple method, since it only involves multiplication by the MASI, $C_{i,t}$. The MASI involves only the population age distribution and the migration age-pattern, R_a, but not the age-specific migration rates themselves, which cancel. This is an important feature since past age-specific migration rates are not available for most countries.

We can age-standardize in-migration by viewing it as out-migration to country i from the rest of the world. We approximate the age distribution of the rest of the world by the age distribution of the world as a whole. This yields an age-standardized in-migration rate, $\text{IMR}^*_{i,t}$. Finally, we obtain the age-standardized net migration rate as $\text{NMR}^*_{i,t} = \text{IMR}^*_{i,t} - \text{OMR}^*_{i,t}$.

Typically the reference population age distribution will be the distribution in a particular year. We standardize to the age pattern of 2020, since this is the most recent census year for many countries. To calculate the age-standardized in- and out-migration rate for period t in terms of the 2020 population age structure, remove the population age structure effects

in period t, and scale the migration rate in terms of the 2020 reference population:

$$\begin{aligned} \text{IMR}^{\star}_{i,t} &= \text{IMR}_{i,t} \frac{\check{C}_{2020}}{\check{C}_t} \\ \text{OMR}^{\star}_{i,t} &= \text{OMR}_{i,t} \frac{C_{i,2020}}{C_{i,t}} \end{aligned} \tag{3.7}$$

where $\check{C}_t$ denotes the MASI for the world at time t.

This implies that the age-standardized net migration rate is given by $\text{NMR}^{\star}_{i,t} = \text{IMR}^{\star}_{i,t} - \text{OMR}^{\star}_{i,t}$. This measure places historic and forecast net migration rates in the context of the 2020 population age structure, removing variation in net migration rates attributable to differences in population age structure across periods. Additionally, this specification demonstrates how to convert between age-standardized rates and historic or future rates relative to the reference population.

After estimating historic in- and out-migration rates calculated from model (3.1), age-standardized in-migration, out-migration, and net migration rates can be calculated for forecasting. To estimate each model, we use total migration across both sexes, accounting for differences in migration propensities between males and females in the forecasting procedure, following the approach outlined in Azose et al. (2016b). While there is inherent uncertainty in past estimates of both net migration and bilateral flows, we do not address data quality differences across countries or unmeasured uncertainty in historical migration estimates.

Figure 3.3 summarizes the key steps to calculating the age-standardized net migration rate for one country. The top left portion of the diagram shows the mixed effects model (3.1) estimation using periods where inflow, outflow, and net flow estimates are all available (dashed black lines). Model-based estimates of historic in-migration and out-migration rates are then generated from this model (solid blue lines). The top right portion of the diagram shows how the MASI is constructed from a Rogers-Castro-like migration age schedule and the population age distribution at the start of each period. The bottom portion of the figure shows how age-standardized migration rates are calculated for base year 2020.

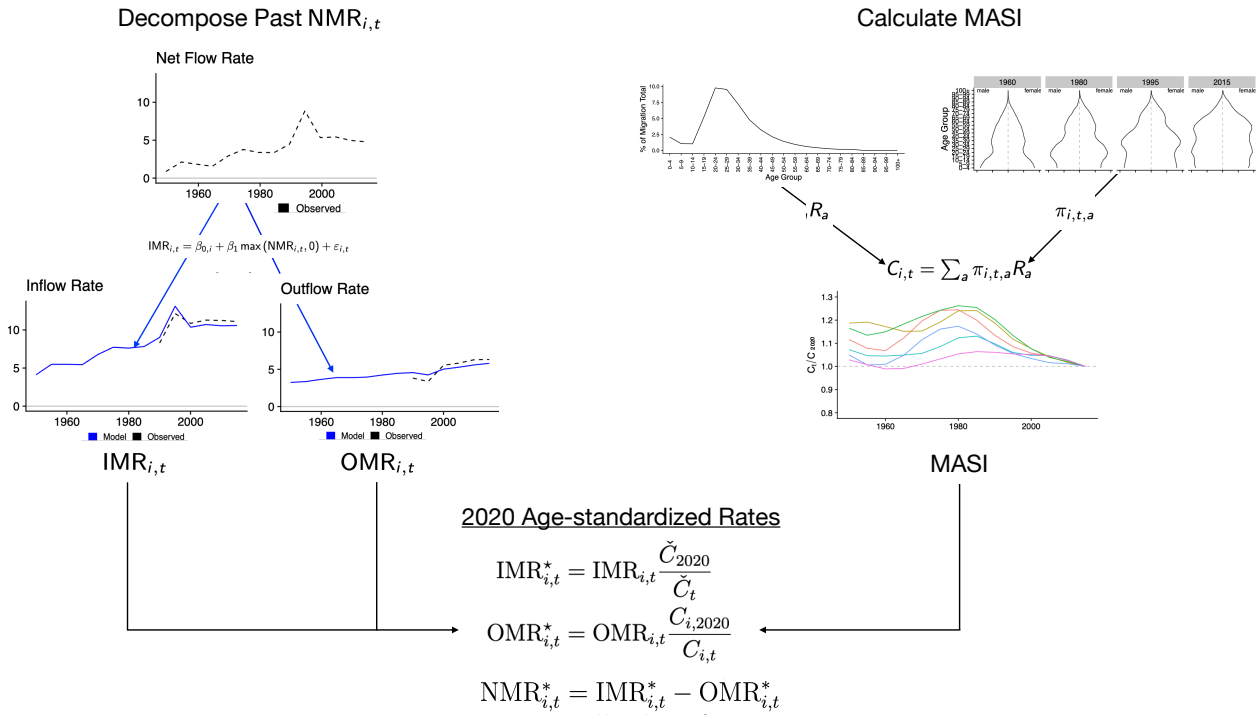


Figure 3.3: Age-standardized net migration rate process diagram showing the net migration rate decomposition method (top left), Migration Age Standardization Index (MASI) calculation using a model migration age schedule and population age distribution over time (top right) and 2020 base year age-standardized rate calculations (bottom). Migration flow estimates are shown at the midpoint of each 5-year period.

Net Migration Rate Model

We use the resulting rates to develop a probabilistic model for age-adjusted net migration rates. Using the above estimates of in-migration and out-migration rates, we compute the age-standardized net migration rates, $\text{NMR}^{\star}_{i,t}$, for all countries i and for t from 1950 through 2020, using the 2020 population age structures as the baseline population, as shown in Section 3.3. While any baseline year is valid, the most recent population estimate is used as the baseline. We then fit the Bayesian hierarchical model of Azose & Raftery (2015a). The top level of this Bayesian hierarchical model for the age-standardized net migration rate for

country i in time period t is as follows.

$$\textbf{Level 1} \begin{cases} \left(\text{NMR}^{\star}_{i,t} - \mu_i\right) = \phi_i \left(\text{NMR}^{\star}_{i,t-1} - \mu_i\right) + \xi_{i,t} \\ \xi_{i,t} \overset{\text{ind}}{\sim} \text{Normal}\left(0, \psi_i^2\right) \end{cases} \tag{3.8}$$

Table 3.2 defines each term in Level 1 of the net migration rate model.

Table 3.2: Notation for Level 1 of Azose & Raftery (2015a) net migration rate model.

$\text{NMR}^{\star}_{i,t}$	Age-standardized net migration rate for country i for period starting in year t
μ_i	Theoretical long-term average migration rate for country i
ϕ_i	Autoregressive parameter for age-standardized net migration rate for country i
$\xi_{i,t}$	Normally distributed random deviation term
ψ_i^2	Variance of random deviations for country i

Hyperparameter values for this model required no adjustment since the default specifications were broadly defined and our specification of the net rate is on the same scale as that of Azose & Raftery (2015a). The full model is estimated using a Markov chain Monte Carlo (MCMC) algorithm. This yields a sample (indexed by $j \in (1, \ldots, J)$) of model parameter vectors, $\left(\mu_i^{(j)}, \phi_i^{(j)}, \psi_i^{2(j)}\right)$, for each country $i \in (1, \ldots, 200)$.

Forecasting

The goal is to forecast migration and population probabilistically by jointly generating a set of future migration and population trajectories. Smith et al. (2001) outline the features distinguishing a forecast from a projection. A projection describes a future outcome based on specific assumptions without assigning probabilities of realizing these outcomes. A forecast, however, arises from a statistical method that includes model-based probabilities of realizing these outcomes.

We use the same approach to forecasting fertility and mortality as the WPP 2019 (United Nations. Department of Economic and Social Affairs, 2019). Our migration forecasting method builds on the approaches of Azose & Raftery (2015a) and Azose et al. (2016b), with

modifications to account for differences in population age distributions over time and across countries. Forecasts are generated through 2100 to compare the deterministic migration method from WPP 2019, the probabilistic age-agnostic migration method of Azose et al. (2016b), and our probabilistic age-standardized migration method.

For each future trajectory $j \in \{1, \ldots, J\}$, we independently project the population for each country i by 5-year age group a through $a+4$ and sex s for year $t+5$, assuming no migration. A total of 21 age groups are included with age bins [0,5), [5,9), ..., [95,100), [100, ∞). The resulting projection, $\tilde{P}^{(j)}_{i,t+5,a,s}$, denotes a realization of the total population that would be observed in year $t+5$ had no one in the population migrated in or out of each country. This approximates the population at risk of migrating for the period t–$t+5$, stratified by age and sex.

We then independently generate an age-standardized net migration rate for the period t–$t+5$, $\text{NMR}^{\star(j)}_{i,t}$, by sampling from the previously estimated Bayesian hierarchical model of Azose & Raftery (2015a). The sampled age-standardized net migration rate is subsequently decomposed into the age-standardized in-migration rate, $\text{IMR}^{\star(j)}_{i,t}$, and age-standardized out-migration rate, $\text{OMR}^{\star(j)}_{i,t}$, using the country-specific posterior predictive mean derived from model (3.1). Note that age-standardized coefficients for model (3.1), $\beta^{\star}_0$ and $\beta^{\star}_1$, are estimated from historical age-standardized rates ($\text{IMR}^{\star}$ and $\text{OMR}^{\star}$), ensuring that all rates are expressed on the same scale.

The mean age-standardized in-migration rate for trajectory j and country i is calculated as

$$\text{IMR}^{\star(j)}_{i,t} = \beta^{\star}_{0,i} + \beta^{\star}_1 \max\left(\text{NMR}^{\star(j)}_{i,t}, 0\right). \tag{3.9}$$

The corresponding out-migration rate for the period starting in year t is given by

$$\text{OMR}^{\star(j)}_{i,t} = \text{IMR}^{\star(j)}_{i,t} - \text{NMR}^{\star(j)}_{i,t}. \tag{3.10}$$

Age-standardized in-migration and out-migration rates then need to be converted back to period-specific rates to calculate inflow and outflow counts corresponding to the projected

population age structure:

$$\begin{aligned} \mathrm{IMR}_{i,t}^{(j)} &= \mathrm{IMR}_{i,t}^{\star(j)} \times \check{C}_t^{(j)} / \check{C}_{2020}, \\ \mathrm{OMR}_{i,t}^{(j)} &= \mathrm{OMR}_{i,t}^{\star(j)} \times C_{i,t}^{(j)} / C_{i,2020}. \end{aligned} \tag{3.11}$$

In-migration rates then are converted to a total inflow counts by multiplying the in-migration rate by the total country population for trajectory j at time t. These inflow totals are disaggregated by sex in proportion to the male and female population, following the approach of Azose et al. (2016b), and by age using a Rogers-Castro-like migration age schedule. This yields the total inflow to country i in period t–$t+5$ for age group a and sex s for trajectory j, denoted $I_{i,t,a,s}^{(j)}$. Out-migration counts by age and sex, $O_{i,t,a,s}^{(j)}$, are calculated similarly.

While a Rogers-Castro-like age schedule is suitable for out-migration in most countries, the unique migration patterns in GCC countries require a different approach. In recent decades, GCC migration has been dominated by large inflows of male migrant workers employed on temporary visas, typically lasting two years or less, with little opportunity for long-term residency or citizenship. As these workers depart and are replaced by new arrivals, the overall age structure of the population remains relatively stable despite changes in migration volume.

To reflect this dynamic, we adjusted the migration age schedule to maintain the foreign-worker-dominated age structure in GCC countries. Rather than applying the standard Rogers-Castro-like model, we modified the out-migration schedule to emphasize outflows of older working-age migrants, aligning with the temporary nature of these populations. Specifically, the out-migration age schedule was derived as the difference between the age distribution of the population and the normalized Rogers-Castro-like schedule, effectively re-weighting outflows to reflect older workers leaving these countries.

This adjustment ensures that GCC migration patterns preserve the prime working-age structure (15–65) of the migrant population, preventing the aging-in-place of large migrant-worker cohorts. While out-migration rates have historically been much smaller than in-

migration rates, these modifications better capture the dynamics of GCC migration and yield more plausible forecasts of population and migration trends through 2100 compared to using the approach used for all other countries.

Rebalancing global inflows and outflows each period was necessary to maintain global net-zero migration. While Azose & Raftery (2015a) and Azose et al. (2016b) proposed methods for this, their approaches do not directly apply to forecasts that generate inflow and outflow counts for each country. Adjusting global net migration counts without specifying the proportion allocated to inflows or outflows creates inconsistencies in the inflow and outflow trajectories relative to the global net flow. To address this, we resolved global net-zero deviations by adjusting country-level inflow and outflow totals in proportion to each country's population share of the global total.

The global rebalancing procedure is as follows. Let $\tilde{I}^{(j)}_{i,t,a,s}$ and $\tilde{N}^{(j)}_{i,t,a,s}$ denote the j^{th} inflow count and net migration count trajectories before global rebalancing, respectively, for country i, period t–$t+5$, age group a–$a+4$, and sex s. Then the adjusted inflow and outflow counts were calculated as

$$
\begin{aligned}
I^{(j)}_{i,t,a,s} &= \tilde{I}^{(j)}_{i,t,a,s} - w \left(\sum_{k \in K} \tilde{N}^{(j)}_{k,t,a,s} \right) \frac{\tilde{P}^{(j)}_{i,t,a,s}}{\sum_{k \in K} \tilde{P}^{(j)}_{k,t,a,s}}, \\
O^{(j)}_{i,t,a,s} &= \tilde{O}^{(j)}_{i,t,a,s} + (1-w) \left(\sum_{k \in K} \tilde{N}^{(j)}_{k,t,a,s} \right) \frac{\tilde{P}^{(j)}_{i,t,a,s}}{\sum_{k \in K} \tilde{P}^{(j)}_{k,t,a,s}},
\end{aligned}
\tag{3.12}
$$

where $w = 0.5$, evenly splitting the global net-zero deviations between inflows and outflows. The index K represents two normalization groups: GCC countries and their labor-supplying origins (Bangladesh, India, Indonesia, Philippines, Pakistan), and all other countries outside the GCC labor corridor.

This adjustment ensures global net migration for each age group and sex equals zero, with net migration for trajectory j in country i during t–$t+5$ given by $N^{(j)}_{i,t,a,s} = I^{(j)}_{i,t,a,s} - O^{(j)}_{i,t,a,s}$. As in Azose & Raftery (2015a), these adjustments led to minimal changes to the raw forecasts.

We then recalculated the balanced age-standardized net migration rate, $\text{NMR}_{i,t}^{\star(j)}$, as

$$\text{NMR}_{i,t}^{\star(j)} = \left(\frac{I_{i,t,+,+}^{(j)}}{\tilde{P}_{i,t+5,+,+}^{(j)}} \right) \times \check{C}_{2020}/\check{C}_{t}^{(j)} - \left(\frac{O_{i,t,+,+}^{(j)}}{\tilde{P}_{i,t+5,+,+}^{(j)}} \right) \times C_{i,2020}/C_{i,t}^{(j)}, \tag{3.13}$$

which will be the rate jump-off for sampling in the next time period. Finally, we account for net migration in the population projection generated at the start of the forecasting routine,

$$P_{i,t+5,a,s}^{(j)} = \tilde{P}_{i,t+5,a,s}^{(j)} + N_{i,t,a,s}^{(j)}, \tag{3.14}$$

before moving to the next forecast period.

3.4 Validation

Age-standardized forecast efficacy was evaluated in terms of out-of-sample performance. Figure 3.4 shows the observed and four-period migration and population forecasts for the United States, El Salvador, South Africa, and Saudi Arabia starting in the 2000-2005 period through the 2015-2020 period. Forecasting methods include 1995-2000 migration rate persistence (effectively the current UN method) for every period, the age-agnostic method of Azose et al. (2016b), and the age-standardized method described above.

Rows of the figure show the migration age structure index (MASI) ratio, i.e. the ratio of the MASI for the period to the MASI for the baseline year 2000, net migration rate, and population for each country, while columns correspond to the four countries. Prediction intervals for both the age-agnostic and age-standardized method included the observed population and net migration rates in all cases. Population and migration forecasts using the age-standardized approach were as good or better than the age-agnostic forecasts by the end of the period for all cases except in El Salvador in terms of agreement with the observed quantities. Persistence of 1995-2000 migration rates performed especially poorly for El Salvador and Saudi Arabia.

The selected countries illustrate the differences in migration and population forecasts with

and without age-standardization. The United States was included for its scale of migration, El Salvador for its rapidly aging population and high out-migration rate, South Africa for the relatively minor contribution migration makes to its population dynamics compared to fertility and mortality, and Saudi Arabia as a representative GCC country.

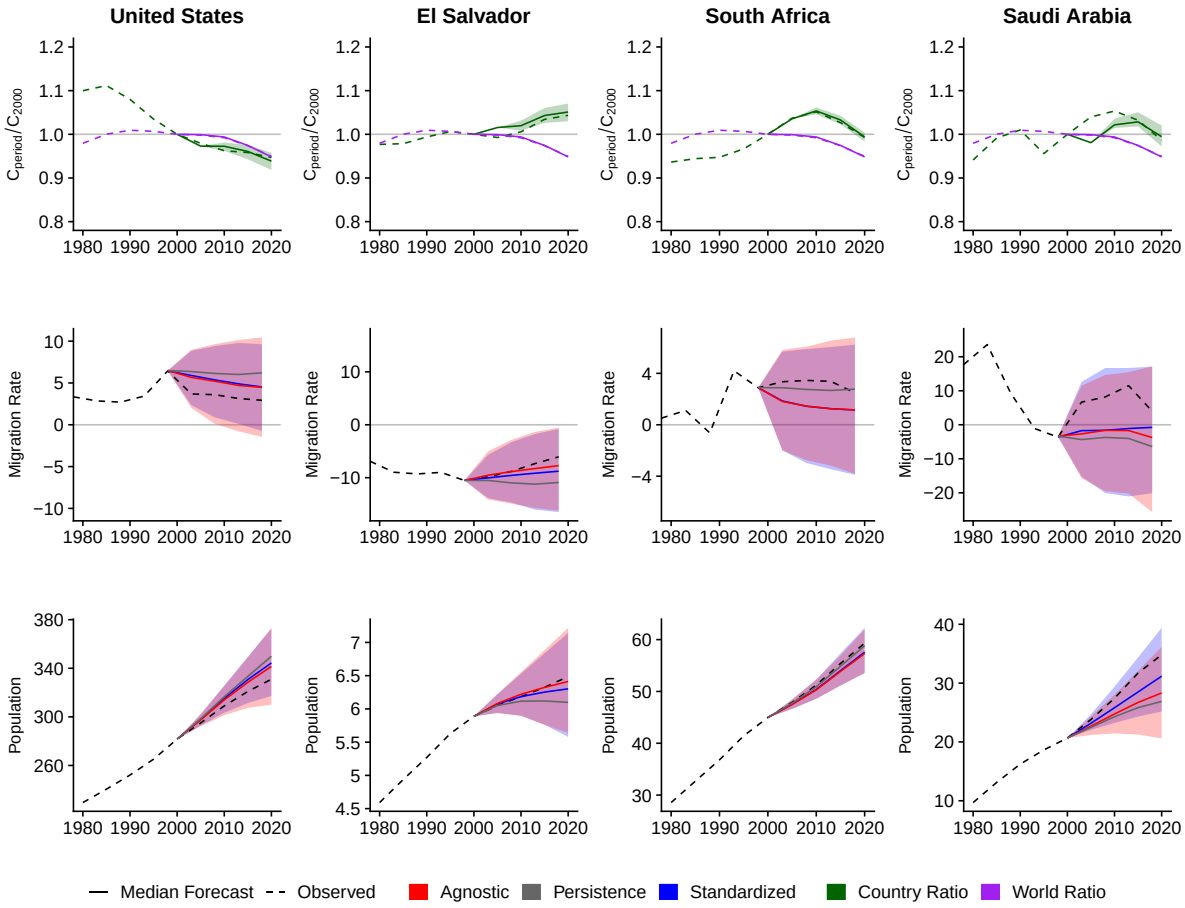


Figure 3.4: Migration age structure index (MASI) ratio for each country (■) and the globe (■) with base-year 2000, out-of-sample validation forecast of population (millions of people), and age-standardized and age-agnostic net migration rate (net annual migrants per thousand), for four countries. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), observed fertility, and observed mortality. Dashed lines in each plot indicate the observed values. Solid lines indicate the median forecast. Shaded regions show the 80% prediction interval. Migration models were fit to 1950–2000 data. Forecasts are for the 2000–2005, 2005–2010, 2010–2015 and 2015–2020 periods. Migration flow estimates and forecasts are shown at the midpoint of each 5-year period.

Age-standardized and age-agnostic forecasts largely agree with one another since population age structure differences make little impact over such a short period of time. Note that the proportion of the migration age population increased relative to the baseline population in El Salvador. The population aged 0–15 was the largest proportion of the population in 2000; however, the proportion of the population aged 0–15 fell during the validation period and the large 0–15 population cohort aged into the prime migration age cohorts by the end of the 2015–2020 period. As a result, El Salvador's age structure ratio increased from the 2000 baseline period. The age structure ratio forecast outpaced the observed age structure ratio for the first two periods (2000–2010), but the prediction interval contained the observed age structure ratio for the last two periods (2010–2020).

Global age ratio forecasts closely matched the true global age structure, but country-level population age ratio forecasts were mixed. Prediction intervals for the MASI ratio captured all or most of the observed values for the United States and South Africa. For Saudi Arabia, the forecast MASI ratio was lower than the true value for the first two periods but included the observed value by the end of the forecast period. In El Salvador, the forecast MASI ratio consistently exceeded observed values from 2000 to 2010. These discrepancies are explained by the higher net migration observations compared to the forecasts for those countries.

Global and country-level MASI ratio values equal one for the baseline population structure (2000 in this case) by definition. When the global MASI ratio exceeds the country-level MASI ratio, then we expect higher in-migration using the age-standardized method compared to the age-agnostic method. Conversely, when the country-level MASI ratio is below 1, we expect lower net out-migration compared to the baseline period using the age-standardized method and a model that does not account for population age structure.

The United States MASI ratio remained below both the global MASI ratio and baseline MASI ratio throughout the forecast period. This combination should correspond to less out-migration from the United States as the global population had a net positive supply of migration-age people and the United States had a shrinking net positive supply of migration age people compared to the baseline population. Indeed, the age-standardized net migration

forecast for the United States was shifted towards higher net in-migration compared to the age-agnostic forecast.

The forecast MASI ratios for El Salvador and South Africa exceeded both the global MASI ratio and their respective baseline ratios, indicating that the age-standardized migration forecast should be shifted towards more net out-migration compared to the age-agnostic forecast. The age-standardized net migration rate was indeed more negative than the age-agnostic migration rate for El Salvador. However, age-standardized and age-agnostic net migration forecasts were indistinguishable for South Africa, reflecting the relatively small share of population change attributable to migration there.

The forecast MASI ratio for Saudi Arabia exceeded the global MASI ratio and the baseline ratio, indicating that the age-standardized migration forecast should be shifted towards more net out-migration compared to the age-agnostic forecast if the forecasting methods were otherwise identical. However, the age-standardized approach for GCC countries uses a distinct methodology from all other countries, making direct comparisons to the age-agnostic approach less relevant. Still, net migration forecasts generated from the age-standardized method were more similar to the observed rates. Since migration is the main contributor to population change in GCC countries, the small improvement in the age-standardized forecast led to a substantial improvement in the population forecast compared to the age-agnostic approach.

Multiple forecast horizons were evaluated, generating predictions one to four periods ahead of the last observed data. Out-of-sample forecasts used the last observed population age structure as the baseline, fit each model to data available prior to the first forecast period, and generated forecasts for each period through the 2015-2020 period. One-period-ahead forecasts were generated for 2000, 2005, 2010, and 2015; two-period-ahead forecasts for 2000, 2005, and 2010; three-period-ahead forecasts for 2000 and 2005; and the four-period-ahead forecast for 2000 only. Five-period-ahead forecasts were not evaluated as it is not possible to fit the mixed-effects model (3.1) with only one period of in-migration and out-migration rate data.

We evaluated the median of the J country-specific forecasts in terms of the Mean Absolute Error (MAE), the Log Mean Absolute Error (LMAE), and the Mean Absolute Scaled Error (MASE). Let $f_{i,t}$ denote the net migration rate forecast for country i and period starting in year t, and $r_{i,t}$ the true net migration rate. The LMAE is defined as follows:

$$l(y) = \text{sign}(y)\left[\log\left(|y| + c\right) - \log c\right] \text{ with } c > 0$$

$$\text{sign}(y) = \begin{cases} 1 & y > 0 \\ 0 & y = 0 \\ -1 & y < 0 \end{cases}$$

$$LMAE_t = \frac{1}{200}\sum_{i=1}^{200}\left|l(f_{i,t}) - l(r_{i,t})\right|. \tag{3.15}$$

We used $c = 1$. The LMAE formulation prevents large errors in a few forecasts from dominating the error metric.

The MASE, recommended by Hyndman & Koehler (2006), evaluates forecast performance in terms of the mean ratio of errors from a proposed method (numerator) and errors from a naïve method (denominator), such as persistence forecast. Denominator errors are calculated from in-sample data, e.g. mean net migration rate errors generated from the 1950-2000 data using the naïve method, and compared to errors calculated from out-of-sample forecasts from a proposed method, e.g. mean net migration rate forecast errors in the 2000-2020 data.

The MASE for k-period-ahead forecasts, based on T_k forecast periods starting with period t_0–$t_0 + 5$ and S_k in-sample periods starting in period s_0–$s_0 + 5$, is defined as

$$\text{MASE}_k = \frac{\frac{1}{200\times T_k}\sum_{i=1}^{200}\sum_{\{t_0\}_k}\left|r_{i,t_0+5(k-1)} - f_{i,t_0+5(k-1)}\right|}{\frac{1}{200\times S_k}\sum_{i=1}^{200}\sum_{\{s_0\}_k}\left|r_{i,s_0+5k} - r_{i,s_0}\right|}. \tag{3.16}$$

This formula computes the ratio of average k-period ahead forecast errors for all possible horizons k in the in-sample and out-of-sample periods. For the $k = 1$ (one-period-ahead forecasts), $\{t_0\}_1 = \{2000, 2005, 2010, 2015\}$, $T_1 = 4$, $\{s_0\}_1 = \{1950, 1955, \ldots, 1990\}$, and

$S_1 = 9$. For the $k = 4$ four-period-ahead forecast, $\{t_0\}_4 = \{2000\}$, $T_4 = 1$, $\{s_0\}_4 = \{1950, 1955, \ldots, 1975\}$, and $S_4 = 6$.

Forecast calibration was evaluated in terms of 95% prediction interval coverage and average prediction half-interval width. A well-calibrated model's 95% prediction intervals should include approximately 95% of the true net migration rate observations. Intervals containing fewer than 95% of true values are too narrow, underestimating forecast variation, while those containing more than 95% are too wide, overestimating variation. The prediction half-interval width is half the difference between the upper and lower prediction interval quantiles. Models that are accurate, well-calibrated, and have narrower half-interval widths are preferred over those with wider intervals but similar accuracy and calibration.

Table 3.3: Mean predictive performance of different methods for net migration rate (migrants per thousand period person years): mean absolute error (MAE), log mean absolute error (LMAE) with $c = 1$, mean absolute scaled error (MASE), 95% prediction interval coverage, and mean prediction interval half-width (HW).

	Method	MAE	LMAE	MASE	Cover	HW
5 Years	Persistence	4.02	0.68	1.42	—	—
	Agnostic	**3.44**	**0.65**	**1.25**	**93**	10.47
	Standardized	3.54	**0.65**	1.28	**93**	**10.39**
10 Years	Persistence	5.33	0.88	1.63	—	—
	Agnostic	**3.86**	**0.76**	**1.16**	**91**	11.63
	Standardized	3.93	0.77	1.17	90	**11.44**
15 Years	Persistence	5.16	1.00	1.44	—	—
	Agnostic	**3.49**	0.83	**1.10**	**92**	12.32
	Standardized	3.51	**0.82**	1.11	**92**	**11.85**
20 Years	Persistence	4.77	1.02	1.27	—	—
	Agnostic	2.91	0.82	1.00	**94**	12.54
	Standardized	**2.86**	**0.79**	**0.98**	**94**	**12.04**

Table 3.3 summarizes out-of-sample predictive performance of the age-standardized, age-

agnostic, and net rate persistence methods. Errors are calculated as the difference between the observed value and the median of 2,000 posterior predictive distribution draws. Rate persistence uses the last observed net rate in each country as the forecast. The best score for each method is shown in bold font.

The age-standardized and the age-agnostic model accuracy were similar across all horizons, but age-standardized forecast accuracy overtook age-agnostic forecast accuracy as the number of periods between the last observations and the forecast period increased for all metrics. The 95% prediction interval coverages were effectively indistinguishable across all forecast horizons, except the two-period-ahead forecast. Both methods led to slight under-coverage compared to the 95% nominal value. However, the age-standardized prediction interval half-widths were narrower for every forecast horizon. The similarity between the age-agnostic and age-standardized models over short horizons is expected, as population age structure changes gradually. Both probabilistic models outperformed the rate persistence model in all horizons.

The age-agnostic and net rate persistence models were fit using the *bayesPop* R package (Ševčíková & Raftery, 2016b), while the age-standardized model was implemented using a custom implementation of the *bayesPop* R package. See Appendix A.5 for software age-standardized model implementation details.

3.5 Results

Figure 3.5 summarizes probabilistic population forecasts in four countries using the age-agnostic migration model from Azose et al. (2016b), the population forecasts using the age-standardized migration model, and the projections from the WPP 2019 (United Nations. Department of Economic and Social Affairs, 2019). Columns of the figure correspond to the same country. Rows of the figure show forecasts through 2100 for the MASI ratio, the net migration rate, and total population.

The MASI ratios in the first row of Figure 3.5 summarize the differences in the population age structure from the 2020 baseline for each country (green) compared to the population

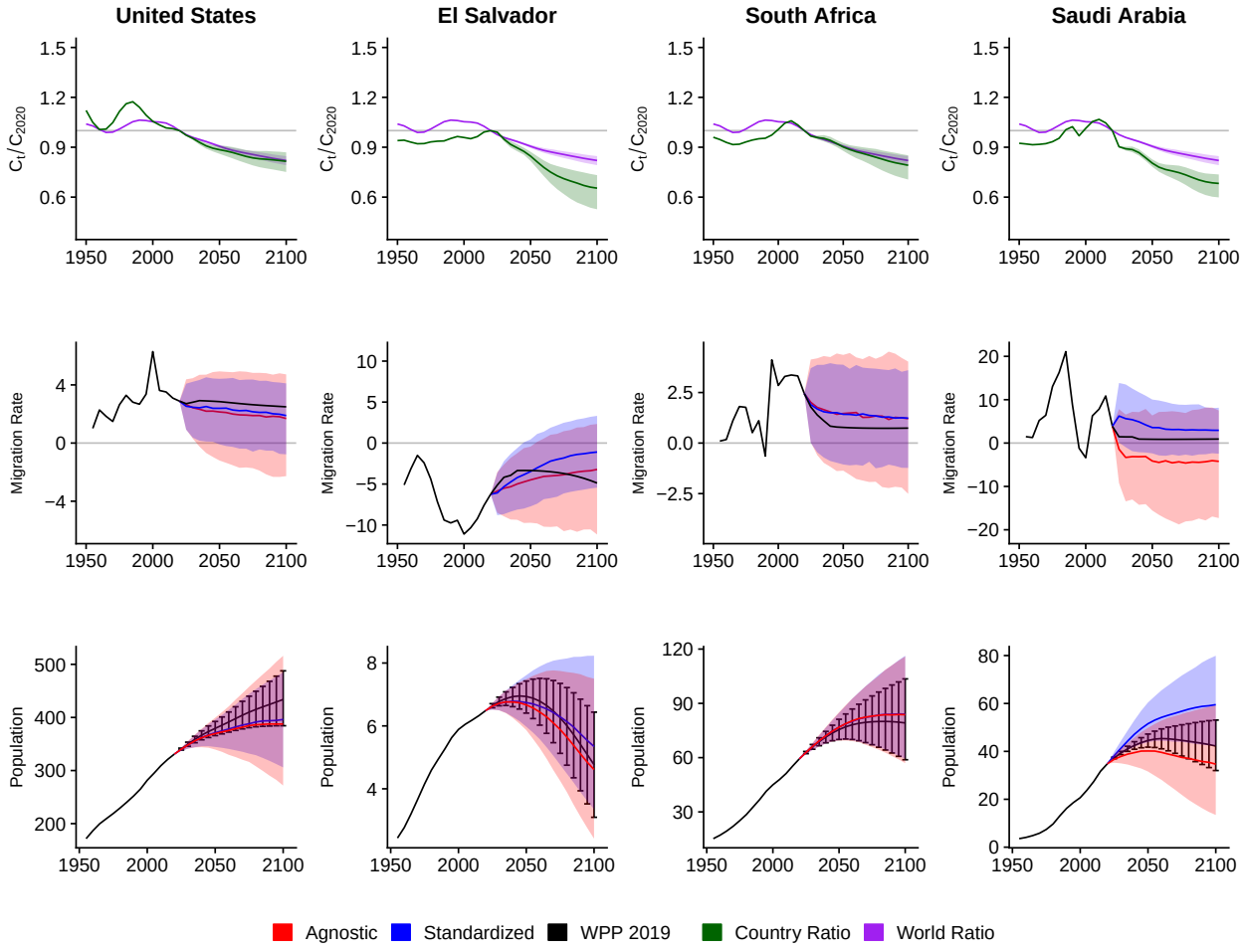


Figure 3.5: 2020 MASI ratios (top row), net migration rate as net annual migrants per thousand (middle row), and probabilistic forecast of population (in millions) age-standardized and age-agnostic (bottom row) for the United States, El Salvador, South Africa, and Saudi Arabia. Forecasts use probabilistic net migration (■=age-standardized and ■=age-agnostic), as well as probabilistic fertility, and mortality. Migration flow estimates and forecasts are shown at the midpoint of each 5-year period. The MASI ratio plots show the values for each country (■) and the world (■). Solid lines in each plot indicate the observed and median forecast. Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

age structure for the world (purple). Countries with a higher migration-age population than the 2020 population baseline rise above 1. Countries aging faster than the world population fall below the global index. Forecast population age structures in the United States and

South Africa generally follow the global aging trend.

When the country-level MASI ratio is similar to the global ratio, the median net migration forecasts from the age-standardized model and age-agnostic model should be more similar compared to countries where the country-level and global MASI ratios diverge. The population age structures of El Salvador and Saudi Arabia are forecast to age much faster than the global average. In these cases, the age adjustment shifts the migration forecast towards higher net inflows.

Median net migration forecasts from both probabilistic net migration models are similar in many countries, but long-term age-standardized net migration forecast intervals (shaded blue regions) tend to be narrower than age-agnostic (shaded red regions) intervals. This is due to the population age normalization step used in the age-standardized method, which removes variation in historic net migration rates caused by differences in population age structure before fitting the Azose & Raftery (2015a) model. Removing variation in past net migration estimates attributable to population age structure reduces the variance in the net migration model parameter estimates and hence the migration and population forecasts. Prediction intervals (80%) for both probabilistic net migration models include the United Nations' WPP 2019 projections through 2100 for all countries shown in Figure 3.5.

Median net migration forecasts from the age-agnostic and age-standardized methods are most similar in countries where population age indices align closely with the global index. The United States and South Africa demonstrate this trend in Figure 3.5.

In countries with population age structures that diverge from the global norm, such as El Salvador, the age-standardized method should produce markedly different forecasts. El Salvador's population is projected to age more rapidly than the global average from 2020 to 2100, resulting in a median age-standardized forecast with less net out-migration compared to the age-agnostic model. This adjustment yields a higher population forecast for the age-standardized model. Additionally, the age-standardized 80% prediction interval shows greater uncertainty about whether El Salvador's population will peak by 2100, with a higher upper bound compared to the age-agnostic forecast.

Figure 3.5 also shows that Saudi Arabia's population age structure is forecast to change about as fast as El Salvador's compared to the 2020 baseline population age structure. Saudi Arabia's net migration forecast is also shifted towards higher net in-migration, but this leads to a higher net positive migration forecast compared to the age-agnostic model. The age-standardized net positive migration forecast is better aligned to the historic net rate data in Saudi Arabia and the UN's WPP 2019 net migration forecast.

Differences in Saudi Arabia's migration and population forecast using the age-standardized approach reflect both population age structure effects and adjustments to modeling unique features of migration in GCC countries. These adjustments yield a substantially higher median population forecast and prediction interval. The UN WPP 2019 population forecast falls to the bottom of the age-standardized prediction interval, compared to its position near the upper middle of the age-agnostic model interval. Detailed forecasts for all countries are provided in Appendix A.6.

Figure 3.6 presents past and median forecast MASI ratios for the globe and selected countries by UN region. The MASI ratio provides a concise summary of historical and forecast population age structure dynamics. The MASI ratio contextualizes a country's structural migration potential relative to the reference population and global trends by weighting age groups using a Rogers-Castro-like migration age schedule. As a result, the MASI ratio provides a more relevant measure of migration-related population age dynamics than metrics like average age, which are less directly tied to structural migration potential, or summaries like population pyramids, which are harder to interpret across periods and countries.

Figure 3.6 shows that the median population age structure forecasts in Western countries align with the global aging trend. This indicates that the that the underlying force of out-migration in Western countries will closely match the global average relative to 2020, leading to similar net migration forecasts from the age-standardized and age-agnostic methods.

In contrast, the migration force from Latin America is projected to decline sharply in the coming decades compared to the 2020 baseline and the global trend, resulting in higher net migration forecasts under the age-standardized method. For some countries, such as Nigeria

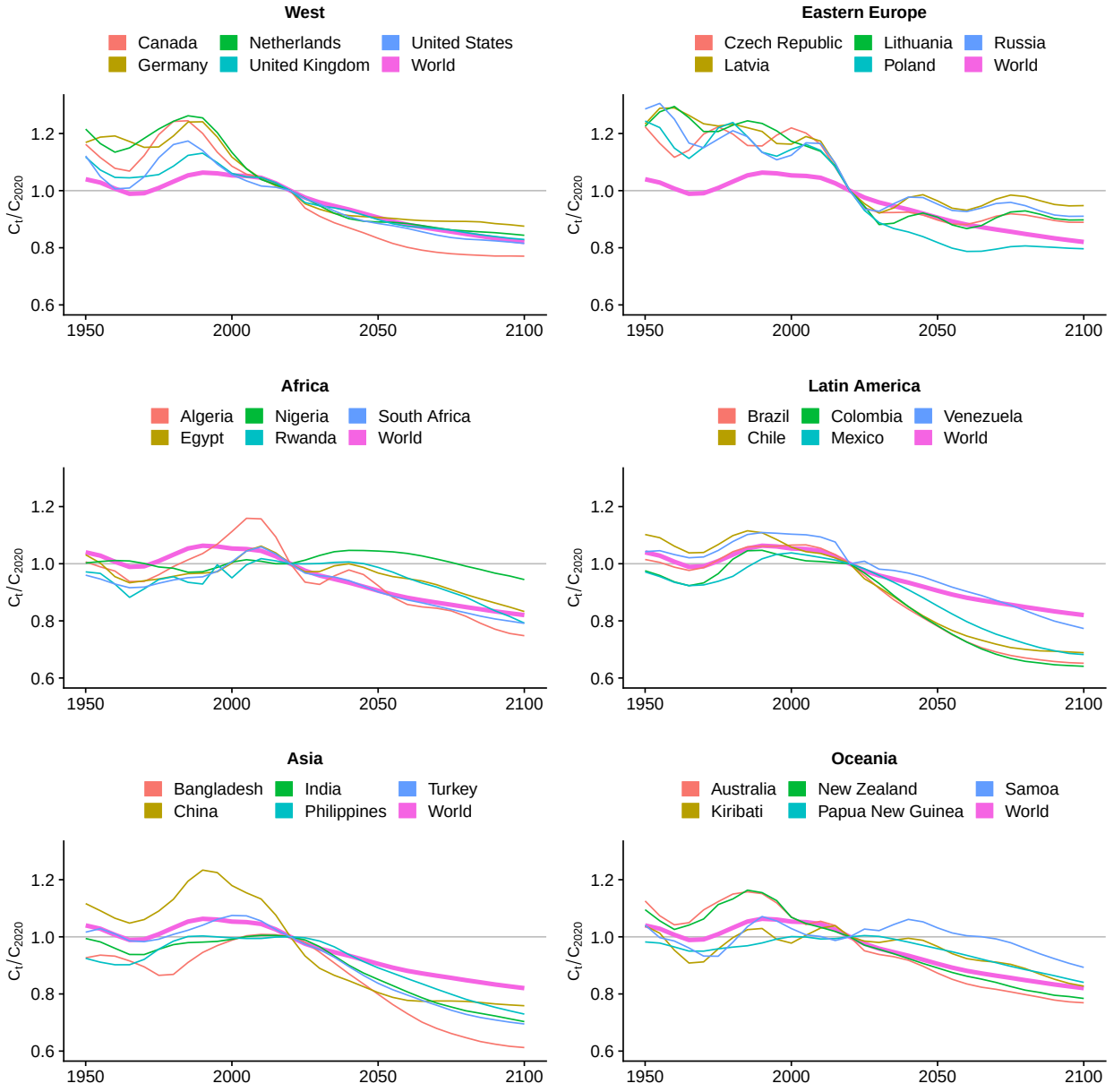


Figure 3.6: Historic and median country-level forecasts of MASI ratios for 2020 baseline by region compared to the world.

(Africa) and Samoa (Oceania), the population age index has yet to peak, indicating that out-migration forecasts from these countries are underestimated in the age-agnostic method.

Age-standardized migration forecasts result in less severe population declines in countries facing significant demographic challenges. Figure 3.7 highlights the age index, net migration, and population forecasts for large countries expected to experience drastic population contractions by 2100.

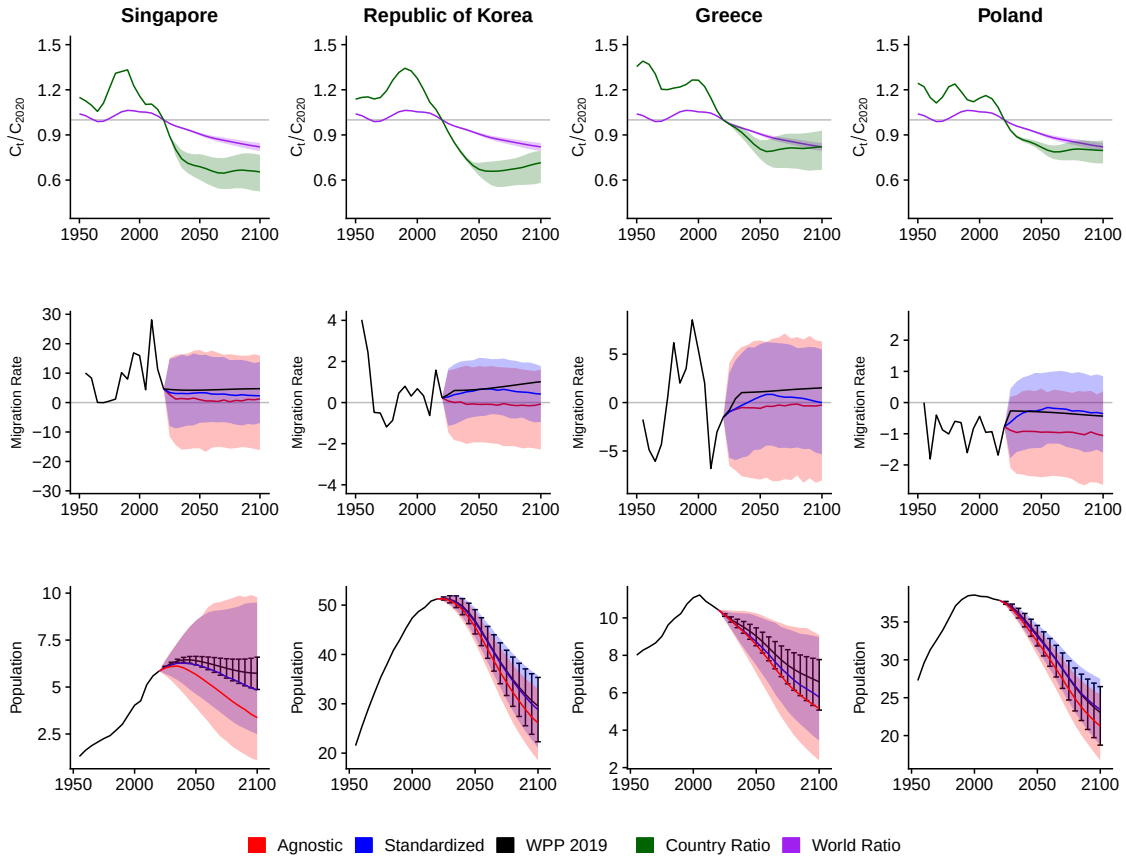


Figure 3.7: 2020 base-year MASI ratios (top row), age-standardized and age-agnostic net migration rate as net annual migrants per thousand (middle row), and probabilistic forecast of population (in millions) age-standardized and age-agnostic (bottom row) for Singapore, South Korea, Greece, Poland. Forecasts use probabilistic net migration (■=age-standardized and ■=age-agnostic), fertility, and mortality. Migration flow estimates and forecasts are shown at the midpoint of each 5-year period. Age-index ratio plots show the age structure ratios for each country (■) and world (■). Solid lines in each plot indicate the observed and median forecast. Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

The age-standardized forecast predicts less steep population declines compared to the age-agnostic model, as aging populations have fewer individuals of prime migration age. The age index indicates that the force of out-migration decreases as these countries age relative to the global population, shifting net migration forecasts toward higher in-migration. Median age-standardized migration forecasts align more closely with the UN's WPP 2019 migration and population projections than the age-agnostic forecasts.

Figure 3.8 shows the population forecast differences by 2100 for each country across three different migration models: age-agnostic and age-standardized probabilistic forecasts and the deterministic method from WPP 2019. Each point shows the deviation of the median age-agnostic or age-standardized forecast from the median WPP 2019 projection, allowing for simultaneous comparison of all three models.

Points near the origin indicate close agreement among all three methods. Points along the horizontal dashed line indicate agreement between median age-agnostic forecast and WPP 2019 that differ from the median age-standardized forecast. Points on vertical dashed line indicate agreement between the median age-standardized forecast and WPP 2019 that differ from the median age-agnostic forecast. Points on diagonal dashed line, far from the origin, indicate agreement between the median age-agnostic and age-standardized forecasts that differ from WPP 2019. Regions bounded by the dashed lines, labeled I–VI, summarize other forecast combinations.

Figure 3.8 shows that all three methods differ in Western Sahara (ESH). For the United Arab Emirates (ARE) and all GCC countries, the median age-agnostic and age-standardized methods forecasts are larger than WPP 2019, with the median age-standardized forecast being the highest.

The age-standardized method generates fewer negative net migration trajectories than the age-agnostic method, aligning with three decades of sustained positive net migration dominated by foreign workers supporting rapidly expanding economies. Net migration rate forecasts under the age-standardized approach maintain this positive trend, driving population growth in GCC countries.

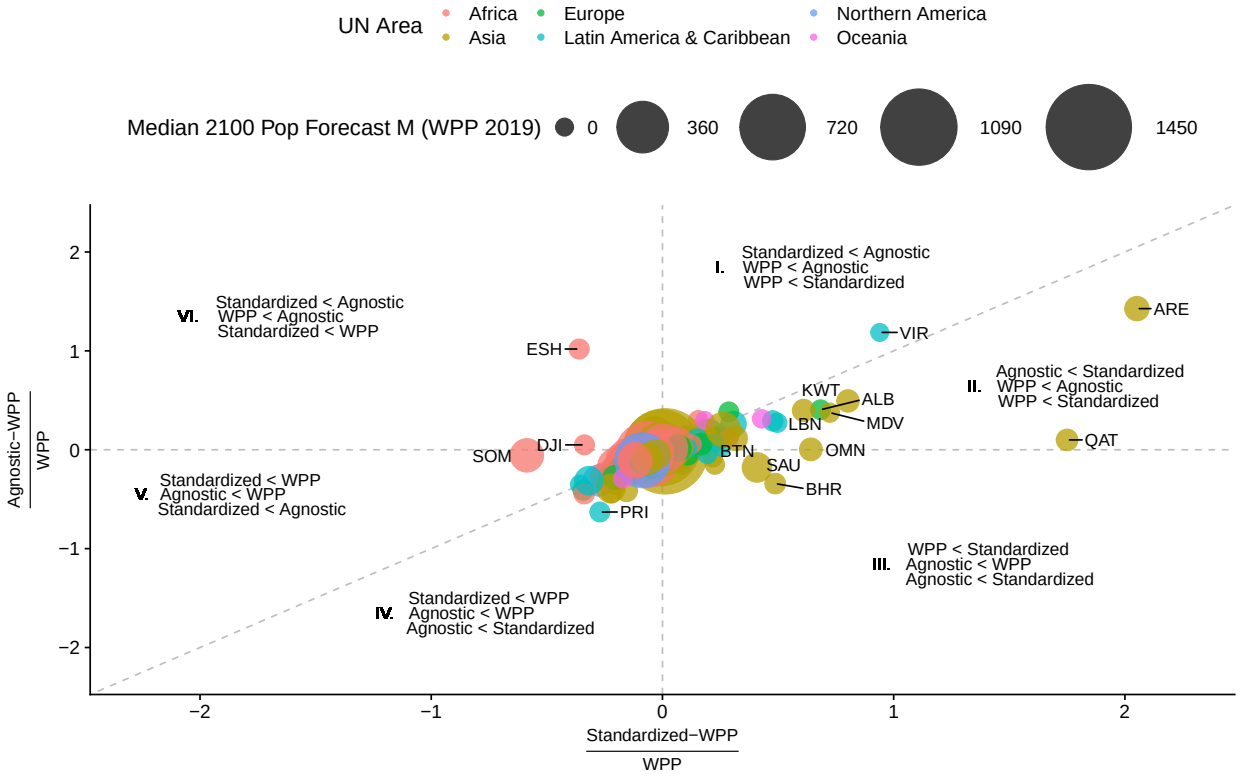


Figure 3.8: Three-way comparison of median 2100 population forecasts using age-standardized migration model, age-agnostic migration model, and median WPP 2019 forecast. Point color indicates UN Area, point diameter indicates WPP 2019 median population forecast size (millions), and regions I-VI show direction of forecast differences using the age-agnostic and age-standardized migration models.

While the age-standardized forecasts are higher than the age-agnostic forecasts, the resulting population age structures in most GCC countries are more plausible and closely align with WPP 2019 projections.

See Appendix A.4 for forecast quantile documentation and Appendix A.6 for forecast summaries of all countries.

3.6 Discussion

The link between migration and age is well established, yet few forecasting methods explicitly account for overall population age structure. We propose a statistical approach that incorporates population age structure changes into net migration forecasts for all countries, building on and addressing key limitations of Raftery & Ševčíková (2023).

Our approach addresses several criticisms of using net migration as a unit of analysis. By decomposing net flow rates into in-migration and out-migration rates, our model alleviates the migration age schedule issues identified by Rogers (1990), who noted that small net migration counts can obscure large, offsetting inflows and outflows. While net migration remains the unit of analysis, we use historic country-specific data to estimate in- and out-migration rates.

A bilateral migration flow model among all countries as in Welch & Raftery (2022) eliminates theoretical challenges induced by the use of net migration rate as the unit of analysis. However, the computational complexity of migration flow forecasting and limited historic bilateral flow data reduce the viability of this alternative for longer-term forecasting.

Probabilistic forecasts provide a systematic approach to incorporating uncertainty, but unexpected shocks to long-term trends can still arise without warning. Over the last decade alone, events such as the COVID-19 pandemic, Syria's civil war, Russia's invasion of Ukraine, and Venezuela's economic collapse have disrupted long-term norms. While new demographic patterns may emerge from these crises, their full effects should be integrated into updated forecasts as data become available.

Our proposed methods address several limitations, though others remain for future research. Using the global population age structure to normalize in-migration rates could be refined to better reflect the age structures of the most important origins. We adopted the global average for analytical tractability and to account for the potential evolution of migration corridors. Out-of-sample validation supports the reasonableness of this approximation.

We do not account for uncertainty in historic migration flow estimates used to calculate

age-standardized net migration rates. The posterior predictive mean from model (3.1) was used to decompose net migration rates into in- and out-migration rates, as using the full posterior distribution added significant computational complexity without improving validation metrics or forecasts. Additionally, we apply a standard Rogers-Castro migration age schedule for non-GCC countries, which minimizes forecast uncertainty from schedule variation. Future work could address these approximations.

Our methods build on recent advances (Azose & Raftery, 2015a), address key challenges of using net migration as the unit of analysis (Rogers, 1990), and provide a computationally efficient alternative to migration flow forecasting (Welch & Raftery, 2022). We also propose a method tailored to the unique dynamics of Gulf Cooperation Council countries, offering more accurate net flow forecasts for countries in this region. Lastly, the MASI ratio serves as a concise measure of population age structure, capturing the pace of aging over the coming decades.

Chapter 4

BAYESIAN PERSON-PLACE MODEL: PROBABILISTIC ADDRESS IMPUTATION FROM ADMINISTRATIVE DATA

4.1 Introduction

Over the past 70 years, tabulations of demographic trends in the United States have predominantly been derived from censuses and carefully designed surveys (F. C. Billari & Zagheni, 2017). While censuses offer wide coverage, they are infrequent. Surveys are more timely but often suffer from low response rates and small sample sizes. Statistical offices around the world have begun collating data from government agencies and private data brokers to narrow the gap between infrequent, broad-coverage census and more timely surveys (Keller et al., 2018; National Academies of Sciences, Engineering, and Medicine, 2023).

Counties without population registers rely on surveys and population-level models between census enumerations for a number of consequential operations and services. Surveys make longitudinal analysis challenging since different people are sampled from round to round. Furthermore, surveys are often inadequate for demographic research among smaller subgroups, as these groups typically constitute a smaller share of the survey sample. While survey methodologies account for such effects, the single best estimates might be too uncertain to be practically useful. In such cases, changes observed from period to period may reflect variation in the estimates rather than actual changes in circumstances.

Administrative records (AR) filed by large portions of the population between censuses, when combined with well-crafted surveys, offer a powerful third alternative. This approach leverages the detailed responses provided by surveys alongside the broad, albeit messy, population-scale observational records collected by state and federal agencies during the intercensal period. To fully realize this potential, statistical models are required to establish

relationships between these vast AR datasets and the vetted survey responses, enabling more timely and accurate population-scale demographic measurements.

The National Academies Report on the US 2020 Census highlighted that the 2020 Census was the first to use administrative records data to enumerate nonresponding households. The report concluded that AR usage in supplementing census operations was a successful innovation and recommended further research. One significant area of such research at the US Census Bureau includes the Person-Place Model (PPM) (Keller et al., 2018; Brown et al., 2023).

Published PPM approaches aim to impute an individual's usual address on a particular date based on their AR data (Brown et al., 2023). These models attempt to estimate the probability that a residence provided in an AR filing matches the residence where the individual would respond to a survey or census on a specified date. Over the past decade, multiple publications and US Census technical reports have demonstrated the efficacy of various PPM specifications (Keller et al., 2018; Murly et al., 2021; Brown et al., 2023).

Following the 2010 US Census, various PPM implementations became a critical component of AR-based research and operations. In the 2020 Census, PPM probability estimates were integral to a multistep algorithm determining AR data reliability for operations like nonresponse follow-up (NRFU) (Keller et al., 2018; Murly et al., 2021). More recently, PPM implementations have been integrated into a broader AR-based research program as a standalone model in the Census Bureau's Demographic Frame program (Ortman, 2024), which establishes a universal standard data set for both operations and research. Hence, advancements in the PPM methodology yield asymmetric benefits.

Naïve alternatives to PPM modeling, such as assigning a person to their *Last Observed* address, intuitively seem well-suited for modeling person-place dynamics by prioritizing the most recent data point. However, as demonstrated by its poor performance across key metrics (see Table 24), such simplistic methods often fail to account for the reliability of different administrative record observations for the same person across multiple addresses. Specifically, the *Last Observed* method's inability to differentiate between the reliability

of various data sources or consider nuanced temporal patterns overlooks complexities of administrative records data.

In contrast, PPMs offer a more sophisticated framework by integrating multiple data sources and dynamically weighting their contributions based on reliability and temporal relevance. This ability to leverage the strengths of diverse data sources while minimizing the impact of less reliable information underscores the need for PPMs. By addressing the limitations of simpler methods like *Last Observed*, PPMs provide a more robust and flexible approach to modeling complex person-place relationships, ultimately yielding more accurate and actionable results.

Several challenges are shared among all published PPM specifications. For example, most people have just one possible AR address assignment. Administrative records data for other people offer many viable candidate addresses over the same period. In either of these two cases, it could be that none of the AR addresses are correct; however, most PPM specifications fail to account for this outcome. Finally, published reports focus on the accuracy of mean probabilities (i.e, point estimates) and offer no assessments of the uncertainty around these measures.

The appeal of extant PPM specifications lies in the simplicity and ease of estimation with off-the-shelf software. Multiple national statistics offices implemented their own PPM approaches for AR-based operations in recent years. Both Canada and New Zealand implemented a PPM for their census operations and research programs (Stats, NZ, 2019b; Yoon et al., 2022). However, all these PPM specifications use a binary outcome variable, giving rise to many undesirable limitations. Most critically, none of the PPM specifications are valid statistical models–except for people with only two possible AR addresses. Current model specifications implicitly encode an assumption that one of a person's AR addresses is the best match for them, but it could be that *none* of their AR addresses is best. In general, approximating the AR address and census/survey match probability with a binary model leads to invalid probability estimates, misspecified uncertainties around these estimates, and intractable statistical inference.

We propose the Bayesian Person-Place Model (Bayes PPM) to close these gaps. The Bayes PPM eliminates challenges associated with extant PPM specifications. We will refer to these specifications as the Binary PPM. The Bayes PPM also yields probability estimates based on a well-defined statistical model that accommodates individual- and residence-level variations, including AR address counts that vary from individual to individual. More robust probability estimates could enable more sophisticated NRFU operations in the US 2030 Census as well as a more expansive research agenda in the US Census Bureau's Demographic Frame program. As more countries integrate AR data into their operations and research, Bayes PPM offers a well-defined statistical model capable of supporting a wider range of analyses.

Administrative records data are strictly controlled to protect privacy, which constrains the pace of methodological advancements using such data. However, the National Academies of Sciences, Engineering, and Medicine (2023) review of the statistical research methods for the 2020 US Census explicitly recommends using simulated data to demonstrate the efficacy of new statistical methods for the 2030 Census. This recommendation recognizes that administrative records data are not collected for PPM estimation and include sources of variation that would be otherwise minimized in a designed survey or study. Following this recommendation, we demonstrate the Bayes PPM on simulated data. Brown et al. (2023) includes AR data summary statistics for the data used to estimate the Binary PPM model. We simulate a data set with characteristics consistent with the contours of this data, which includes continuous, categorical, and indicator predictor variables.

The remaining sections are organized as follows. Section 4.2 summarizes past administrative records research in general and the most relevant discrete choice model literature underpinning the Bayes PPM approach. Section 4.3 describes the data simulation methods. Section 4.4 details the Bayes PPM specification. Section 4.5 summarizes Bayes PPM evaluations compared to extant PPM specifications. Section 4.6 outlines our contributions to this class of models and limitations to the Bayes PPM.

4.2 Background

Census Person-Place Model Rastogi & O'Hara (2012) provide a foundational assessment of the AR-based enumeration in the United States, particularly in light of the increasing costs associated with NRFU operations. Building on the US Census 2000's Statistical Administrative Records System (StARS), this report is among the earliest to propose using a logistic model to select the optimal person-address pair for individuals with multiple possible addresses. This model was estimated using responses from the 2010 Census and contemporaneous AR data.

Morris (2014) offer one of the first comprehensive descriptions of the Binary PPM specification. However, their methodology does not employ PPM probabilities directly. Instead, PPM probabilities serve as inputs to an algorithm that establishes a decision rule for determining the reliability of a household's AR data for enumeration purposes. This two-model approach mitigates the impact of any Binary PPM misspecifications, because PPM probabilities serve as leading indicators for a heuristic approach. Both Morris (2014) and Rastogi & O'Hara (2012) note that early versions of the extant PPM specifications included multiple stages to account for different AR address categories, but neither offer details on these components.

Keller et al. (2018) detail an algorithm that evaluates the feasibility of AR enumeration by integrating results from two models: a Binary PPM specification similar to Morris (2014) and a more sophisticated multinomial occupancy model initially outlined by Rastogi & O'Hara (2012). Their work includes a comprehensive summary of US Census research on this topic from 2000 to 2018, providing insightful historical context for AR-based census methodologies. See (Mule & Keller, 2021; Murly et al., 2021, 2023) for additional details using the methods in (Keller et al., 2018).

Brown et al. (2023) is most relevant to the Bayes PPM, as it explicitly utilizes PPM probabilities to conduct a *Simulated Census* using AR data alone. This approach employs a Binary PPM similar to the specification in Morris (2014) and Keller et al. (2018), but AR

data and two different years of American Community Survey (ACS) responses determine the ground truth used for model estimation instead of census responses. To address the multinomial nature of the outcome variable, Binary PPM probability estimates for each person-address pair are normalized to sum to one for each individual.

AR Address Imputation Outside the United States Stats, NZ (2019b) describe New Zealand's approach to individual-level address imputation. New Zealand adapts the two-model approach outlined in Keller et al. (2018) for NRFU operations, including a Binary PPM specification. Statistics New Zealand (2023) outline the Administrative Population Census (APC), similar to the Brown et al. (2023) US *Simulated Census*. For an overview of AR data's role in Statistics New Zealand, see (Stats, NZ, 2019a; Bycroft et al., 2021).

Canada's approach, described in Lundy (2022) and Yoon et al. (2022), implements a Binary PPM similar to the specification outlined in Keller et al. (2018). Statistics Canada reports that their AR-based implementation of the *Simulated Census* placed individuals in the correct household 80% of the time and successfully recovered the census enumerated household composition approximately 60% of the time. For a summary of the Binary PPM person-level and household-level variables used by Statistics Canada, see Yoon et al. (2022).

Australia's 2016 Census was the first time ARs were used for their census operations and only included the Australian Address Register (Harding et al., 2022). The 2021 census subsequently added person-level AR data. Australian Bureau of Statistics (2020) use AR data to generate a total population estimate and compare these results to the official population count, Estimated Resident Population (ERP), similar to Brown et al. (2023). Currently, AR data play an indirect role in Australia's census operations, but there are plans to increase the use of AR data for a combined AR-traditional census on a small scale in 2026 and potentially at a large scale in 2031 (Harding et al., 2022).

Elliott & Blackwell (2023) describe the United Kingdom's use of AR data for intercensal demographic estimation; however, AR data are not used for person- and household-level inferences. Instead, AR data are used to estimate demographic trends aggregated over

specified geographic units and conduct longitudinal analyses.

Nordic countries have developed the most advanced AR-based demographic estimation infrastructure of any country in the world (Economic Commission for Europe and others, 2007; F. C. Billari & Zagheni, 2017). Denmark was the first to transition from a traditional census to a population register-based census in 1981 (Economic Commission for Europe and others, 2007). The register-based approach to population estimation requires limited statistical modeling compared to countries without population registers. In countries with population registers, modeling is primarily used to place individuals lacking any dwelling identifiers and typically uses a rules-based method instead of statistical modeling. For example, a person missing a dwelling identifier might be assigned to an address according to their family associations or last known dwelling (Holmberg & Watmuff, 2023). Sweden uses a model-based approach to predict whether a person out-migrated. In this case, a person's lack of AR data establishes a leading indicator that someone might have left the country (Holmberg & Watmuff, 2023). For a summary of Sweden's population register census operations, see Axelson et al. (2021), and for a summary of register-based census best practices in Denmark, Finland, Norway, and Sweden, see Economic Commission for Europe and others (2007).

Methodological research in countries that use population registers for census operations is mainly focused on data sources, best practices for merging multiple sources, and evaluation of merged AR data (Economic Commission for Europe and others, 2007). For AR data integration from a statistical point of view and a concise, more recent summary of AR-based population estimation efforts in Nordic countries, see Zhang (2012).

Economic and Social Commission for Asia and the Pacific (2022) offer an overview of different census-taking frameworks and modern terminology. They also describe the growing use of ARs in Asia and the Pacific, highlighting modernization efforts in four countries: Australia, Indonesia, South Korea, and Turkey. Most are transitioning to a combined-census approach similar to Canada's plan for the 2031 census. Dias et al. (2016) summarize Portugal's work transitioning to a combined-census model similar to the one targeted by

most other countries outlined in Economic and Social Commission for Asia and the Pacific (2022) and Canada.

Chun et al. (2021) provide a comprehensive overview of AR research around the world, focusing on data quality limitations rather than person-level modeling. For a detailed history of census-taking around the world, see Skinner (2018).

Alternative AR Imputation Methods Alternative modeling approaches to the Binary PPM are limited. Long et al. (2022) propose an innovative AR-based method and estimation approach for timely person-level imputations of the population in Washington state. They utilize driving license, voter, social services, and birth records to estimate the probability that an individual continues to reside at their last observed AR address. The inclusion of driving license data is a unique contribution to the literature that could be beneficial for state governments; however, these data are not readily available to US federal agencies. Long et al. (2022) note that many prior studies use AR-derived vital statistics for social science research (pp. 254); however, these studies aim to answer well-defined health, income, and education questions. In contrast, Long et al. (2022) aim to generate a timely snapshot of society at the population scale that can be used to answer social science and policy questions yet to be defined at the time of estimation. They argue that population-scale AR-based datasets provide unique benefits that are not available from decennial census data (which are not timely) or from the American Community Survey (ACS) microdata samples (which are not longitudinally linked, not individual-specific, and too sparse for detailed analysis).

Dunne & Zhang (2023) propose a capture-recapture statistical model fit to multiple AR databases as an alternative intercensal population estimation approach and census post-enumeration survey. They use the Republic of Ireland's driving license database and Quarterly National Household Survey to estimate the degree of under- or over-coverage using AR data contained in the Person Activity Register alone. However, the method described in Dunne & Zhang (2023) is not intended for use as an individual-level model. Instead, the approach in Dunne & Zhang (2023) is offered as an alternative approach to estimating ag-

gregated population counts between censuses, in contrast to the cohort component method (Preston et al., 2001) or Hamilton-Perry method (Baker et al., 2021).

Discrete Choice Modeling The Bayes PPM is based on a discrete choice model. Discrete choice modeling has a rich history and wide application across various disciplines. McFadden (1973) introduces the conditional logit model, a seminal contribution to econometrics and transportation research. This model has become a standard tool for analyzing decision-making processes where individuals choose one option from a set of alternatives. Haghani et al. (2021) provide a comprehensive overview of the discrete choice modeling literature and research community. For more accessible summaries of discrete choice modeling, see Jain et al. (1994); Train (2009); Bruch & Mare (2012); Elshiewy et al. (2017).

Hoffman & Duncan (1988) argue for broader adoption of discrete choice modeling for demographic processes and contrast the limitations of a multinomial logistic model with the flexibility of a discrete choice model. They highlight the benefits of parameter interpretation using a discrete choice model approach compared to an individual-focused multinomial model—specifically, the interpretation of parameters in terms of choices of an individual rather than the average response over a group of individuals. However, Hoffman & Duncan (1988) is just one example that calls attention to the uneven response in the research community. Thomas et al. (2006) propose a spatial discrete choice model for caribou environmental preferences. The model is specified to properly estimate population- and individual-level effects. The author describes how others have modeled the same process with independent logistic regression models fit to specific animals and why resulting parameter estimates are both unreliable and fail to properly estimate standard errors.

Discrete choice models have seen wider adoption in estimating individual-level migration destination preferences (Spring et al., 2017; Kucheva, 2022). Inoa et al. (2015) use a discrete choice model for residential, workplace, and employment preferences using a nested logit model framework. See Train (2009) for an overview of individual-level discrete choice models, of the benefits and challenges of Bayesian estimation, and the basis for individual-

specific choice categories. Lenk (2014) provide an extensive overview of Bayesian discrete choice model history, a high-level overview of estimation techniques, and hyperparameter specifications.

Hoffman & Duncan (1988) note that discrete choice models assume choices are made independently and based on a utility-maximizing process, which may not always hold true. McFadden (1978) explicitly outline how to navigate oversimplifications caused by using the multinomial likelihood in its simplest form. Dai (1998) evaluate the impacts of ignoring relationships between the choice set that may be correlated with the utility an individual might assign to that set. Hausman & McFadden (1984) and Cheng & Long (2007) discuss practical implications of the Independence of Irrelevant Alternatives (IIA) assumption in applied research. Bhat & Guo (2004) present a mixed spatial discrete choice model framework for residential preferences along with a brief series of suggestions to relax IIA constraints. Sener et al. (2011) build on this work with a more flexible model for correlations in the choice set. Bruch & Mare (2012) discuss common approaches to navigating IIA violations, noting that a mixed logit model is one way to handle IIA issues. Krueger et al. (2023) discuss robust error distributions for discrete choice modeling when outliers are present.

While there are well-known limitations (Dai, 1998), discrete choice models provide a framework flexible enough to specify a well-defined statistical Person-Place Model that avoids more serious challenges to the existing approaches outlined in Keller et al. (2018) and Brown et al. (2023). In particular, a valid choice set must satisfy the following conditions (Train, 2009): (1) alternatives are mutually exclusive, (2) the choice set must be exhaustive, and (3) the number of alternatives must be finite. In the Person-Place Model context, the choice set is either *None* or one AR address matches the address where a person would respond to a census or survey. Each person has a variable number of candidate addresses, which is finite. Hence, this choice set satisfies the three discrete choice set assumptions needed to formulate a valid model.

Another key feature of the discrete choice model specification is that it relates explanatory variables to the response variable without knowing the exact process that makes one "choice"

preferable to another (Train, 2009). This is an important property as there is no well-defined address categorization in a Person-Place Model context. Instead, each address category is determined by the characteristics of each AR address.

Discrete choice model theory is the basis of the Bayesian Person-Place Model (Bayes PPM) proposed here. Restrictive category labels required for a standard multinomial logistic model are eliminated with the discrete choice model framework. The Bayesian specification avoids challenges to model estimation using a likelihood optimization approach (Train, 2009). A Bayesian specification also shares information across different AR address counts among individuals (Gelman & Hill, 2007). This feature overcomes challenges associated with estimating separate models for each group of people with the same number of AR addresses, e.g., small sample size for large AR address counts.

4.3 Data

Administrative records data are tightly controlled to safeguard individuals' privacy and to maintain the public's trust in the institutions serving them. These safeguards are crucial to carrying out the missions of the institutions generating administrative records data; however, these controls also slow the pace of research and innovation. We use publicly released data and technical reports from the last decade to simulate data as similar as possible to the characteristics found in real-world data. This simulated data set is then used to demonstrate how the Bayes PPM might perform compared to published PPM specifications.

Evaluating PPM specifications with simulated data offers advantages. For example, performance assessments are carried out when the true underlying data generating mechanism is known. However, complex interdependencies found in real data are difficult to replicate. Even with such limitations, demonstrating Bayes PPM using simulated data is consistent with recommendations found in chapter 8 of (National Academies of Sciences, Engineering, and Medicine, 2023). In particular, page 217 notes:

> *The use of artificial (simulated) populations with a variety of properties—for*

> *example, various relations of administrative values to true values—is a valuable component of a comprehensive assessment of the performance of administrative records and should be pursued. The idea is to simply make a simulated census data set [...], and then test the performance of various models as the simulated dataset is adjusted for various assumed properties [...]. The disadvantage is all the assumptions one must make, but the benefit is that one knows what the true values are for the simulation.*

The remainder of this section outlines our approach to simulating a data set that shares as many characteristics as possible with real world administrative records data used to estimate the Binary PPM. Summary statistics from (Brown et al., 2023) and US population age and sex data from the 2020 Census are the primary sources of data for our simulation.

4.3.1 Real-Time 2020 Administrative Record Census Simulation Features

Publicly released summary statistics of AR data are reported in Brown et al. (2023). These include the distribution of the number of AR sources where people were observed in one or more ARs, of the percent of AR source frequency, and of the number of AR address candidates. Table 4.1 summarizes the AR feature summaries from (Brown et al., 2023) needed to construct a minimally viable simulated data set.

Table 4.1: Administrative records tables adapted from Brown et al. (2023) for data simulation.

Table	Description
26	Distribution of Number of AR Data Sources Where AR Census People Were Observed
27	Percent of AR Census People Observed in Each AR Data Source
38	Distribution of Number of MAFIDs Per AR Census Person

The distribution of AR source counts represents how often one or more sources were observed at the same candidate address. The distribution of AR source frequency represents how often a particular source is observed. The number of AR address candidates represents

how often an individual had 0-5+ candidate addresses. The joint distribution of these three metrics is not reported in (Brown et al., 2023). However, marginal distributions were used to estimate approximations to the joint distributions of these quantities as needed.

Number of Administrative Record Sources Table 4.2 shows the percent of people with 1-5+ AR observations from (Brown et al., 2023) table 26. This summary determines the simulation parameter, p_T, used to sample the total number of AR sources observed for each simulated individual.

Table 4.2: Distribution of number of AR data reported in Brown et al. (2023) table 26 and simulation probability vector p_T.

Number of AR data sources	Percent of people	p_T
1	35.8	0.358
2	19.2	0.192
3	20.0	0.200
4	11.4	0.114
5+	13.7	0.137

Number of Administrative Record Addresses Table 4.3 shows the distribution of AR addresses from (Brown et al., 2023) table 38. This summary determines the simulation parameter, p_J, used to sample the total number of AR addresses observed for each individual. Brown et al. (2023) table 38 indicates that 2.57% of people reported zero addresses; however, only AR sources that include both a person and address identifier can be used for model estimation. Simulation parameter p_J was defined as the normalized empirical distribution of AR address counts 1-5+.

Selected Administrative Record Sources Table 4.4 shows the distribution of the population observed in a subset of the 30+ sources used in (Brown et al., 2023) table 27. We chose a subset of these sources for both practical and technical reasons. Using a subset of the sources summarized in (Brown et al., 2023) reduced the computational and inference

Table 4.3: Distribution of number of addresses reported in Brown et al. (2023) table 38 and simulation probability vector p_J with $\max J[i] = 5$ simulated addresses.

Number of addresses	Percent of people	p_J
0	2.57	0.000
1	51.59	0.530
2	27.51	0.282
3	11.47	0.118
4	4.48	0.046
5+	2.38	0.024

burdens. Furthermore, we wanted to simulate AR sources most likely to be used for the 2030 decennial Census. For example, the National Academies of Sciences, Engineering, and Medicine (2023) report recommends eliminating data sources generated by law enforcement agencies.

Tax agency data provide the broadest coverage and most reliable sources among multiple national statistical agencies (Holmberg & Watmuff, 2023). The most frequent Internal Revenue Service (IRS) sources, forms 1040 and 1099, were included as a result.

Some AR sources are filed by a small segment of the overall population but highly associated with the best address for members of certain subgroups. For example, the Centers for Medicare and Medicaid (CMS) Member Enrollment Database (MEDB) should be highly associated with the best single address for a person age 65 or older. While CMS MEDB observations are generated by a subset of the population, this source offers an example of a low-coverage source that might be highly predictive of a person's single best address if filed.

Third party data sources supplement ARs generated by government agencies. Data from third party sources may be the only available record for some individuals. Source data characteristics associated with VSGI (a 3^{rd} party data source) were included in the data simulation as one such example.

The US Postal Service National Change of Address (NCOA) form is filed by an individual who wants to change their mailing address. This form includes a person's old address where they no longer want to receive mail and their new address where they want to receive mail.

This source was included to assess each model's ability to resolve the association between an infrequently filed AR source compared to more frequently observed AR sources.

Table 4.4 shows two simulation parameters derived from Brown et al. (2023) table 27. Parameter vector $p_{F(s)}$ determines the probability that source s is ever observed. This *Percent in source* column to was converted to independent probabilities that each source s is observed at any AR address for an individual. Parameter vector $p_{G(s)}$ is an approximation to the probability of observing one or more additional AR sources given that AR source s was observed, i.e., 1 - *Percent only source*. Values of $p_{G(s)}$ closer to 1 indicate that AR source s is less likely to be the only AR source for a person.

Table 4.4: Distribution of percent of people observed in each AR source from Brown et al. (2023) table 27 and simulation probability vectors $p_{F(s)}$ denoting the probability that a person is observed in AR source s and $p_{G(s)}$ denoting the probability that a person is observed in at least one more source given that they were observed in s.

s	Source	Percent in source	Percent source only	$p_{F(s)}$	$p_{G(s)}$
1	IRS 1040	76.37	21.28	0.764	0.787
2	IRS 1099	60.37	4.36	0.604	0.956
3	3^{rd} Party (VSGI)	41.62	4.86	0.416	0.951
4	CMS MEDB	15.11	0.09	0.151	0.999
5	USPS NCOA	4.66	0.17	0.047	0.983

Demographic Variables The US population age and sex distribution from the 2020 Census is used to simulate each individual's age and sex. Population totals are reported by 5 year age group, $(0-4, 5-9, \ldots, 80-84, 85+)$, and sex, (female, male). These totals were used to generate a vector of the population age and sex distribution.

In particular, let D denote the 1×36 dimensional categorical random defined by concatenating the 18 female age groups and 18 male age groups from the 2020 Census. Let p_D denote the probability distribution defined by the number of people in each sex-age group divided

by the total population according to the 2020 US Census. Then parameter p_D, definded as

$$p_D = \{P_{\text{female,age group=1:18}}, P_{\text{male,age group=1:18}}\} / \sum_{\text{sex, age group}} P_{\text{sex, age group}}, \tag{4.1}$$

determine the individual-level demographic attributes of simulated individuals.

4.3.2 Data Representation and Notation

Person-Place data include person-level attributes, address-level attributes, and source-specific attributes. Encoding the interconnected relationships among all these variables poses a challenge that receives limited attention in existing descriptions of the models fit to these data. The complexity introduced by an unambiguous data structure may have been a distraction to the main objectives and findings in past studies; however, existing descriptions of the Person-Place data obscure Binary PPM nuances. This section outlines a well-defined Person-Place data structure for each group of attributes.

Person, Address, and Source Indices We use index $i \in \{1, \ldots, N\}$ to refer to the unique anonymized person IDs and index $j[i] \in \{1, \ldots, J[i]\}$ to refer to the unique address IDs associated with person index i. The index $j[i]$ denotes the j^{th} address index associated with person i among $J[i]$ total addresses that vary by individual. This makes the person index i interchangeable with their anonymized ID. In contrast, the anonymized address ID is determined by the key-value pair $\{(i, j[i]), \text{address ID}\}$ with *key* $(i, j[i])$ and *value* address ID. This abstraction is essential to the Bayes PPM specification and implementation.

Source indices, $s \in \{1, \ldots, S\}$, are fixed and shown in Table 4.4. For example, Table 4.4 shows source index $s = 1$ corresponds to the IRS 1040 administrative record. The value S denotes the total number of sources, and $S = 5$ in our case. Table 4.5 summarizes the person, address, and source indices.

Table 4.5: Person, address, and source index definition summary.

Index	Definition
i	Person index with $i \in \{1, \ldots, N\}$ for N total people
$j[i]$	Address in position $j[i] \in \{1, \ldots, J[i]\}$ among $J[i]$ addresses for person i
$J[i]$	Total number of unique AR addresses for person i
N	Total number of unique individuals
s	Source index with $s \in \{1, \ldots, S\}$ corresponding to the AR source for S total sources
S	Total number of AR sources

Person-Level Attributes Person-level attributes include the attributes of a person that are fixed across households and AR sources. These could include age, sex, race, birth country, or any other demographic characteristics of an individual. We simulate age and sex demographic variables only.

The age and sex for individual i are sampled from $d[i] \sim \text{Categorical}(p_D)$ and stored in the $N \times 3$ dimensional array with entry $P_{i,a,l} = \{i, \text{age group} = 5 \times [(d[i] - 1) \mod 18], \text{male} = d[i] > 18\}$. For example, suppose $d[i] = 5$, then $P_{i,a,l} = \{i, \text{age group} = 20, \text{male} = 0\}$ to indicate that the individual index i represents a 20-24 year old female. Similarly, if $d[i] = 30$, then $P_{i,a,l} = \{i, \text{age group} = 60, \text{male} = 1\}$ denotes a 60-64 year old male. Table 4.6 summarizes the person-level attribute variables and data objects.

Address-Level Attributes Address-level attributes could include information about the housing unit itself (e.g., single unit versus multifamily), occupancy status (e.g., occupied versus vacant), classification (e.g., residential versus nonresidential), or any other characteristic about the house that does not vary on the address dimension.

We simulated a single address-level categorical variable, $v[i, j[i]] \sim \text{Categorical}(p_V)$, with three levels for each person's AR addresses. The simulated category $v[i, j[i]]$ for address ID associated with key $(i, j[i])$ is stored in the $N \times J[i]$ ragged array entry $U_{i,j[i]}$. Address-level features simulated for demonstration purposes only and not used in our analysis. Table 4.7 summarizes the person-level attribute variables and data objects.

Table 4.6: Person-level data attribute summary.

Index	Definition
i	Person index with $i \in \{1, \ldots, N\}$ for N total people
a	Age group index $a \in \{1, \ldots, 18\}$ corresponding to 5 year age groups $0-4, \ldots, 80-84, 85+$
l	Indicator variable denoting person i is female (0) or male (1)
$d[i]$	Sex and age category for person i sampled from D
p_D	Observed distribution of the number of people in female age groups 0-4,..., 80-84, 85+ followed by the number of people in male age groups 0-4,..., 80-84, 85+ divided by the total population size of the United States according to the 2020 Census
D	Random vector of indicators over the age and sex categories defined by the demographic characteristics of the United States according to the 2020 US Census with dimensions 1×36 over 5 year age categories for females and males. Age groups include 0-4, ..., 80-84, 85+, yielding 18 female age group categories followed by 18 male age group categories.
$P_{i,a,l}$	Person-level array with entries i denoting the person index, a the age group, and l an indicator variable denoting male (1) or female (0)

Source-Level Attributes Source-level attributes could include an indicator that a person-address pair is observed in an AR source, the number of times a person-address pair is observed in an AR source, or a wide range of metrics representing temporal characteristics of AR sources for the same person-address pair. Well-defined, interpretable source variable specifications are inseparably linked all PPM specifications and subsequent statistical inference.

We simulate AR source-specific indicator variables along with the absolute difference in months between a person's AR source observation and ground truth event (e.g., American Community Survey interview date). The set of addresses for person i is determined by the unique addresses found in their AR sources. This set defines the person-address key-value pairs for person i. Given the person-address key value pairs for person i, the source indicator

Table 4.7: Address-level data attribute summary.

Index	Definition
i	Person index with $i \in \{1, \ldots, N\}$ for N total people
$v[i, j[i]]$	Address category label for address ID associated with key $(i, j[i])$
p_V	Probability of address category 1, 2, or 3
V	Random vector of indicators over three address categories, $\{1, 2, 3\}$
$U_{i,j[i]}$	Address attribute array with address category entry at key $(i, j[i])$

variable for source s observed at address $(i, j[i])$ is defined as follows.

$$\mathbb{1}_s(i, j[i]) = \begin{cases} 1 & \text{if individual } i \text{ observed at address index } (i, j[i]) \text{ for AR index } s \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

A person-address pair can be observed in one or more of the S total AR sources. A vector of indicator variables is used to denote the AR sources where a person was observed with the same address index among all available sources. Person-address source indicator vectors for the same person are arranged in a ragged array with a person-specific total number of addresses, $J[i]$. We denote this ragged array as $A_{i=1:N,j[i]=1:J[i],s=1:S}$. For example, the address index and source indicator matrix for person i over all $J[i]$ addresses and S sources is as follows:

$$A_{i,1:J[i],1:S} = \begin{array}{c|cccc} & \text{AR 1} & \text{AR 2} & \ldots & \text{AR S} \\ \hline \text{j[i]=1} & \mathbb{1}_1(i,1) & \mathbb{1}_2(i,1) & \ldots & \mathbb{1}_S(i,1) \\ \text{j[i]=2} & \mathbb{1}_1(i,2) & \mathbb{1}_2(i,2) & \ldots & \mathbb{1}_S(i,2) \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ \text{j[i]=J[i]} & \mathbb{1}_1(i,J[i]) & \mathbb{1}_2(i,J[i]) & \ldots & \mathbb{1}_S(i,J[i]). \end{array} \tag{4.3}$$

AR sources observed closer to a ground truth event or reference date should be more relevant than those far before or afterwards. Let O_i denote the date of a ground truth event for person i. Let $H_{i,j[i],s}$ denote the date person i was observed at address $(i, j[i])$ in AR

source s no longer than 24 months prior to O_i or 12 months after O_i as done in Brown et al. (2023). Then the AR source date difference in integer months corresponding to entry $A_{i,j[i],s}$ are calculated as follows:

$$B_{i,j[i],s} = \begin{cases} -\lceil \frac{|O_i - H_{i,j[i],s}|}{30} \rceil & \text{if } A_{i,j[i],s} = 1, O_i \neq H_{i,j[i],s} \\ -1 & \text{if } A_{i,j[i],s} = 1, O_i = H_{i,j[i],s} \\ 0 & \text{otherwise.} \end{cases} \tag{4.4}$$

Date differences are negatively oriented to encode the fact that an AR source observation from the same source at a different address farther away a ground truth event should be less relevant to selecting the most likely address–all else equal. An AR source observation in the same month as the ground truth event is not distinguished from an observation in the month before or after this event. This is sensible considering the precision of source observation dates. Sources that were never filed by person i at address $(i, j[i])$ have source date differences equaling zero. Observations in array $B_{i,j[i],s}$ are simulated by randomly sampling (-24, -23, ..., -1, 1, ..., 12) for $A_{i,j[i],s} = 1$. Table 4.8 summarizes the source-level attribute variables and data objects.

Table 4.8: Source-level data attribute summary.

Index	Definition
i	Person index with $i \in \{1, \ldots, N\}$ for N total people
$J[i]$	Realization of total number of unique AR addresses for individual i draw from the distribution of random variable J
O_i	Date of ground truth event for person i
$H_{i,j[i],s}$	Date person i observed in AR source s at address $(i, j[i])$
$A_{i,j[i],s}$	Ragged array of indicators denoting person i observed at address index $(i, j[i])$ in AR source s
$B_{i,j[i],s}$	Ragged array of negative absolute difference in months separating person i ground truth event and AR source s observation at address index $(i, j[i])$

Response Variable The Binary PPM specification uses an indicator variable to denote instances where a person is enumerated (i.e., "placed") at the same address using AR source data or a survey or census (Brown et al., 2023; Keller et al., 2018; Morris, 2014). Thus, for person i observed at addresses $(i, j[i])$ with $j[i] = 1, \ldots, J[i]$, the response is the random variable

$$Y_{i,j[i]} = \begin{cases} 1 & \text{Address ID } (i, j[i]) = \text{Address ID survey or census} \\ 0 & \text{otherwise.} \end{cases} \tag{4.5}$$

This response specification implicitly defines $J[i]+1$ possible outcomes for person i: a match among one of the $j[i] = 1, \ldots, J[i]$ AR addresses and no match for any address when

$$\sum_{j[i]=1}^{J[i]} Y_{i,j[i]} = 0. \tag{4.6}$$

A Two Stage PPM models the no match or *None* case explicitly. Let Z_i denote the binary random variable with $Z_i = 1$ if any of the AR source addresses for person i match the census or survey enumerated address for that person. Given $Z_i = 1$ in stage 1, stage 2 can be used to enumerate the portion of the population as before. The following specification summarizes the Two Stage PPM.

$$\textbf{Stage 1: } Z_i = \begin{cases} 1 & \text{Any AR address ID for person } i \text{ matches census/survey} \\ 0 & \text{otherwise.} \end{cases} \tag{4.7}$$

$$\textbf{Stage 2: } Y_{i,j[i]} | Z_i = 1 = \begin{cases} 1 & \text{Address ID } (i, j[i]) = \text{Address ID survey or census} \\ 0 & \text{otherwise.} \end{cases} \tag{4.8}$$

The final response specification encodes the *None* category explicitly by prepending address index $j[i] = 0$ to the person-address key value pair for each individual so that the

$(i, j[i] = 0)$ key encodes the *None* address ID, yielding

$$Y_{i,0:J[i]} = (1 - Z_i) \bigcup Y_{i,1:J[i]}. \tag{4.9}$$

Table 4.9 summarizes the source-level attribute variables and data objects.

Table 4.9: Response variable notation summary.

Index	Definition
i	Person index with $i \in \{1, \ldots, N\}$ for N total people
$J[i]$	Realization of total number of unique AR addresses for individual i draw from the distribution of random variable J
$Y_{i,0:J[i]}$	Indicator random vector of 0s for all $j[i] \in \{0, \ldots, J[i]\}$ except for the $(i, j[i])$ address ID matching the survey or census address ID for person i with $Y_{i,j[i]=0} = 1$ when no AR address ID match exists
Z_i	Binary random variable with $Z_i = 1$ if any AR source address ID of person i matches the census or survey enumerated address ID for person i

Section 4.4 details the simulation algorithm used to populate each element of the simulated Person-Place data set.

4.4 Methods

This section describes the algorithm used to simulate the Person-Place data set and all PPM specifications evaluated.

4.4.1 Data Simulation Algorithm

Tables 26, 27, and 38 in (Brown et al., 2023) provide broad contours of the Person-Place data used for estimation. Each of these tables describe a marginal distribution associated with some aspect of the Person-Place data. We use a combination of statistical methods and physical constraints to simulate a synthetic data set of plausible realizations aligned as

closely as possible to the summary statistics reported in (Brown et al., 2023). The key steps of the simulation are described below.

For person $i = 1, \ldots, N$:

1. Jointly sample the demographic category (i.e., age and sex) for individual i, $d[i] \sim$ Categorical(p_D), and record the results in $P_{i,a,l}$.

2. Sample an approximation to the joint distribution of total AR source addresses per person and total number of sources as follows. We use iterative proportional fitting (IPF) (Rüschendorf, 1995) to estimate the joint distribution of the total number of addresses for person i, $J[i] \sim$ Categorical(p_J), and total number of sources filed, $t[i] \sim$ Categorical(p_T), when only p_J and p_T are known. Note that $J[i] \leq t[i]$ by definition since there must be at least one AR source observation for each address. This is essential to properly initialize the estimate of the joint probability distribution, $\hat{p}_{TJ}$.

   Initialize the joint probability matrix with $T = 5$ rows, $\max J[i] = 5$ columns, and entries

$$p^0_{TJ}(t, j) = \begin{cases} 1 & \text{if } j \leq t \\ 0 & \text{otherwise.} \end{cases} \tag{4.10}$$

   Use the IPF algorithm initialized with matrix p^0_{TJ} to find the closest approximation, $\hat{p}_{TJ}$, to the matrix with column sums $p_J = (0.530, 0.282, 0.118, 0.046, 0.024)$ and row sums $p_T = (0.385, 0.192, 0.200, 0.114, 0.137)$. Samples from an approximation to the joint distribution were generated via $(t[i], J[i]) \sim$ Categorical$(\hat{p}_{TJ})$. Table 4.10 shows the estimated $\hat{p}_{TJ}$.

Table 4.10: Approximation to joint distribution of (T, J) (rounded).

T	**J**					
	1	2	3	4	5	Total
1	0.376	0.000	0.000	0.000	0.000	0.376
2	0.066	0.120	0.000	0.000	0.000	0.186
3	0.047	0.085	0.062	0.000	0.000	0.194
4	0.021	0.039	0.028	0.023	0.000	0.111
5	0.021	0.038	0.028	0.023	0.024	0.134
Total	0.531	0.281	0.118	0.046	0.024	1.000

The IPF algorithm must be initialized with positive values to converge to the maximum likelihood estimate, which is not plausible in this case. The matrix of joint probabilities using IPF in this case is only helpful to the extent that it improves agreement between the simulated Person-Place data and the marginal summary statistics for Tables 26, 27, and 38 in (Brown et al., 2023). We evaluated the agreement between the Person-Place data simulated by sampling $T|J[i]$. The IPF approximation outperformed the alternatives in terms of recovering the Person-Place data summary statistics for Tables 26, 27, and 38 in (Brown et al., 2023).

3. Sample the presence of each source indicator, $s = 1, \ldots, S$, across any of the $J[i]$ addresses. The previous step determines the distribution of the total number of AR observations for person i given the total number of unique addresses, $t[i]$, but makes no distinction between multiple AR observations from the same source and observations from unique AR sources. This step determines which of the AR sources are observed for person i across their $J[i]$ distinct addresses.

   Table 4.4 shows the empirical probability of observing an individual in each source ($p_F(s)$) along with the empirical probability of observing one or more additional AR sources given a person is observed in source s ($p_G(s)$). These empirical probabilities were used to estimate the distinct AR source observations and conditional probability

of observing a person in an AR source given that they appear in another source, denoted $p_{F|G=1}(s)$.

We make several simplifying assumptions in this step. First, we approximate the unique sources observed for person i with the total number of AR source addresses, $J[i]$. Second, if person i is under age 65 and has $J[i] = 5$, then we revise $J[i] = 4$ since this would imply a person younger than 65 would be enrolled in Medicare Part B. If person i is over 64 with $J[i] = 5$, then we require CMS as one of the AR observations. If the total number of addresses for person i equals 1 (i.e., $t[i] = 1$), then we sample the source indicator in proportion to the probability of appearing in this source alone, (i.e., $1 - p_G(s)$). Otherwise, we sample the $J[i]$ distinct AR source observations in proportion to

$$p_{F|G=1}(s) = \frac{p_F(s)/p_G(s)}{\sum_{s=1}^{S} p_F(s)/p_G(s)}. \tag{4.11}$$

This step yields a $1 \times S$ dimensional indicator vector $I_{i,1:S}$ with entries

$$I_{i,s} = \begin{cases} 1 & \text{if source } s \text{ sampled} \\ 0 & \text{otherwise} \end{cases} \tag{4.12}$$

denoting the which AR sources are observed at least once among person i's $J[i]$ distinct addresses.

4. Let $A_{i,j,s}$ denote the array of indicator variables (initialized with 0s for each person index i) that individual i was observed at AR address ID $(i, j[i])$ in source s. Initialize variable $r[i] = t[i]$ to track the number of AR source observations remaining after the following three-step process.

   4(a) Ensure each address of person i has an AR source. For k in $1, \ldots, J[i]$, draw $s^\star$ by randomly sampling the source index of $I_{i,1:S} > 0$, set $A_{i,k,s^\star} = 1$, and decrement $r[i] = r[i] - 1$. At the end of this step, each address for person i will have exactly

one AR source.

4(b) Distribute the second AR source for each address in order of $1 - p_G(s)$ for each AR address of person i if $r[i] > 0$ as follows. For k in $1, \ldots, J[i]$, select $s^\star$ from the index of $I_{i,1:S} > 0$ with the largest $1 - p_G(s)$, set $A_{i,k,s^\star} = 1$, and decrement $r[i] = r[i] - 1$. At the end of this step, each address for person i will have no more than two AR sources.

4(c) Distribute the remaining AR source counts if $r[i] > 0$ and any $A_{i,1:J[i],1:S} = 0$. In this step, we attempt to distribute the remaining $r[i]$ sources not allocated in the previous steps. We iterate through $A_{i,1:J[i],1:S}$ along index $s = 1, \ldots, S$ to search for any AR address to allocate the remaining $I_{i,1:S} > 0$ sources. If all $A_{i,1:J[i],s} = 1$ for s in $I_{i,1:S} > 0$, then we discard the remaining $r[i]$ sources as duplicate observations from the same AR source at the same address.

At the end of this step, the ragged array $A_{i,1:J[i],1:S}$ is fully populated for person i with indicators of the AR sources observed at each address.

5. Now sample the date difference variables person i's AR source observations. Let $B_{i,1:J[i],1:S}$ denote the ragged array of the number of months each source observation was filed before or after a ground truth event (e.g., census or survey response). Entries of $B_{i,1:J[i],1:S}$ are defined as

$$B_{i,j[i],s} = \begin{cases} b_{i,j[i],s} & \text{if } A_{i,j[i],s} = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{4.13}$$

with $b_{i,j[i],s}$ randomly sampled from $(-24, -23, \ldots, -1, -1, -2, \ldots, -12)$.

6. Sample address-level covariates for each address associated with person index i. Address-level categorical covariates could represent a range of characterizations for a candidate address. We randomly sample a categorical variable with three levels, $(1, 2, 3)$, with

probability vector p_V for each candidate residence of individual i and store the results in the vector $U_{i,1:J[i]}$.

7. Return to 1. until $i = N$.

This procedure yields four sets of predictor variables:

1. $P_{i,a,l}$ is an individual-level predictors for age and sex of person i

2. $A_{i,1:J[i],1:S}$ is an indicator matrix of address indices (rows) and AR source indices (columns) observed for individual i

3. $B_{i,1:J[i],1:S}$ is a matrix of AR date differences from the target date over address indices (rows) and AR source date differences (columns) observed for individual i

4. $U_{i,1:J[i]}$ is the address-level categorical variable vector.

4.4.2 Bayesian Person-Place Model

Unambiguous indexing is critical to the Bayes PPM specification. A key feature of the Bayesian PPM specification is that it relates explanatory variables to the response variable without knowing the exact process that makes one AR address category preferable to another. Each address category is determined by the characteristics of the AR address and these characteristics are organized so that their position in the outcome vector is aligned to the position of the address covariates in the covariate array. This specification renders the Bayesian PPM a flexible, well-defined statistical model. The Bayesian PPM is specified as

follows:

$$Y_{i,0:J[i]} \mid \pi_{i,0:J[i]} \sim \text{Multinomial}\left(1, \pi_{i,0:J[i]}\right) \tag{4.14}$$

$$\pi_{i,j[i]} = \exp \eta_{i,j[i]} / \sum_{j[i]=0}^{J[i]} \exp \eta_{i,j[i]} \tag{4.15}$$

$$\eta_{i,j[i]} \sim \text{Normal}\left(\alpha_{J[i]} I_{\{j[i]=0\}} + \sum_{s=1}^{S} w_{i,j[i],s} \beta_s + \sum_{s=1}^{S} m_{i,j[i],s} \gamma_s, \sigma^2\right) \tag{4.16}$$

$$\beta_s \sim \text{Normal}\left(\beta_0, 10^2\right) \tag{4.17}$$

$$\alpha_{J[i]}, \beta_0, \gamma_s \sim \text{Normal}\left(0, 10^2\right) \tag{4.18}$$

$$\sigma \sim \text{Uniform}(a, b) \tag{4.19}$$

The Bayes PPM specification includes three sets of parameters and an overdispersion parameter. The intercept vector, $\alpha_{1:\max J[i]}$, is used to estimate the association between the number of unique AR addresses and the *None* address category. Parameter vector $\beta_{1:S}$ captures the overall level of association between each AR source and the probability of a match at the address where each source was observed all else equal. The condition $\sum_{s=1}^{S} \beta_s = 0$ makes the model identifiable. Parameter vector $\gamma_{1:S}$ is used to estimate the rate of change from the baseline AR association (established by $\beta_{1:S}$) as the negative absolute number of AR source observation months difference approaches zero from below. The overdispersion parameter, σ, provides a mechanism to estimate variation in addition to the variance of a multinomial likelihood. Table 4.11 describes each term of the model.

Table 4.12 shows the structure of the training data used to estimate the Bayes PPM.

Outcome Variable Simulation We generate outcomes from the Bayesian discrete choice PPM with simulated true parameters: $\alpha_{1:J[i]}$, $\gamma_{1:S}$, $\beta_{1:S}$, and σ. Table 4.13 summarizes the parameter simulation distributions used.

Simulated $\alpha_{1:2}$ parameters were sampled from Uniform(-5, -1) distribution and sorted in increasing order. This implies that individuals with 1-2 candidate addresses were less likely

Table 4.11: Bayes PPM definitions.

Parameter	Definition
$Y_{i,0:J[i]}$	Indicator random vector denoting address ID $(i, j[i])$ matching a survey or census address ID for person i with $Y_{i,0}$ the *None* address ID category
$\pi_{i,j[i]}$	Probability person i matches address ID $(i, j[i])$ with $\sum_{j[i]=0}^{J[i]} \pi_{i,j[i]} = 1$
$\eta_{i,j[i]}$	Weight calculated from source indicator vector $w_{i,j[i],1:S}$, negative absolute months difference vector $m_{i,j[i],s}$, & model parameters
$\alpha_{J[i]}$	$J[i]$-specific weight associated with *None* category assignment
β_s	Baseline association between address match and source s with $\sum_{s=1}^{S} \beta_s = 0$
γ_s	Monthly rate of change from baseline association between address match for source s
$w_{i,j[i],s}$	Indicator vector denoting that person i observed at address $(i, j[i])$ in source s with $w_{i,j[i]=0,s} = 0$ for all i and s
$m_{i,j[i],s}$	Integer vector of negative absolute number of months between person i observed at AR address $(i, j[i])$ for source s with 0 denoting instances where person i was never observed at address $(i, j[i])$ and with $w_{i,j[i]=0,s} = 0$ for all i and s
$J[i]$	Total number of unique AR addresses for individual i
i	Person index with $i = 1, \ldots, N$
a, b	User-specified tuning parameters for overdispersion parameter

than average to have ground truth addresses that match none of the AR addresses all else equal. Simulated $\alpha_{3:5}$ parameters were sampled from Uniform(1,3) distribution and sorted in increasing order too. This implies that individuals with 3-5 AR addresses are less likely than average to have ground truth addresses that match none of the AR addresses. These settings lead to 10-20% of the population being assigned to category *None*. The ranges for $\alpha_{1:5}$ were selected so that simulated $\beta_{1:5}$ and $\gamma_{1:5}$ can overtake the magnitude of $\alpha_{1:5}$ in establishing the probability of an address match compared to the *None* category.

Simulated $\beta_{1:5}$ parameters were randomly sampled from a Truncated Normal distribution. The truncation lower bound was 0.95 to encode the fact that an AR source observation should always be associated with an increased probability of assignment. The upper bound was 4 to avoid having any one source completely dominate other sources.

Simulated $\gamma_{1:S}$ parameters were randomly sampled from a Truncated Normal distribution. The lower bound was set to 0 to allow the possibility that some sources experience little

Table 4.12: Example Bayes PPM outcome and design matrix for five individuals, S=3 AR sources and associated observation month differences, address-level categorical variable, and two individual-level variables (integer age and male indicator variable).

i	$J[i]$	$Y^T_{i,0:J[i]}$	$w_{i,0:J[i],1:S} \mid m_{i,0:J[i],1:S} \mid u_{i,j[i]}$	Age_i	Male_i
1	1	$\begin{bmatrix}0\\1\end{bmatrix}$	$\left[\begin{array}{ccc\|ccc\|c} 0&0&0&0&0&0&0\\ 1&0&0&-12&0&0&1 \end{array}\right]$	18	1
2	2	$\begin{bmatrix}0\\0\\1\end{bmatrix}$	$\left[\begin{array}{ccc\|ccc\|c} 0&0&0&0&0&0&0\\ 0&1&0&0&-18&0&3\\ 1&1&0&-2&-2&0&2 \end{array}\right]$	37	0
3	3	$\begin{bmatrix}0\\0\\1\\0\end{bmatrix}$	$\left[\begin{array}{ccc\|ccc\|c} 0&0&0&0&0&0&0\\ 0&1&0&0&-12&0&1\\ 1&1&1&-3&-8&-13&1\\ 1&0&0&-1&0&0&2 \end{array}\right]$	66	1
4	1	$\begin{bmatrix}0\\1\end{bmatrix}$	$\left[\begin{array}{ccc\|ccc\|c} 0&0&0&0&0&0&0\\ 1&0&0&-12&0&0&1 \end{array}\right]$	2	1
5	1	$\begin{bmatrix}0\\1\end{bmatrix}$	$\left[\begin{array}{ccc\|ccc\|c} 0&0&0&0&0&0&0\\ 1&0&0&-12&0&0&1 \end{array}\right]$	28	0

change in their association with a person's true address. These parameters must be positive since they reflect the change in association between an AR observation and the negative absolute date difference in months between observation and a ground truth event. The upper bound was selected to avoid massive degradation of the association as the negative absolute date difference in months grows more negative.

The overdispersion parameter, σ, was sampled from a Uniform$(0, 1)$ distribution. Large values of σ obscure the association between AR source observations and date differences.

Response vectors for each person, $Y_{i,0:J[i]}$, were simulated by sampling the Bayes PPM using the parameters shown in Table 4.13.

Table 4.13: Person-Place data parameter distribution summary.

Parameter	Distribution	Parameter Index 1	2	3	4	5
α	Uniform$(-5,-1)$, Uniform$(1,3)$	-3.404	-2.241	1.640	2.015	2.887
β	Normal$(1.5, 3^2, \min = 0.95, \max = 4)$	2.994	3.744	1.355	2.785	2.285
γ	Normal$(1, 3^2, \min = 0, \max = 3)$	2.079	2.341	1.465	2.108	0.373
σ	Uniform$(0,1)$	0.398	0.398	0.398	0.398	0.398

4.4.3 Extant Person-Place Model Specifications

Binary Person-Place Model

We use logistic regression to estimate the Binary PPM as in (Brown et al., 2023). Simulated person- and address-level features described in Section 4.3 are translated to long format, $\{Y_{i,j[i]}, X_{i,j[i],1:K}\}$, yielding 2 individual-level (age and sex) + S AR source indicators + S AR date differences + 1 address-level features = $3 + 2S = K$ total features. Table 4.14 shows the structure of the binary PPM outcome and features for $S = 3$ AR sources and corresponding date differences for those sources.

Table 4.14: Example Binary PPM outcome and design matrix for five individuals, S=3 AR sources, negative absolute month differences, two individual-level variables (integer age and male indicator variable), and address-level categorical variable.

i	$j[i]$	$Y_{i,j[i]}$	Age$_i$	Male$_i$	$w_{i,j[i],1:S}$ 1	2	3	$m_{i,j[i],1:S}$ 1	2	3	$u_{i,j[i]}$
1	1	0	18	1	1	0	0	-12	0	0	1
2	1	0	37	0	0	1	0	0	-18	0	3
2	2	1	37	0	1	1	0	-2	2	0	2
3	1	0	66	1	1	0	0	-12	0	0	1
3	2	1	66	1	1	1	1	-3	8	-13	1
3	3	0	66	1	1	0	0	-1	0	0	2
4	1	1	2	1	1	0	0	5	0	0	1
5	1	1	28	0	1	0	0	5	0	0	1

The binary outcome variable and data in Table 4.14 are used to estimate probability

$P\left\{Y_{i,j[i]} = 1 \mid X_{i,j[i],1:K}\right\}$, which represents the probability that the respondent's true address is the same as the AR address $j[i]$ for person i on a specified date. The Binary PPM is fit as if the data are independent binary outcomes. Estimated probabilities are then normalized by person to approximate multinomial probabilities. However, this approach is not a valid statistical model except for $J[i] = 2$. This model is also not able to estimate the probability that none of the AR addresses are a match as specified.

Logistic regression is used to estimate the strength of association between the AR sources, AR source observation date differences, and address categories. The model is as follows

$$\text{logit}P\{Y_{i,j[i]} = 1|w_{i,j[i],1:S}, m_{i,j[i],1:S}\} = \theta_0 + \lambda_{J[i]} + \sum_{s=1}^{S} \rho_k w_{i,j[i],s} + \sum_{s=1}^{S} \mu_k m_{i,j[i],s}. \quad (4.20)$$

Table 4.15 defines each term of this model. Though individual-level attributes are available, they are not used in this analysis.

Table 4.15: Binary PPM model specification.

Parameter	Definition
$Y_{i,j[i]}$	Indicator variable denoting person i address matches AR source address ID $(i, j[i])$
$w_{i,j[i],1:S}$	Vector of indicator variables indicating person i observed at address ID $(i, j[i])$ in AR source $s = 1, \ldots, S$
$m_{i,j[i],1:S}$	Vector of negative integer absolute date differences when person i observed in AR source $s = 1, \ldots, S$ relative to their ground truth event
θ_0	Binary PPM intercept parameter
$\lambda_{1:\max J[i]}$	Binary PPM parameter vector denoting the association between total number of AR addresses for person i, $J[i]$, and $\text{logit}P\left\{Y_{i,j[i],s} = 1\|w_{i,j[i],s}, m_{i,j[i],s}\right\}$
$\rho_{1:S}$	Binary PPM parameter vector denoting the strength of association between AR source s observation and ground truth event for $\text{logit}P\left\{Y_{i,j[i],s} = 1\|w_{i,j[i],s}, m_{i,j[i],s}\right\}$
$\mu_{1:S}$	Binary PPM parameter vector denoting the strength of association between AR source s negative absolute difference in months between AR source s observation and ground truth event for $\text{logit}P\left\{Y_{i,j[i],s} = 1\|w_{i,j[i],s}, m_{i,j[i],s}\right\}$

Once fit, Binary PPM probability estimates, $\hat{P}\left\{Y_{i,j[i]} = 1|w_{i,j[i],1:S}, m_{i,j[i],1:S}\right\}$, are used to approximate multinomial probabilities by dividing each person's estimated probabilities

by the sum of their probabilities across all AR addresses, i.e.

$$\tilde{P}\left\{Y_{i,j[i]} = 1 | w_{i,j[i],1:S}, m_{i,j[i],1:S}\right\} = \frac{\hat{P}\left\{Y_{i,j[i]} = 1 | w_{i,j[i],1:S}, m_{i,j[i],1:S}\right\}}{\sum_{j[i]=1}^{J[i]} \hat{P}\left\{Y_{i,j[i]} = 1 | w_{i,j[i],1:S}, m_{i,j[i],1:S}\right\}}. \tag{4.21}$$

This approximation introduces a well known misspecification of the true multinomial probability point estimates and associated uncertainty around these point estimates (MCCullagh & Nelder, 1989; Agresti, 2002).

The Binary PPM was fit using logistic regression as if a person-address match is an independent binary outcome, replicating the approach described in (Brown et al., 2023; Keller et al., 2018; Morris, 2014).

Two Stage Person-Place Model

Rastogi & O'Hara (2012) and Morris (2014) note that early versions of the Binary PPM included multiple stages to account for different AR address categories. We specify the Two Stage PPM to explicitly estimate the probability that none of the AR source addresses for an individual match the address a person might list when responding to a census or survey. While this category is implicitly encoded in the Binary PPM specification, the Two Stage PPM estimates the probability that none match in Stage 1 and, given that a match exists, estimates the probability that one of the AR addresses match the address one might list when responding to a census or survey. The Two Stage PPM estimates the Stage 1 probability in terms of individual-level attributes and Stage 2 is estimated in the same way as the Binary PPM.

We use a simple model to demonstrate the Two Stage PPM. We simulate the strength of the association between $\text{logit}P(Z_i = 1)$ in terms of the number of AR addresses for person i so that larger $J[i]$ is associated with larger $\text{logit}P(Z_i = 1)$. The total number of AR addresses for person i, $J[i]$, is coded as a categorical variable with $\max J[i] = 5$ levels in total. Stage 2

uses the same model specification as the Binary PPM. The full Two Stage PPM is as follows.

$$\textbf{Stage 1: } \text{logit}P\left\{Z_i = 1\right\} = \delta_0 + \delta_{J[i]} \tag{4.22}$$

$$\textbf{Stage 2: } \text{logit}P\left\{Y_{i,j[i]} = 1 | Z_i = 1\right\} = \omega_0 + \kappa_{J[i]} + \sum_{s=1}^{S} \tau_s w_{i,j[i],s} + \sum_{s=1}^{S} \psi_s m_{i,j[i],s} \tag{4.23}$$

The Two Stage PPM is estimated from subsets of the full Person-Place data. Stage 1 is estimated at the individual-level with outcome variable Z_i denoting that person i's true address is present in the AR source addresses. Stage 2 is estimated from the subset of Person-Place data where each person's true address is known to be in the AR source addresses.

Adding Stage 1 in the Two Stage PPM improves on the Binary PPM specification (explicit estimation of the *None* category), but retains the binary misspecification issues of the Binary PPM in Stage 2. When the use case calls for it, approximating a multinomial probability at the person-level becomes more complicated. For example, suppose person i has $J[i] = 2$. Then the normalized probability of matching *None* and addresses (i,j[i]=1) and (i,j[i]=2) is as follows:

$$\left(\hat{P}\left\{Z_i = 0\right\}, \frac{\hat{P}\left\{Y_{i,j[i]=1} = 1 | Z_i = 1\right\} \hat{P}\left\{Z_i = 1\right\}}{\sum_{j[i]=1}^{J[i]} \hat{P}\left\{Y_{i,j[i]} = 1 | Z_i = 1\right\}}, \frac{\hat{P}\left\{Y_{i,j[i]=2} = 1 | Z_i = 1\right\} \hat{P}\left\{Z_i = 1\right\}}{\sum_{j[i]=1}^{J[i]} \hat{P}\left\{Y_{i,j[i]} = 1 | Z_i = 1\right\}} \right). \tag{4.24}$$

Table 4.16 defines each term of the Two Stage PPM model.

The Two Stage PPM uses logistic regression to estimate the probability that any AR address matches simulated census/survey responses. The second stage is estimated using the Binary PPM.

4.5 Results

All PPM model formulations were evaluated using the simulated Person-Place data with 50,000 simulated individuals and 87,594 person-address pairs. Section 4.5.1 assesses the level of agreement between the simulated and observed data characteristics used to fit the

Table 4.16: Two Stage PPM model specification.

Parameter	Definition
Z_i	Binary random variable denoting true address is present in AR sources observed for person i
$Y_{i,j[i]}$	Indicator variable denoting person i survey or census address matches AR source address ID $(i, j[i])$
$w_{i,j[i],1:S}$	Vector of indicator variables indicating person i observed at address ID $(i, j[i])$ in AR source $s = 1, \ldots, S$
$m_{i,j[i],1:S}$	Vector of negative integer absolute date differences when person i observed in AR source $s = 1, \ldots, S$ relative to their ground truth event
ω_0	Stage 2 PPM intercept parameter
$\kappa_{1:\max J[i]}$	Stage 2 PPM parameter vector denoting the association between total number of AR addresses for person i, $J[i]$, and $\text{logit}P\left\{Y_{i,j[i],s} = 1 \mid Z_i = 1, w_{i,j[i],s}, m_{i,j[i],s}\right\}$
$\tau_{1:S}$	Stage 2 PPM parameter vector denoting the strength of association between AR source s observation and ground truth event for $\text{logit}P\left\{Y_{i,j[i],s} = 1 \mid Z_i = 1, w_{i,j[i],s}, m_{i,j[i],s}\right\}$
$\psi_{1:S}$	Stage 2 PPM parameter vector denoting the strength of association between AR source s negative absolute difference in months between AR source s observation and ground truth event for $\text{logit}P\left\{Y_{i,j[i],s} = 1 \mid Z_i = 1, w_{i,j[i],s}, m_{i,j[i],s}\right\}$
$\delta_{0:\max J[i]}$	Stage 1 parameter vector

models. Section 4.5.2 assesses the performance of all model PPM specifications.

4.5.1 Simulated Data

Table 4.17 shows the simulated distribution of the number of AR data sources compared to the percentage of the population observed in (Brown et al., 2023). The greatest departures occur for the percent of the population with two and three AR sources. Our simulation led to approximately eight percent more of the population with two AR sources compared to the real records and approximately six percent less of the population with three AR sources. This is sensible considering that we only simulate five AR sources where the real data include more than thirty sources. All others were within three percentage points of the observed data.

Table 4.18 shows the simulated distribution of the number of AR data sources compared to the percentage of the population observed in (Brown et al., 2023). The simulated percentage of the population found in each AR source was within three percentage points for all sources

Table 4.17: Observed distribution of number of AR data sources from (Brown et al., 2023) table 26 compared to simulated data in percent of people.

Number of AR data sources	Observed	Simulated
1	38.5	37.7
2	19.2	27.5
3	20.0	14.8
4	11.4	8.8
5+	13.7	11.2

except IRS 1099. Our simulation led to approximately thirteen percent less of the population in source IRS 1099 compared to the observed data. The relative order of both the Percent in source and Percent in source only remain unchanged, however. The percentage of the population simulated in IRS 1040 only was more than ten percent higher in our simulation compared to the observed data, but the share of sole AR source filings was within two percentage points of the observed data in all other cases. This outcome is sensible since we simulate a subset of the AR sources available in (Brown et al., 2023).

Table 4.18: Observed distribution of AR observations from (Brown et al., 2023) table 27 compared to simulated data.

		Percent in source		Percent in source only	
s	Source	Observed	Simulated	Observed	Simulated
1	IRS 1040	76.4	79.6	21.3	31.6
2	IRS 1099	60.4	47.3	4.4	4.7
3	3^{rd} Party	41.6	38.1	4.9	3.0
4	CMS	15.1	13.5	0.1	0.2
5	NCOA	4.7	7.9	1.7	0.5

Table 4.19 shows the simulated percentage of the population with 1-5+ addresses. The simulated distribution of the population's address counts were all within 0.2 percentage points of the observed data and best approximation to the observed data compared to the other tables used for simulation.

Table 4.20 shows the simulated distribution of address counts and percentage of no AR

Table 4.19: Observed distribution of number of addresses from (Brown et al., 2023) table 38 compared to simulated data in percent of people.

Number of addresses	Observed	Simulated
1	53.0	53.2
2	28.2	28.0
3	11.8	11.7
4	4.6	4.6
5+	2.4	2.5

address match (i.e., *None* category) in each group. Though Brown et al. (2023) does not include an estimate of the no match rates, it was simulated to roughly increase with the number of AR addresses.

Table 4.20: Simulated distribution of AR address count and percentage of population with address count $J[i]$ in AR address category *None*.

$J[i]$	$J[i]$ Distribution	Percent No Match | $J[i]$
1	53	14
2	28	12
3	12	45
4	5	44
5	2	53
All	100	19

Tables 4.17 - 4.19 shows our simulated data resemble the real data summary statistics published in (Brown et al., 2023)

4.5.2 Person-Place Model Performance

Well-calibated, interpretable Person-Place Model probabilities are essential to policymaking and operational awareness. Reliability diagrams are used to assess the calibration of a probabilistic model (Berrocal et al., 2008). Reliability diagrams are created by binning estimated PPM probabilities and computing the fraction of AR address matches for each probability

bin. If the model is well-calibrated, then the midpoint of each bin should approximate the fraction of address matches in its bin.

Figure 4.1 shows the reliability diagram for each PPM specification and estimated R^2 between the binned probability and probability bin midpoint. Binary PPM reliability diagrams before and after normalization are shown in the first row (a-b). We refer to Binary PPM probabilities that are not rescaled to sum to one within an individual as the *raw* probabilities. Binary PPM probabilities that have been rescaled to sum to one are referred to as the *normalized* probabilities. Raw Binary PPM probabilities were well-calibrated except for the bin that included probability zero. The Binary PPM assigns category *None* probability zero, and this misspecification explains the departure from the y=x line near Binary PPM probability 0. Variation in the normalized Binary PPM probabilities increased because the normalization fails to account for people assigned the *None* category.

The Two Stage PPM reliability diagram is shown in Figure 4.1(c). Including the *None* category in Stage 1 of the model and normalizing the estimated model probabilities as described in Section 4.4 improves probability calibration of probabilities less than 0.5; however, the Two Stage PPM underestimates most probabilities larger than 0.5.

The Bayes PPM reliability diagram is shown in Figure 4.1(d) and was best calibrated among the four specifications. The Bayes PPM was fit with and without the overdispersion parameter, σ, to assess the impact of ignoring overdispersion when it is present. Figure 4.1 suggests that ignoring overdispersion leads to nearly indistinguishable calibration compared to the more complicated implementation that properly accounted for it. However, Table 4.21 and Table 4.22 show that inference without accounting for overdisperson in the data results in incorrect parameter estimates and inference.

Table 4.21 shows the true parameter value, posterior quantiles, and an indicator of whether the true parameter is contained in the 90% posterior sample quantile for the Bayes PPM implementation that includes an overdispersion parameter, σ. The Bayes PPM model successfully recovered all model parameters.

Table 4.22 summarizes the the Bayes PPM implementation results without including the

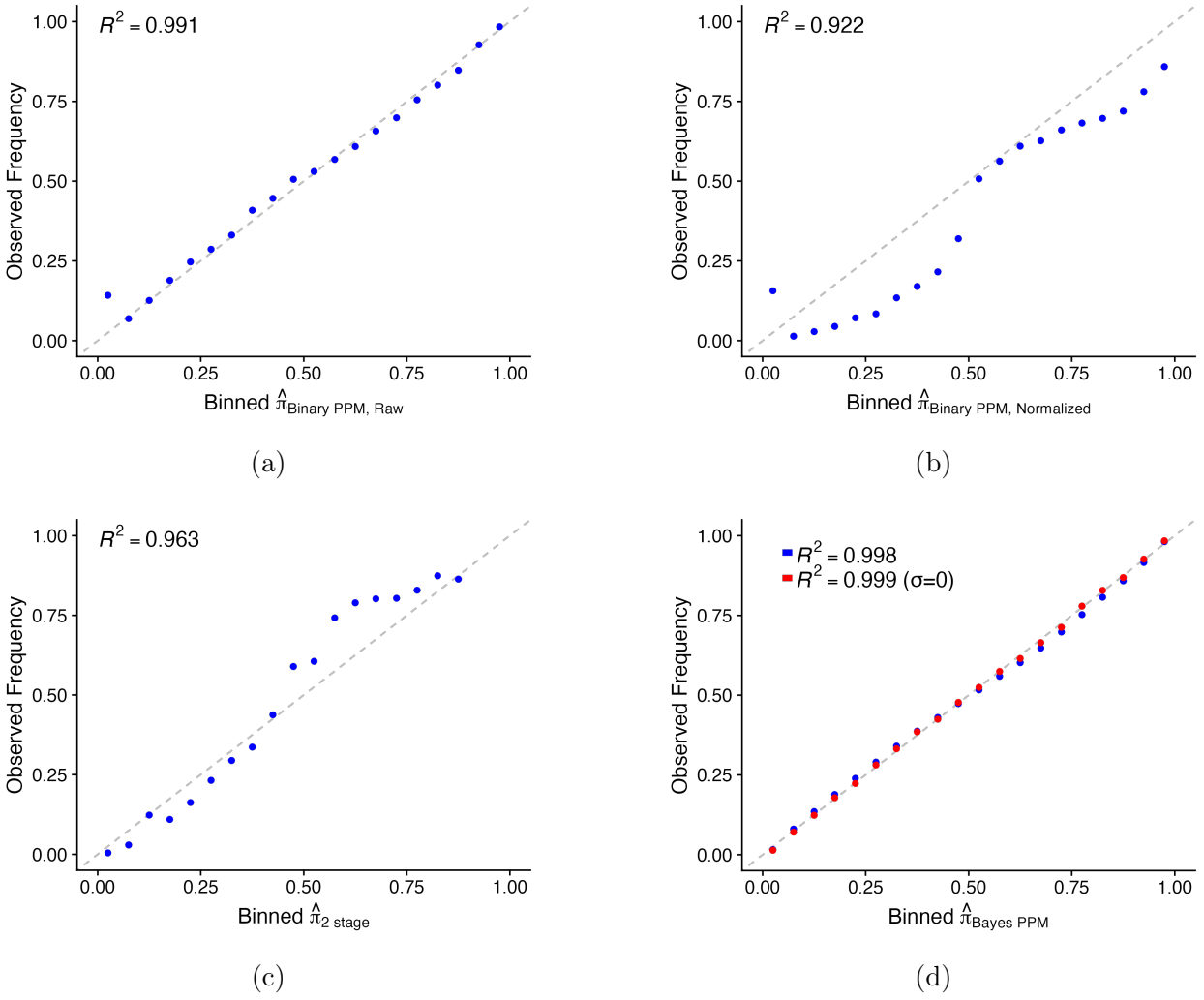


Figure 4.1: Reliability diagram for (a) Binary PPM probabilities without normalization, (b) Binary PPM probabilities with normalization so that the probabilities for a person sum to one across all AR source addresses, (c) Two Stage PPM, and (d) Bayes PPM with and without accounting for overdispersion for $\max J[i] = 5$ and $N = 50,000$.

overdispersion parameter, σ. Less than a third (7 of 15) of the 90% posterior sample quantiles include the true Bayes PPM parameter value; however, the misspecified implementation leads to median parameter estimates that were close to the true values. The parameter estimates without overdispersion leads to *None* category estimates that are too large compared to the

Table 4.21: Simulated Bayes PPM parameter values, Bayes PPM posterior medians, 5 and 95 percent posterior sample quantiles, and an indication of whether the simulated parameter value is contained in the 90 percent posterior sample quantile.

Parameter	Value	50%	5%	95%	Recovered
α_1	-3.40	-3.27	-3.43	-3.17	True
α_2	-2.24	-2.17	-2.29	-2.07	True
α_3	1.64	1.66	1.56	1.77	True
α_4	2.02	2.05	1.93	2.18	True
α_5	2.89	2.79	2.64	2.94	True
β_1	2.99	2.96	2.86	3.09	True
β_2	3.74	3.74	3.63	3.90	True
β_3	1.35	1.34	1.26	1.43	True
β_4	2.79	2.67	2.53	2.84	True
β_5	2.29	2.16	2.00	2.31	True
γ_1	2.08	2.04	1.98	2.12	True
γ_2	2.34	2.32	2.25	2.41	True
γ_3	1.47	1.44	1.39	1.51	True
γ_4	2.11	2.03	1.95	2.13	True
γ_5	0.37	0.35	0.32	0.39	True
σ	0.40	0.27	0.03	0.47	True

true values for individuals with 1-2 AR addresses (α_1 and α_2) and too large for individuals with 3 AR addresses (α_3), estimates that are too small compared to the true values for AR sources 1-2 (β_1 and β_2), and estimates of the change from baseline for a one month date difference decrease that are too small for AR sources 1-3 ($\gamma_1 - \gamma_3$). Ignoring overdisperson (σ) did not drastically change the inference or probability calibration in this case, but both the median parameter estimates and posterior quantile estimates were impacted.

Although the Binary PPM offers a useful approximation to model-based probabilities, it is inherently misspecified. Consequently, parameter interpretations from the Binary PPM should be treated with caution. Due to the interpretive limitations arising from this misspecification, Binary PPM parameter estimates are not evaluated.

Figure 4.2 shows the Markov Chain Monte Carlo (MCMC) posterior sample trace plots for each Bayes PPM parameter. Four chains were randomly initialized and run for a total of

Table 4.22: Simulated Bayes PPM parameter values, Bayes PPM posterior medians without including overdispersion parameter, 5 and 95 percent posterior sample quantiles, and an indication of whether the simulated parameter value is contained in the 90 percent posterior sample quantile.

Parameter	Value	50%	5%	95%	Recovered
α_1	-3.40	-3.20	-3.28	-3.13	False
α_2	-2.24	-2.12	-2.21	-2.03	False
α_3	1.64	1.63	1.54	1.72	True
α_4	2.02	2.01	1.90	2.12	True
α_5	2.89	2.73	2.60	2.85	False
β_1	2.99	2.90	2.83	2.97	False
β_2	3.74	3.66	3.58	3.74	False
β_3	1.35	1.31	1.24	1.39	True
β_4	2.79	2.62	2.50	2.75	False
β_5	2.29	2.11	1.97	2.25	False
γ_1	2.08	2.00	1.96	2.03	False
γ_2	2.34	2.27	2.23	2.31	False
γ_3	1.47	1.41	1.38	1.45	False
γ_4	2.11	1.99	1.93	2.05	False
γ_5	0.37	0.34	0.31	0.38	True

55,000 iterations. The first 5,000 iterations were discarded as burn-in. The remaining 50,000 iterations were thinned to 5,000 iterations to reduce autocorrelation. Trace plots indicate slow mixing the overdispersion parameter, σ. Potential scale reduction factors estimated with the Gelman-Rubin Diagnostic were all smaller than 1.1 for parameters except σ (1.39) (Gelman & Rubin, 1992). The Raftery-Lewis Diagnostic indicated that the 5,000 burn-in iterations were sufficient and that a minimum of 3,800 iterations were sufficient to estimate the 0.025 quantile with an accuracy of $\pm$ 0.005 with 0.95 probability for all parameters (Raftery & Lewis, 1992). However, the Raftery-Lewis diagnostic suggested the sampler should be run for another 60,000 or more iterations for most parameters.

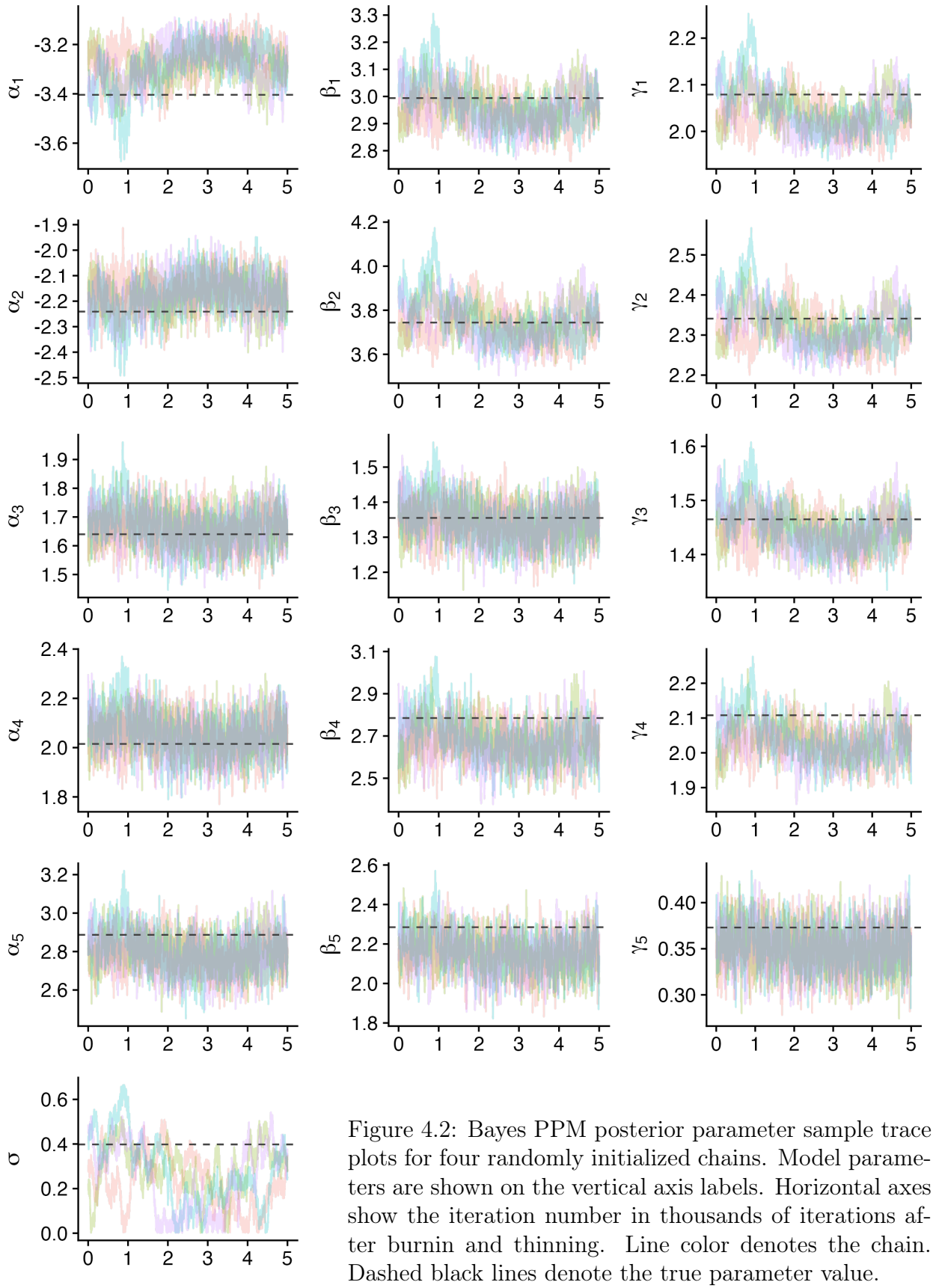
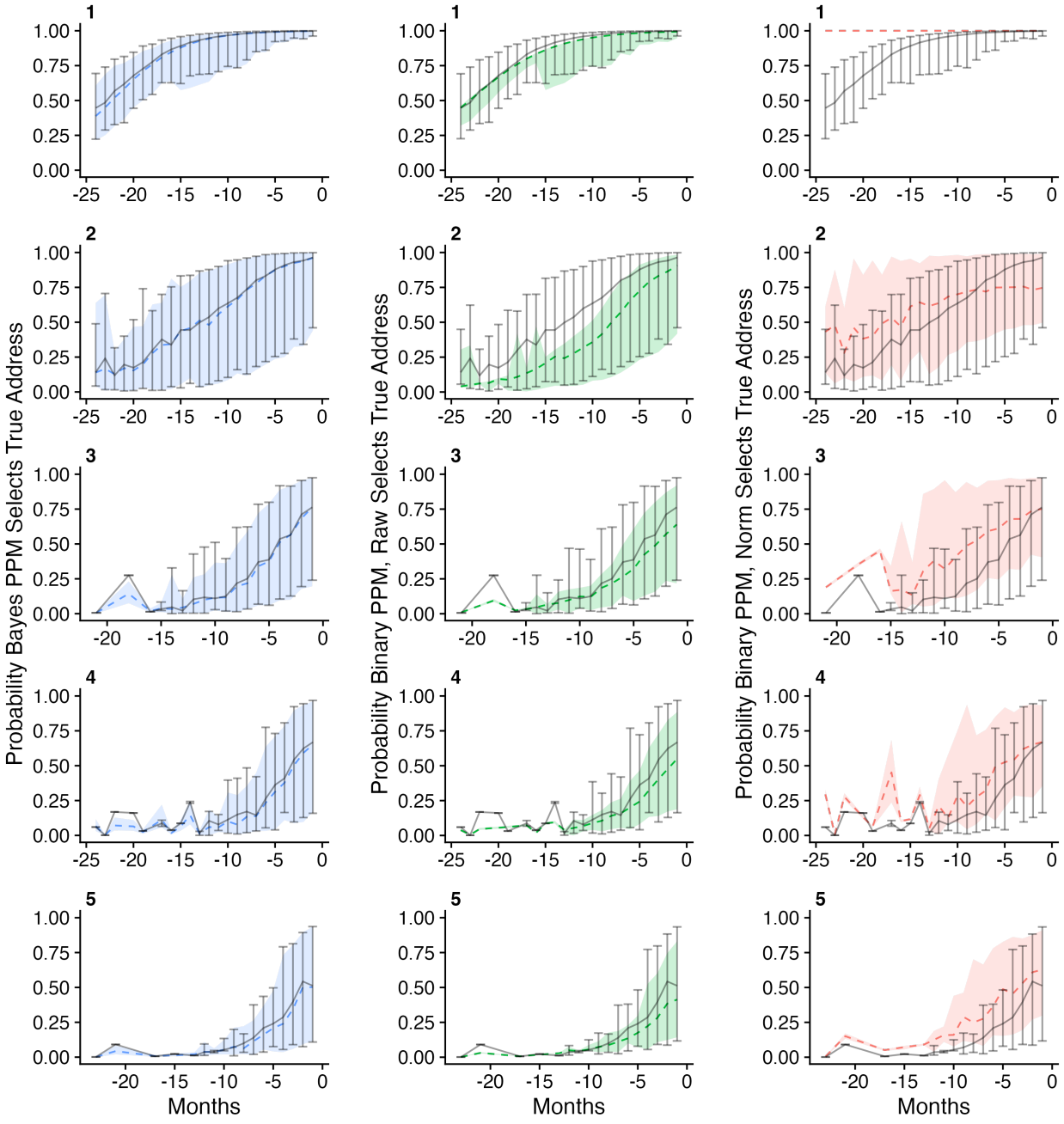


Figure 4.2: Bayes PPM posterior parameter sample trace plots for four randomly initialized chains. Model parameters are shown on the vertical axis labels. Horizontal axes show the iteration number in thousands of iterations after burnin and thinning. Line color denotes the chain. Dashed black lines denote the true parameter value.

Bayes PPM was implemented with the *NIMBLE* package (de Valpine et al., 2024a,b) in *R* (R Core Team, 2023). Parameters α, β, and γ were each sampled with independent Random Walk Block Metropolis-Hastings samplers. All other parameters were sampled with independent Random Walk Metropolis-Hastings samplers. Convergence, chain mixing, and successful recovery of all other simulated parameters were improved by the non-centered parameterization for σ (Papaspiliopoulos et al., 2003); however, this parameterization increased the sampler complexity and runtime substantially. All samplers used *Nimble*'s default settings.

Figure 4.3 summarizes the probability of selecting an individual's true address by the last AR observation's number of months prior to a ground truth event such as a census or survey response. Columns show the median (dashed colored lines) and range of the 90% probability quantiles (shaded regions) each month for the Bayes PPM, the Binary PPM without probability normalization, and the Binary PPM with normalized probabilities compared to the true probabilities. Probabilities are summarized by the number of AR addresses for each person in the simulated data. The solid line shows the median of the true probabilities by month and number of AR addresses. Vertical error bars show the true distribution of probabilities for each group. If a model is accurate, then the median probability (dashed line) should fall near the true probability median (grey solid line). If a model is well-calibrated, then the range of probabilities around the median (shaded regions) should span the true 90% quantiles for each group (vertical lines).

Figure 4.3(a) shows the distribution of the Bayes PPM probabilities matches both the true median match probability as well as the true 5% and 95% match probability quantiles across all time horizons. Match probabilities tended to increase as the negative absolute date difference in the last AR observation approached zero.



(a) Bayes PPM (b) Binary PPM, Raw (c) Binary PPM, Normalized

Figure 4.3: Probability of selecting an individual's true address (vertical axis) by the number of months prior to a ground truth event (horizontal axis) and number of AR addresses (rows) for the (a) Bayes PPM, (b) Binary PPM, Raw, and (c) Binary PPM, Normalized. Solid gray lines show the median true probabilities for each address number-month group. Dashed lines show the median estimated probability for each model. Shaded regions of each plot show the range of the 5%-95% model quantiles. Error bars denote the 5%-95% true probability quantiles for each address number-month group.

Figure 4.3(b) shows the distribution of the Binary PPM raw probabilities compared to the true median match probability as well as the true 5% and 95% match probability quantiles. The median Binary PPM probability estimate approximated the true probability medians for all address counts and last AR observed date differences except for the two AR address case prior to -10 months. However, upper quantiles for individuals with 1-2 addresses tended to be underestimated using the Binary PPM raw probabilities, impacting more than 80% of the population. Upper and lower quantiles were estimated by sampling the joint distribution of the Binary PPM model parameters estimated on the logit scale, calculating the logit match probability for each observation, and transforming the results back to the probability scale. A total of 1,000 samples were used to generate the 5%, 50%, and 95% quantiles shown in the figure.

Figure 4.3(c) shows the distribution of the Binary PPM normalized probabilities compared to the true median match probability as well as the true 5% and 95% match probability quantiles. The normalized Binary PPM probability equals 1 when there is only one address for any AR source date difference. This result is a vast departure from the true median probability and associated uncertainty shown by the grey lines on the figure. This error affects the majority of individuals since most (53% in Table 4.17) have just one unique AR address. There is considerable overlap in true and estimated quantiles for all other address counts, but the median probabilities estimated tended to be far from the median true probabilities across all date differences. Upper and lower quantiles were estimated using the same procedure as the raw probabilities, but quantiles were calculated after normalizing each sample to sum to one.

Figure 4.4 shows the range of the 5%-95% probability quantiles and median probability for individual's true address by number of unique AR addresses for the Bayes PPM, Binary PPM without normalization, and Binary PPM with normalization compared to the true probabilities. Median probability estimates from the Bayes PPM and the Binary PPM without normalization were closest to the true median probability across all address counts, except in the case of two addresses. In this case, the Binary PPM normalized probability

is closer to the true value than the Binary PPM raw probability but remains farther from the true value compared to the Bayes PPM. Bayes PPM 5% and 95% quantiles were also best aligned to the true values among the three methods. Normalized Binary PPM quantiles and medians were worst aligned to the true median, lower, and upper quantile probabilities. Normalized probabilities for individuals with one AR address equaled one for all quantiles, too narrow compared to the true range, and diverged from the true median as the number of AR addresses increased. Table 4.23 shows value of each model quantile by address count shown in Figure 4.4.

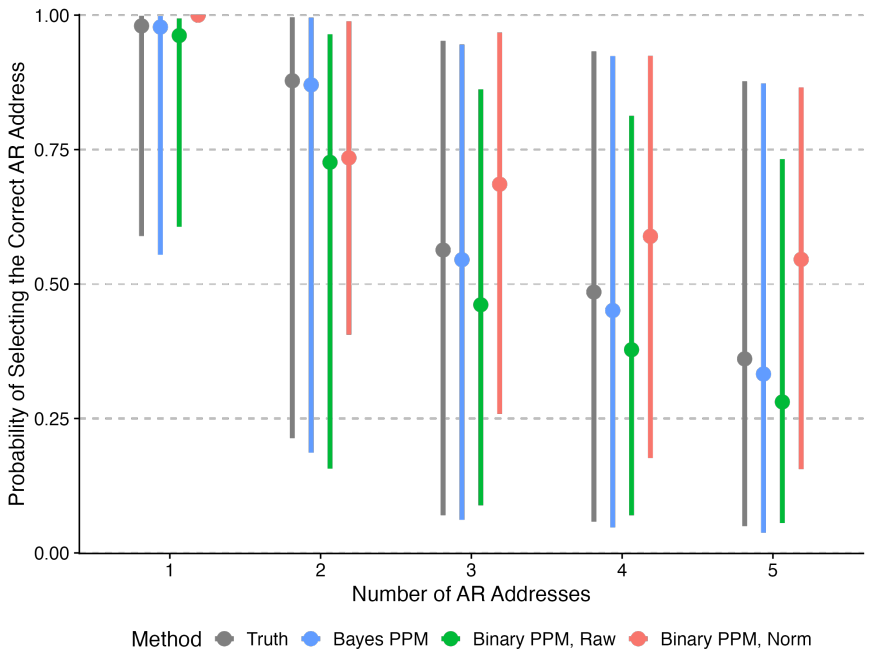


Figure 4.4: Probability of selecting the true address quantiles for Bayes PPM, Binary PPM without normalization, and Binary PPM with normalization compared to the true probabilities for each AR address count. Points show the median probability for each method. Lines show 5%-95% probability quantile range.

Table 4.23: Distribution of probabilities (as percentages) for the matched address using the true probabilities (Truth) compared to probability quantiles estimated using the Bayes PPM, Binary PPM without normalization, and Binary PPM with normalized probabilities by number of AR addresses.

	J=1			J=2			J=3			J=4			J=5		
Method	5%	50%	95%	5%	50%	95%	5%	50%	95%	5%	50%	95%	5%	50%	95%
Truth	59	98	100	21	88	100	7	56	95	6	49	93	5	36	88
Bayes PPM	55	98	100	19	87	100	6	55	95	5	45	92	4	33	87
Binary PPM, Raw	61	96	99	16	73	96	9	46	86	7	38	81	6	28	73
Binary PPM, Norm	100	100	100	41	73	99	26	69	97	18	59	92	16	55	87

Note that two-stage model results are omitted from Figure 4.3 and Figure 4.4 to maintain figure legibility and focus on the Binary PPM as the method of record. However, for a basis of comparison, two-stage results are included in Figure 4.1 and Table 4.24. A more comprehensive analysis of two-stage model results may be explored in future work.

Table 4.24 shows the in-sample accuracy, precision, recall, mean log loss (MLL), mean Brier Score, and Mean Absolute Error (MAE) for the Bayesian PPM, Two Stage PPM, and Binary PPM with and without normalization. Accuracy, precision, and recall were determined by the maximum predicted address probability for each individual. Binary PPM metrics were calculated using the raw and normalized probabilities. All other metrics were calculated using the estimated probabilities for each model. Accuracy, recall, and precision for the raw Binary PPM probabilities were calculated using a threshold value of 0.5, allowing for no matches if no raw probability is greater than 0.5 for any of an individual's addresses or multiple address assignments for an individual with more than one address with raw probability greater than 0.5 to demonstrate a commonly used threshold approach to calculating these metrics in this context. (Note that the max probability approach used for the raw probabilities would yield the same accuracy, recall, and precision for the normalized model since scaling does not change the probability order.) The Binary PPM assigns probability 0 to the *None* category, rendering MLL values equal to negative infinity. Table 4.24 also calculates accuracy, precision, and recall that results from assigning individuals to the address with the smallest absolute difference between their ground truth event and AR observations

(Closest in Time). The *Last Observed* method had the worst performance across all metrics except total compute time. This result is noteworthy, as one might intuitively expect that prioritizing the most recent observation would perform well. Instead, this finding highlights a key strength of the PPM specifications: their ability to weigh different sources appropriately, even when a more recent observation is available, thereby improving overall accuracy.

Table 4.24: In-sample accuracy, precision, recall, mean log loss (MLL), mean Brier Score, Mean Absolute Error (MAE), and runtime in minutes for the Bayesian PPM, Two Stage PPM, and Binary PPM with and without normalization. The 'Last Observed' method places a person at the address where the most recent AR was observed.

Method	Accuracy	Recall	Precision	MLL	Brier	MAE	Time
Bayes PPM	87	83	83	-0.15	0.25	0.04	110
Bayes PPM, $\sigma = 0$	87	83	83	-0.14	0.24	0.04	49
Two Stage PPM	83	77	77	-0.21	0.34	0.15	2
Binary PPM, Norm	82	75	75	-Inf	0.42	0.15	2
Binary, Raw	83	67	84	-Inf	0.38	0.12	2
Last Observed	51	40	71				<1

The Bayes PPM estimated with and without accounting for overdispersion performed best across all evaluation metrics except runtime. Accuracy, recall, and precision were nearly indistinguishable with and without estimating the overdispersion parameter. Perturbations around the mean probabilities induced by overdispersion in the full Bayes PPM led to slightly worse MLL, Brier Score, and MAE estimates compared to the Bayes PPM estimated without overdispersion. However, Table 4.21 and Table 4.22 show that ignoring overdispersion in this case would yield slightly different inferences. Furthermore, recovery of the correct uncertainty intervals around the median probability shown in Figure 4.3 and Figure 4.4 show that including a parameter to model overdispersion was neccessary since it was present by construction. In summary, overdispersion led to more variable probabilities around the mean and hence contributed a wider range of errors to the MLL, Brier Score, and MAE metrics as a result. Ignoring overdispersion might yield more concentrated probabilities around the true value, but inferences and uncertainty around the mean were misspecified as a result.

Including an overdispersion term is a key strategy for addressing potential violations of the Independence of Irrelevant Alternatives (IIA) assumption in discrete choice models. This addition enhances model flexibility by allowing for greater heterogeneity in choice probabilities. However, the overdispersion term can be omitted to reduce computational complexity if there is strong justification for the IIA assumption, such as context-specific considerations or consistent parameter estimates obtained through alternative approaches like nested logits. For a comprehensive discussion of these issues, see (Bruch & Mare, 2012).

4.6 Discussion

The Binary PPM and Two Stage PPM offer fast approximations to a well-defined statistical model. However, the binary approximations currently in use are incapable of yielding valid statistical inference or correct uncertainty intervals around these approximations. The Bayes PPM provides one approach to overcome these limitations.

The Bayes PPM eliminates the need to normalize probabilities. Normalizing probabilities estimated from a model that assumes independence between person-address pairs leads to distortions in estimated probabilities. However, raw probabilities estimated from a binary model are not interpretable at the person-level. Still, estimates of uncertainty around raw Binary PPM probabilities are massively understated as estimated, leaving researchers and policymakers alike with no measure of variation around point estimates. The Bayes PPM yields the correct probability point estimates and measures of uncertainty by design (as well as a representation of the full generative distribution). These advantages offer the potential for deeper insights into estimates of Person-Place probabilities and lay the foundation for a method capable of generating forecasts of probabilities with the correct levels of uncertainty.

Estimating the *None* category alongside specific AR addresses represents another benefit of the Bayes PPM. We did not include individual- or house- level covariates in our model other than the number of AR addresses observed for an individual. More sophisticated Bayes PPM specifications could do so, however. Such specifications are possible, but could become difficult for a broad user base to reason through. A multi-stage approach with an

individual-level model in stage 1 with its own set of covariates could be specified to estimate the probability that an AR address match exists. The Bayes PPM could then be used in stage 2 to estimate the probability of the AR match given the outcome of stage 1. Such a model would require careful derivations of the joint model implied by each stage though and could be difficult to maintain.

Computational complexity and prior distribution tuning nuances are the primary disadvantages of the Bayes PPM. We demonstrated Bayes PPM with 50,000 simulated individuals, but millions of records could be used to estimate these models. Fitting the Bayes PPM to millions of records is not practical. It is also unlikely to be necessary unless many more covariates are needed to adequately estimate the AR source address match probability, leading to ever smaller subgroups with associated interdependencies to include. A model using only a handful of AR sources avoids these challenges and the need for sample sizes on the order of millions for accurate estimation. Deploying the Bayes PPM for real-world use would require an assessment of the sample size needed to effectively model probabilities for the subgroups of interest.

We proposed the Bayes PPM to address deficiencies in the Binary PPM, but recently developed alternatives could be used under a different formulation of the AR match use case. For example, Long et al. (2022) propose a model to estimate the probability that an individual remains at their last observed AR address. This approach is promising since it effectively estimates the probability that a person remains at the last observed AR address or not (i.e., the *None* category). However, there are advantages to estimating the reliability of an AR address match and source-specific rates of decay jointly since some AR sources are more predictive and longer-lasting than others. Such advantages should be retained.

The Bayes PPM builds on extant PPM specifications with more accurate probability estimates, properly calibrated uncertainty intervals, and a valid model that can support statistical inference with AR data. For a single-day-focused enumeration like the decennial census, this yields more trustworthy probability estimates for operations. For broader applications, such as expansive research programs like the Demographic Frame, the Bayes

PPM facilitates more robust longitudinal analyses that can now include more reliable estimates of uncertainty across such analyses. This also applies to demographic insights for subpopulations and smaller geographic units, which are often underrepresented or less reliably captured by survey operations. Bayes PPM enhancements pave the way for more robust demographic forecasting and policy analysis, supporting both immediate enumeration efforts and longer-term research objectives.

Following (National Academies of Sciences, Engineering, and Medicine, 2023) recommendations, we started by simulating a data set with characteristics as similar as possible to the data used to fit the Binary PPM, generated outcomes from known parameters, and demonstrated that our method/implementation recovered the true parameters of interest. Our simulation shows that the Binary PPM and Two Stage PPM methods led to tractable approximations of the full generative model, but failed to recover the characteristics of the underlying data generating mechanism. We demonstrated the viability of a simple Bayes PPM specification with simulated data, but a more robust evaluation with real-world data should be carried out in the future.

There are sources of variability not fully addressed by our data simulation, particularly differences in the timing of source data updates and the seasonality inherent in certain data sources, such as IRS filings concentrated around April each year. While the Bayes PPM specification accounts for the baseline association and the rate of change from this baseline, our simulation does not explicitly model variability in the timing of source updates. To address this in practice, predictor variable codings should be designed to align with the least frequent update schedule among all sources. This ensures that the linear slope parameters, which capture differences in the baseline associations for source indicators, remain comparable across all sources by operating on a consistent temporal scale.

The Binary PPM leads to a fast approximation to the mean probabilities. However, this approximation to the data generating process leads to critical limitations that can only be overcome by a more robust model specification. The Bayes PPM specification recovers the correct probabilities, associated uncertainty, and interpretable probabilities without

post-processing. By addressing the Binary PPM limitations and providing a more flexible and accurate representation of the data, Bayes PPM offers one approach to more robust demographic estimates and administrative records-based research.

Chapter 5

CONCLUSION

5.1 Contributions to Research

This dissertation makes several contributions to statistical demography, particularly in the areas of migration forecasting and demographic estimation.

In Chapter 2, a Bayesian hierarchical model is introduced for forecasting global bilateral migration flows among the 200 most populous countries. This model leverages historical data to produce well-calibrated probabilistic projections of migration flows, accounting for uncertainty. The model's appeal lies in its focus on the population at risk of migration, separating outflow and inflow model components to enable parallel processing, and pooling information among 39,800 bilateral flows to overcome the challenge of having only a few periods of data. These insights led to improved predictive accuracy and uncertainty intervals compared to existing models. This advancement provides policymakers and researchers with a more reliable method for understanding and forecasting international migration patterns.

Chapter 3 addresses the critical gap in migration modeling by incorporating the role of population age structure in forecasting migration flows. The Migration Age Structure Index (MASI) is introduced as a novel method for adjusting net migration rates based on the age distribution of past and future populations. By accounting for the varying migration propensity across different age groups, the MASI offers a novel single-number summary of the whole force of migration distribution over time. Removing variation in historic and forecast migration rates using the MASI reduces net migration forecast uncertainty and surprisingly showed that current projections of steep population contraction are somewhat overstated.

Chapter 4 develops the Bayesian Person-Place Model (Bayes PPM) for improving residency estimation using administrative records, addressing the limitations of existing models

and providing a robust statistical foundation for demographic analyses. This model enhances the accuracy of demographic estimates and the reliability of uncertainty intervals, making it particularly valuable for population estimation in countries without population registers. The Bayes PPM offers a much more robust model as for countries expanding their use of administrative records to enhance intercensal population estimation, improve current survey and census operations, or even transition to an administrative record first enumeration design for certain subgroups or writ large.

Collectively, these contributions advance the state of the art in migration forecasting and demographic estimation, providing new methods to improve the accuracy and reliability of population estimates and projections.

5.2 Future Work

This dissertation addresses fundamental challenges in migration forecasting and residency imputation. Many directions, however, were left unexplored.

Migration Flow Forecasting The Bayesian Flow Model provides well-calibrated probabilistic forecasts based on 5-year bilateral flow estimates among the 200 largest countries. While these forecasts improve on the alternative methods, forecasts were generally concentrated around the long-term mean flow between two countries with wide uncertainty intervals. This is reassuring in some ways, but the Bayesian Flow Model offered no mechanism to integrate information within a period to sharpen the forecast. Furthermore, the Bayesian Flow Model does not include any predictor variables to help reduce forecast uncertainty in origin country outflow magnitudes or changes in destination country desirability or policy changes that could increase or constrain flows. Furthermore, the 5-year intervals may prove less helpful to near-term decisions. Future research should address these constraints as they make forecasts more timely and better reflect regional shifts in migration propensity due to economic, political, or environmental changes/shocks.

Age-Standardized Migration Forecasting The age-standardized net migration forecasts used a single Rogers & Castro-like migration age schedule for all countries outside of Gulf Cooperation Council (GCC) countries and calculate the MASI for inflows in terms of the global age distribution. These are both sensible in light of the slow rate of change in the population age structure and unknown changes in the origin country composition far in the future. However, calculating the MASI in terms of country-specific in-migration age schedules and the age composition of likely origin countries could offer further benefits.

Residency Imputation While the Bayesian Person-Place Model (Bayes PPM) shows promising results using simulated data, future research should apply this model to real-world administrative records to validate its performance in operational settings. The computational complexity of the Bayes PPM limits its scalability when applied to large datasets. Future work could focus on optimizing the model for use with much larger administrative record data sets. More sophisticated model specifications should also be explored for the *None* category assignment. This dissertation only focused on the probability of the *None* category in terms of the number of unique administrative records addresses observed for an individual, but higher fidelity model specifications should be explored.

These areas represent exciting opportunities for further advancing migration forecasting and demographic estimation, contributing to more flexible policymaking tools in an increasingly complex global landscape.

REFERENCES

Abel, G. J. (2013a). Estimating Global Migration Flow Tables Using Place of Birth Data. *Demographic Research*, *28*, 505–546.

Abel, G. J. (2013b, January-June). Estimating Global Migration Flow Tables Using Place of Birth Data. *Demographic Research*, *28*, 505-546.

Abel, G. J., & Cohen, J. E. (2019, June). Bilateral International Migration Flow Estimates for 200 Countries. *Nature*, *6*(82), 1-13.

Abel, G. J., DeWaard, J., Ha, J. T., & Almquist, Z. W. (2021). The Form and Evolution of International Migration Networks, 1990–2015. *Population, Space and Place*, *27*(3), e2432. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/psp.2432 doi: https://doi.org/10.1002/psp.2432

Acosta, R. J., Kishore, N., Irizarry, R. A., & Buckee, C. O. (2020). Quantifying the dynamics of migration after Hurricane Maria in Puerto Rico. *Proceedings of the National Academy of Sciences*, *117*(51), 32772–32778.

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). Hoboken, New Jersey: John Wiley & Sons.

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. UK: Chapman and Hall, Ltd.

Alexander, M., Polimis, K., & Zagheni, E. (2020). Combining Social Media and Survey Data to Nowcast Migrant Stocks in the United States. *Population Research and Policy Review*, *41*. Retrieved from https://doi.org/10.1007/s11113-020-09599-3 doi: 10.1007/s11113-020-09599-3

Alkema, L., Raftery, A. E., Gerland, P., Clark, S. J., Pelletier, F., Buettner, T., & Heilig, G. K. (2011). Probabilistic Projections of the Total Fertility Rate for All Countries. *Demography*, *48*(3), 815–839.

Australian Bureau of Statistics. (2020, October). *Assessing Administrative Data Quality to Enhance the 2021 Census.* online. Retrieved from https://www.abs.gov.au/statistics/research/assessing-administrative-data-quality-enhance-2021-census

Axelson, M., Holmberg, A., Jansson, I., & Westling, S. (2021). A register-based census: The swedish experience. In (pp. 179–204). Wiley Online Library.

Azose, J. J., & Raftery, A. E. (2015a). Bayesian Probabilistic Projection of International Migration. *Demography*, *52*(5), 1627–1650.

Azose, J. J., & Raftery, A. E. (2015b). Bayesian Probabilistic Projection of International Migration. *Demography*, *52*(5), 1627–1650.

Azose, J. J., & Raftery, A. E. (2019a). Estimation of Emigration, Return Migration, and Transit Migration Between All Pairs of Countries. *Proceedings of the National Academy of Sciences*, *116*(1), 116–122.

Azose, J. J., & Raftery, A. E. (2019b). Estimation of Emigration, Return Migration, and Transit Migration Between All Pairs of Countries. *Proceedings of the National Academy of Sciences*, *116*(1), 116–122.

Azose, J. J., Ševčíková, H., & Raftery, A. E. (2016a). Probabilistic Population Projections with Migration Uncertainty. *Proceedings of the National Academy of Sciences*, *113*(23), 6460–6465.

Azose, J. J., Ševčíková, H., & Raftery, A. E. (2016b). Probabilistic Population Projections with Migration Uncertainty. *Proceedings of the National Academy of Sciences*, *113*(23), 6460–6465.

Bahar, D., Piccone, T., & Trinkunas, H. (2018). Venezuela: A Path Out of Misery. *Brookings Policy Brief*.

Baker, J., Swanson, D., & Tayman, J. (2021). The Accuracy of Hamilton–Perry Population Projections for Census Tracts in the United States. *Population Research and Policy Review*, *40*, 1341–1354.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Berrocal, V. J., Raftery, A. E., & Gneiting, T. (2008). Probabilistic Quantitative Precipitation Field Forecasting Using a Two-stage Spatial Model. *The Annals of Applied Statistics*, *2*(4), 1170 – 1193. Retrieved from https://doi.org/10.1214/08-AOAS203 doi: 10.1214/08-AOAS203

Bhat, C. R., & Guo, J. (2004). A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling. *Transportation Research Part B: Methodological*, *38*(2), 147–168.

Bijak, J. (2006). *Forecasting International Migration: Selected Theories, Models, and Methods* (CEFMR Working Paper No. 4). Warsaw: Central European Forum for Migration Research.

Bijak, J., Disney, G., Findlay, A. M., Forster, J. J., Smith, P. W., & Wiśniowski, A. (2019). Assessing Time Series Models for Forecasting International Migration: Lessons from the United Kingdom. *Journal of Forecasting*, *38*(5), 470–487.

Bijak, J., & Wisniowski, A. (2010). *Forecasting International Migration in Europe: A Bayesian View*. Springer Netherlands. Retrieved from https://books.google.com/books?id=XGiApQtIt2QC

Billari, F., & Zagheni, E. (2017). Big Data and Population Processes: A Revolution? In *Proceedings of the conference of the italian statistical society.*

Billari, F. C., & Zagheni, E. (2017). Big Data and Population Processes: A Revolution? In A. Petrucci & R. Verde (Eds.), *Proceedings of the Conference of the Italian Statistical Society* (pp. 167–178). Firenze University Press.

Bloom, D. E., & Zucker, L. M. (2023, June). *Aging is the Real Population Bomb* (Tech. Rep.). International Monetary Fund.

Bongaarts, J. (2009). Human Population Growth and the Demographic Transition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1532), 2985–2990.

Brown, J. D., Cohen, S. R., Denoeux, G., Dorinski, S., Heggeness, M. L., Lieberman, C., ... Yi, M. (2023, May). *Real-Time 2020 Administrative Record Census Simulation: A New Design for the 21st Century* (Technical Report No. CBDRB-FY23-0253/CBDRB-FY23-0255). U.S. Census Bureau. Retrieved from https://www2.census.gov/programs-surveys/decennial/2020/program-management/evaluate-docs/EAE-2020-admin-records-experiment.pdf

Bruch, E. E., & Mare, R. D. (2012). Methodological Issues in the Analysis of Residential Preferences, Residential Mobility, and Neighborhood Change. *Sociological Methodology*, *42*(1), 103–154.

Bycroft, C., Miller, S., Gath, M., Matheson-Dunning, N., Simpson, K., & Das, S. (2021). *The Quality of Administrative Data for Census Variables: Strengths, Limitations, and Opportunities* (Tech. Rep.). Statistics New Zealand.

Cesare, N., Lee, H., McCormick, T., Spiro, E., & Zagheni, E. (2018). Promises and Pitfalls of Using Digital Traces for Demographic Research. *Demography*, *55*(5), 1979–1999. Retrieved from https://doi.org/10.1007/s13524-018-0715-2 doi: 10.1007/s13524-018-0715-2

Cheng, S., & Long, J. S. (2007). Testing for IIA in the Multinomial Logit Model. *Sociological Methods & Research*, *35*(4), 583-600. Retrieved from https://doi.org/10.1177/0049124106292361 doi: 10.1177/0049124106292361

Chun, A. Y., Larsen, M. D., Durrant, G., & Reiter, J. P. (Eds.). (2021). *Administrative Records for Survey Methodology*. Hoboken, New Jersey: Wiley.

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, *74*(368), 829–836. doi: 10.2307/2286407

Cohen, J. E., Roig, M., Reuman, D. C., & GoGwilt, C. (2008). International Migration Beyond Gravity: A Statistical Model for Use in Population Projections. *Proceedings of the National Academy of Sciences*, *105*(40), 15269–15274.

Coleman, D. (2008). The Demographic Effects of International Migration in Europe. *Oxford Review of Economic Policy*, *24*(3), 452–476.

Currie, C. (2022). *2017 Hurricanes: Update on FEMA's Disaster Recovery Efforts in Puerto Rico and the U.S. Virgin Islands* (Testimony No. GAO-22-106211). Washington, DC: United States Government Accountability Office.

Czado, C., Gneiting, T., & Held, L. (2009). Predictive Model Assessment for Count Data. *Biometrics*, *65*(4), 1254–1261.

Dai, J. (1998). Calibration and Test of a Discrete Choice Model with Endogenous Choice Sets. *Geographical Analysis*, *30*(2), 95-118. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1538-4632.1998.tb00391.x doi: https://doi.org/10.1111/j.1538-4632.1998.tb00391.x

de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., ... Paganin, S. (2021). NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling [Computer software manual]. Retrieved from https://cran.r-project.org/package=nimble (R package version 0.11.1) doi: 10.5281/zenodo.1211190

de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., ... Paganin, S. (2024a). NIMBLE: MCMC, particle filtering, and programmable hier-

archical modeling [Computer software manual]. Retrieved from https://cran.r-project.org/package=nimble (R package version 1.1.0) doi: 10.5281/zenodo.1211190

de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., ... Paganin, S. (2024b). NIMBLE user manual [Computer software manual]. Retrieved from https://r-nimble.org (R package manual version 1.1.0) doi: 10.5281/zenodo.1211190

de Valpine, P., Turek, D., Paciorek, C., Anderson-Bergman, C., Temple Lang, D., & Bodik, R. (2017). Programming with Models: Writing Statistical Algorithms for General Model Structures with NIMBLE. *Journal of Computational and Graphical Statistics*, *26*, 403-413. doi: 10.1080/10618600.2016.1172487

Dias, C. A., Wallgren, A., Wallgren, B., & Coelho, P. S. (2016). Census Model Transition: Contributions to its Implementation in Portugal. *Journal of Official Statistics*, *32*(1), 93–112.

Director of National Intelligence. (2021). *The Future of Migration* (Assessment No. NIC-2021-02486). United States Director of National Intelligence.

Dunne, J., & Zhang, L.-C. (2023, 04). A System of Population Estimates Compiled from Administrative Data Only. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *187*(1), 3-21. Retrieved from https://doi.org/10.1093/jrsssa/qnad065 doi: 10.1093/jrsssa/qnad065

Economic and Social Commission for Asia and the Pacific. (2022, April). *Tapping into Administrative Data in Census-taking: An Emerging Trend in Asia and the Pacific* (Stats Brief No. 31). United Nations.

Economic Commission for Europe and others. (2007). *Register-based Statistics in the Nordic Countries: Review of Best Practices with Focus on Population and Social Statistics* (Tech. Rep. No. 978-92-1-116963- 8). United Nations.

Elliott, D., & Blackwell, L. (2023, February). *Progress on an Integrated Statistical Design for the Transformed Population and Social Statistics System: Bayesian Methods for Demographic Estimation* (Report). United Kingdom Office of National Statistics.

Elshiewy, O., Guhl, D., & Boztuğ, Y. (2017). Multinomial Logit Models in Marketing–from Fundamentals to State-of-the-Art. *Marketing: ZFP–Journal of Research and Management*, *39*(3), 32–49.

European Commission – Press release. (2021, January). *Migration Statistics Update: The Impact of COVID-19.* Online.

Fertig, M., & Schmidt, C. M. (2005). Aggregate-level Migration Studies as a Tool for Forecasting Future Migration Streams. In *International Migration* (pp. 129–156). Routledge.

Fiorio, L., Zagheni, E., Abel, G., Hill, J., Pestre, G., Letouzé, E., & Cai, J. (2021, 02). Analyzing the Effect of Time in Migration Measurement Using Georeferenced Digital Trace Data. *Demography*, *58*(1), 51-74. Retrieved from https://doi.org/10.1215/00703370-8917630 doi: 10.1215/00703370-8917630

Fontanet, A., Autran, B., Lina, B., Kieny, M. P., Karim, S. S. A., & Sridhar, D. (2021). SARS-CoV-2 Variants and Ending the COVID-19 Pandemic. *The Lancet*, *397*(10278), 952–954.

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge University Press.

Gelman, A., & Hill, J. (2007). *Data Analysis using Regression and Multilevel/Hierarchical Models.* United Kingdom: Cambridge University Press.

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, *7*(4), 457 – 472. Retrieved from https://doi.org/10.1214/ss/1177011136 doi: 10.1214/ss/1177011136

Hackman, M., & Caldwell, A. A. (2021, March). Surge of Migrants at U.S. Southern Border: Biden's Plan and What You Need to Know. *Wall Street Journal*. Online.

Haddad, B. (2011). The Political Economy of Syria: Realities and Challenges. *Middle East Policy*, *18*(2), 46.

Haghani, M., Bliemer, M. C., & Hensher, D. A. (2021). The Landscape of Econometric Discrete Choice Modelling Research. *Journal of Choice Modelling*, *40*, 100303.

Harding, S., Liddle, L., McDonald, P., Morrison, P., Trewin, D., & Walters, S. (2022). *Report on the Quality of 2021 Census Data* (Tech. Rep.). Australian Bureau of Statistics.

Hausman, J., & McFadden, D. (1984). Specification Tests for the Multinomial Logit Model. *Econometrica*, *52*(5), 1219–1240. Retrieved 2024-07-06, from http://www.jstor.org/stable/1910997

Hoekstra, M., & Orozco-Aleman, S. (2021). *Illegal Immigration: The Trump Effect* (Working Paper No. 28909). Cambridge, MA: National Bureau of Economic Research.

Hoffman, S. D., & Duncan, G. J. (1988). Multinomial and Conditional Logit Discrete-Choice Models in Demography. *Demography*, *25*(3), 415–427.

Holmberg, A., & Watmuff, R. (2023, May 31). *International Approaches to Conducting a Census in the 21st Century: Australia.* Presentations in Committee on National Statistics Public Webinar Series. Retrieved from https://www.nationalacademies.org/event/05-31-2023/cnstat-public-webinar-series-international-approaches-to-conducting-a-census-in-the-21st-century-australia

Hyndman, R. J., & Booth, H. (2008). Stochastic Population Forecasts Using Functional Data Models for Mortality, Fertility and Migration. *International Journal of Forecasting*, *24*(3), 323–342.

Hyndman, R. J., & Koehler, A. B. (2006). Another Look at Measures of Forecast Accuracy. *International Journal of Forecasting*, *22*(4), 679–688.

Inoa, I. A., Picard, N., & de Palma, A. (2015). Effect of an Accessibility Measure in a Model for Choice of Residential Location, Workplace, and Type of Employment. *Mathematical Population Studies*, *22*(1), 4–36.

Jackman, S. (2009). *Bayesian Analysis for the Social Sciences* (Vol. 846). John Wiley & Sons.

Jain, D. C., Vilcassim, N. J., & Chintagunta, P. K. (1994). A Random-coefficients Logit Brand-choice Model Applied to Panel Data. *Journal of Business & Economic Statistics*, *12*(3), 317–328.

Keller, A., Mule, V. T., Morris, D. S., & Konicki, S. (2018). A Distance Metric for Modeling the Quality of Administrative Records for Use in the 2020 U.S. Census. *Journal of Official Statistics*, *34*(3), 599-624.

Kim, C. J. (2019). *Aging Societies: Policies and Perspectives* (Tech. Rep.). Asian Development Bank Institute.

Kim, K., & Cohen, J. E. (2010). Determinants of International Migration Flows To and From Industrialized Countries: A Panel Data Approach Beyond Gravity. *International Migration Review*, *44*(4), 899–932.

Kolk, M. (2019). Period and Cohort Measures of Internal Migration. *Population*, *74*(3), 333–350.

Krueger, R., Bierlaire, M., Gasos, T., & Bansal, P. (2023). Robust Discrete Choice Models with t-Distributed Kernel Errors. *Statistics and Computing*, *33*(1), 2.

Kucheva, Y. (2022). What Would it Take to Desegregate US Metropolitan Areas? Pathways to Residential Desegregation by Race. *Demography*, *59*(2), 433–459.

Kupiszewski, M. (2002). How Trustworthy Are Forecasts of International Migration Between Poland and the European Union? *Journal of Ethnic and Migration Studies*, *28*(4), 627–645.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science*, *343*(6176), 1203–1205. Retrieved from https://science.sciencemag.org/content/343/6176/1203 doi: 10.1126/science.1248506

Lazer, D., & Radford, J. (2017). Data ex Machina: Introduction to Big Data. *Annual Review of Sociology*, *43*(1), 19-39. Retrieved from https://doi.org/10.1146/annurev-soc-060116-053457 doi: 10.1146/annurev-soc-060116-053457

Lee, R. (2011). The Outlook for Population Growth. *Science*, *333*(6042), 569–573.

Lenk, P. (2014). Bayesian Estimation of Random Utility Models. In *Handbook of Choice Modelling* (pp. 457–497). Edward Elgar Publishing.

Lerpold, L., Sjöberg, Ö., & Wennberg, K. (2023). *Migration and Integration in a Post-pandemic World: Socioeconomic Opportunities and Challenges*. Springer Nature.

Long, M. C., Pelletier, E., & Romich, J. (2022). Constructing Monthly Residential Locations of Adults Using Merged State Administrative Data. *Population Studies*, *76*(2), 253–272.

Lundy, E. R. (2022). Predicting the Quality and Evaluating the use of Aministrative Data for the 2021 Canadian Census of Population. *Statistical Journal of the International Association for Official Statistics*, *38*, 1177 – 1183.

Martell, P. (2011). After Independence, What Next for South Sudan? *Africa Renewal*, *25*(2), 3–5.

Mayer, T., & Zignago, S. (2011, December). *Notes on CEPII's Distance Measures: The GeoDist database* (Working Papers No. 2011-25). Centre d'Études Prospectives et d'Informations Internationales.

McAuliffe, M., & Oucho, L. A. (Eds.). (2024). *World Migration Report 2024*. Geneva: International Organization for Migration (IOM).

MCCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). London: Chapman and Hall/CRC.

McFadden, D. (1973). Frontiers in econometrics. In (chap. Conditional Logit Analysis of Qualitative Choice Behavior). Wiley.

McFadden, D. (1978). Modeling the Choice of Residential Location. In *Transportation Forecasting and Travel Behavior* (p. 72-77). Transportation Research Board of the National Academy of Sciences.

Morris, D. S. (2014). A Comparison of Methodologies for Classification of Administrative Records Quality for Census Enumeration. *American Statistical Association: 2014 Joint Statistical Meeting*.

Mule, V. T., & Keller, A. (2021). Administrative Records Applicaitons for the 2020 Census. In (pp. 205–230). Wiley Online Library.

Münz, R. (2013, September). *Demography and Migration: An Outlook for the 21st Century* (Policy Brief No. 4). Migration Policy Institute.

Murly, M. H., Mule, V. T., Keller, A., & Konicki, S. (2021, April). *Administrative Record Modeling in the 2020 Census* (Report). Washington, DC: United States Census Bureau.

Murly, M. H., Tello-Trillo, C. J., Mule, V. T., & Keller, A. (2023, February). *Full Report of the Comparisons of Administrative Records Rosters to Census Self-Responses and NRFU Household Member Responses* (Technical Report No. CBDRB-FY22-307). Washington, DC: United States Census Bureau.

National Academies of Sciences, Engineering, and Medicine. (2023). *Assessing the 2020 Census: Final Report*. Washington, DC: The National Academies Press.

Ortman, J. (2024, September). The Demographic Frame. *International Journal of Population Data Science*, *9*(5).

Papaspiliopoulos, O., Roberts, G. O., & Skold, M. (2003). Non-centered Parameterisations for Hierarchical Models and Data Augmentation. In J. Bernardo et al. (Eds.), *Bayesian statistics 7: Proceedings of the seventh valencia international meeting* (p. 307-326). Oxford University Press, USA.

Preston, S. H., Heuveline, P., & Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes*. Malden, Massachusetts: Blackwell.

R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

R Core Team. (2023). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Raftery, A. E., Alkema, L., & Gerland, P. (2014). Bayesian Population Projections for the United Nations. *Statistical Science*, *29*(1), 58–68.

Raftery, A. E., Chunn, J. L., Gerland, P., & Ševčíková, H. (2013). Bayesian Probabilistic Projections of Life Expectancy for All Countries. *Demography*, *50*(3), 777–801.

Raftery, A. E., & Lewis, S. (1992). How Many Iterations in the Gibbs Sampler? In J. Bernardo, J. Berger, A. Dawid, & A. Smith (Eds.), *Bayesian statistics 4: Proceedings of the fourth valencia international meeting* (p. 763-773). Oxford University Press, USA.

Raftery, A. E., Li, N., Ševčíková, H., Gerland, P., & Heilig, G. K. (2012). Bayesian Probabilistic Population Projections for All Countries. *Proceedings of the National Academy of Sciences*, *109*(35), 13915–13921.

Raftery, A. E., & Ševčíková, H. (2023). Probabilistic Population Forecasting: Short to Very Long-Term. *International Journal of Forecasting*, *39*(1), 73-97.

RAND. (2005). *Population Implosion? Low Fertility and Policy Responses in the European Union* (Research Brief). Cambridge, UK: RAND Corporation.

Rastogi, S., & O'Hara, A. (2012, November). *2010 Census Match Study* (Final Report No. 247). Washington, DC: United States Census Bureau.

Raymer, J., Bai, X., & Smith, P. W. (2020). Developments in Demographic Forecasting. In S. Mazzuco & N. Keilman (Eds.), (chap. Forecasting Origin-Destination-Age-Sex Migration Flow Forecasting Origin-Destination-Age-Sex Migration Flow Tables with Multiplicative Components). Springer Nature.

Raymer, J., Guan, Q., Shen, T., Hertog, S., & Gerland, P. (2023, December). *Modelling the Age and Sex Profiles of Net International Migration* (Technical Report No. UN DESA/POP/2023/TP/No. 7). New York: Department of Economic and Social Affairs, Population Division.

Raymer, J., & Willekens, F. (2008). International Migration in Europe: Data, Models and Estimates. In (chap. 6). Wiley. Retrieved from https://books.google.com/books?id=sHCUNd7g-pEC

Raymer, J., & Wiśniowski, A. (2018). Applying and Testing a Forecasting Model for Age and Sex Patterns of Immigration and Emigration. *Population Studies*, *72*(3), 339-355. Retrieved from https://doi.org/10.1080/00324728.2018.1469784 (PMID: 29873285) doi: 10.1080/00324728.2018.1469784

Rogers, A. (1990). Requiem for the Net Migrant. *Geographical Analysis*, *22*(4), 283–300.

Rogers, A., & Castro, L. J. (1981a). *Model Migration Schedules* (Research Report No. 81-30). Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A., & Castro, L. J. (1981b). *Model Migration Schedules* (Tech. Rep. No. RR-81-030). Laxenburg, Austria: International Institute for Applied System Analysis.

Rogers, A., Willekens, F., Little, J., & Raymer, J. (2002). Describing Migration Spatial Structure. *Papers in Regional Science*, *81*(1), 29–48.

Rüschendorf, L. (1995). Convergence of the Iterative Proportional Fitting Procedure. *The Annals of Statistics*, *23*(4), 1160–1174.

Ruths, D., & Pfeffer, J. (2014). Social Media for Large Studies of Behavior. *Science*, *346*(6213), 1063–1064.

Sener, I. N., Pendyala, R. M., & Bhat, C. R. (2011). Accommodating Spatial Correlation Across Choice Alternatives in Discrete Choice Models: An Application to Modeling Residential Location Choice Behavior. *Journal of Transport Geography*, *19*(2), 294–303.

Shiraef, M. A., Hirst, C., Weiss, M. A., Naseer, S., Lazar, N., Beling, E., ... others (2021). COVID Border Accountability Project, a hand-coded global database of border closures introduced during 2020. *Scientific Data*, *8*(1), 253.

Skinner, C. (2018). Issues and Challenges in Census Taking [Journal Article]. *Annual Review of Statistics and Its Application*, *5*, 49-63. Retrieved from https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-041715-033713 doi: https://doi.org/10.1146/annurev-statistics-041715-033713

Skjerpen, T., & Tønnessen, M. (2021). Using Future Age Profiles to Improve Immigration Projections. *Population Studies*, *75*(2), 255-267.

Smith, S. K., Tayman, J., & Swanson, D. A. (2001). *State and Local Population Projections: Methodology and Analysis*. New York: Kluwer Academic/Plenum Publishers.

Spring, A., Ackert, E., Crowder, K., & South, S. J. (2017). Influence of Proximity to Kin on Residential Mobility and Destination Choice: Examining Local Movers in Metropolitan Areas. *Demography*, *54*(4), 1277–1304.

Statistics New Zealand. (2023). *Experimental Administrative Population Census (3rd iteration): Data Sources, Methods, and Quality for Household Information* (Tech. Rep.). Stats NZ.

Stats, NZ. (2019a). *Overview of Statistical Methods for Adding Admin Records to the 2018 Census Dataset.* Wellington, NZ: Statistics New Zealand. Available at: https://www.stats.govt.nz/methods/overview-of-statistical-methods-for-adding-admin-records-to-the-2018-census-dataset/.

Stats, NZ. (2019b). *Predicting the Quality of Admin Location Information for Use in the 2018 Census.* Online. Stats NZ Tatauranga Aotearoa Wellington, New Zealand.

Thomas, D. L., Johnson, D., & Griffith, B. (2006). A Bayesian Random Effects Discrete-Choice Model for Resource Selection: Population-level Selection Inference. *The Journal of Wildlife Management*, *70*(2), 404–412.

Train, K. E. (2009). *Discrete Choice Methods with Simulation.* Cambridge University Press.

United Nations. Department of Economic and Social Affairs. (2015). World Population Prospects: The 2015 Revision, Methodology of the United Nations Population Estimates and Projections. New York, United States of America.

United Nations. Department of Economic and Social Affairs. (2019). World Population Prospects: The 2019 Revision, Methodology of the United Nations Population Estimates and Projections. New York, United States of America.

United Nations High Commissioner for Refugees. (n.d.). *Ukraine Refugee Situation.* Retrieved January 28, 2024, from https://data.unhcr.org/en/situations/ukraine

United Nations. (2022, July). *World Population Prospects 2022: Methodology of the United Nations Population Estimates and Projections* (Tech. Rep. No. UN DESA/POP/2022/TR/NO.6). New York: Department of Economic and Social Affairs, Population Division.

United Nations. (2024, July). *World Population Prospects 2024: Methodology of the United Nations Population Estimates and Projections* (Tech. Rep. No. UN DESA/POP/2024/DC/NO.10). New York: Department of Economic and Social Affairs, Population Division.

Ševčíková, H., & Raftery, A. E. (2016a). bayesPop: Probabilistic Population Projections. *Journal of Statistical Software*, *75*(5), 1–29. doi: 10.18637/jss.v075.i05

Ševčíková, H., & Raftery, A. E. (2016b). bayesPop: Probabilistic Population Projections. *Journal of Statistical Software*, *75*(5), 1–29.

Welch, N. G., & Raftery, A. E. (2022). Probabilistic Forecasts of International Bilateral Migration Flows. *Proceedings of the National Academy of Sciences*, *119*(35), e2203822119.

Willekens, F. (2018). Towards Causal Forecasting of International Migration. *Vienna Yearbook of Population Research*, *16*, 199–218.

Willekens, F., & Baydar, N. (1986). Population Structures and Models. Developments in Spatial Demography/ed. by. In R. Woods & P. Rees (Eds.), (pp. 203–244). Allen & Unwin.

Wiśniowski, A., Smith, P. W., Bijak, J., Raymer, J., & Forster, J. J. (2015a). Bayesian Population Forecasting: Extending the Lee-Carter Method. *Demography*, *52*(3), 1035–1059.

Wiśniowski, A., Smith, P. W., Bijak, J., Raymer, J., & Forster, J. J. (2015b). Bayesian Population Forecasting: Extending the Lee-Carter Method. *Demography*, *52*(3), 1035–1059.

Yoon, T., Davis, K., Lundy, E., & Goussanou, A. (2022, May). Modernization of the Canadian Census: An Administrative Data-Driven Approach to Defining Households. In *Statistical Society of Canada Annual Meeting.*

Yu, C. C., Ševčíková, H., Raftery, A. E., & Curran, S. R. (2023). Probabilistic County-Level Population Projections. *Demography*, *60*(3), 915–937.

Zagheni, E., Garimella, V. R. K., Weber, I., & State, B. (2014). Inferring International and Internal Migration Patterns from Twitter Data. In *Proceedings of the 23rd international conference on world wide web* (pp. 439–444). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2567948.2576930 doi: 10.1145/2567948.2576930

Zagheni, E., & Weber, I. (2012). You Are Where You E-Mail: Using E-Mail Data to Estimate International Migration Rates. In *Proceedings of the 4th annual acm web science conference* (pp. 348–351). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2380718.2380764 doi: 10.1145/2380718.2380764

Zagheni, E., Weber, I., & Gummadi, K. (2017). Leveraging Facebook's Advertising Platform to Monitor Stocks of Migrants. *Population and Development Review*, *43*(4), 721–734. Retrieved from http://www.jstor.org/stable/26622775

Zhang, L.-C. (2012). Topics of Statistical Theory for Register-based Statistics and Data Integration. *Statistica Neerlandica*, *66*(1), 41–63.

Appendix A

APPENDICES TO CHAPTER 3

A.1 Proof of Theorem

Theorem 2. *Consider a population that has the same age-specific out-migration rates as country i in period t, but a different population age distribution given by $\pi^*_{i,t,a}$. Then this population has out-migration rate*

$$OMR^*_{i,t} = OMR_{i,t} \frac{\sum_a \pi^*_{i,t,a} OMR_{i,t,a}}{\sum_a \pi_{i,t,a} OMR_{i,t,a}}. \tag{A.1}$$

If both populations have the same age-specific pattern of migration rates, R_a, then

$$OMR^*_{i,t} = OMR_{i,t} \frac{C^*_{i,t}}{C_{i,t}}, \tag{A.2}$$

where $C_{i,t} = \sum_a \pi_{i,t,a} R_a$ is the migration age structure index (MASI).

Proof. Starting from the definition of $\mathrm{OMR}_{i,t}$,

$$\begin{aligned}
\mathrm{OMR}_{i,t} &= \frac{O_{i,t}}{\tilde{P}_{i,t,+,+}} && \text{(definition of OMR)} \\
&= \sum_a \frac{O_{i,t,a}}{\tilde{P}_{i,t,a,+}} \times \frac{\tilde{P}_{i,t,a,+}}{\tilde{P}_{i,t,+,+}} && \text{(OMR disaggregated by age)} \\
&= G_{i,t} \sum_a \frac{\mathrm{OMR}_{i,t,a}}{G_{i,t}} \times \pi_{i,t,a} && \left(\text{Multiplication by } \frac{G_{i,t}}{G_{i,t}}\right).
\end{aligned} \tag{A.3}$$

The same derivation applies to $\text{OMR}^*_{i,t}$, implying that

$$\text{OMR}^*_{i,t} = G_{i,t} \sum_a \frac{\text{OMR}_{i,t,a}}{G_{i,t}} \times \pi^*_{i,t,a}. \tag{A.4}$$

Rearranging $\text{OMR}_{i,t}$ in terms of $G_{i,t}$ and substituting the equation into (A.4) yields the desired result:

$$\text{OMR}^*_{i,t} = \text{OMR}_{i,t} \frac{\sum_a \text{OMR}_{i,t,a} \pi^*_{i,t,a}}{\sum_a \text{OMR}_{i,t,a} \pi_{i,t,a}}. \tag{A.5}$$

Multiplying (A.5) by $\frac{G_{i,t}}{G_{i,t}}$, assuming the same $R_a = \text{OMR}_{i,t,a}/G_{i,t}$ in both populations, and applying the definition of R_a, we obtain

$$\text{OMR}^*_{i,t} = \text{OMR}_{i,t} \frac{\sum_a R_a \pi^*_{i,t,a}}{\sum_a R_a \pi_{i,t,a}} = \text{OMR}_{i,t} \frac{C^*_{i,t}}{C_{i,t}}, \tag{A.6}$$

as desired. This concludes the proof. □

A.2 Historic Net Migration into In- and Out-migration Model Validation

Table A.1 presents the mean absolute error (MAE), scaled by 1,000, for in-migration rates (IMR) and out-migration rates (OMR) across six 5-year periods. Both IMR and OMR show the lowest MAE in the mid-range periods (1995–2000 for IMR at 3.222 and 2000–2005 for OMR at 3.677) and the highest in 2015–2020 (8.178 for IMR and 8.151 for OMR). The errors are similar during 2005–2015 but increase in the later periods, indicating lower model accuracy in the last period.

Table A.1: Mixed effects model (1) mean absolute error (MAE) × 1,000 of in- and out-migration rates for all countries over all 5-year periods.

	1990–1995	1995–2000	2000–2005	2005–2010	2010–2015	2015–2020
IMR	4.806	3.222	4.138	3.542	5.431	8.178
OMR	5.547	3.294	3.677	3.542	5.434	8.151

Table A.2 presents the root mean squared error (RMSE), scaled by 1,000, for in- and out-migration rates (IMR and OMR) across six 5-year periods. Similar to the MAE findings, the RMSE for both IMR and OMR is lowest in the mid-range periods (1995–2000 for IMR at 5.560 and 2005–2010 for OMR at 5.742) and highest in 2015–2020 (14.784 for IMR and 14.753 for OMR). A notable spike is observed in IMR RMSE during 2000–2005 (11.413), while OMR also exhibits elevated values in 1990–1995 (11.949). Across most periods, OMR tends to have slightly higher RMSE than IMR, particularly in earlier years, though the two metrics converge in 2005–2015. Overall, the RMSE trends align closely with the MAE results, showing lower errors in mid-range periods and higher uncertainty in the earlier and later periods.

Figure A.1 shows the autocorrelation function (ACF) plot for the IMR residuals ($\text{IMR}_{\text{fit}} - \text{IMR}_{\text{obs}}$) from model (1) for each country for 5-year periods during 1990–2020. While some ACF estimates fall far from zero, none of fell beyond the 95% confidence region shown by the blue dashed lines (approximately ± 0.8). This suggests no significant autocorrelation

Table A.2: Mixed effects model (1) root mean squared error (RMSE) × 1,000 of in- and out-migration rates for all countries over all 5-year periods.

	1990–1995	1995–2000	2000–2005	2005–2010	2010–2015	2015–2020
IMR	7.733	5.560	11.413	5.741	9.313	14.784
OMR	11.949	6.273	9.600	5.742	9.314	14.753

in the residuals at any lag, supporting the assumption of model adequacy in capturing the temporal dynamics of IMR for all countries.

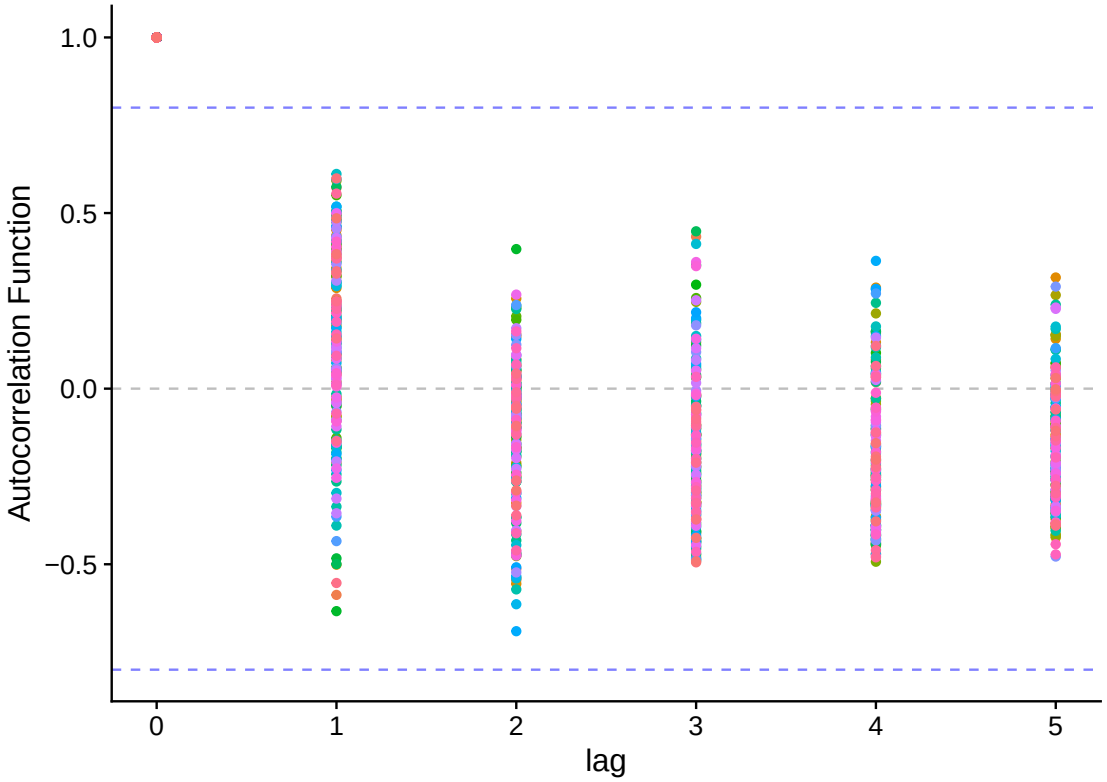


Figure A.1: Model (1) autocorrelation function (ACF) estimates for $\text{IMR}_{\text{fit}} - \text{IMR}_{\text{obs}}$ residuals in each country over the 5-year periods during 1990–2020 with 95% standard error estimate (blue dashed line) over all possible lag values. Grey horizontal dashed line denotes the ACF=0 value.

Figure A.2 shows the distribution of model (1) IMR residuals against the observed net migration rate (NMR) for 5-year periods from 1990 to 2020. Point sizes correspond to each country's population at the start of the period, and colors indicate United Nations (UN) areas. Qatar (QAT) and Kuwait (KWT) consistently exhibit the largest residuals from 1990–1995 through 2010–2015, while Eritrea (ERI) stands out in the 2000–2005 period. For 2010–2020, Luxembourg (LUX) shows significant residuals, along with Iceland (ISL), Malta (MLT), and Aruba (ABW) in 2015–2020. These outliers contribute to the higher errors summarized in Tables A.1 and A.2 for the 2015–2020 period. For all other countries and periods, IMR residuals remain close to zero across observed NMR values.

Figure A.3 shows the distribution of model (1) IMR residuals against the observed fitted IMR values for 5-year periods from 1990 to 2020, with point colors representing United Nations (UN) areas. A Locally Estimated Scatterplot Smoothing (LOESS) curve, a non-parametric regression method for revealing trends without assuming a global functional form (Cleveland, 1979), is shown as a solid blue line in Figure A.3. The LOESS curve remains close to zero for most countries and periods, indicating small residuals and good model performance in these cases. However, significant deviations are observed for three Gulf Cooperation Council (GCC) countries—Bahrain (BHR), United Arab Emirates (ARE), and Qatar (QAT)—during the first four periods (1990–2015). These large residuals suggest that model (1) struggled to account for departures in these GCC countries during this time frame, despite performing well in other contexts.

Figure A.4 shows the distribution of model (1) IMR residuals for 5-year periods from 1990 to 2020 for countries in each UN area. Dashed horizontal grey lines represent $\text{IMR}_{\text{fit}} - \text{IMR}_{\text{obs}} = 0$. The residuals generally have a mean close to zero and appear systematic across most regions, except for European countries during the 2015–2020 period. In this period, the residuals for Europe show a leftward skew, indicating that fitted IMR values are substantially larger than observed values in several European countries.

These evaluations indicate that model (1) performs well overall, with residuals for in-migration rates (IMR) and out-migration rates (OMR) generally small and distributed

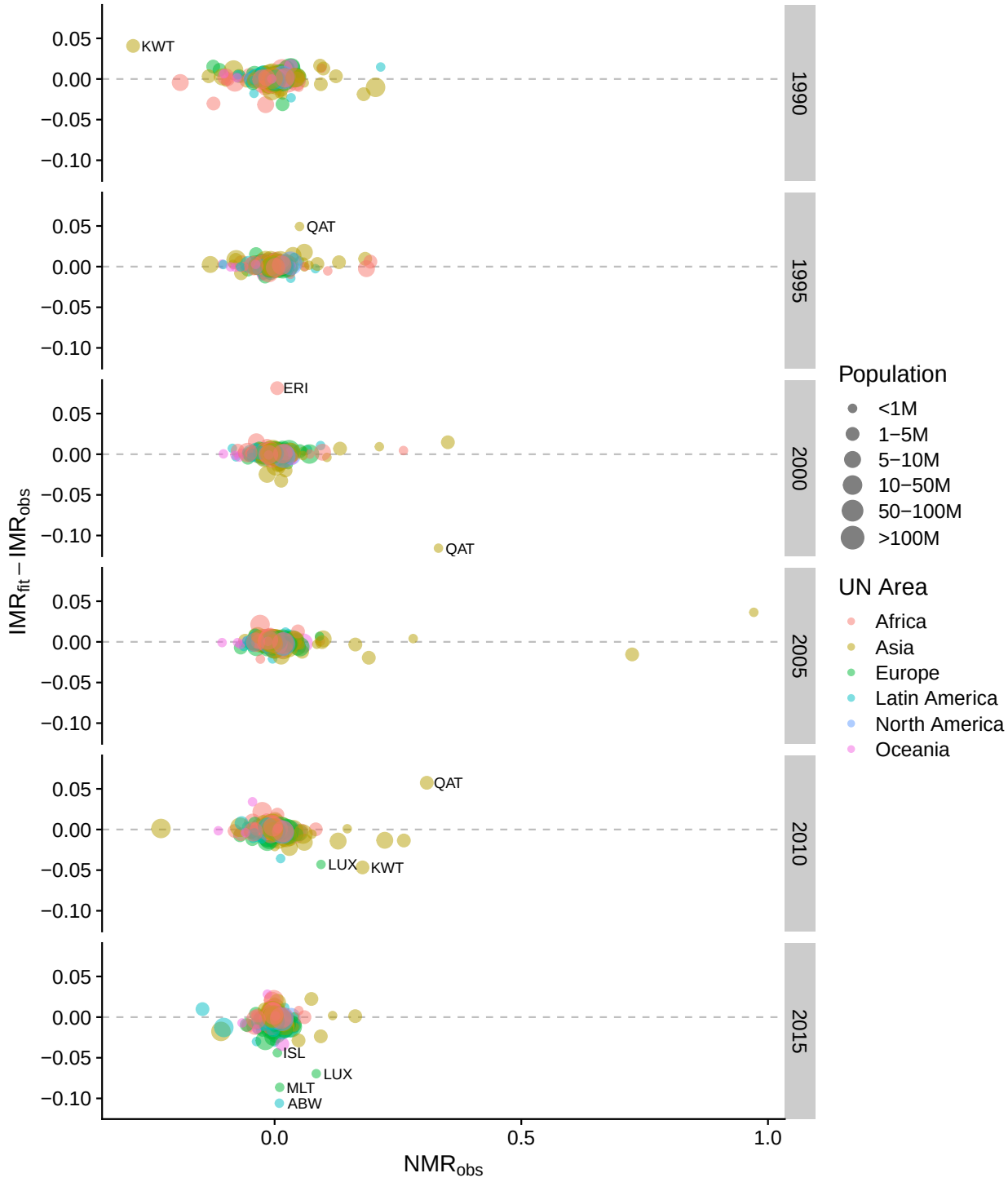


Figure A.2: Model (1) residuals by observed net migration rate (NMR_{obs}) for 5-year periods over 1990-2020 (first year of period shown on left vertical axis). Point sizes denote the size of each country's population at the start of the period. Point colors denote the United Nations area designation. Grey horizontal dashed lines denote $\text{IMR}_{\text{fit}} - \text{IMR}_{\text{obs}} = 0$.

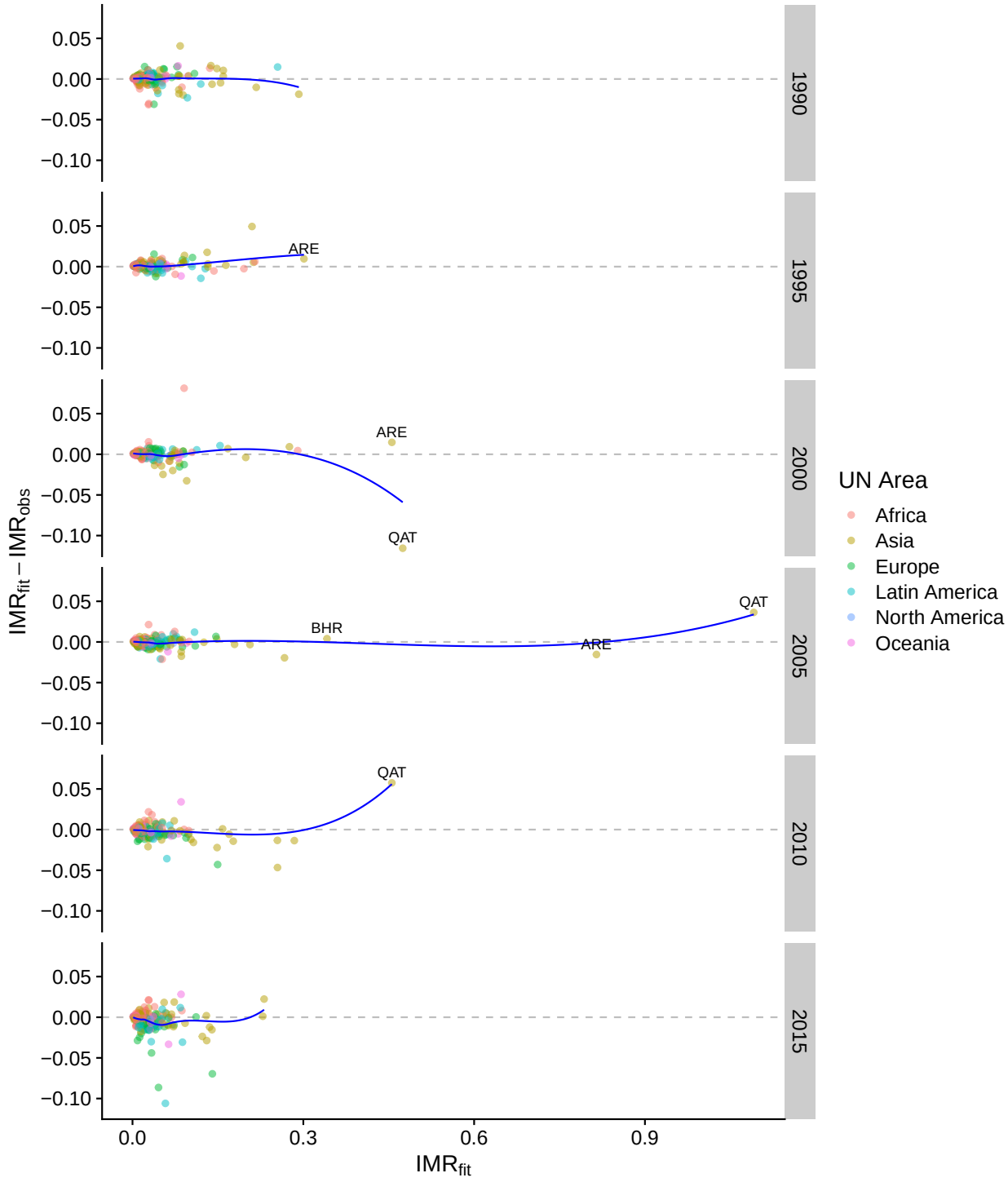


Figure A.3: Model (1) residuals by fitted in-migration rate (IMR_{fit}) for 5-year periods over 1990–2020 (first year of period shown on left vertical axis). Point colors denote the United Nations area designation. The LOESS curve for each period is shown by the solid blue line. Grey horizontal dashed lines denote $\text{IMR}_{\text{fit}} - \text{IMR}_{\text{obs}} = 0$.

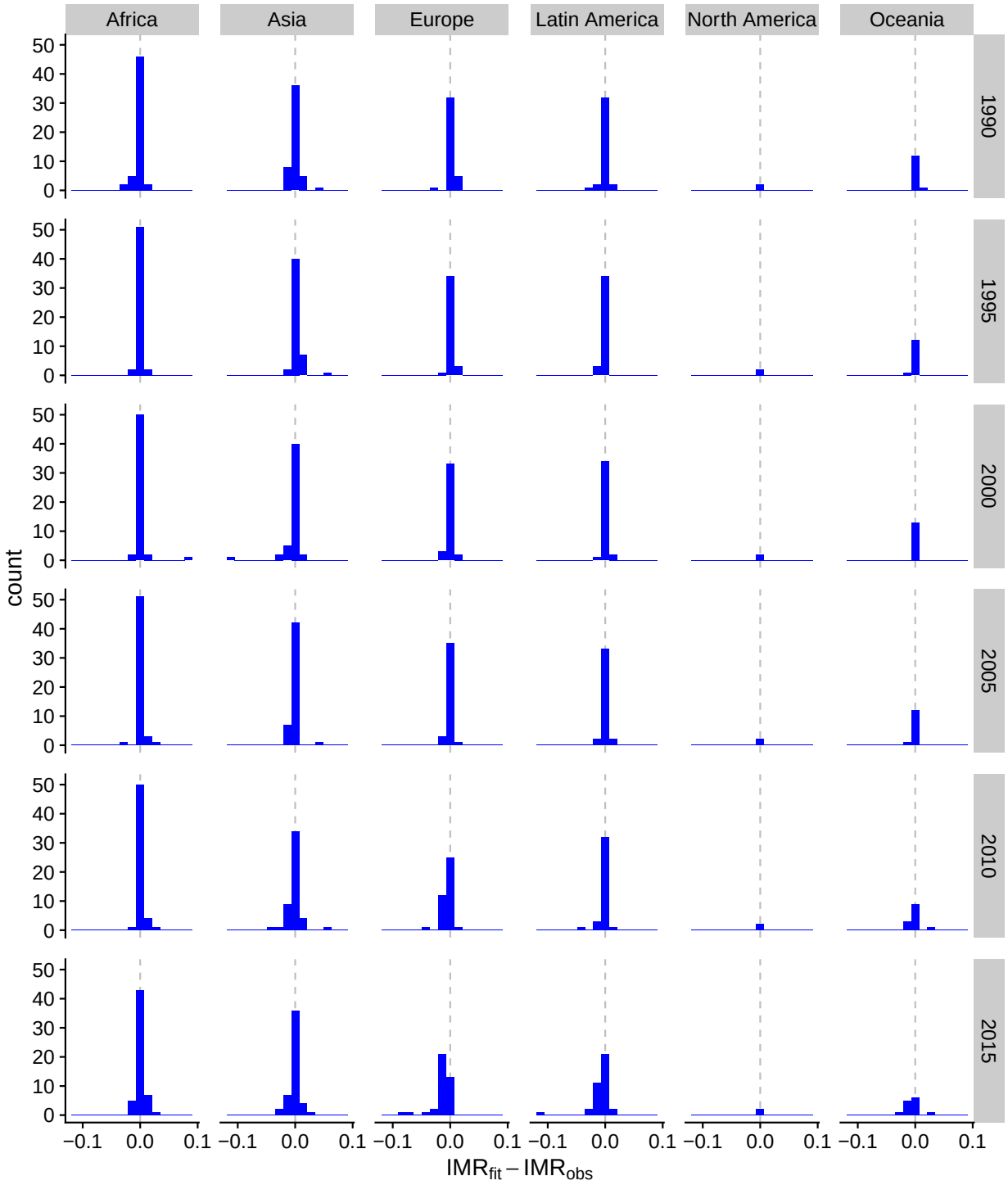


Figure A.4: Model (1) residual histograms by United Nations area and period (first year of period shown on left vertical axis). Grey vertical dashed lines denote $\text{IMR}_{\text{fit}} - \text{IMR}_{\text{obs}} = 0$.

around zero across most countries and periods. Both the mean absolute error (MAE) and root mean squared error (RMSE) are lowest during the mid-range periods (1995–2000 and 2005–2010), but errors increase significantly in the earlier (1990–1995) and later periods (2015–2020), reflecting higher model uncertainty during these times. However, notable regional outliers emerge, particularly among Gulf Cooperation Council (GCC) countries, such as Qatar, United Arab Emirates, and Bahrain, where residuals are consistently large from 1990 to 2015. Additionally, European countries exhibit leftward-skewed residuals during the 2015–2020 period, with fitted IMR values systematically exceeding observed values. While the model captures general trends effectively, as shown by residual distributions and LOESS curves remaining close to zero in most contexts, these regional anomalies suggest potential unmodeled factors or structural differences that the model fails to account for in specific regions and timeframes. Despite its simplicity, model (1) demonstrates robustness across most contexts but has limitations in explaining migration patterns in certain regions or periods.

The figures on the following pages compare the IMR and OMR estimates from model (1) for each country with the IMR and OMR values derived from the pseudo-Bayes estimates by Abel & Cohen (2019), based on the World Population Prospects (WPP) 2019 revision (United Nations. Department of Economic and Social Affairs, 2019). Each country has three plots:

- (left) WPP 2019-derived NMR for 5-year periods from 1950 to 2020
- (middle) IMR values for 1990–2020 derived from Abel & Cohen (2019) and model (1) IMR estimates for 1950-2020
- (right) OMR values for 1990–2020 derived from Abel & Cohen (2019) and model (1) OMR estimates for 1950-2020

Dashed black lines represent the WPP 2019 observations for NMR and the pseudo-Bayes IMR and OMR estimates from Abel & Cohen (2019). Solid blue lines indicate the IMR and OMR estimates from model (1) for all 5-year periods between 1950 and 2020.

Model (1) aligns closely with the estimates from Abel & Cohen (2019) when the solid blue lines and dashed black lines overlap during 1990–2020. Deviations between the two indicate departures from model (1) that remain unexplained. Model (1) estimates for 1950–1990 are based solely on WPP 2019 NMR observations, as there are no corresponding 5-year IMR and OMR values available for comparison during this earlier period.

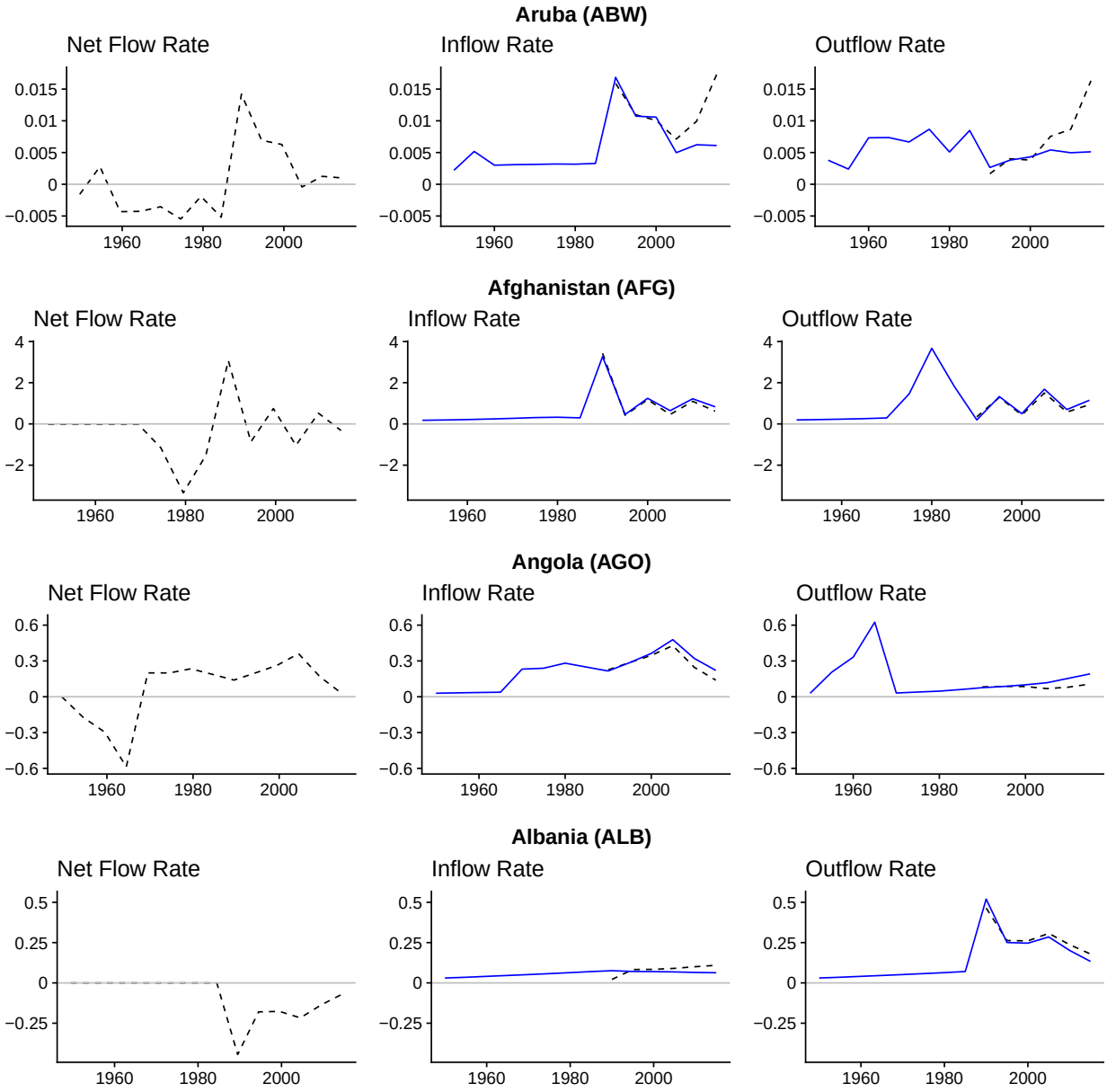


Figure A.5: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

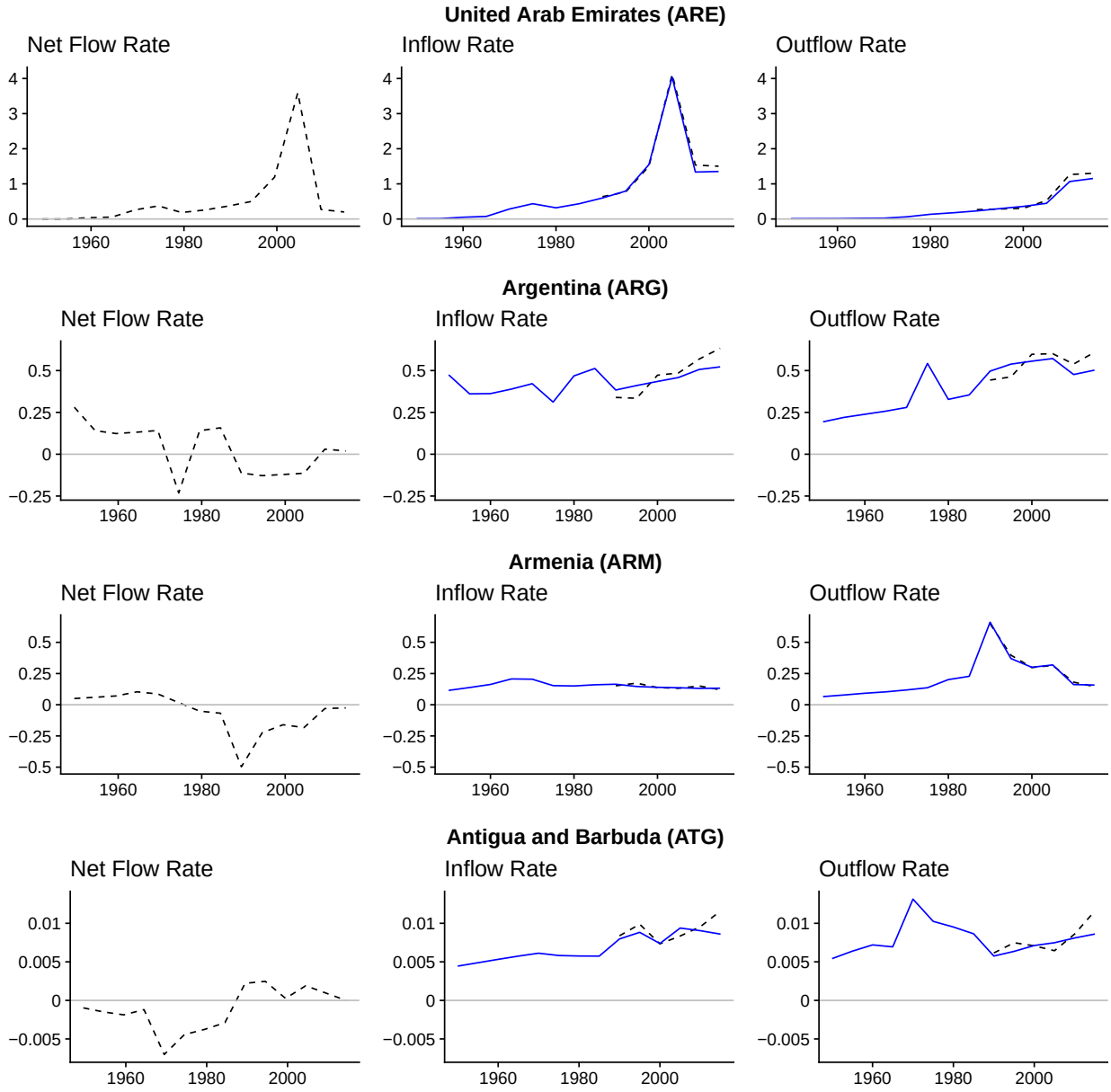


Figure A.6: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

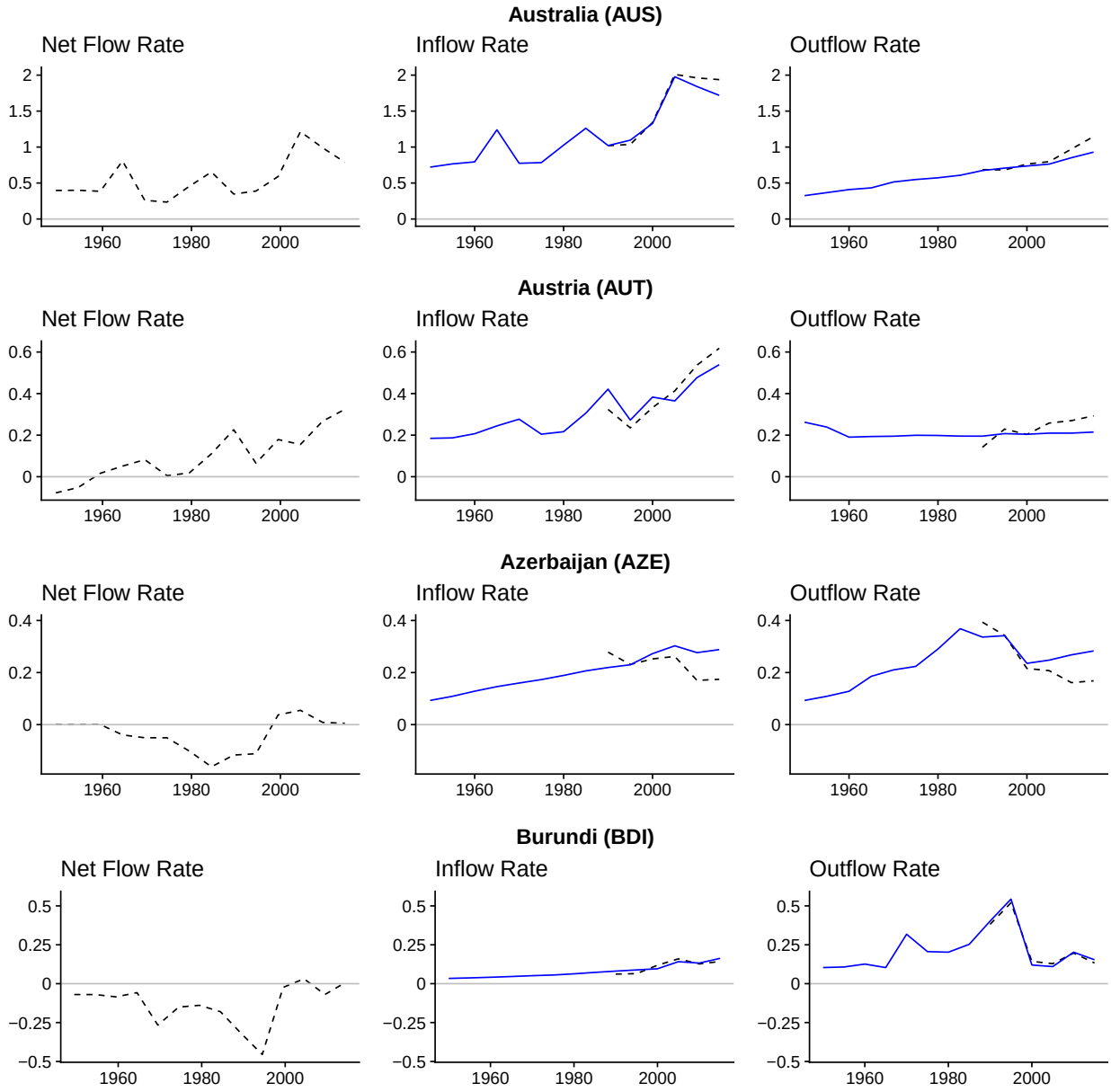


Figure A.7: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

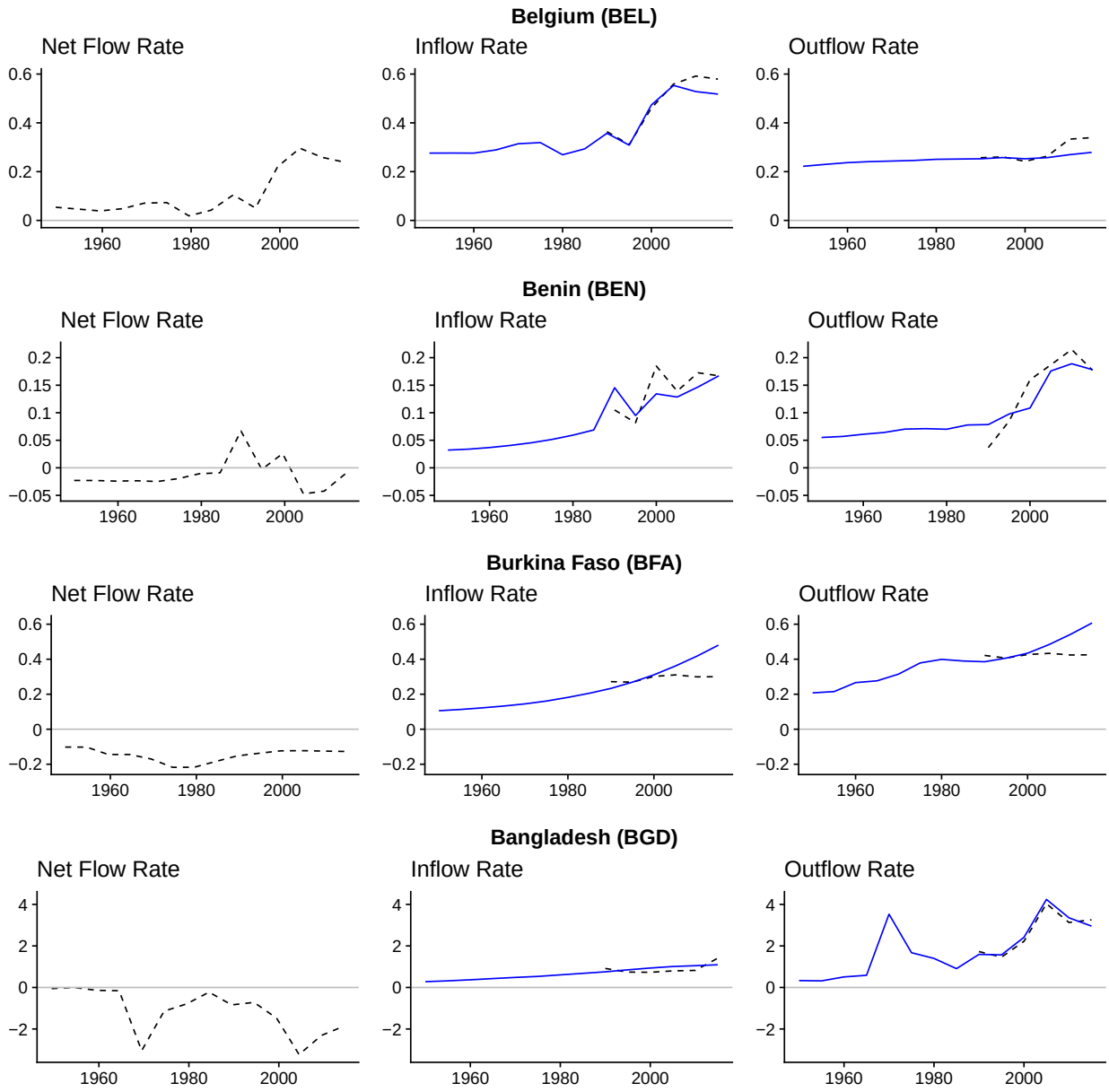


Figure A.8: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

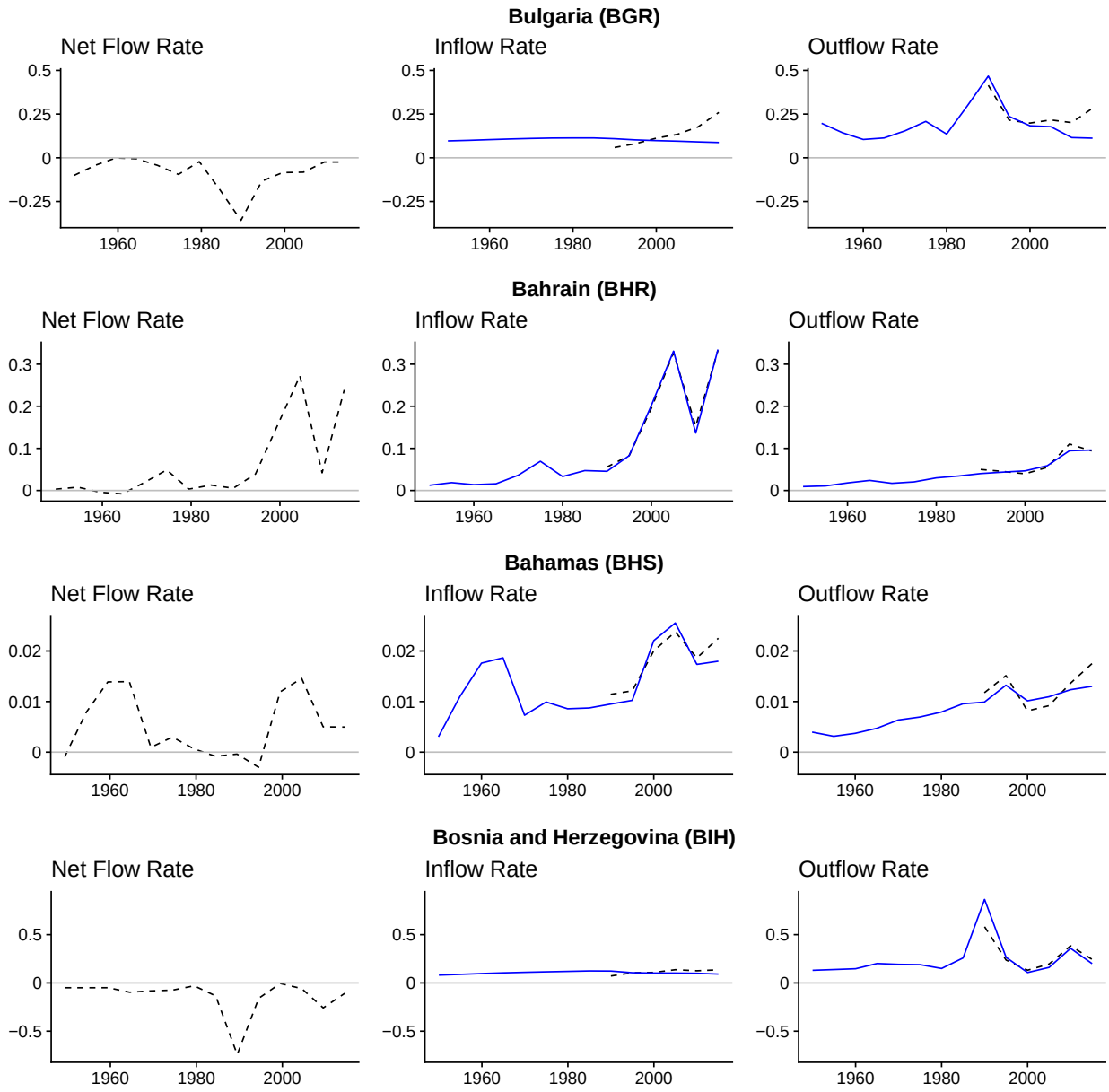


Figure A.9: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

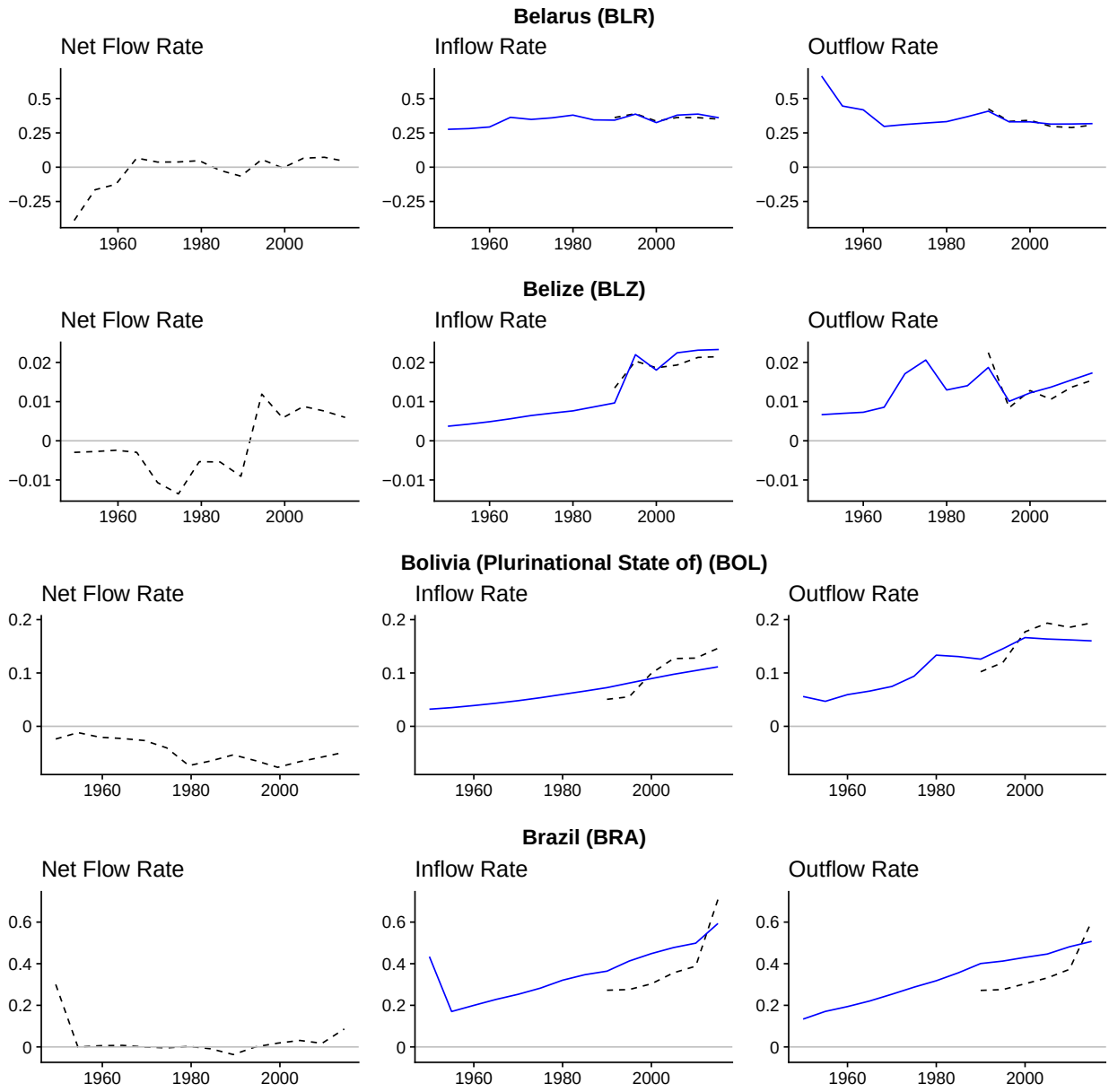


Figure A.10: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

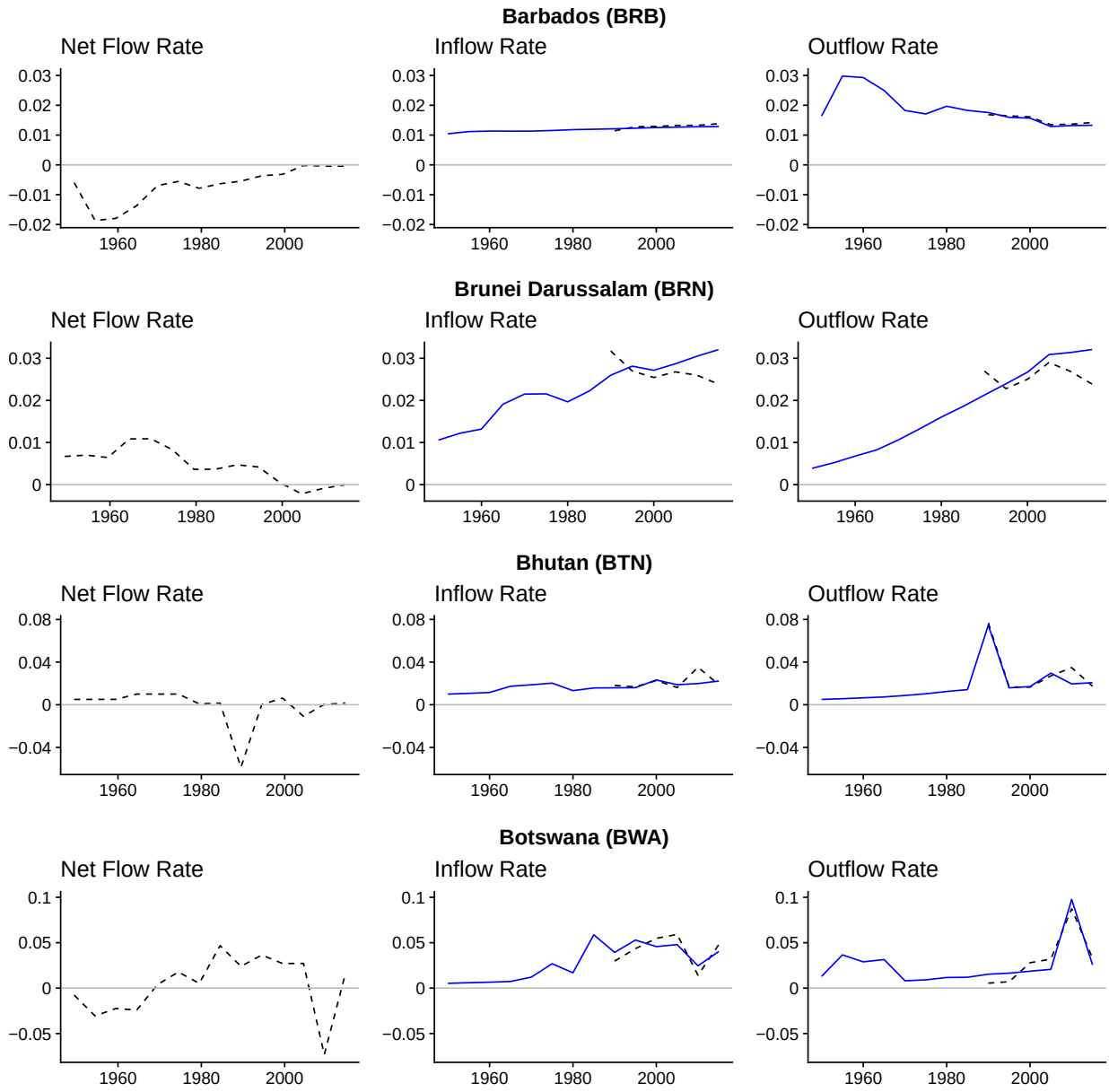


Figure A.11: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

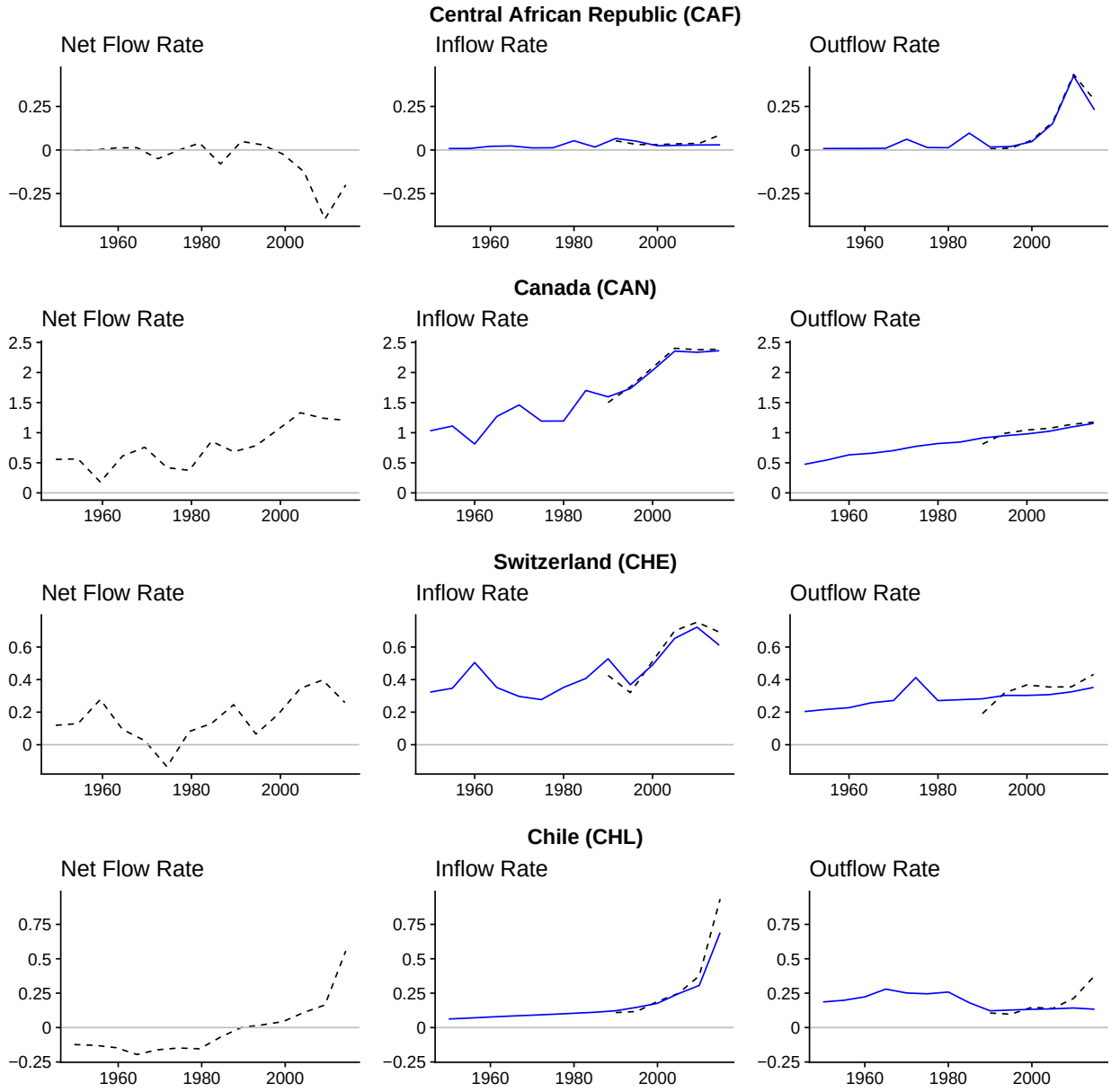


Figure A.12: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

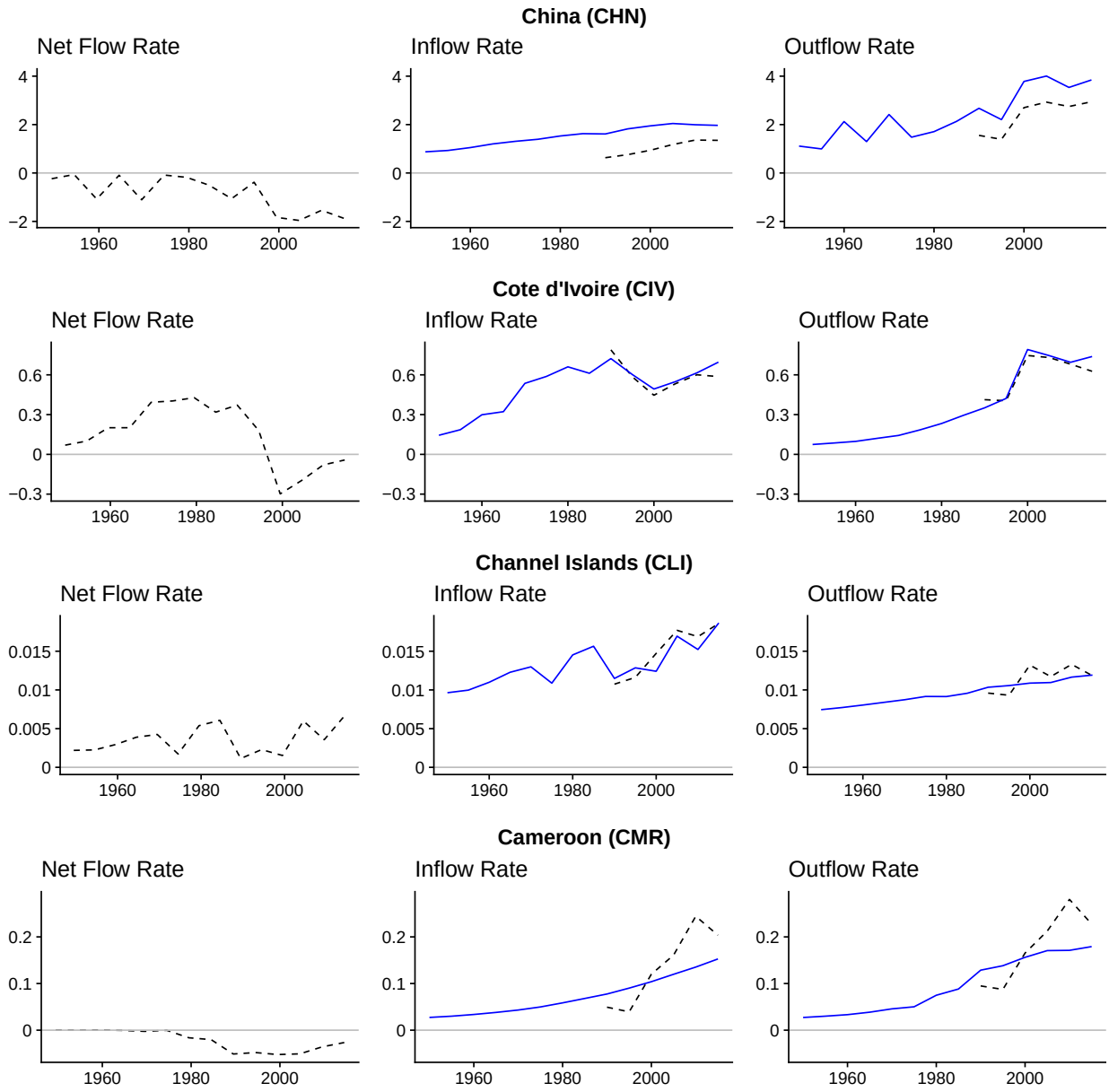


Figure A.13: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

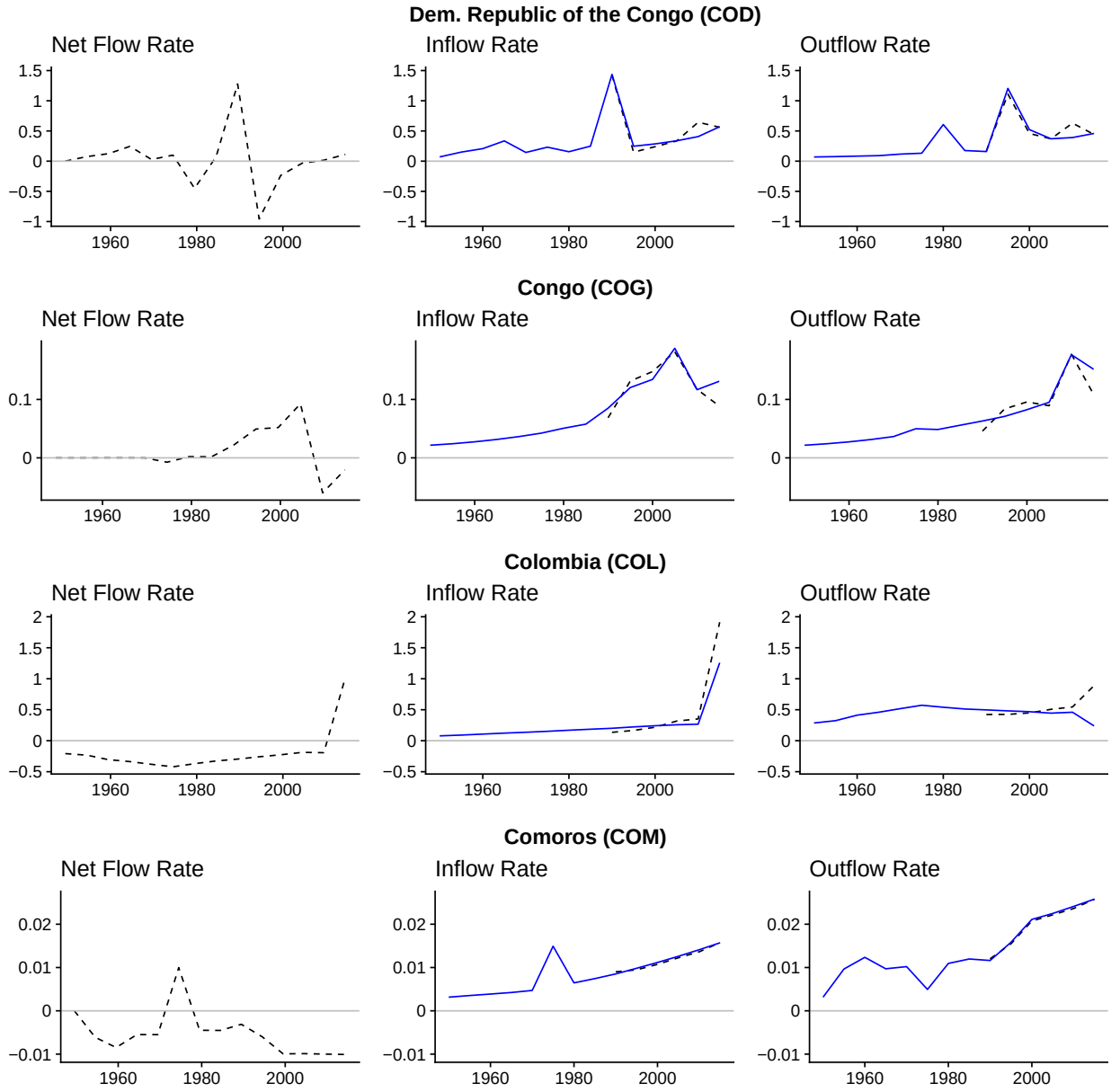


Figure A.14: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

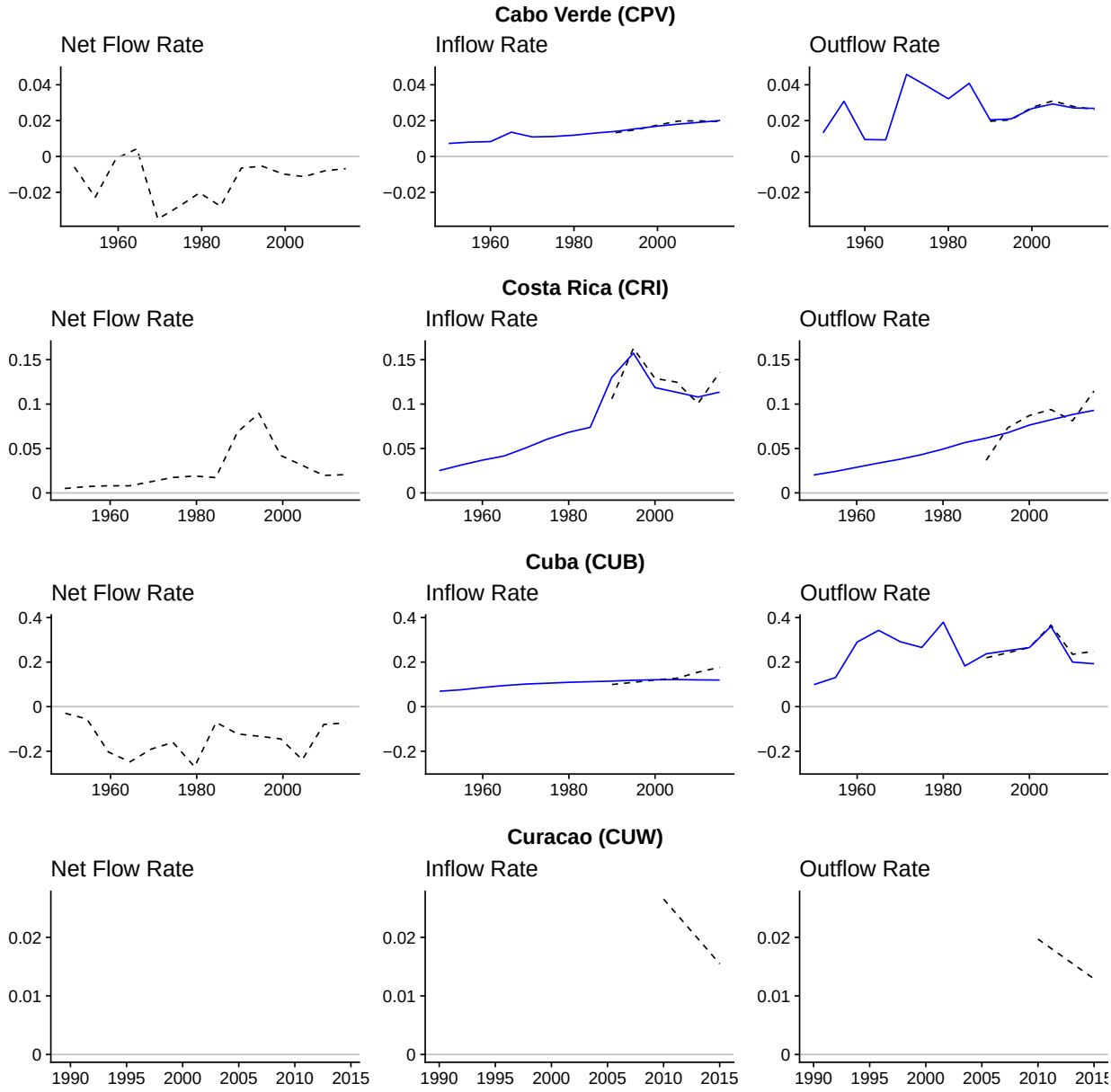


Figure A.15: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

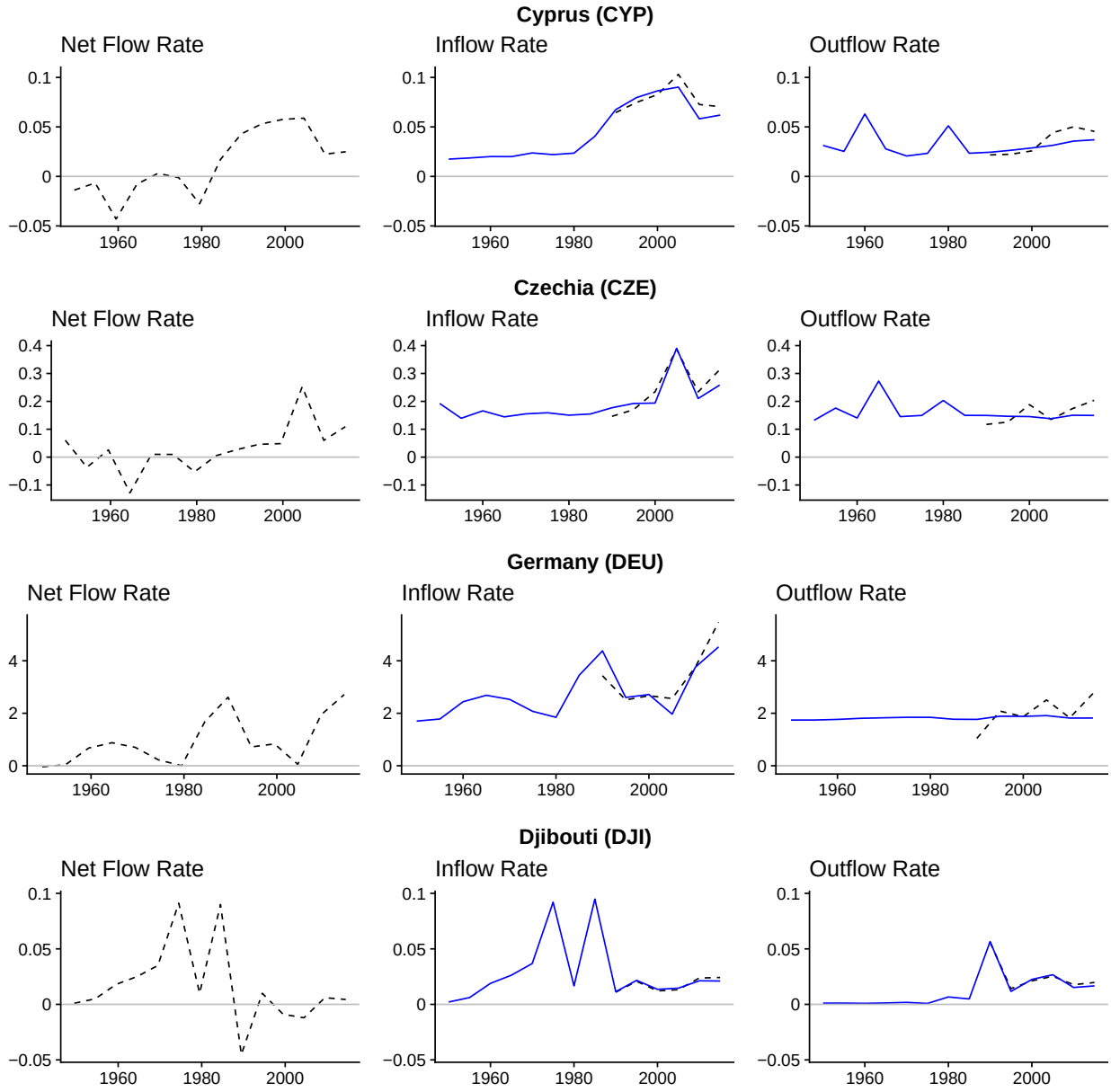


Figure A.16: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

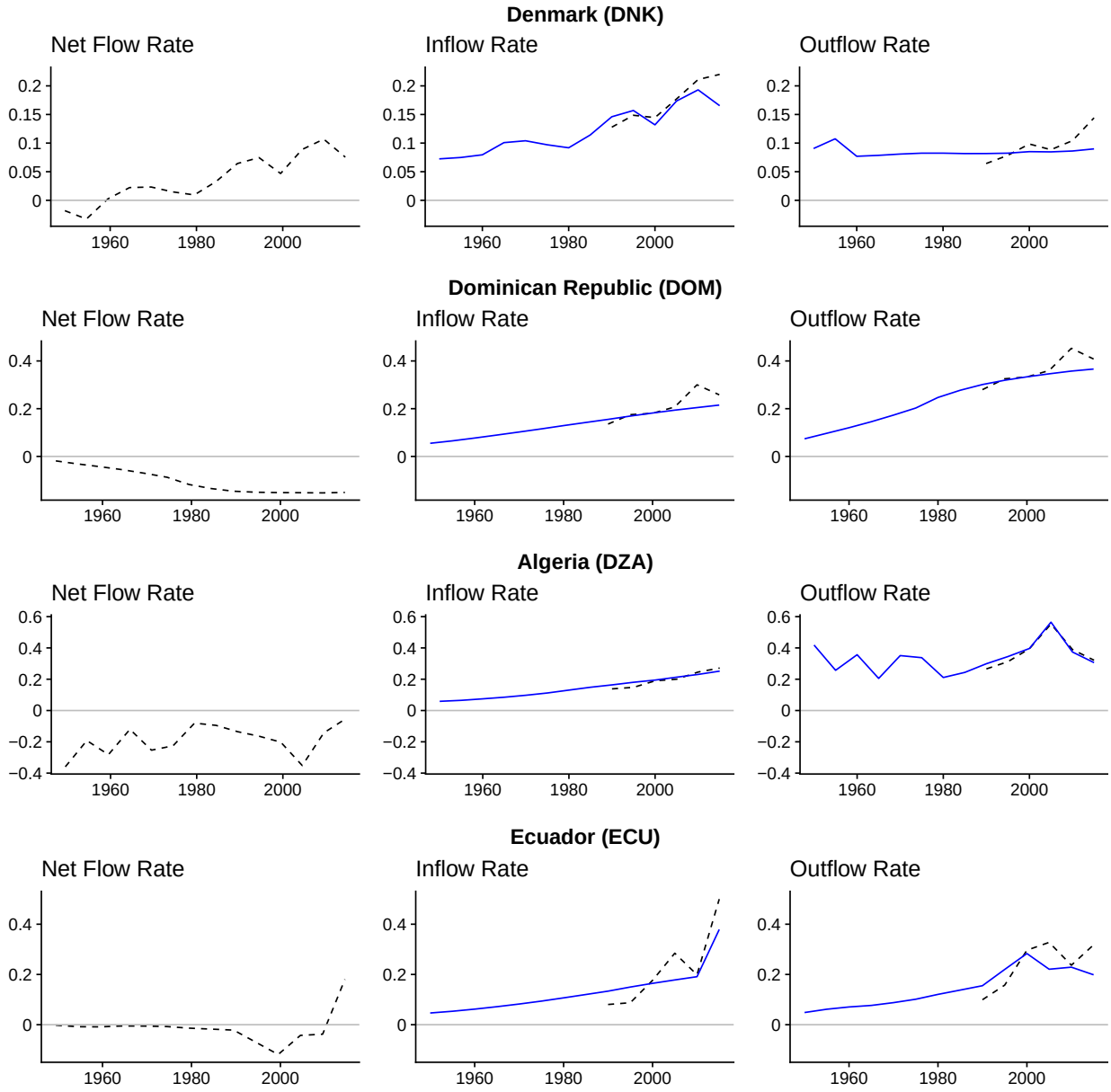


Figure A.17: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

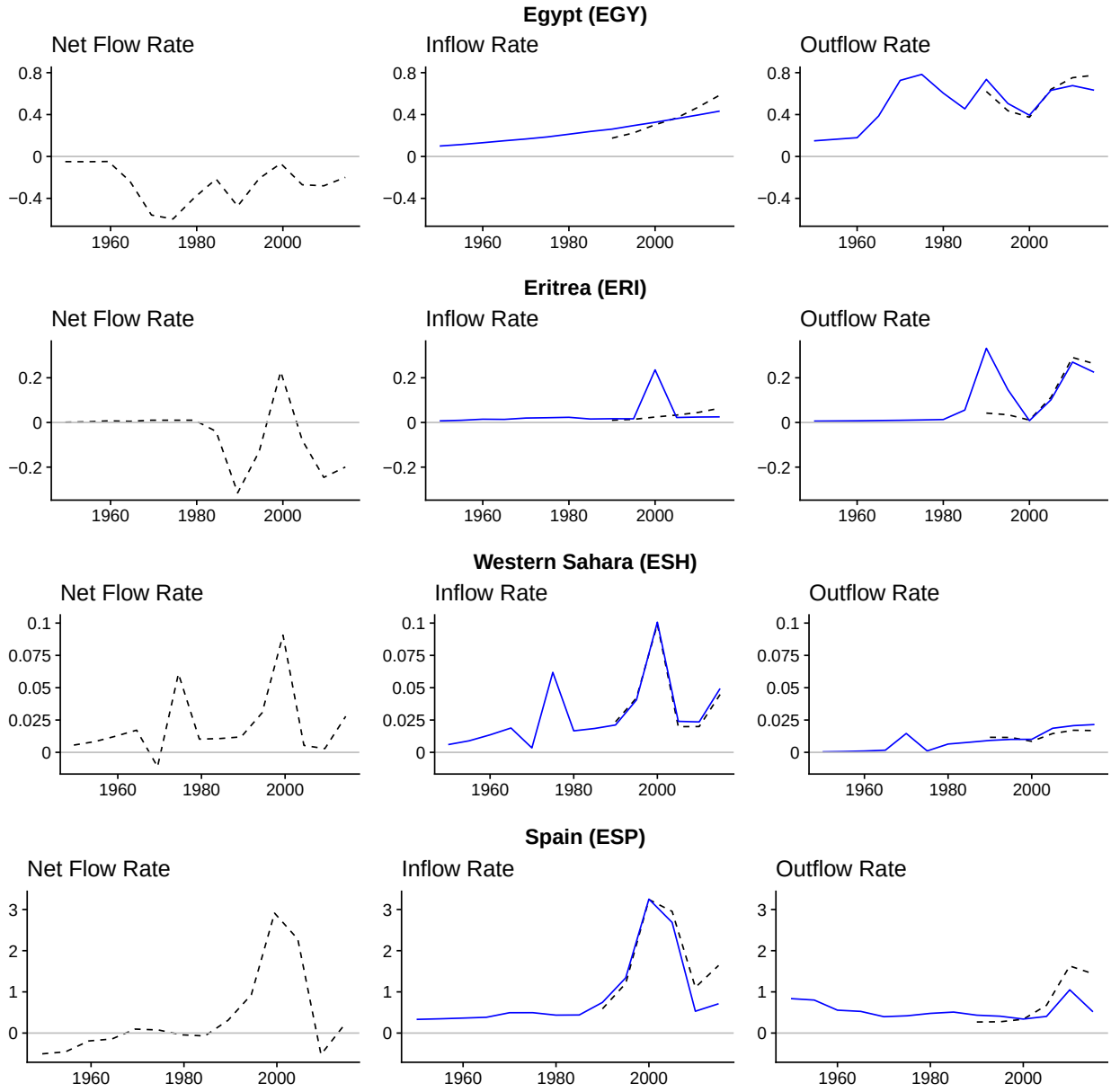


Figure A.18: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

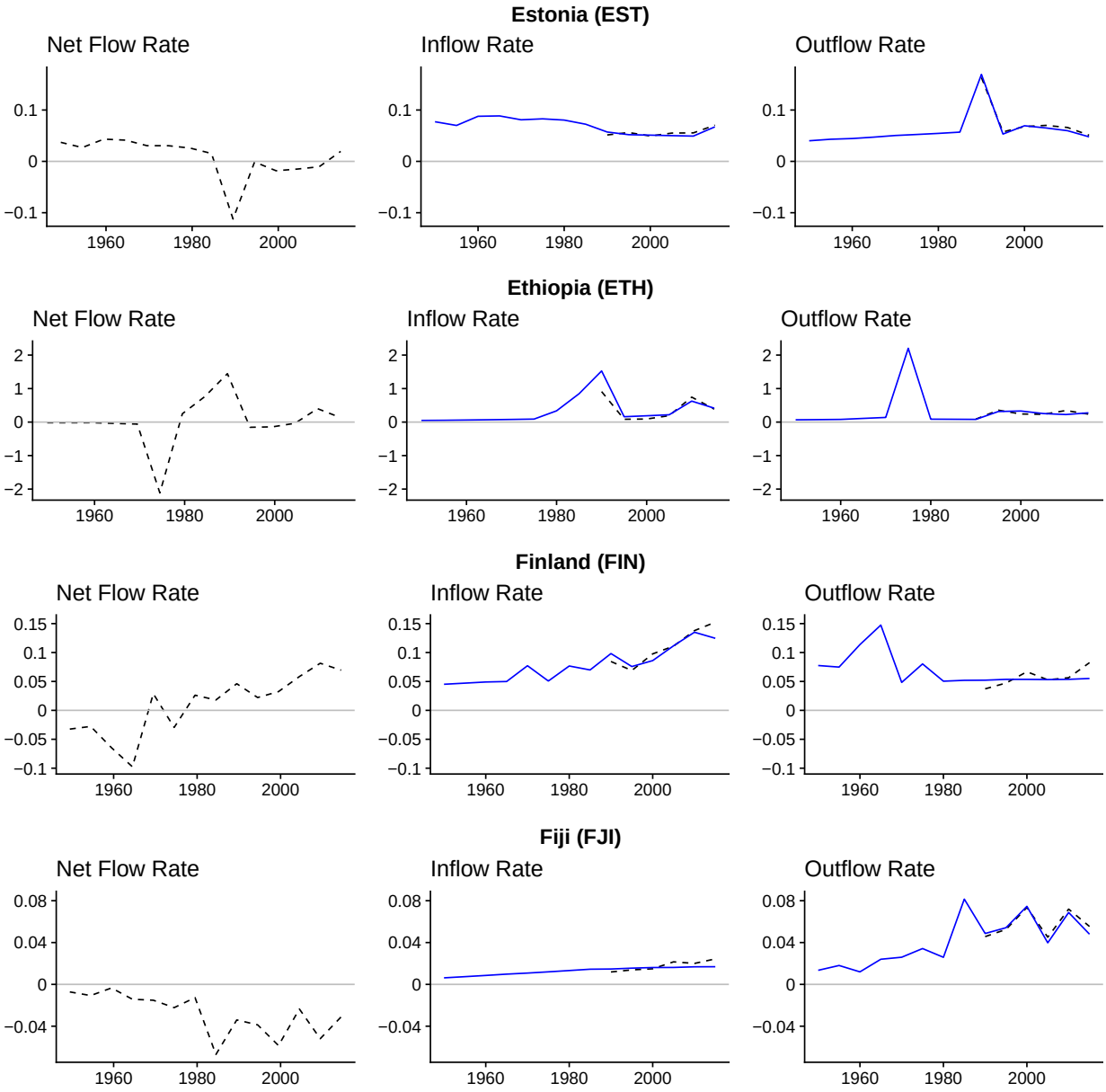


Figure A.19: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

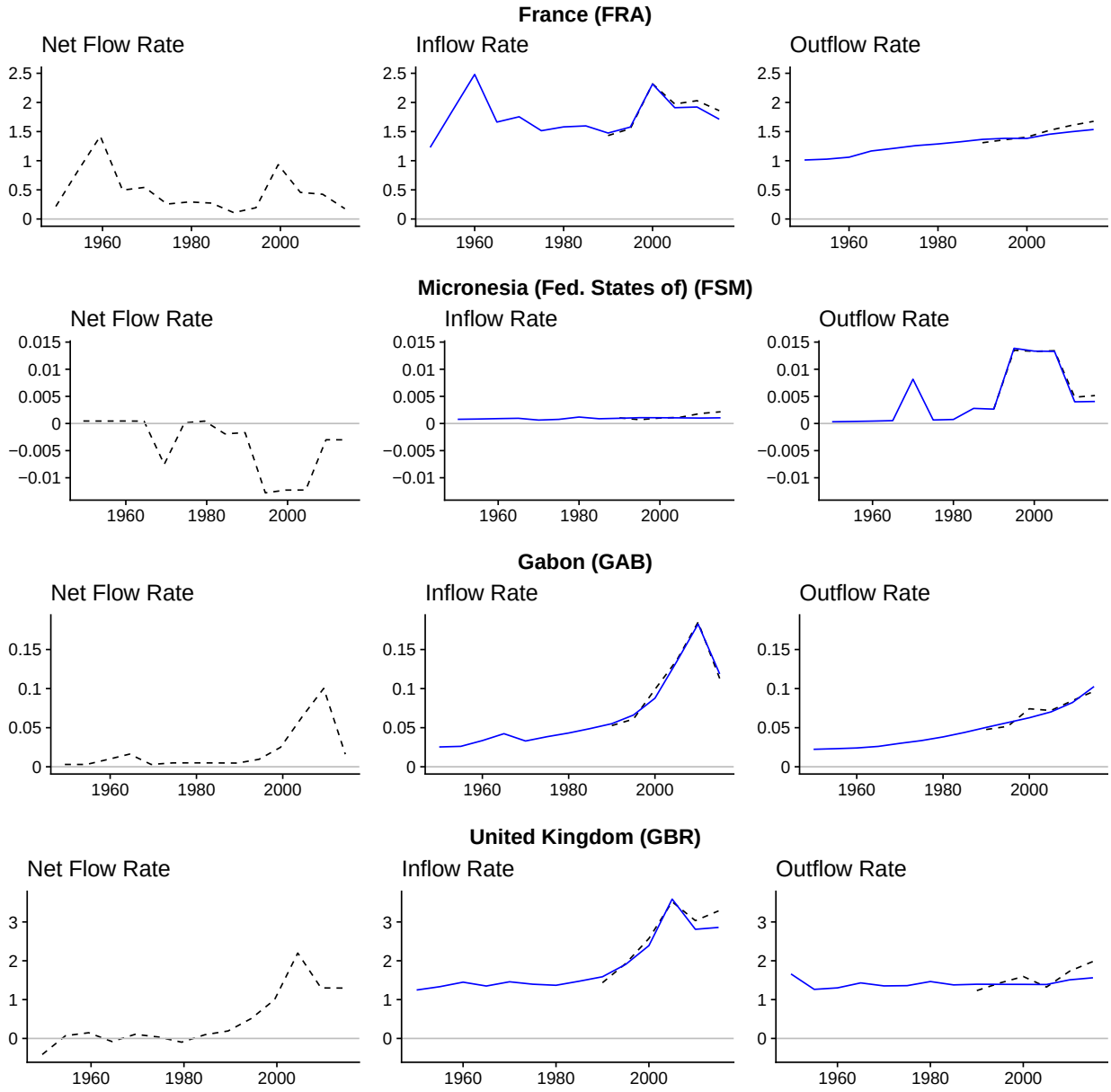


Figure A.20: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

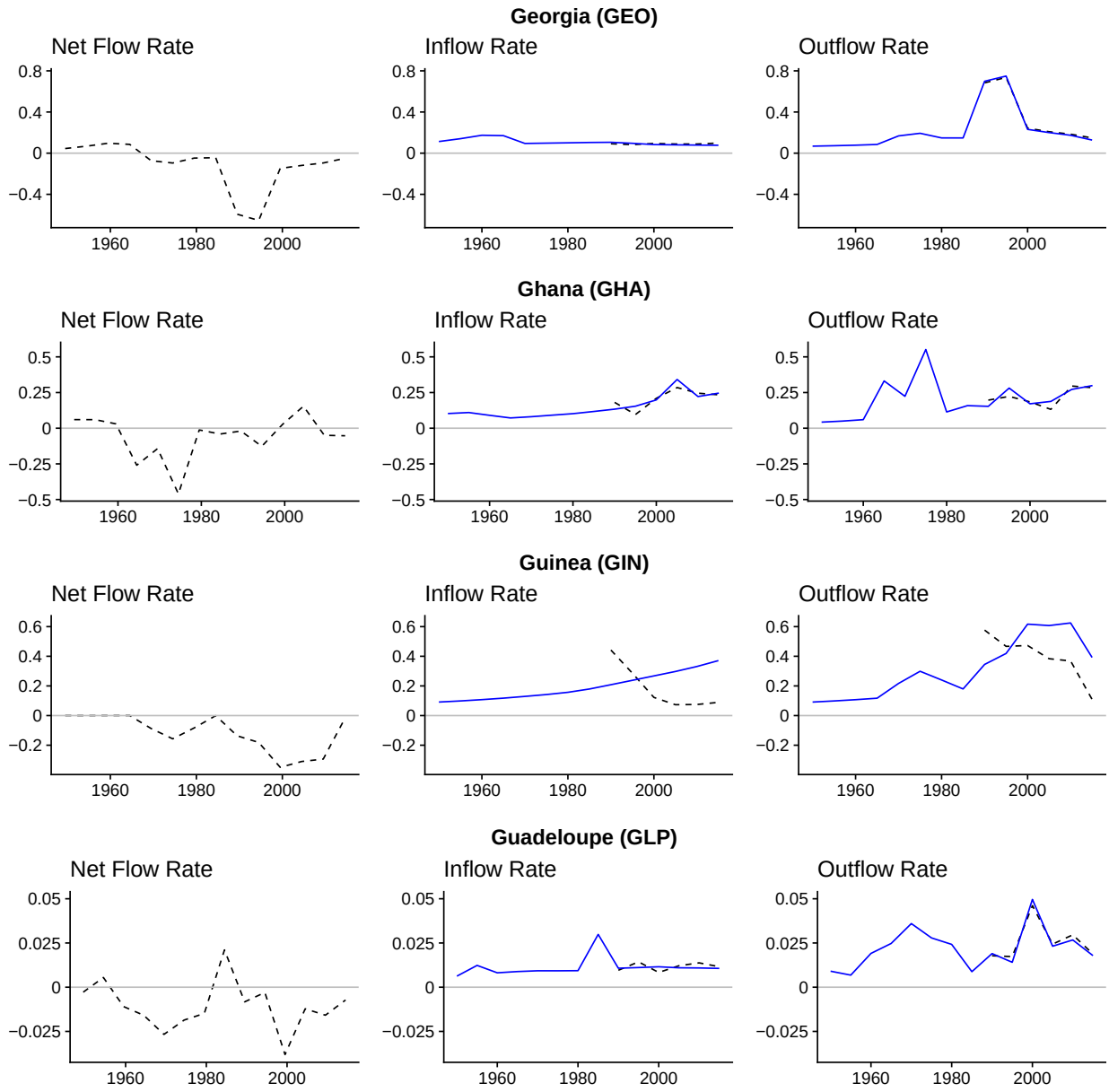


Figure A.21: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

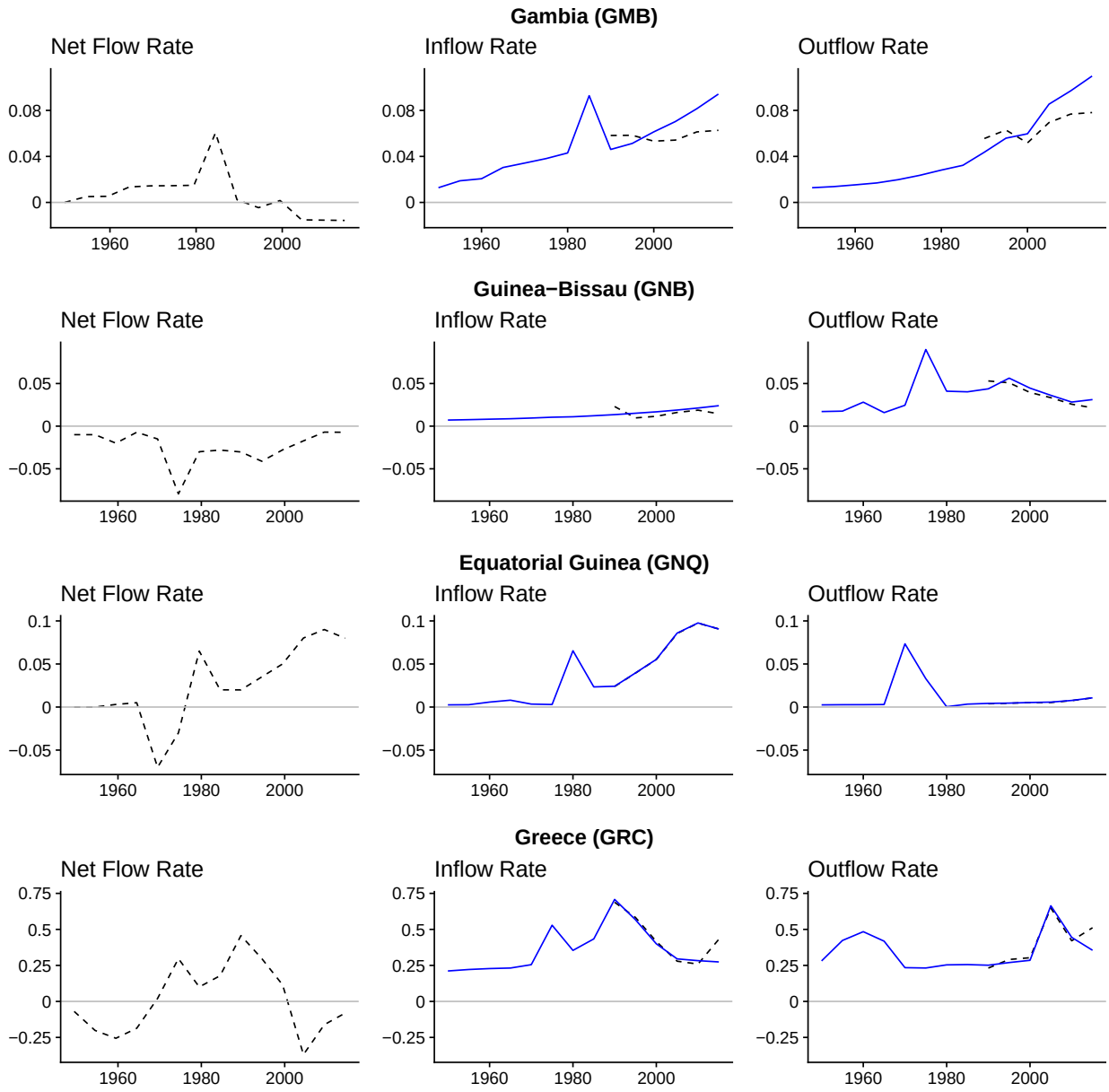


Figure A.22: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

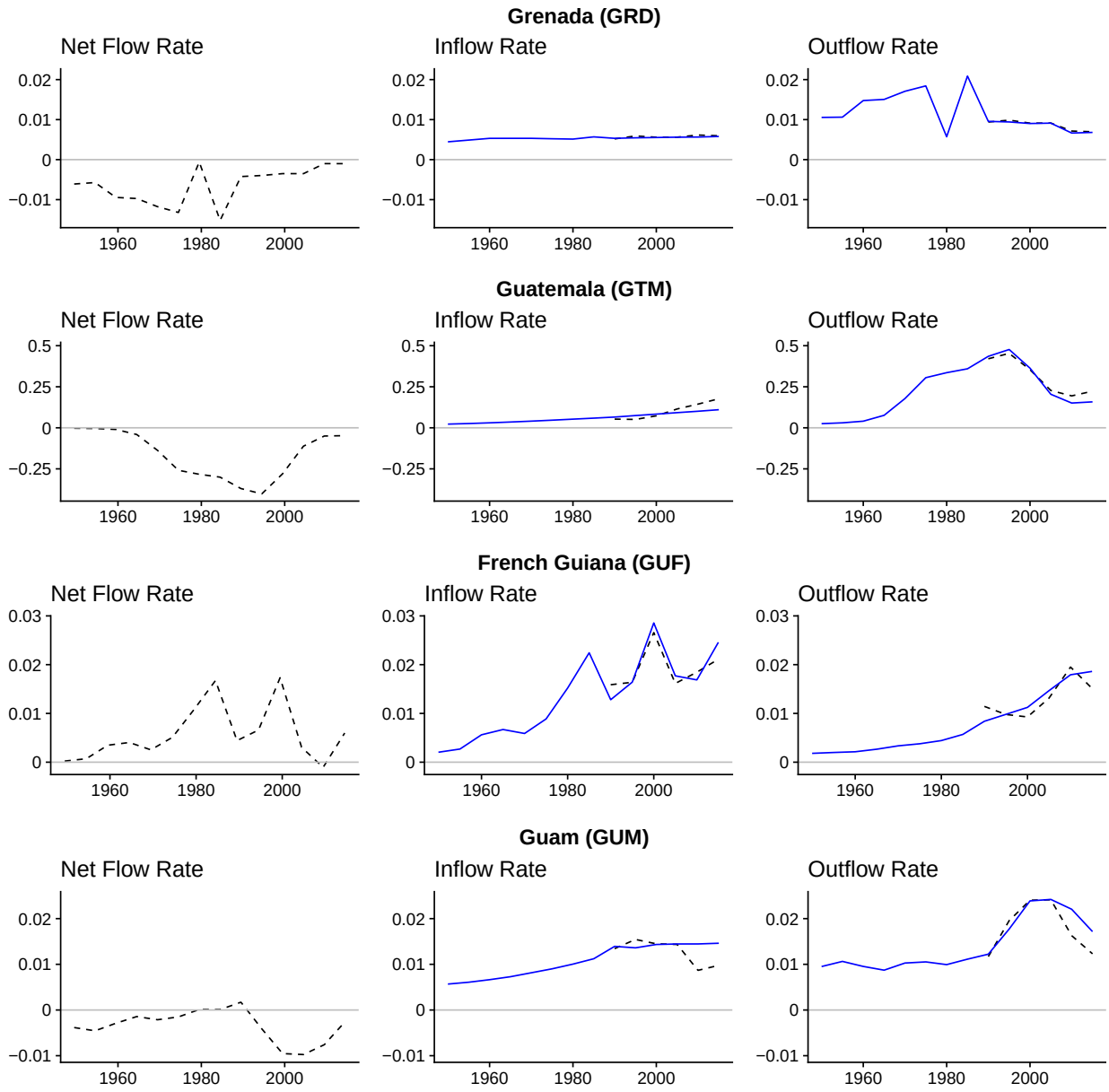


Figure A.23: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

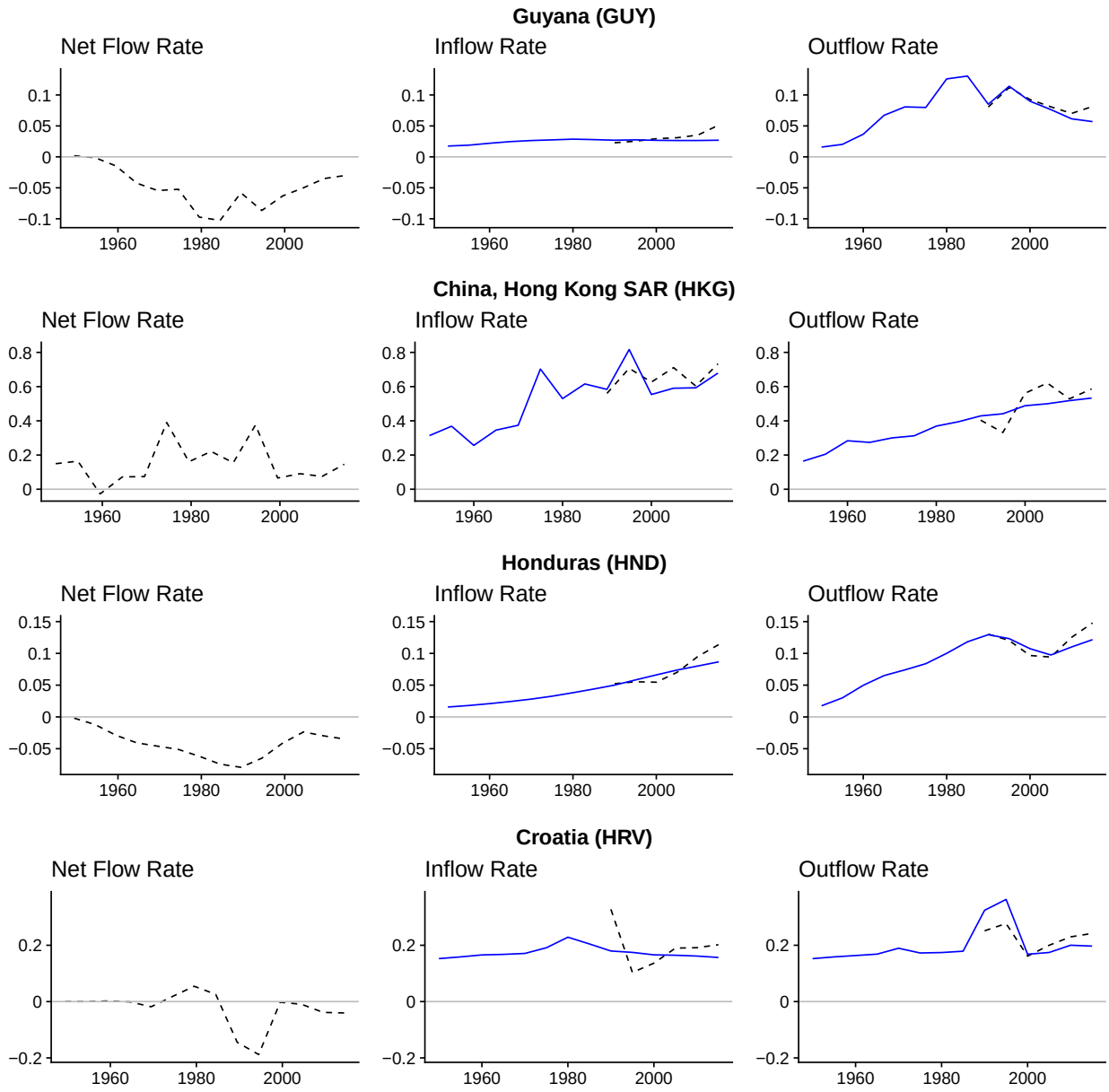


Figure A.24: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

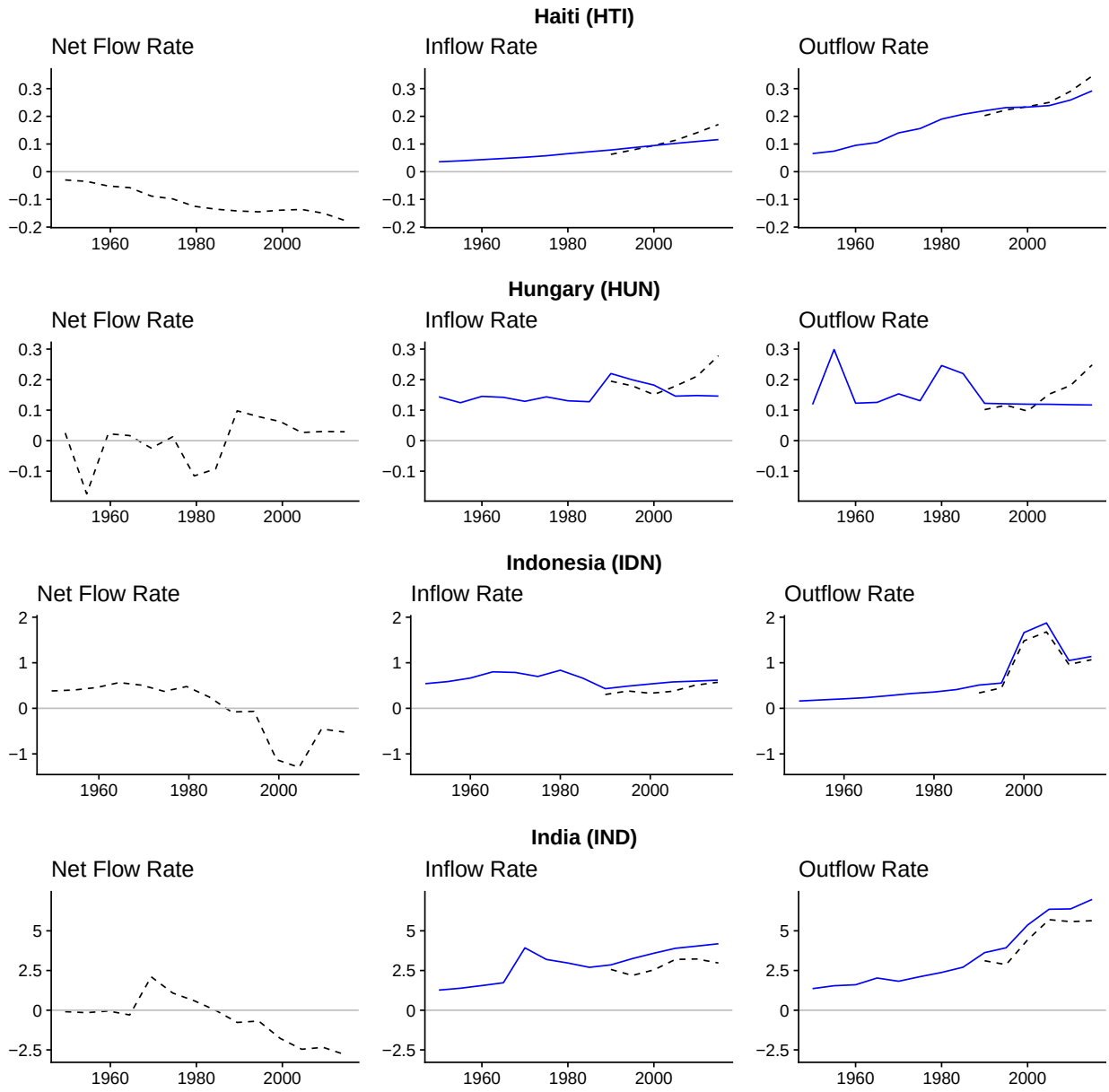


Figure A.25: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

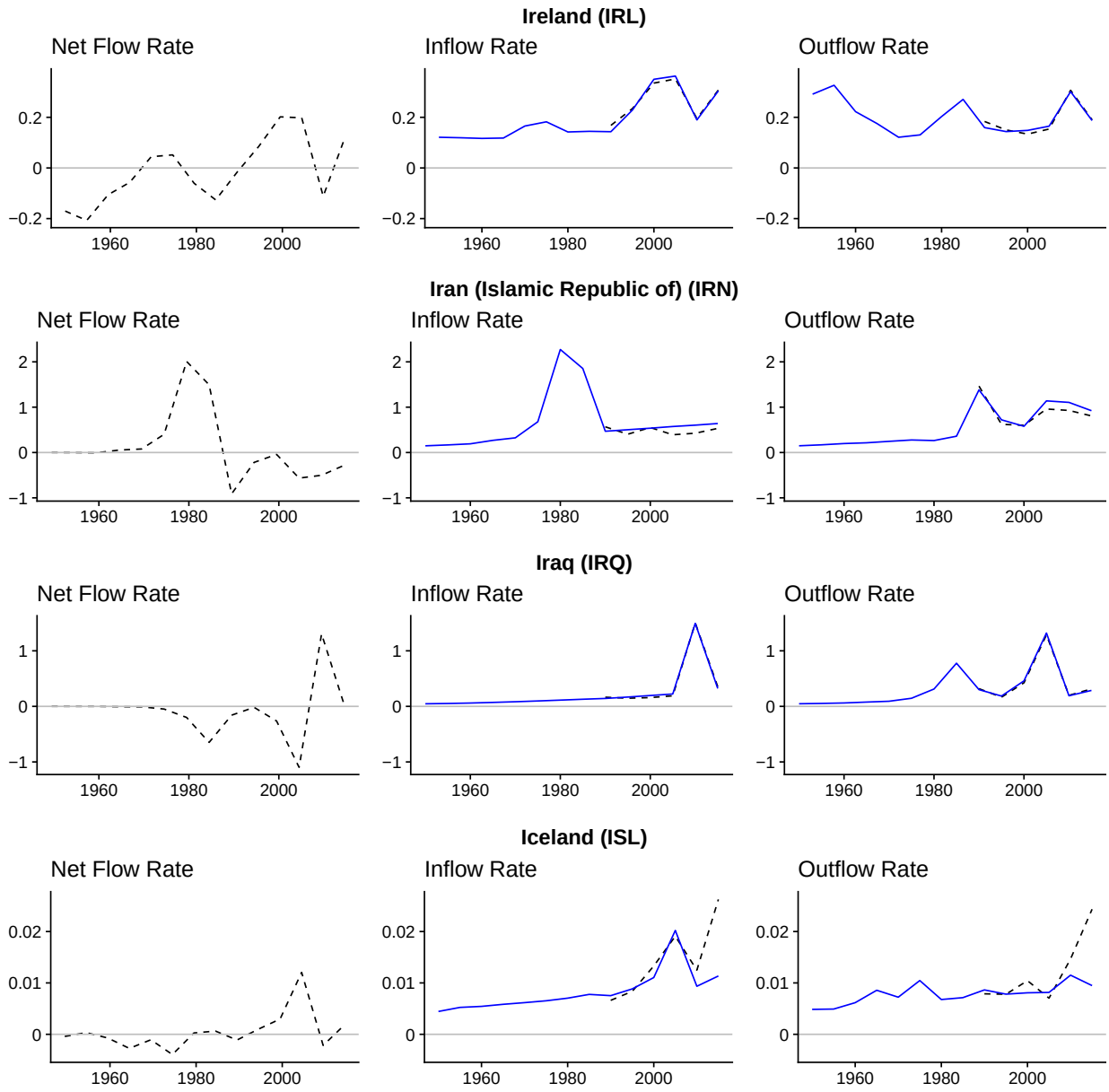


Figure A.26: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

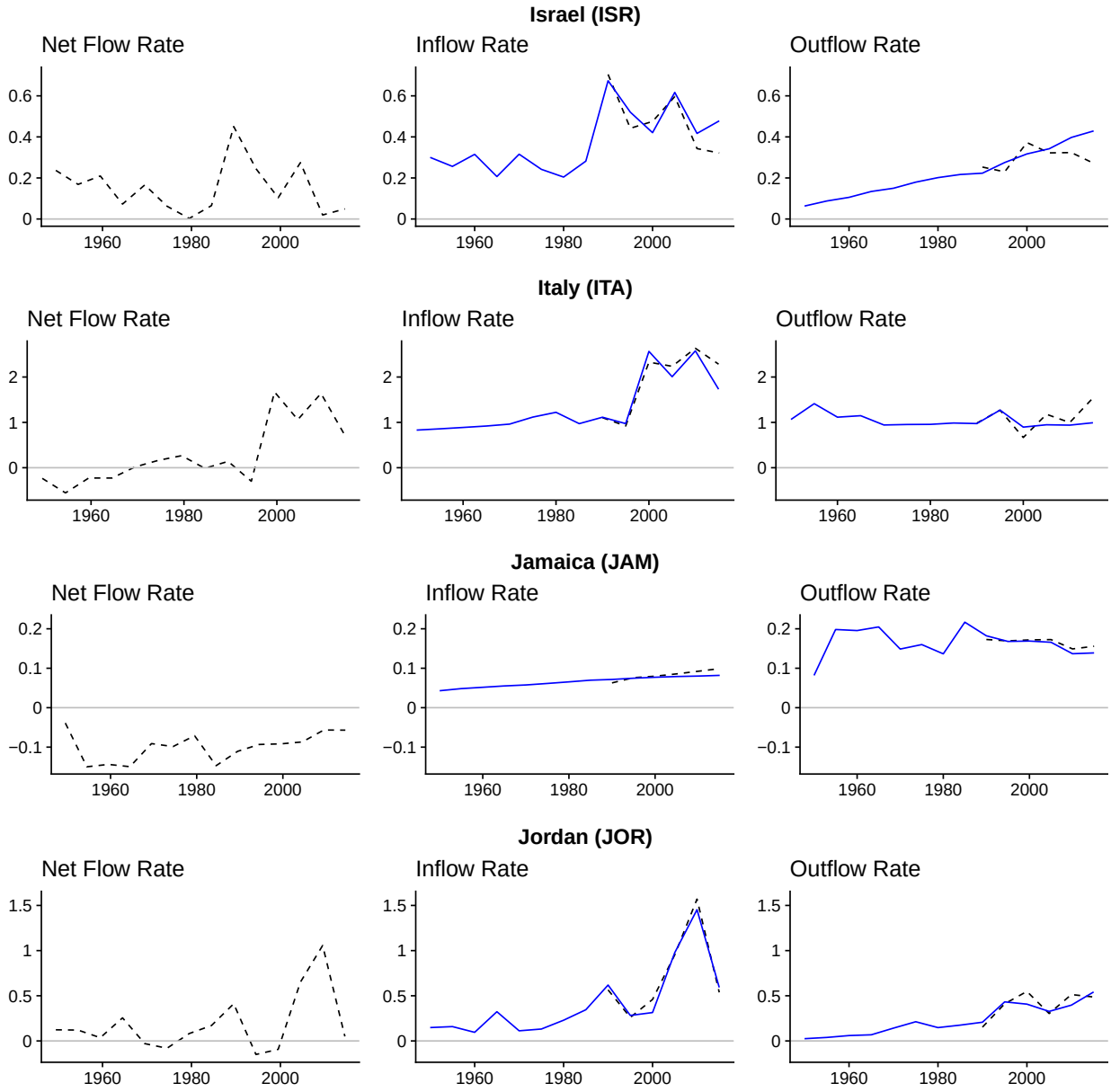


Figure A.27: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

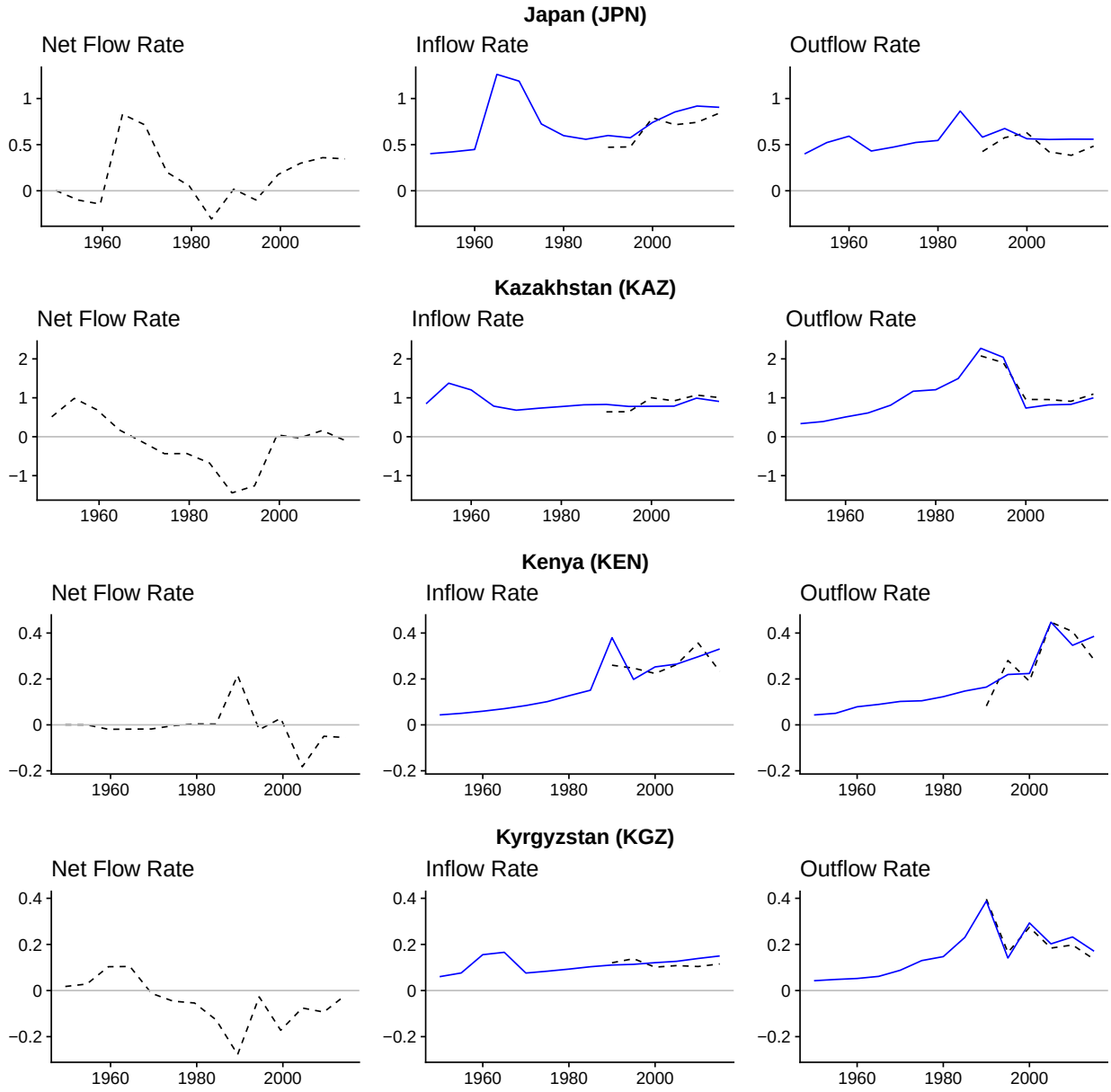


Figure A.28: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

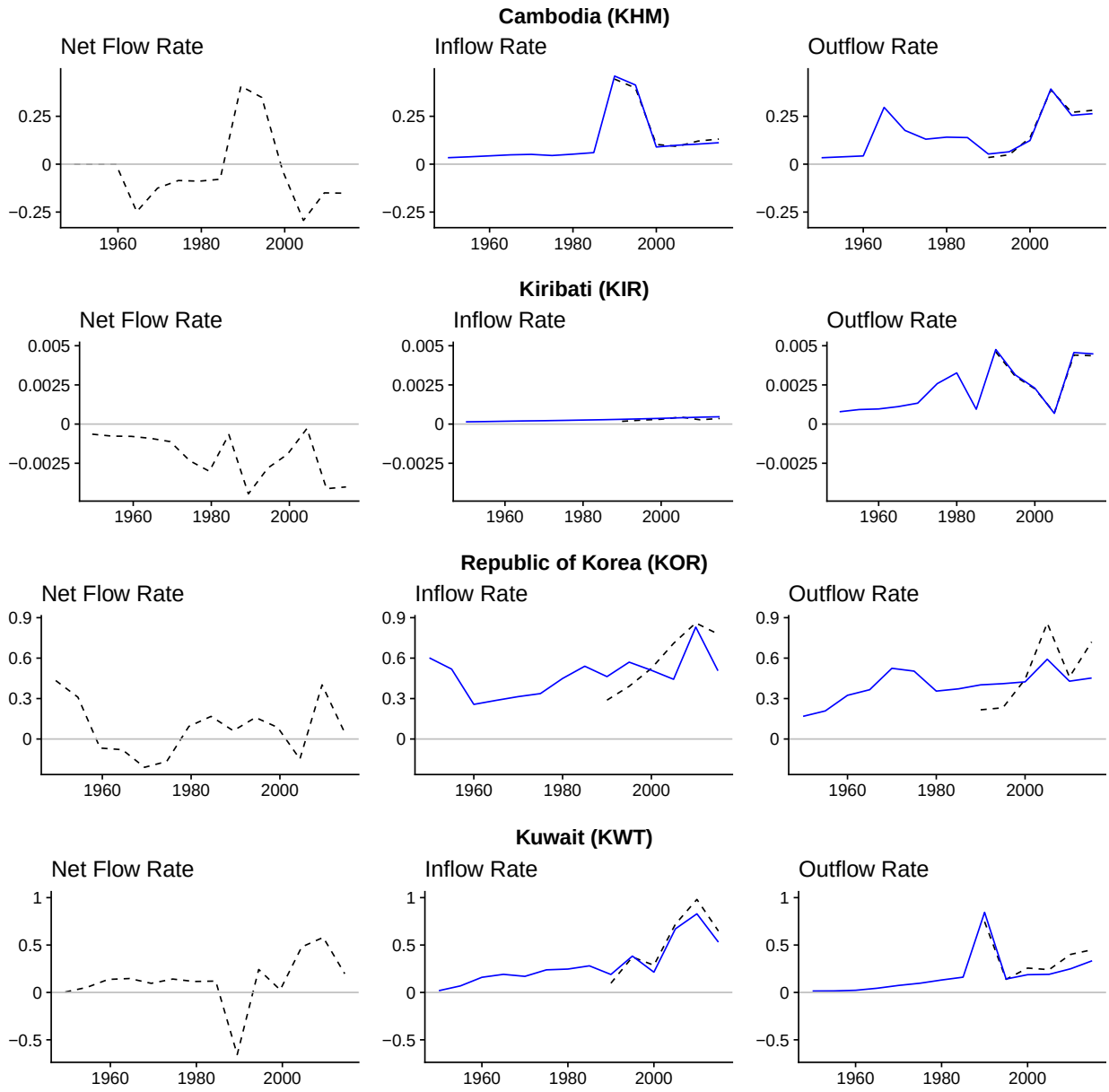


Figure A.29: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

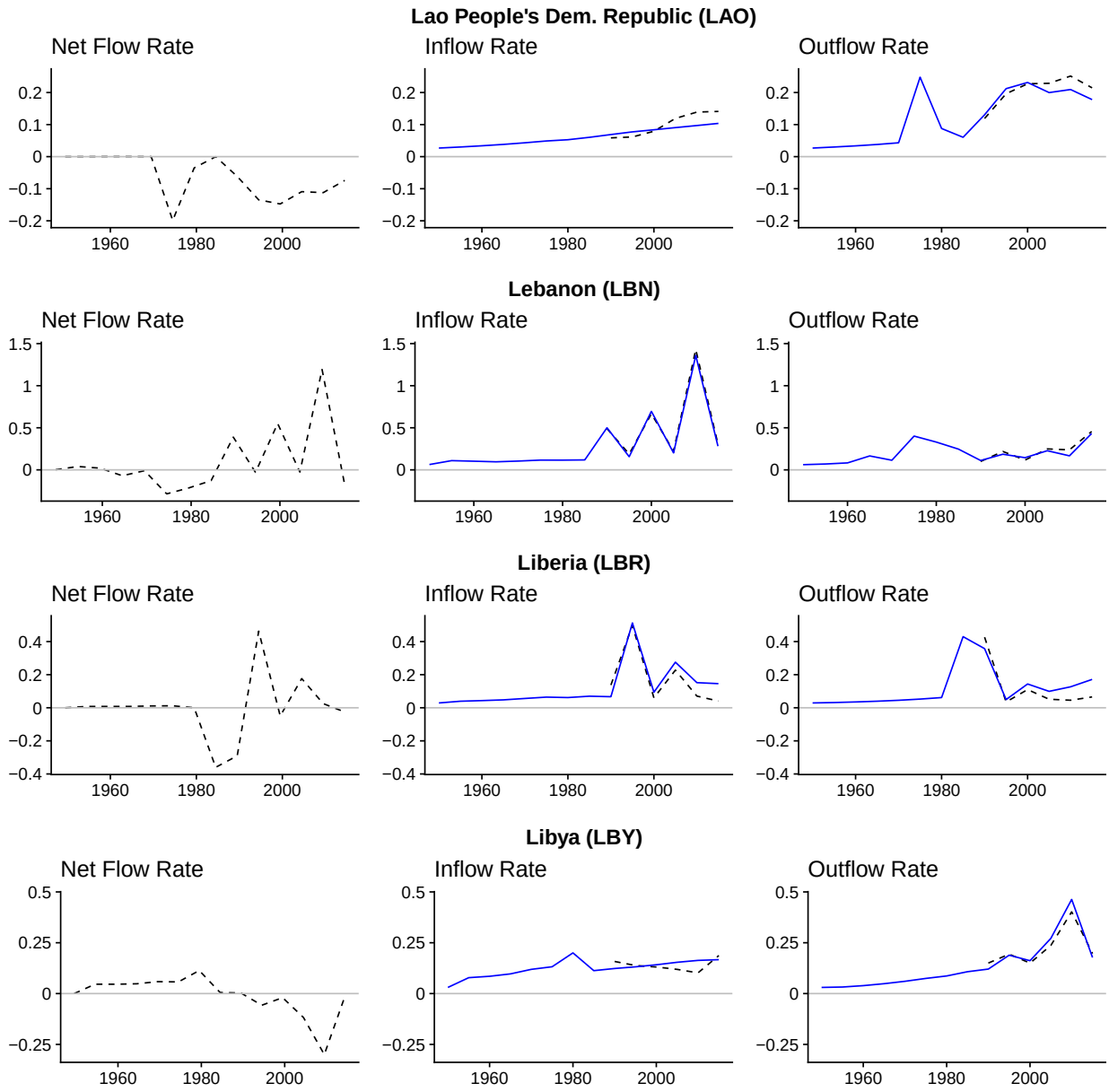


Figure A.30: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

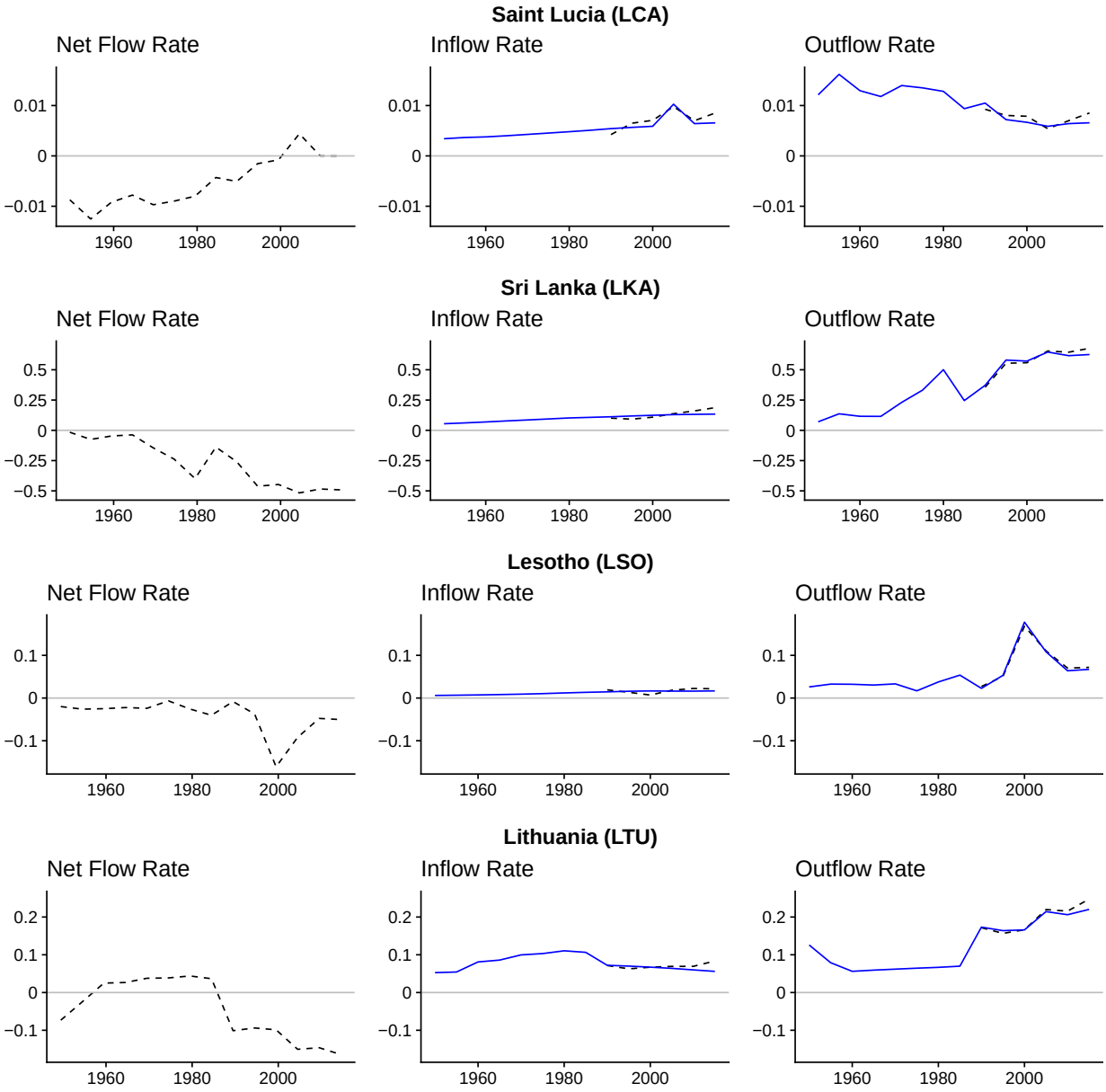


Figure A.31: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

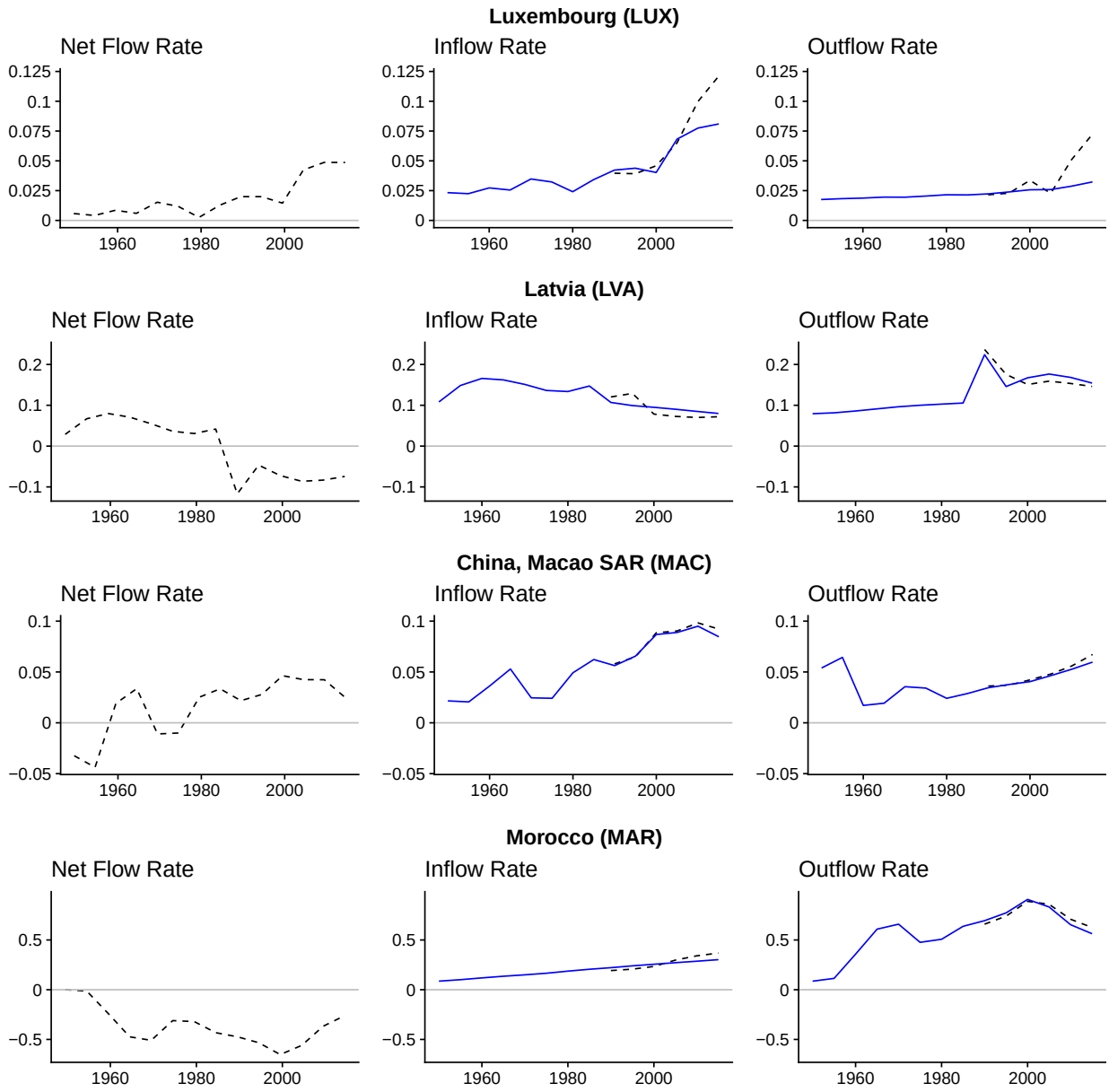


Figure A.32: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

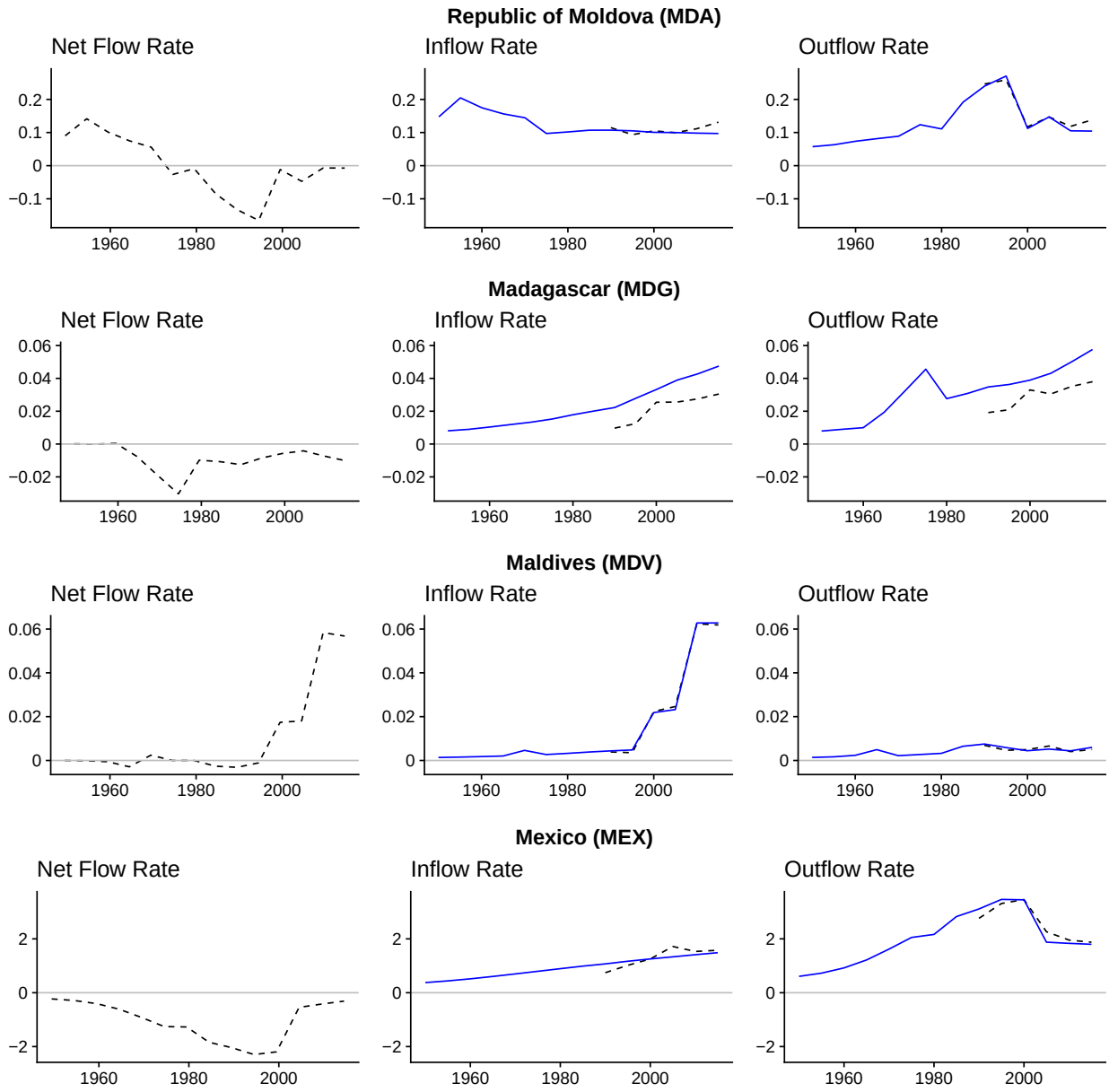


Figure A.33: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

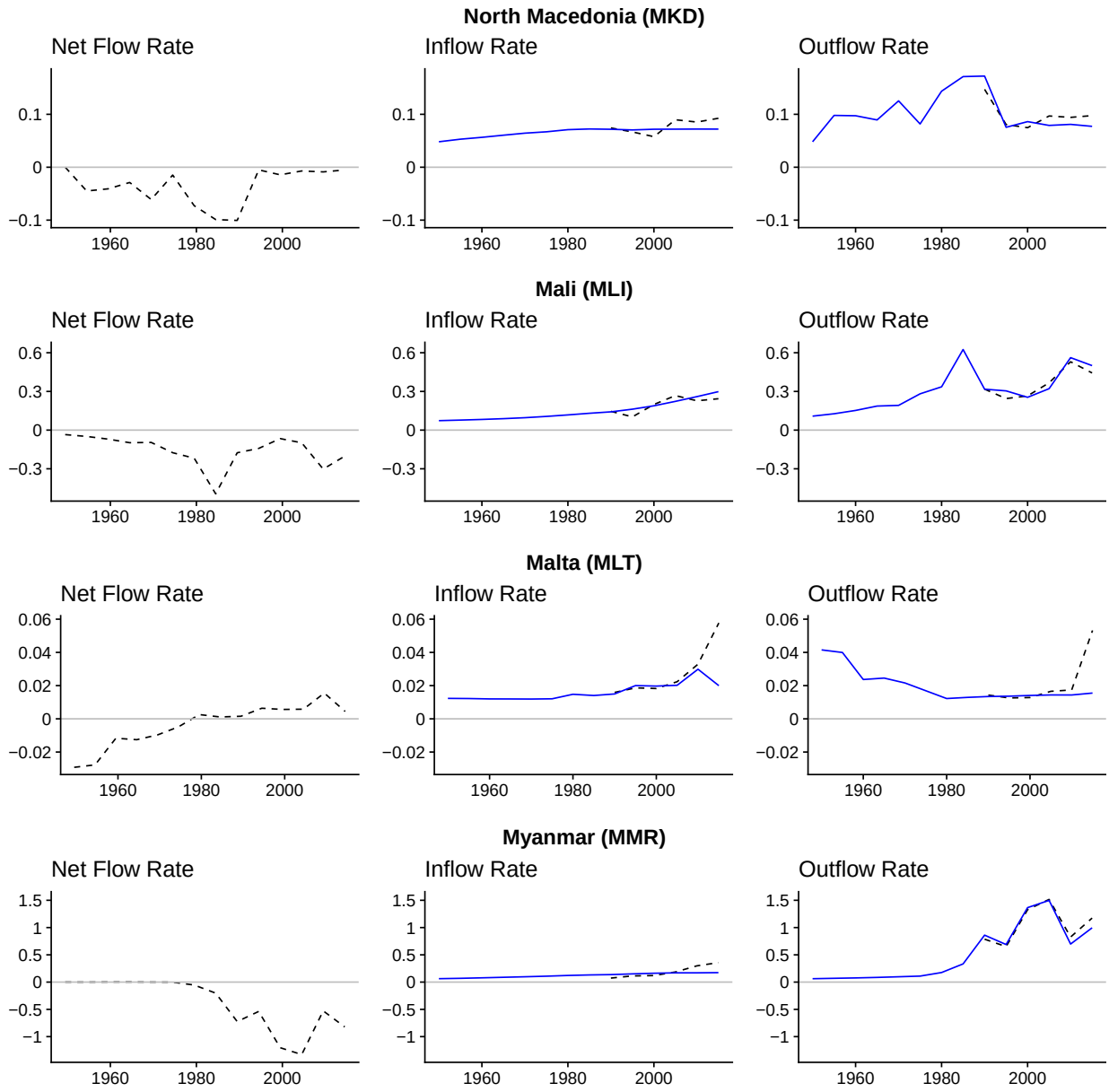


Figure A.34: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

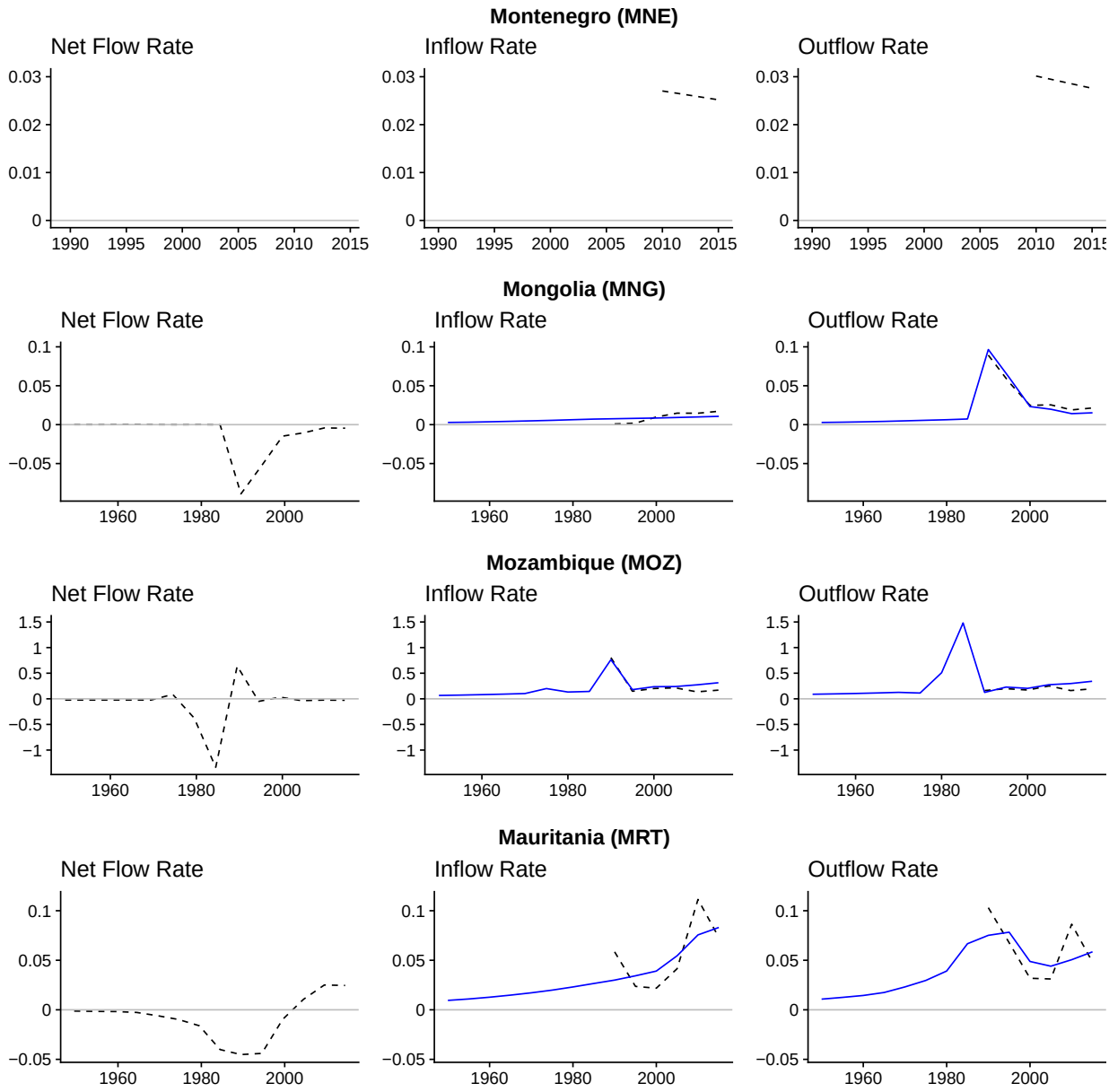


Figure A.35: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

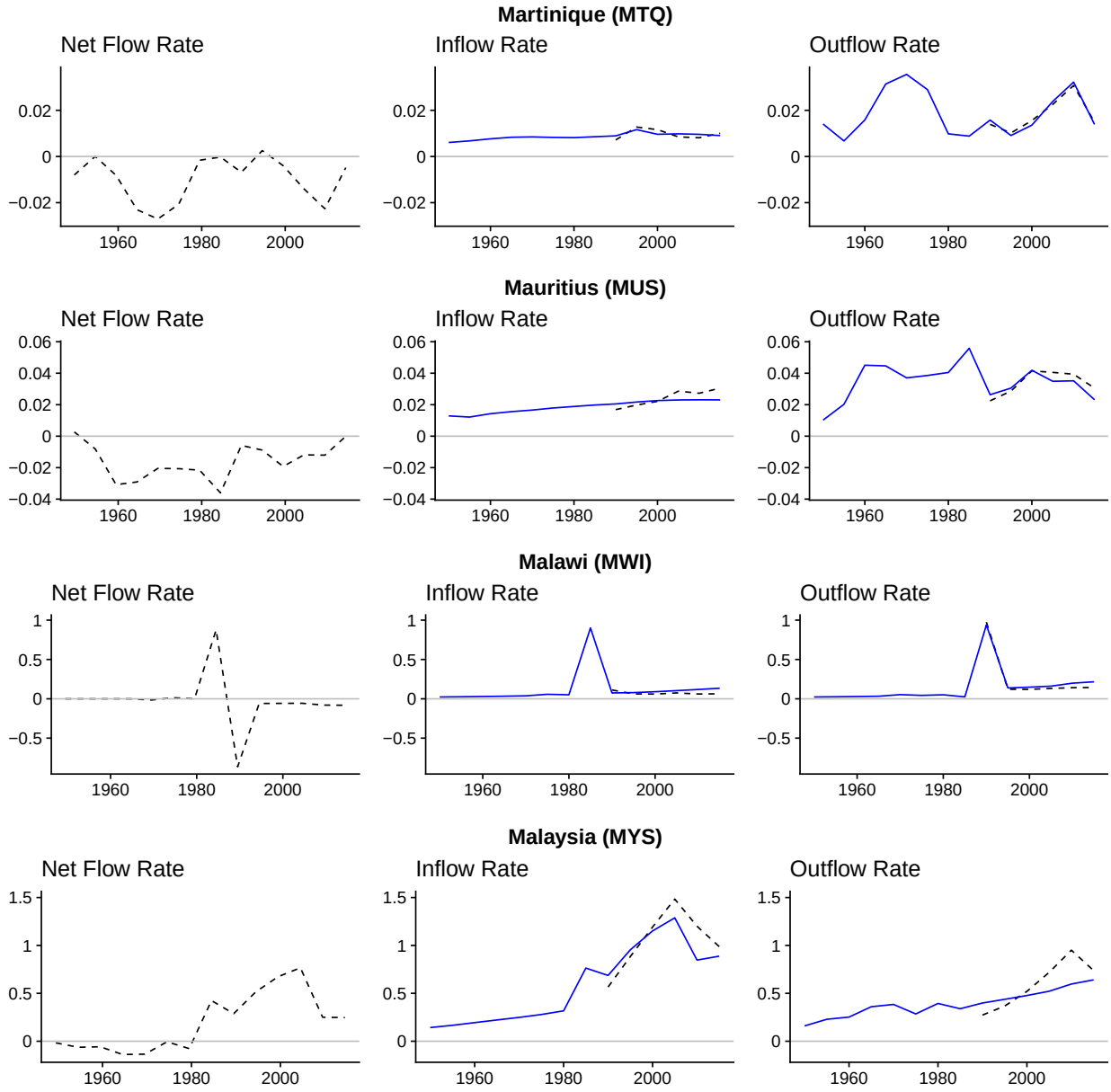


Figure A.36: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

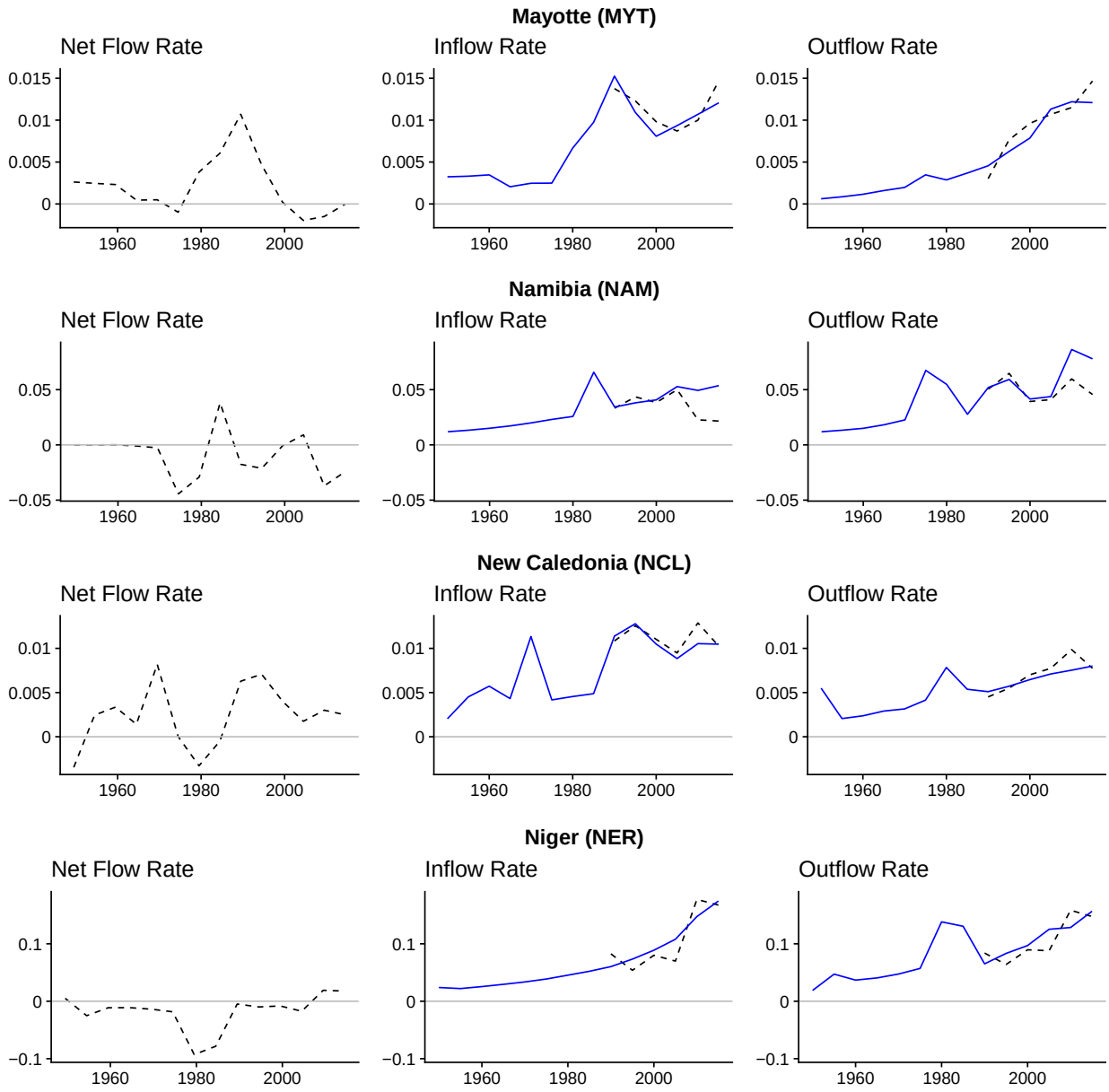


Figure A.37: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

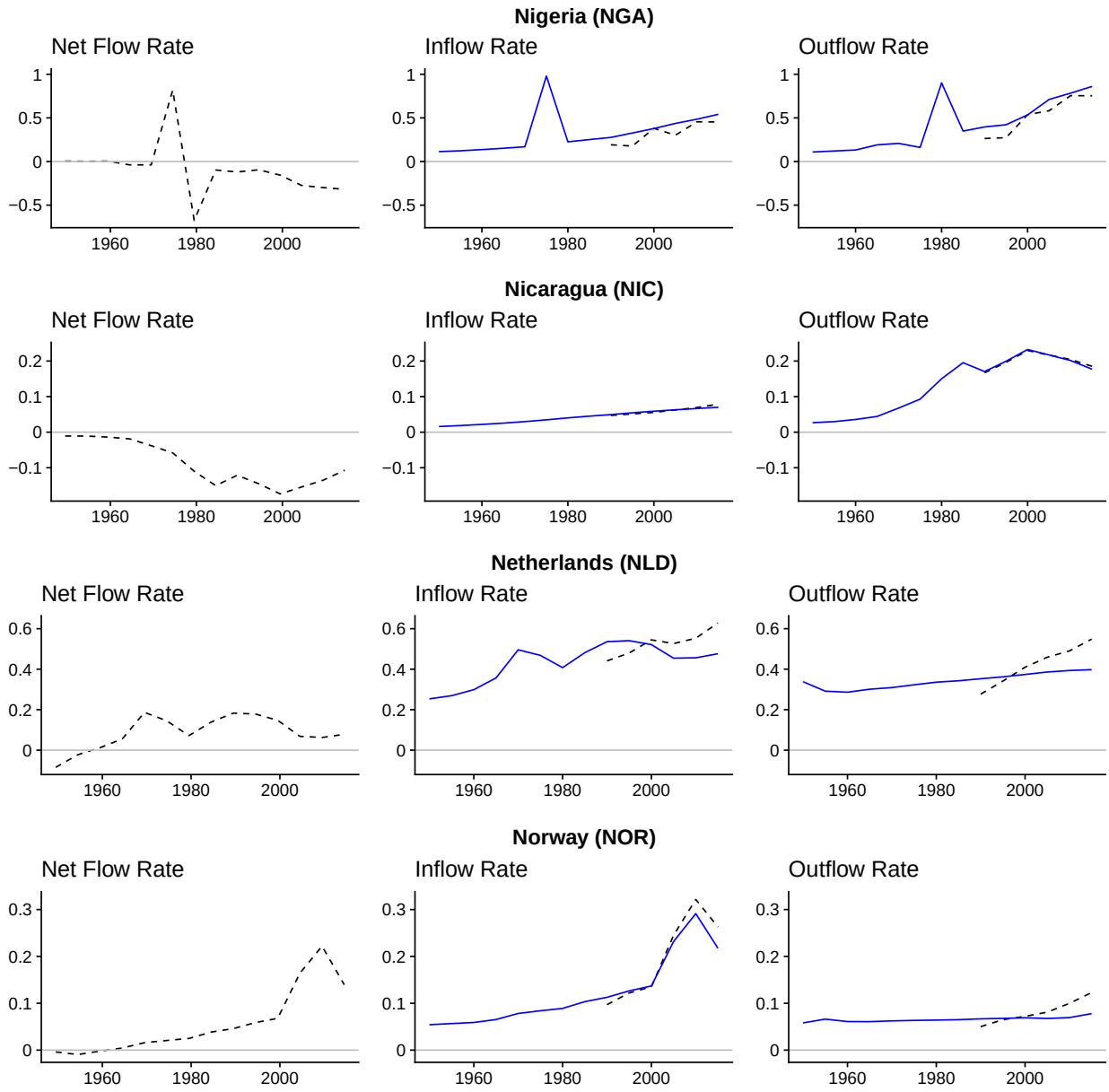


Figure A.38: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

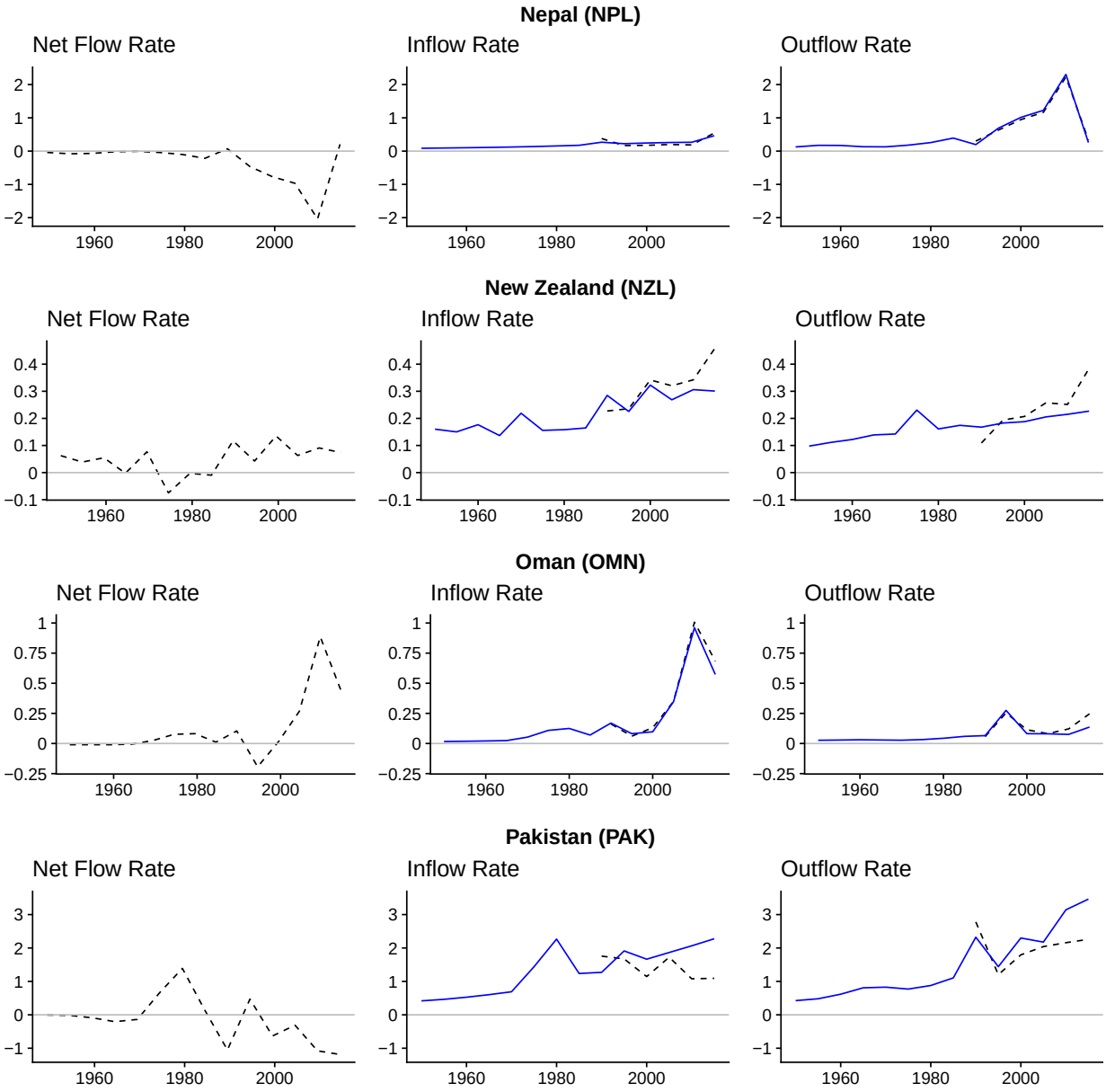


Figure A.39: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

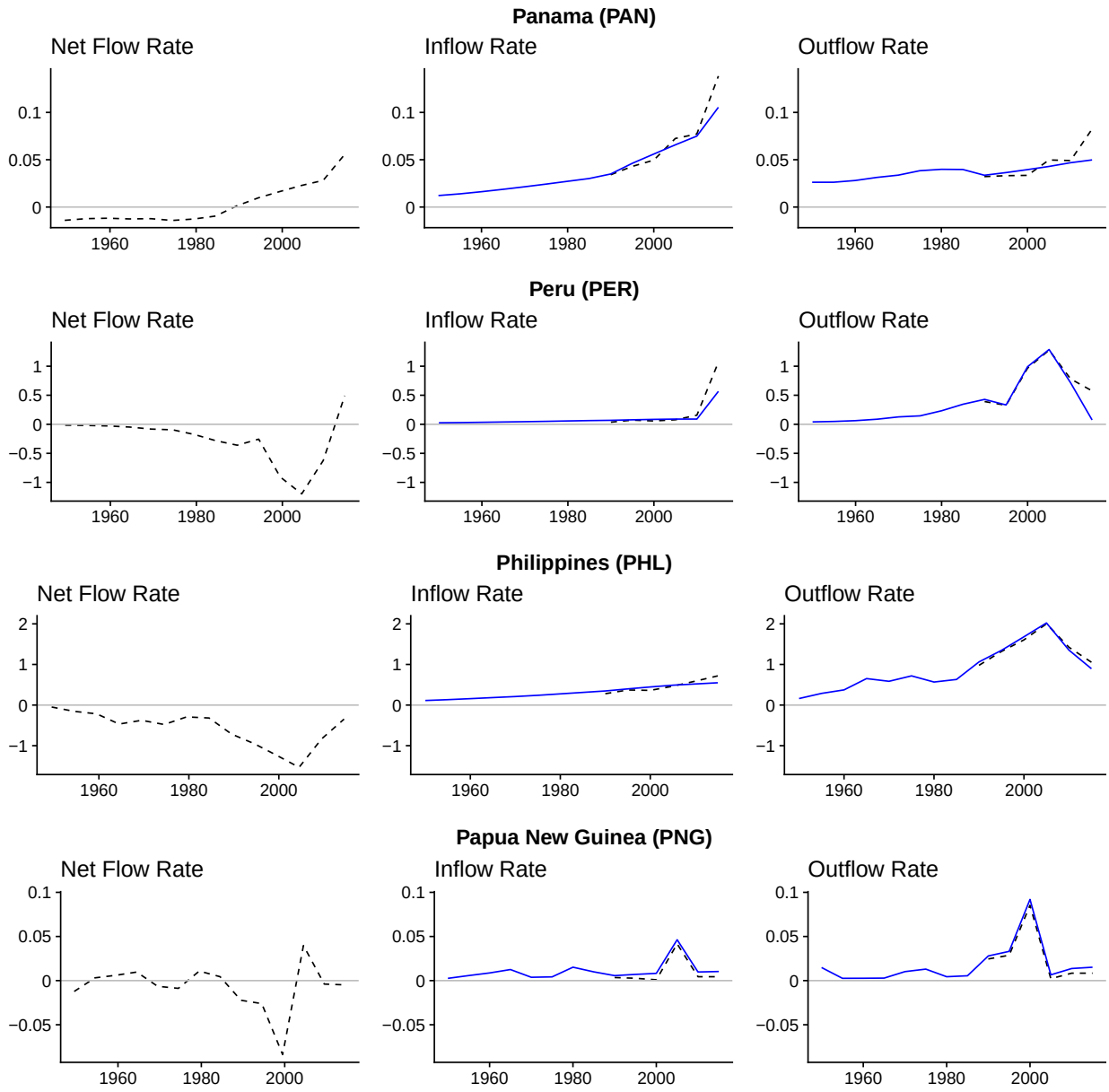


Figure A.40: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

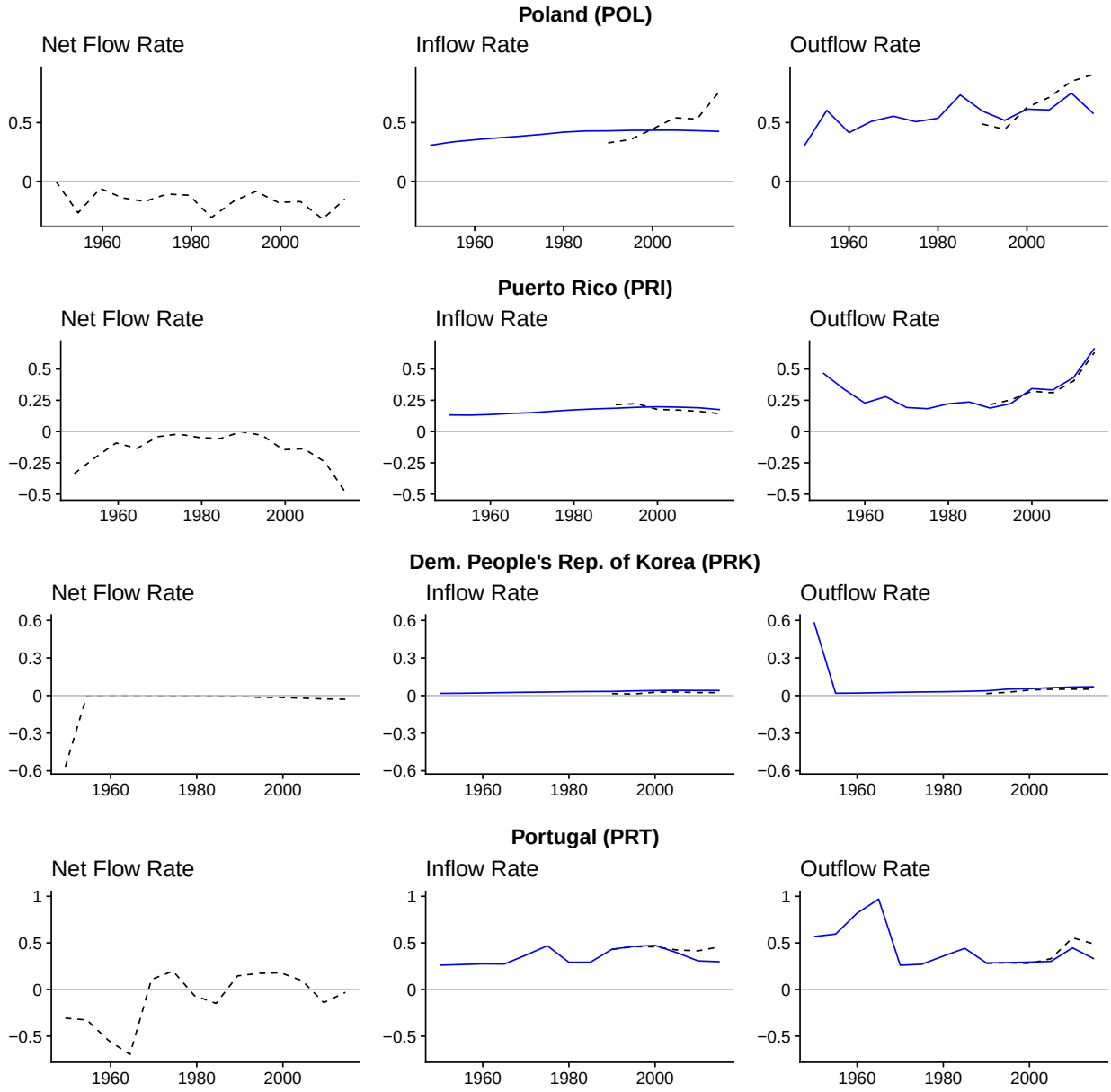


Figure A.41: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

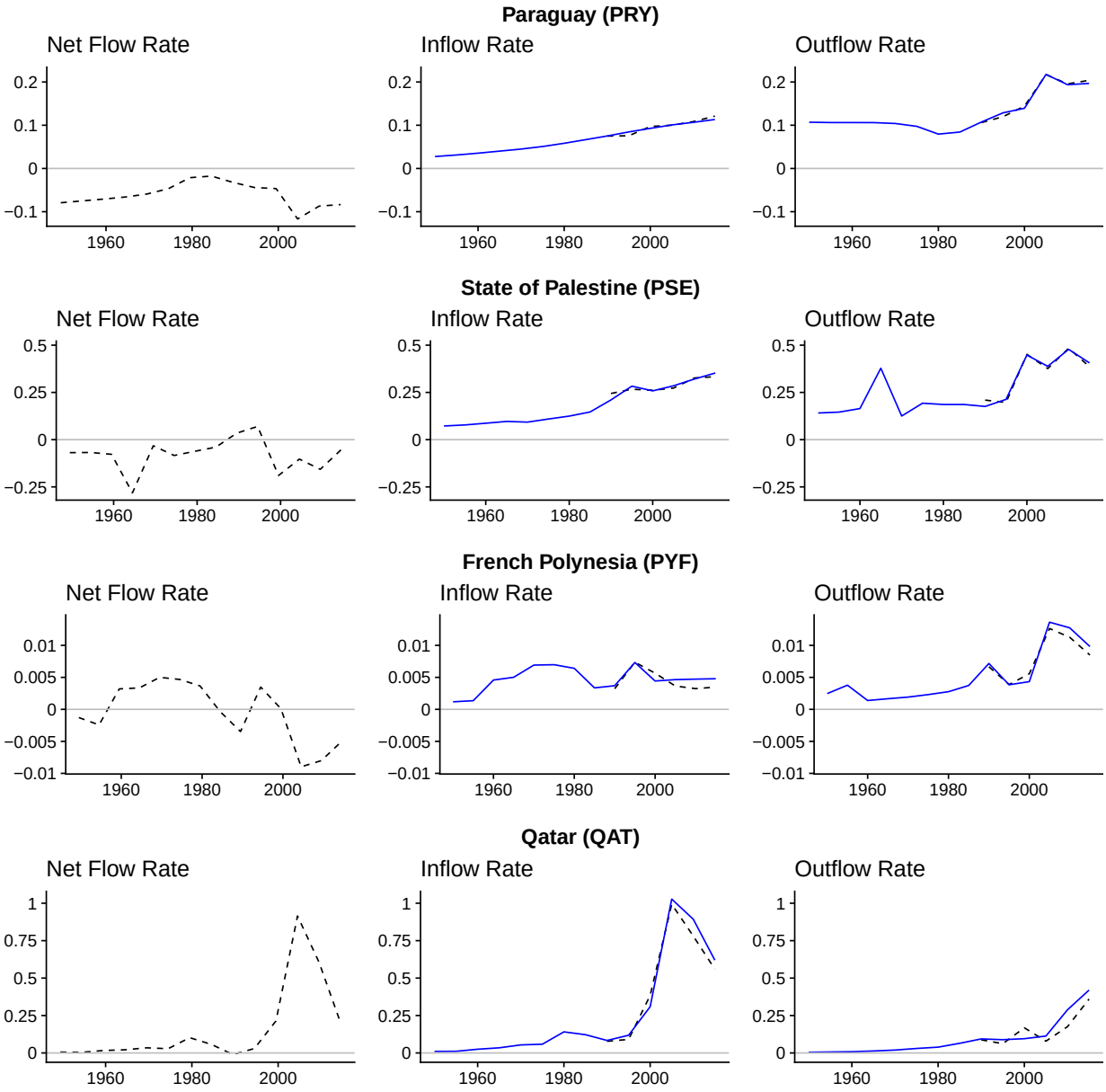


Figure A.42: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

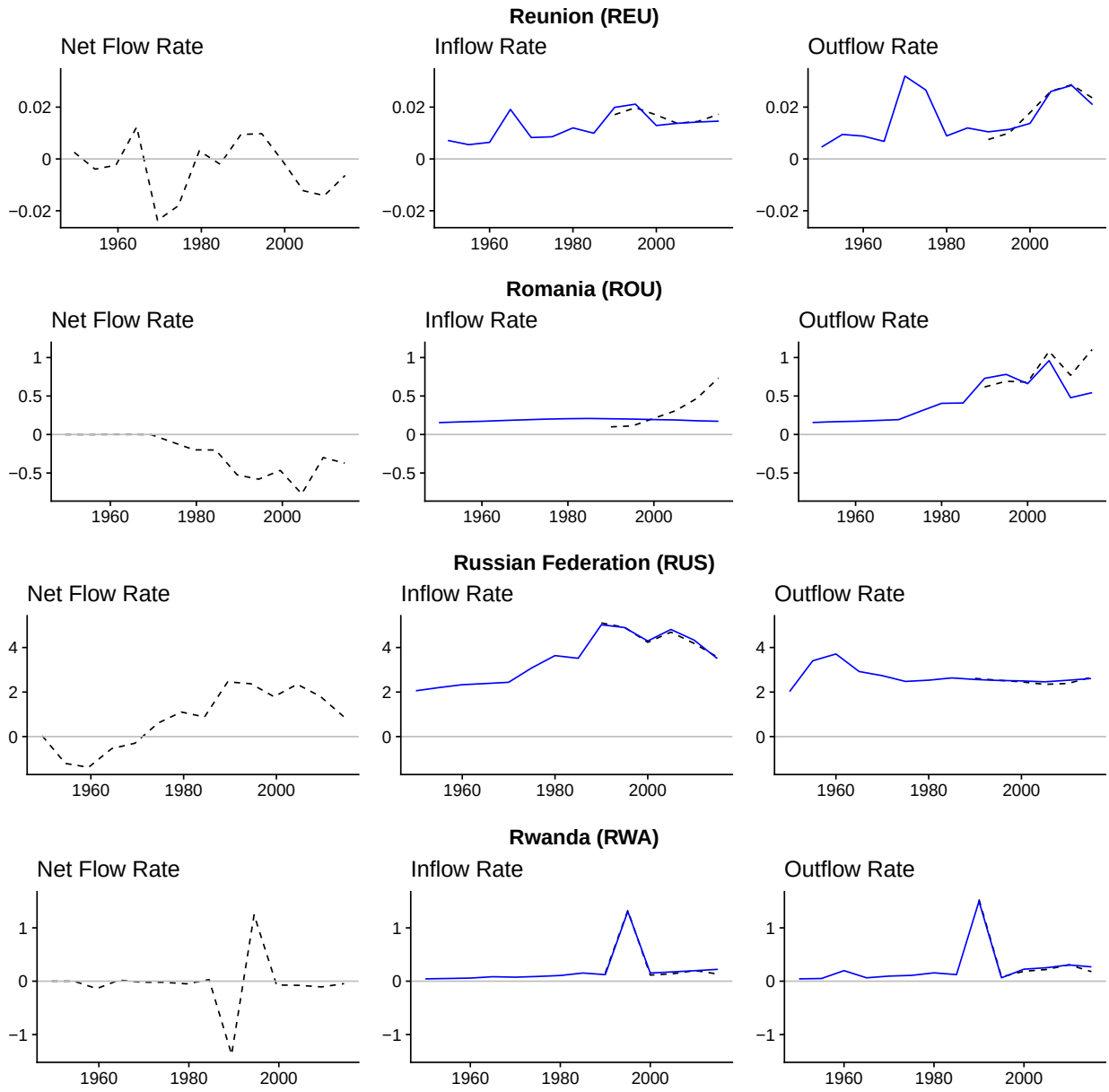


Figure A.43: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

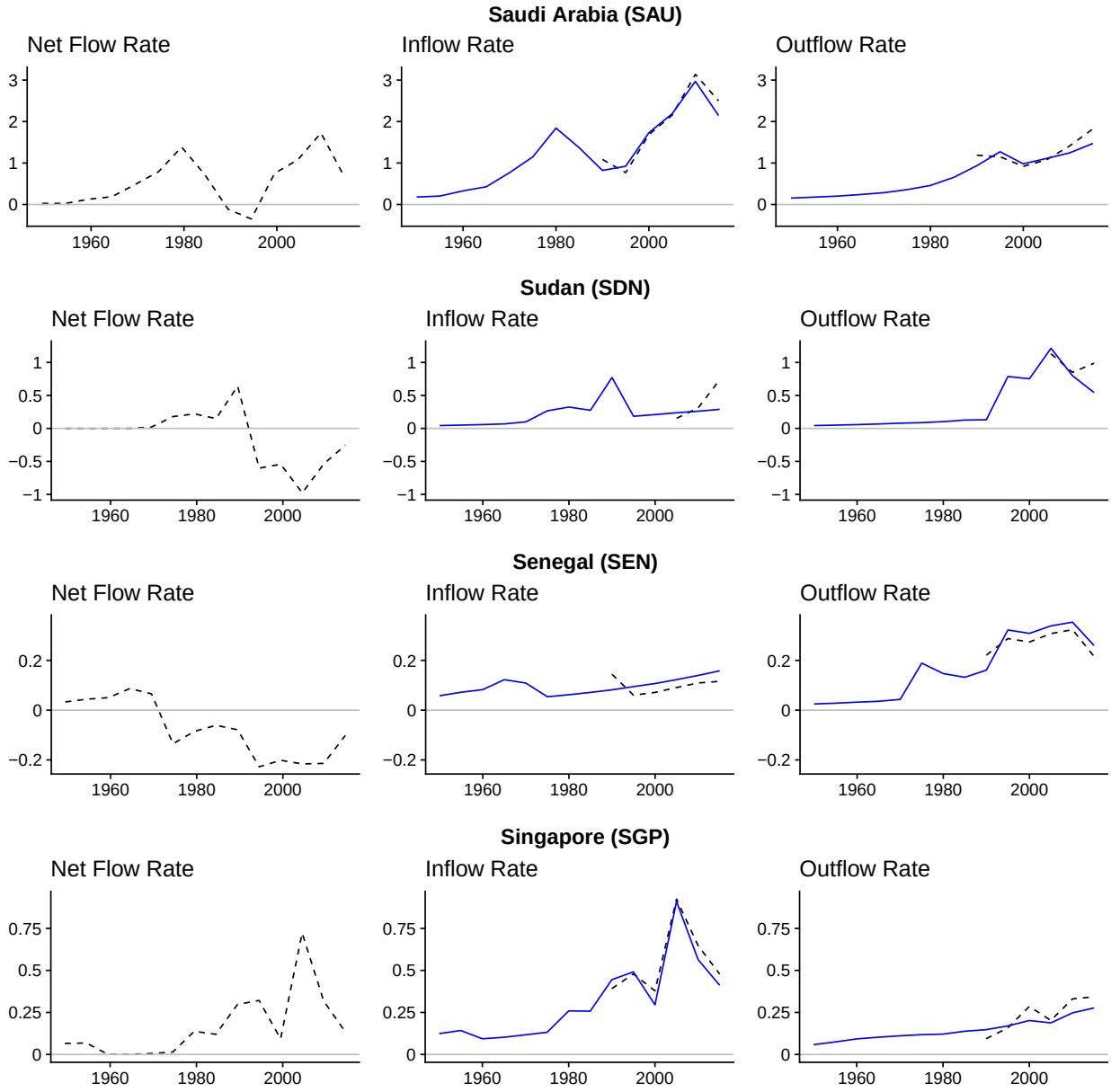


Figure A.44: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

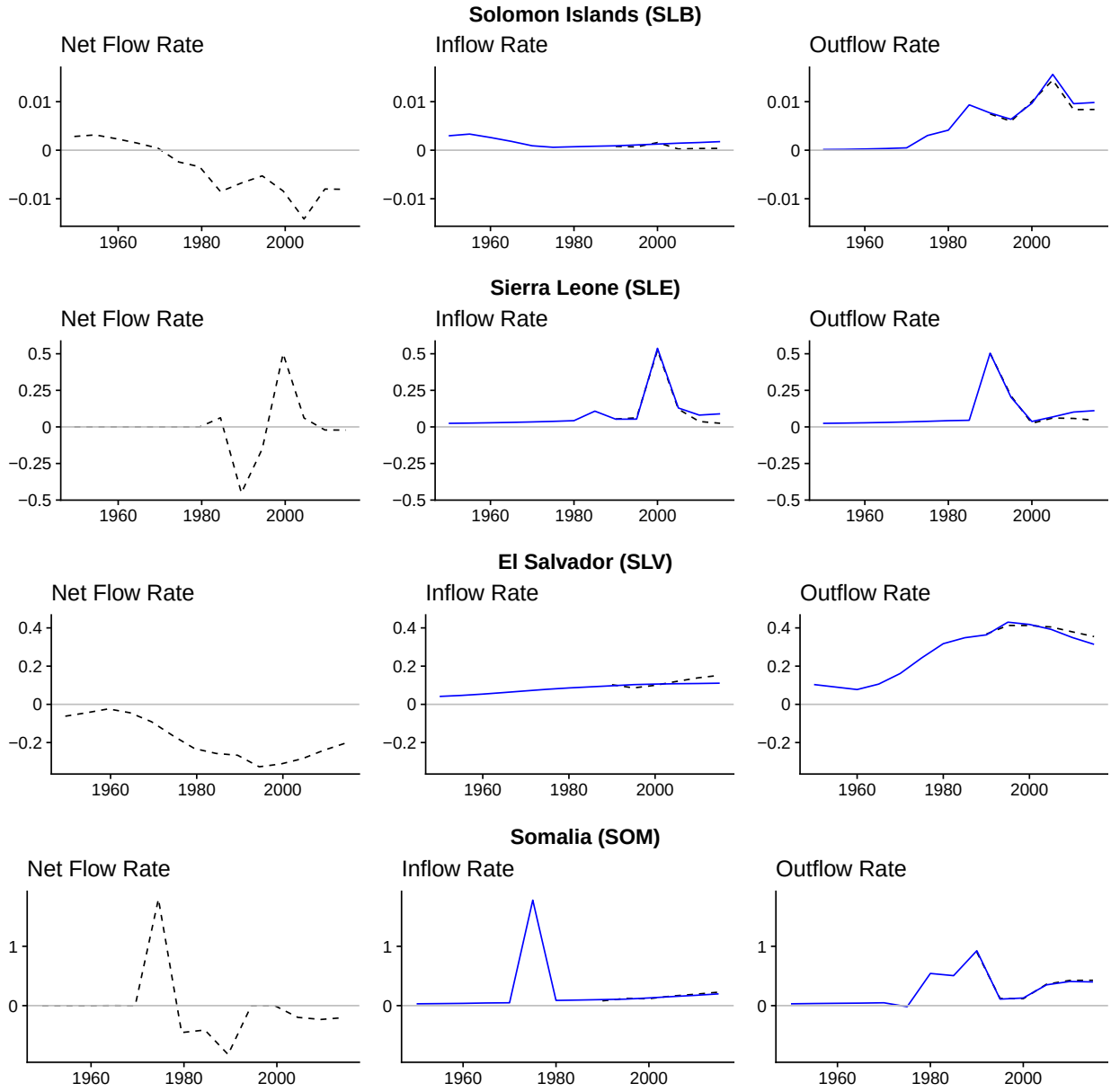


Figure A.45: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

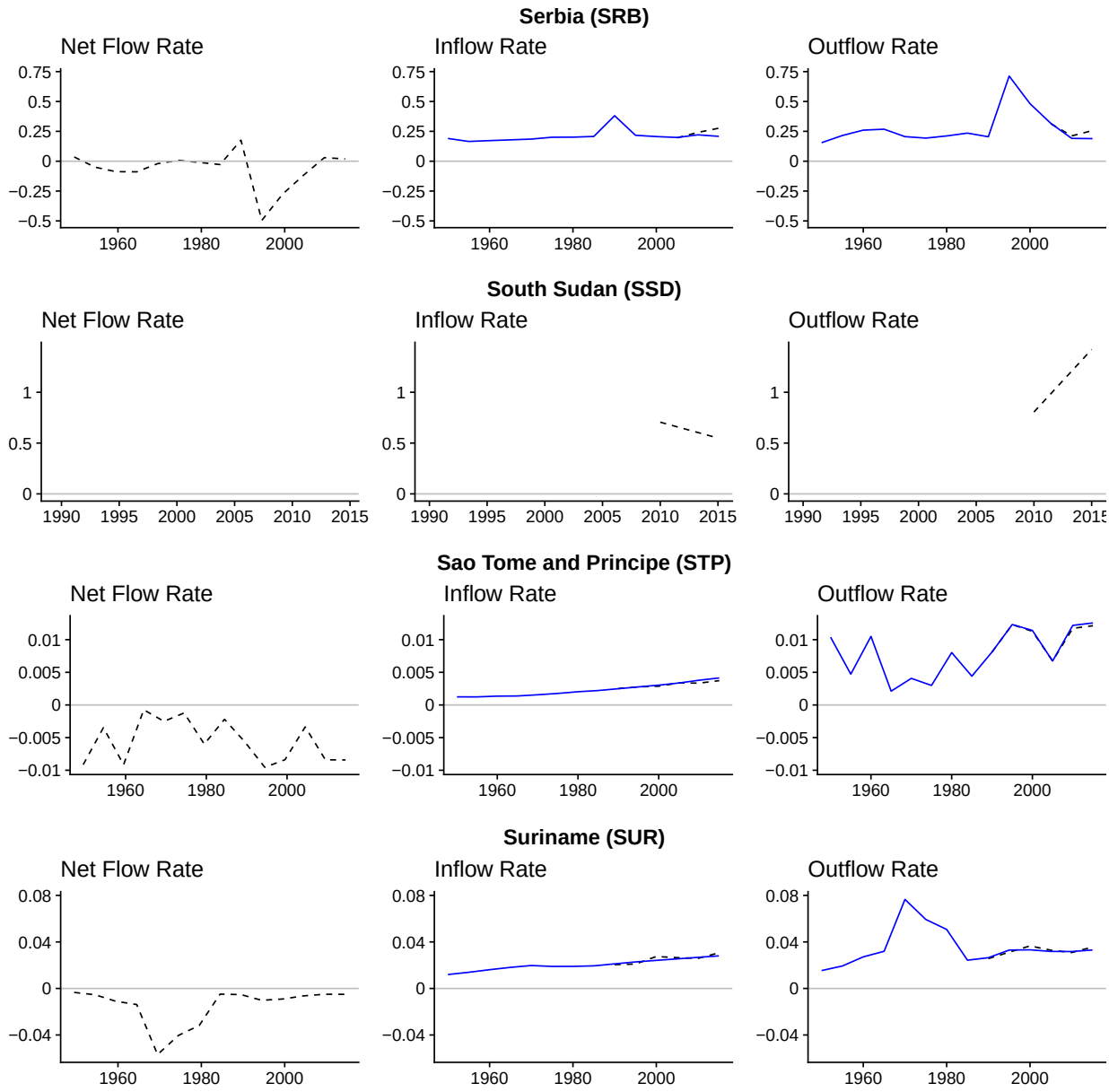


Figure A.46: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

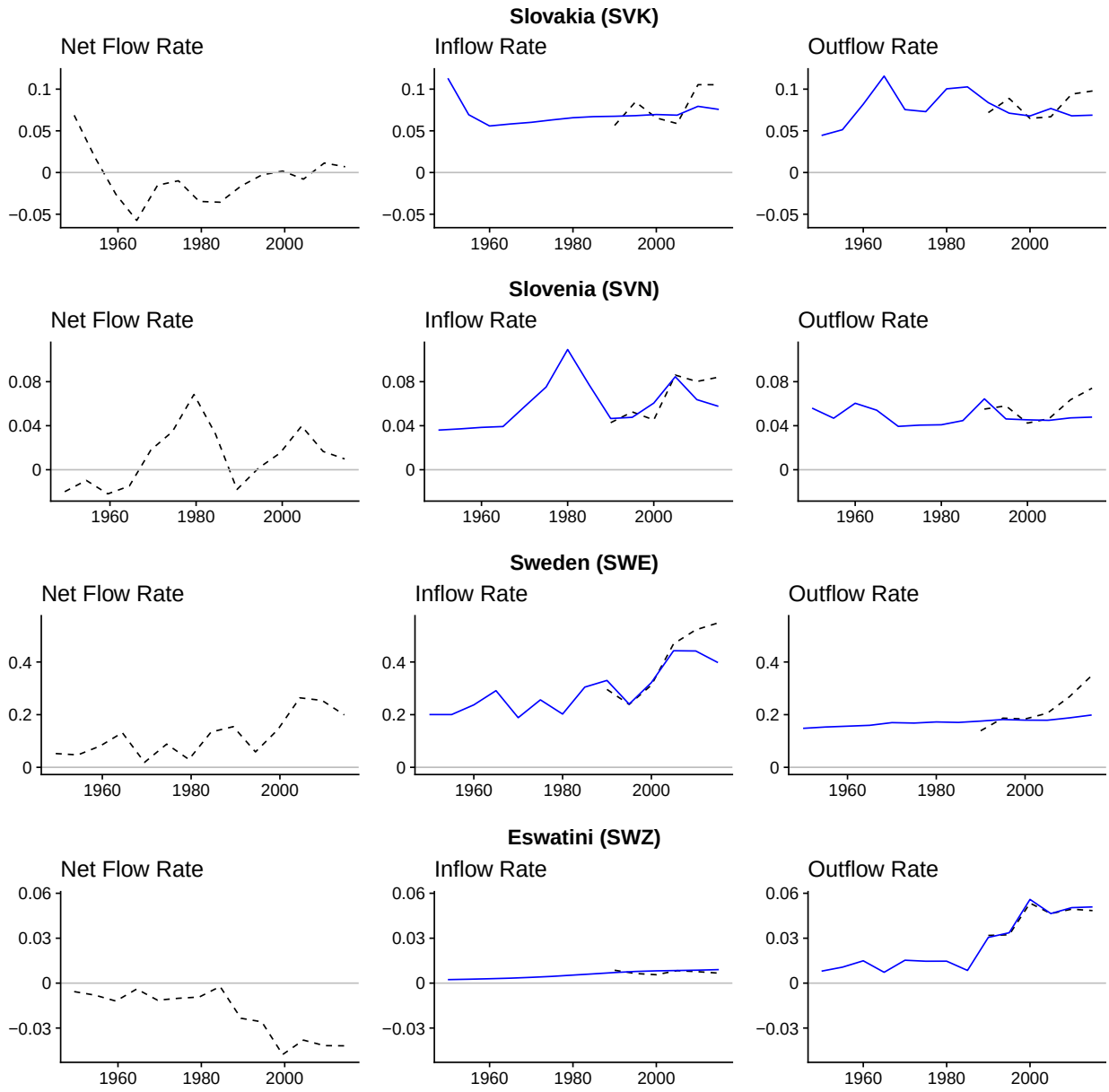


Figure A.47: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

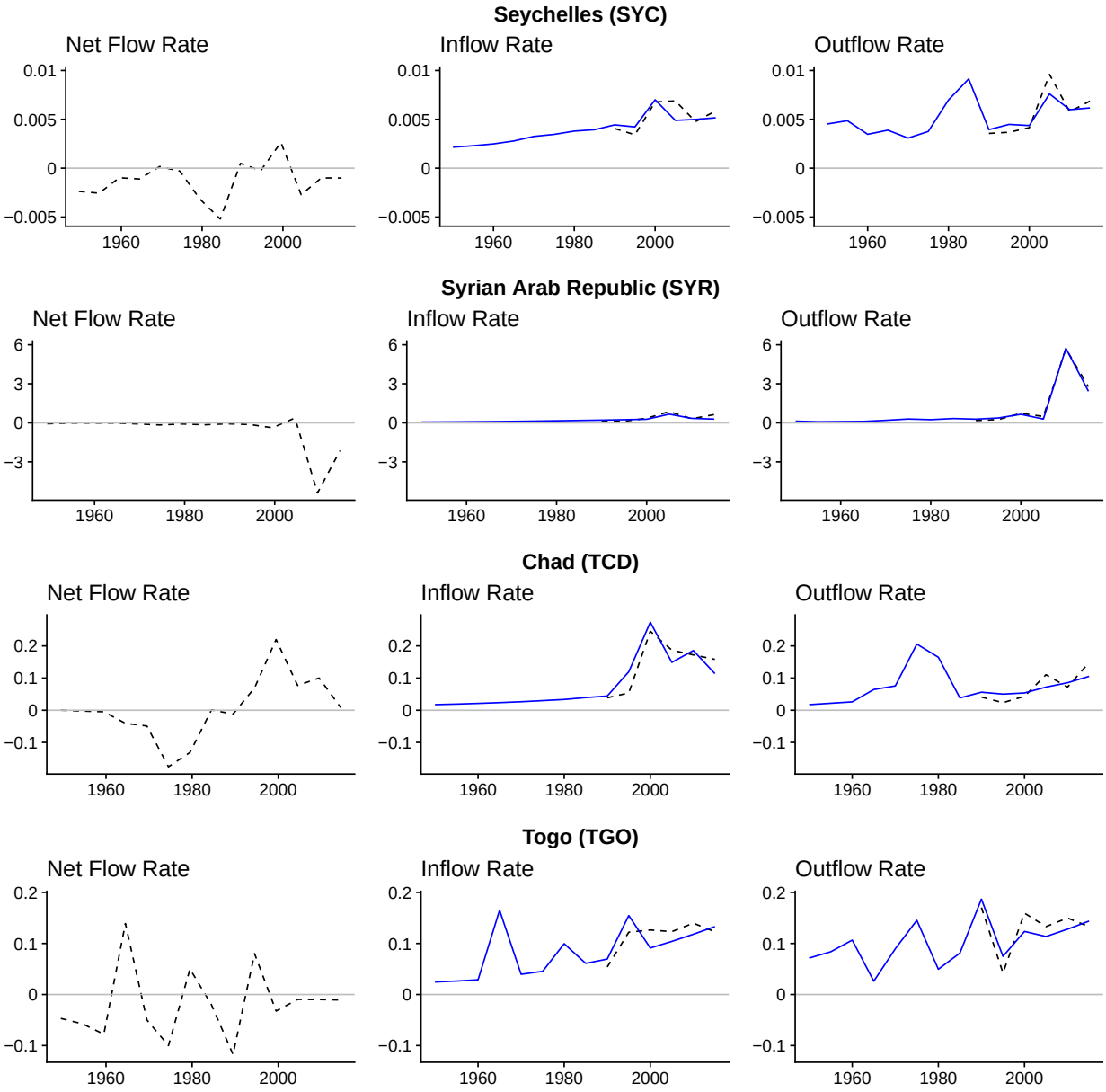


Figure A.48: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

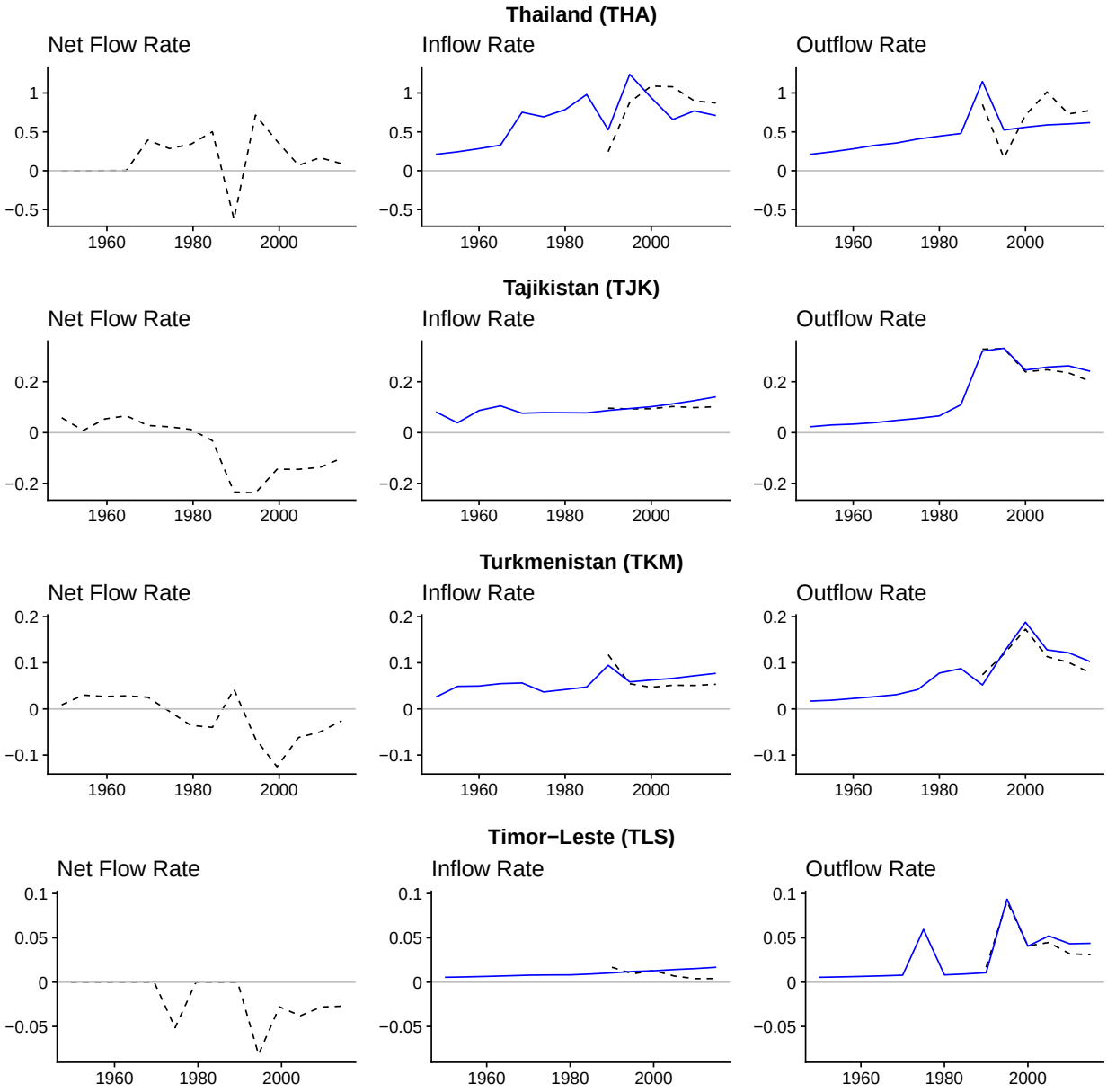


Figure A.49: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

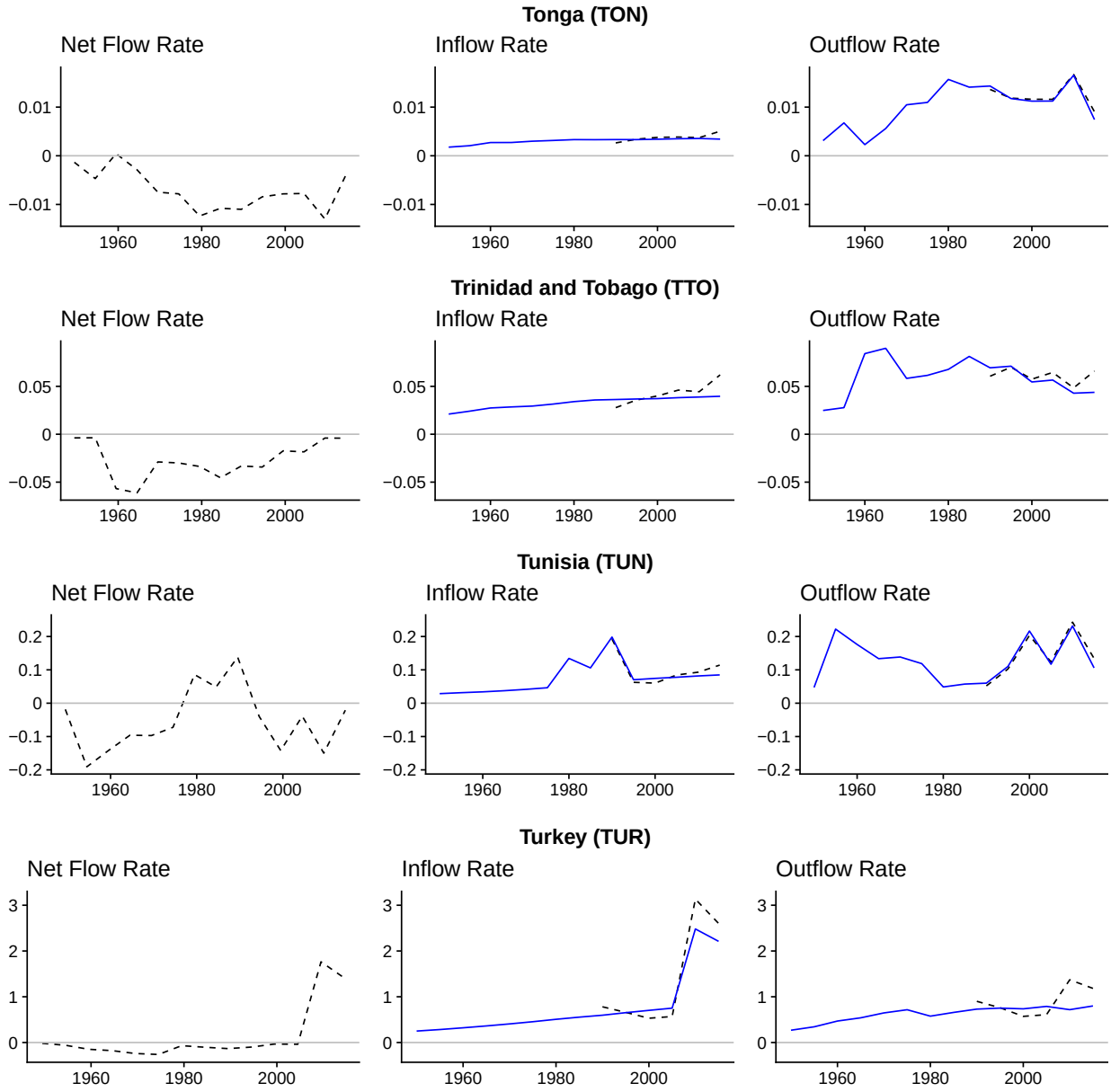


Figure A.50: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

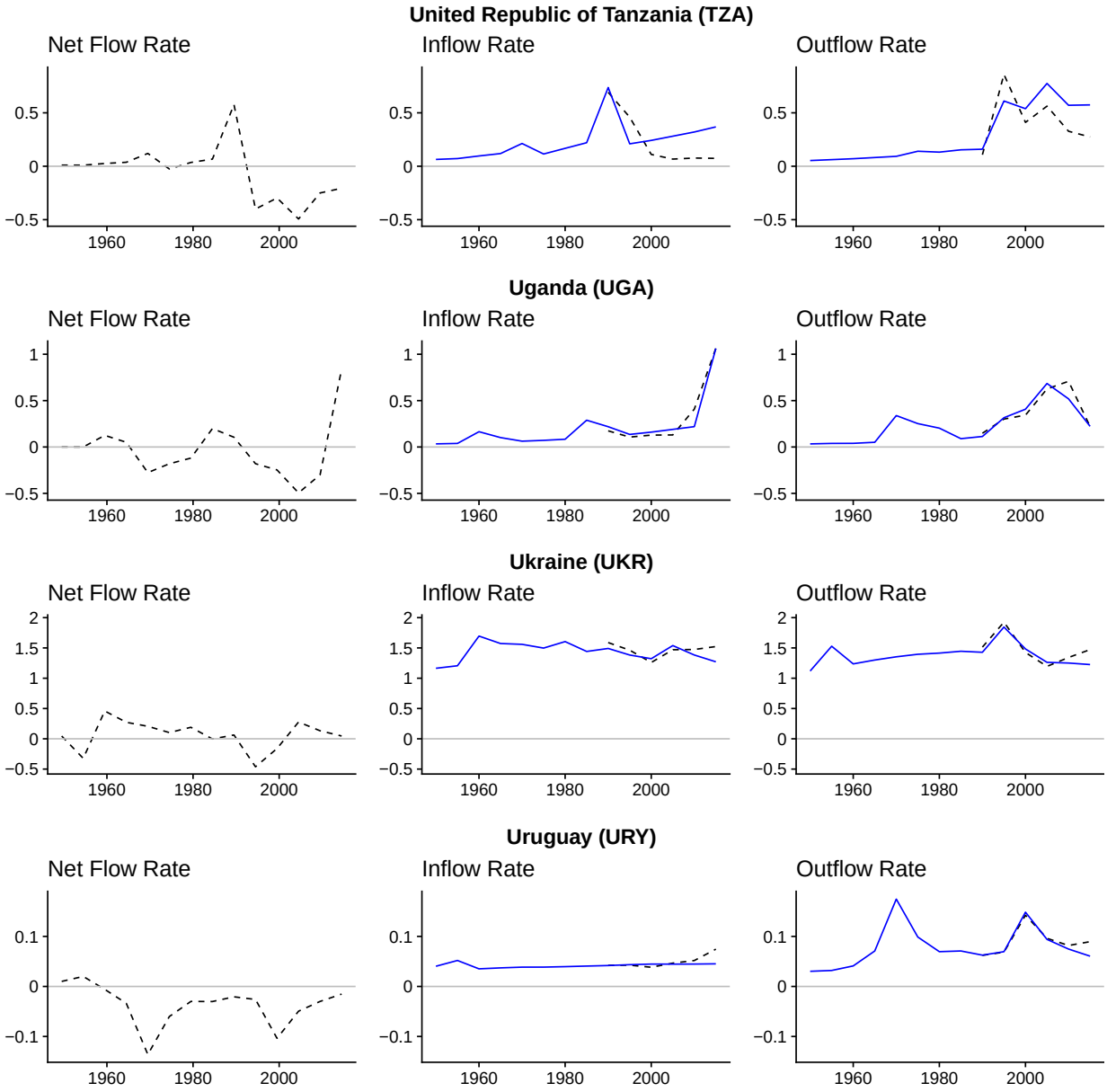


Figure A.51: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

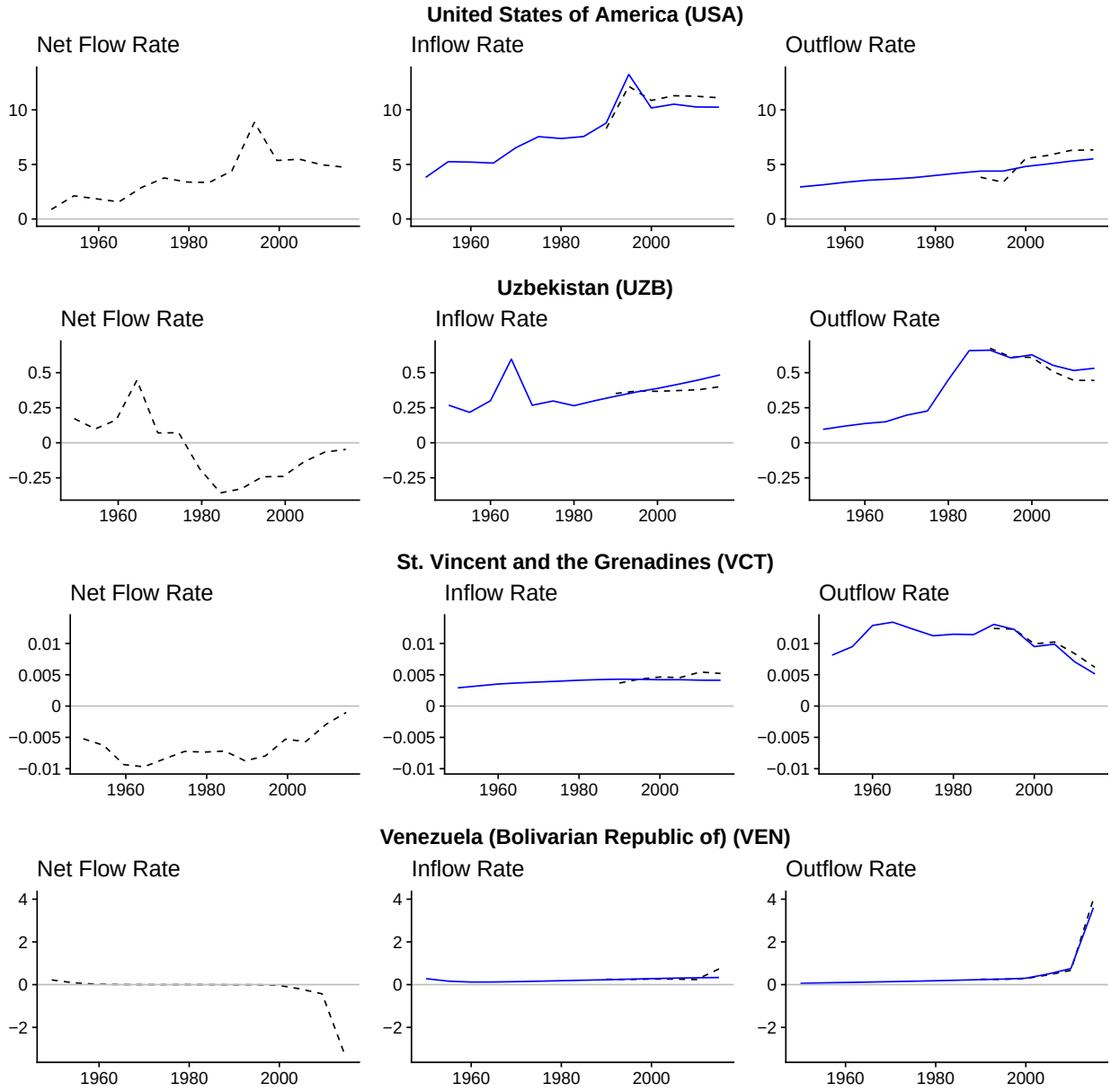


Figure A.52: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

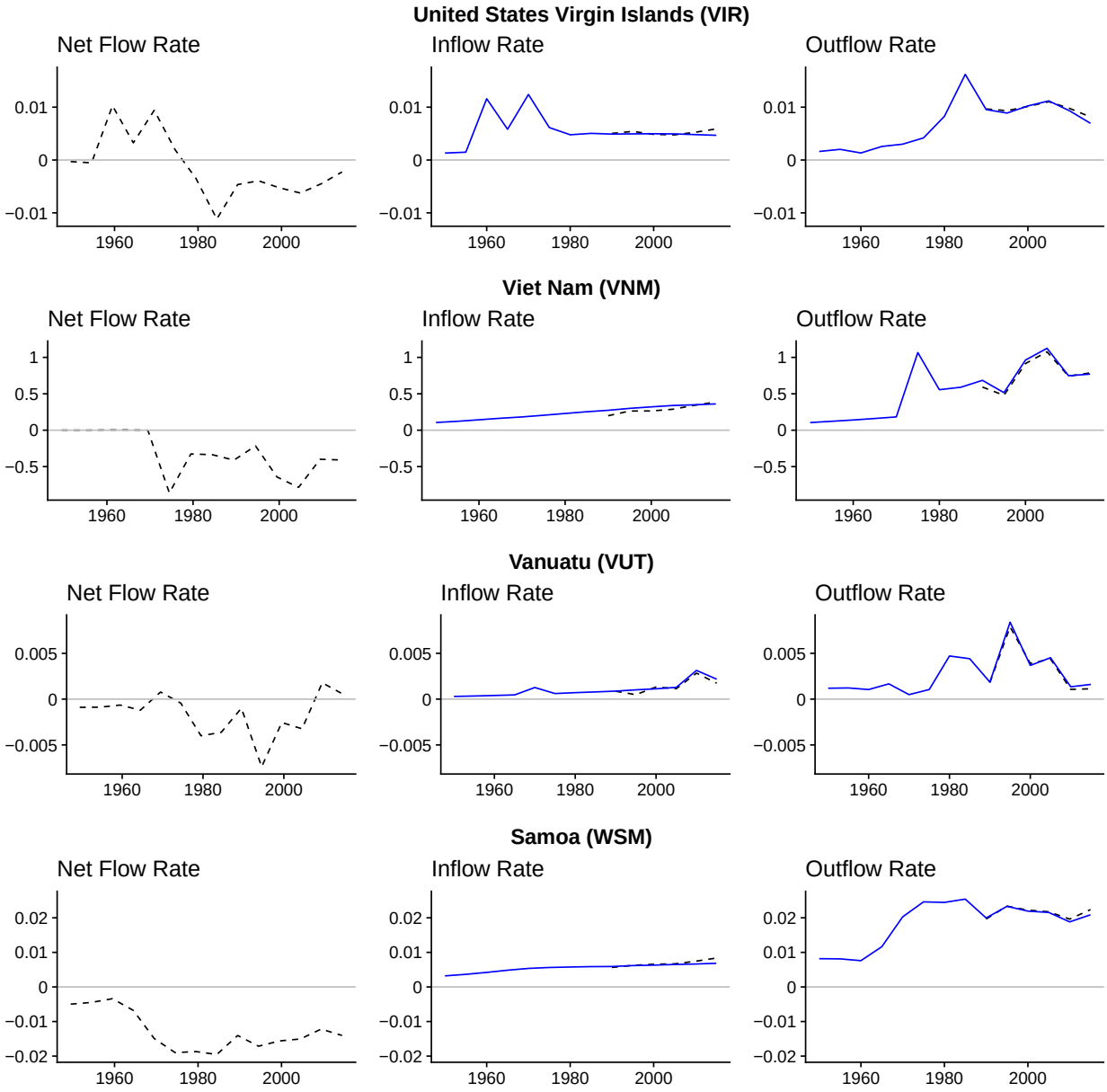


Figure A.53: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

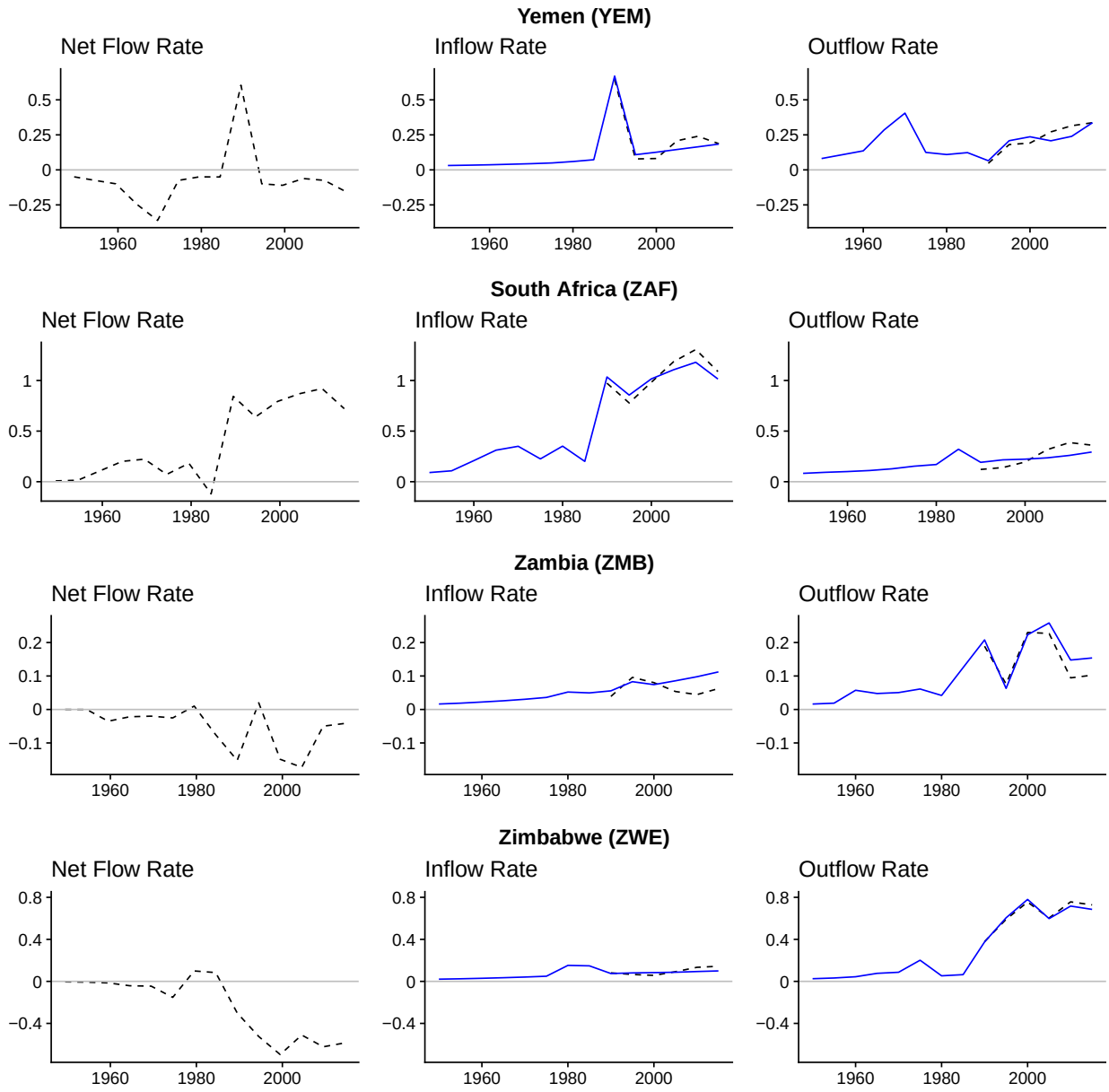


Figure A.54: Observed net migration (left column), decomposed into in-migration (middle column), and out-migration rates (right column), on the scale of annual migrants per thousand people. Solid blue lines show the model-based estimates. Dashed black lines show the observed migration rates used for the estimation. Migration flow estimates are shown at the midpoint of each 5-year period.

A.3 Forecast Summaries by United Nations Area

Figure A.55 shows the world and regional forecasts using the age-agnostic and age-standardized migration forecast models. World forecasts of population are indistinguishable, which makes sense as both methods use the same fertility and mortality models. Differences in the median age-standardized migration forecasts are most pronounced in countries aging faster or more slowly than the global average rate of aging. Furthermore, age-standardized prediction intervals are generally narrower than age-agnostic intervals since variation in past net migration rates attributable to population age structure is removed in the age-standardized forecasts.

Africa's population forecast by the end of the century is lower under the age-standardized model compared to the agnostic migration model. This makes sense because the force of migration associated with population age structure in many African countries is forecast to far exceed the global median for most of the remainder of the century. As a result, long-term forecasts that ignore population age structure for many countries in Africa are slightly overstated compared to the age standardized migration model. End-of-century median population forecasts for all other regions are higher using the age-standardized migration model.

The most pronounced differences are for Europe and North America. Age-standardized prediction intervals for Europe and North America lie within the age-agnostic prediction intervals. The most extreme population declines in Europe and Latin America are less likely under the age-standardized model. Migration forecasts from the age-standardized model are lowered by aging populations in both regions. The most extreme population growth and decline trajectories in North America are less likely under the age-standardized model.

Table A.3 details the differences in the population forecasts using the age-standardized and age-agnostic migration forecasting methods shown in Figure A.55.

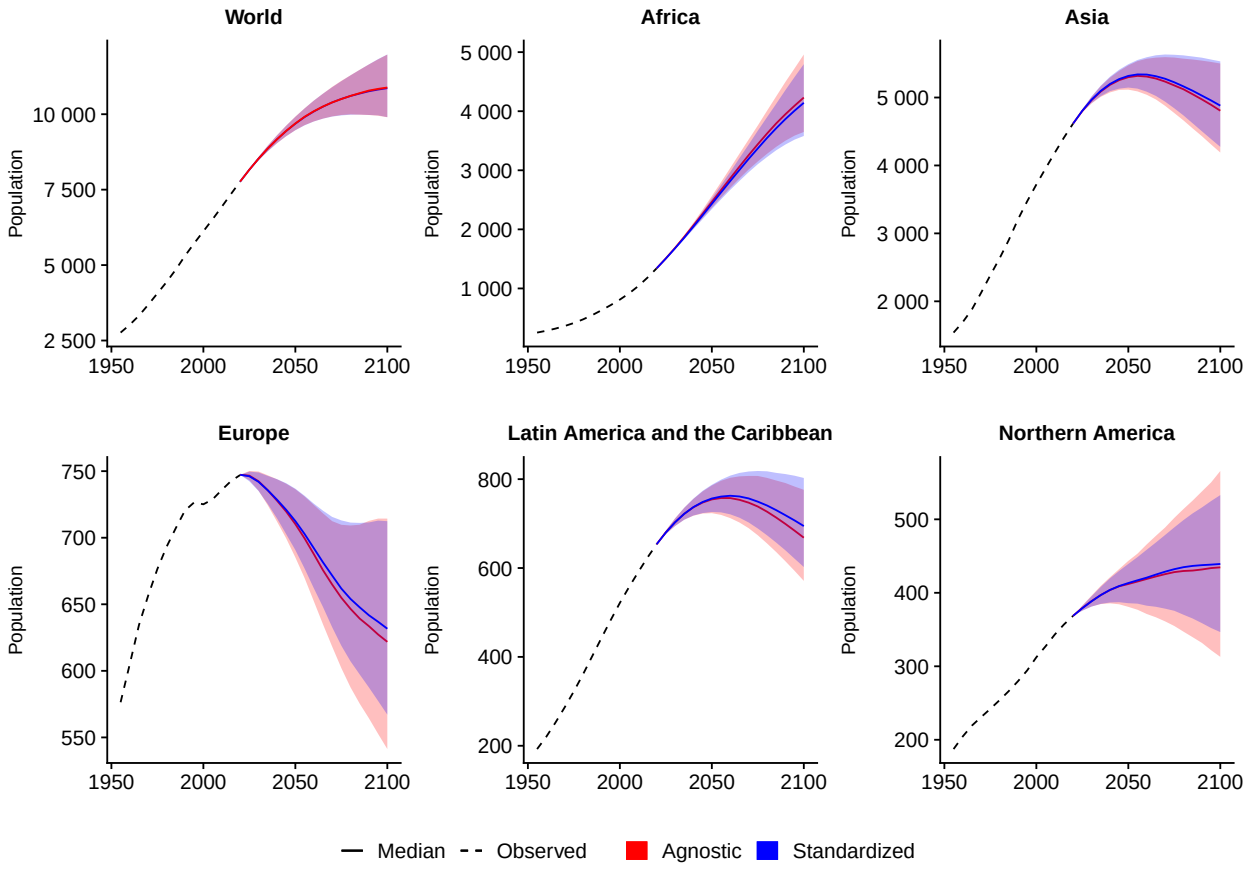


Figure A.55: Probabilistic forecast of population (millions of people) by region. Forecasts use probabilistic net migration (■=age-standardized and ■=age-agnostic), fertility, and mortality. Solid lines indicate the median forecast. Dashed lines indicate the observed population. Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

Table A.3: Median population and net migration forecasts at the end of the century by United Nations Area and the world using age-standardized (S) and age-agnostic (A) migration forecasting in millions of people along with the net difference (S-A) and percent difference $100\times\frac{S-A}{A}$ %.

UN Area	Measure	S	A	S - A	$100\times\frac{S-A}{A}$ %
Africa	population	4142.6	4231.3	-88.7	-2.0
	net migration	-10.5	-6.3	-4.3	68.0
Asia	population	4880.3	4806.0	74.4	2.0
	net migration	0.1	-1.2	1.3	-109.0
Europe	population	631.6	621.8	9.8	2.0
	net migration	3.4	2.8	0.7	24.0
Latin America & the Caribbean	population	694.4	668.3	26.2	4.0
	net migration	0.7	-1.4	2.1	-153.0
Northern America	population	439.3	434.6	4.6	1.0
	net migration	4.4	3.9	0.5	12.0
Oceania	population	65.1	62.6	2.5	4.0
	net migration	0.4	0.2	0.2	132.0
World	population	10863.3	10884.4	-21.1	0.0

A.4 Forecast Quantile Data

File *age_standardized_forecast_data.csv* includes net migration and population forecast quantiles for five-year-periods starting in 2020-2025 through 2095-2100. Table A.4 summarizes each column of the data set.

Table A.4: Age-standardized migration forecast quantile summary data dictionary.

country_code	(character) 3-digit ISO 3166 numeric country code
iso	(character) 3-letter ISO 3166 country abbreviation
name	ISO 3166 country name
year0	(integer) First year of the five-year forecast period
pop0_10	(numeric) 10% quantile age-standardized population forecast at the start of the period in thousands
pop0_50	(numeric) 50% quantile age-standardized population forecast at the start of the period in thousands
pop0_90	(numeric) 90% quantile age-standardized population forecast at the start of the period in thousands
net_10	(numeric) 10% quantile age-standardized net migration forecast for the period starting in year0 in thousands
net_50	(numeric) 10% quantile age-standardized net migration forecast for the period starting in year0 in thousands
net_90	(numeric) 10% quantile age-standardized net migration forecast for the period starting in year0 in thousands

A.5 Age-Standardized Migration Implementation

Computer codes that implement the age-standardized migration method are available on GitHub: `https://github.com/ngwelch/migAge`. Modifications to the *bayesPop* (Ševčíková & Raftery, 2016b) package can be found in the *bayesPopAge* directory of the repository. Codes in the *src* directory execute all methods described in the manuscript using a combination of bash and R (R Core Team, 2021) scripts. Age-standardized migration forecasts were generated with restricted access fertility and mortality files to enhance comparisons to the published WPP 2019 (United Nations. Department of Economic and Social Affairs, 2019) population projections. To generate new forecasts, users must specify (or remove to use default settings) references to the following bayesPopAge *pop.predict* function input setting *tfr.sim.dir, e0F.sim.dir, e0M.sim.dir, migM, migF, patterns*. All other data required to execute the age-specific migration forecasting method are publically available within the WPP 2019 United Nations. Department of Economic and Social Affairs (2019) and from Abel & Cohen (2019).

A.6 Country-level Forecasts Through 2100

Forecast summaries through 2100 for each country are shown on the following pages. Forecasts for some countries were much different from the age-agnostic and/or WPP 2019 forecast. The discussion below highlights the most extreme forecast differences and offers explanations for the sources of the differences where necessary.

Using the Rogers & Castro-like migration age schedule rather than a net migration age schedule approach reduced uncertainty in the population age structure for many countries, including:

- Cape Verde
- Western Sahara
- Equatorial Guinea
- Reunion
- All GCC countries (Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, United Arab Emirates).

Furthermore, age-standardized method leads to less net out-migration from Western Sahara than the age-agnostic method, bringing the age-standardized forecast into line with the WPP 2019 forecast and reducing the wide uncertainty in the age-agnostic forecast.

The age-standardized migration method led to much narrower prediction intervals for Luxembourg compared to the age-agnostic method. Median forecasts for both the age-agnostic and age-standardized methods are similar and slightly higher than the WPP 2019 forecast.

The Maldives 2020 population age structure bears striking similarities to the migrant worker dominated populations of the GCC countries. Both the age-agnostic and age-standardized methods imply that the large population of working-age males will remain in the Maldives

and age in place, which may not be appropriate. However, the WPP 2019 forecast reverses the large influx of migrant workers in the next period. All three methods lead to a similar median population age distributions by 2100.

Long-term migration trends in Puerto Rico were disrupted in September 2017 by two Category 5 hurricanes that hit the island as many weeks. The scale of damages and slow recovery from Hurricanes Irma and Maria help explain the massive departure in 2015-2020 in Puerto Rico's long-term net migration trend (Acosta et al., 2020; Currie, 2022). As US citizens, Puerto Rican migration from the US Territory to other parts of the United States presents a much lower barrier compared to international migration. The WPP 2019 forecast suggests large net out-migration in 2015-2020 will reverse to net in-migration in 2020-2025. Both the age-standardized and age-agnostic models indicate a much slower return to the long-term net-migration trend.

Age-standardized migration forecasts for GCC countries indicates sustained net positive migration through 2100 while the age-agnostic forecast has many more negative net migration trajectories. The age-standardized forecast leads to sustained population growth in the United Arab Emirates (ARE) and Qatar (QAT) compared to the WPP 2019 and age-agnostic methods. Age-standardized forecasts for Bahrain (BHR), Kuwait (KWT), Oman (OMN), and Saudi Arabia (SAU) are also higher than the other methods, but resulting population differences by 2100 are less extreme in these cases. Population age structure in GCC countries is universally more plausible using the age-standardized approach compared to the age-agnostic method. The age-standardized population age structure forecast is also more similar to the WPP 2019 forecasts in GCC countries. Sustained net positive migration in GCC countries using the age-standardized method is consistent with the historic net positive migration trend in those countries. Age-standardized migration forecasts continue the more recent trend in population growth in GCC countries, but sustained population growth in the United Arab Emirates and Qatar through 2100 suggests additional refinements to the age-standardized forecasting method may be justified in these cases.

The migration shock in Ukraine resulting from Russia's 2022 invasion in Ukraine fell after

the last period of observed demographic data. All forecasts for Ukraine fail to account for the unprecedented net out-migration due to Russia's invasion. As of December 31, 2023, the United Nations High Commission for Refugees estimated that approximately 6,357,100 Ukrainians sought asylum due to the war (United Nations High Commissioner for Refugees, n.d.). This estimate alone would land Ukraine's 2020-2025 net out migration rate around -28.6 migrants per 1000 people annually–well below the 10% quantile forecast of -1.5. Secondary impacts such as increased mortality and lower fertility will also substantially alter Ukraine's population forecast. While these impacts can be integrated in future forecasts, it is too early to account for the full scale of the war with the data currently available. Furthermore, it is impossible to anticipate the rate at which asylum seekers who fled the war zone will return to Ukraine after the war concludes. The forecast for Ukraine will undoubtably change drastically after reliable data from the 2020-2025 period become available.

Finally, none of these forecasts account for the effects of the global COVID-19 pandemic. The World Health Organization classified COVID-19 as a global pandemic on March 11, 2020 (Shiraef et al., 2021). The pandemic subsequently disrupted economic and demographic norms around the world. Most international borders closed to all movement–including migration–for most of 2020 (Shiraef et al., 2021). While migration was not possible during this time, the usual forces that motivate people to move across international boundaries surely did not evaporate completely (Lerpold et al., 2023). The full effect of the pandemic will emerge once 2020-2025 (and perhaps 2025-2030) data are integrated into long-term forecasts. When the full impact of the 2020 pandemic emerges in the data, then our methods will integrate those effects in new forecasts.

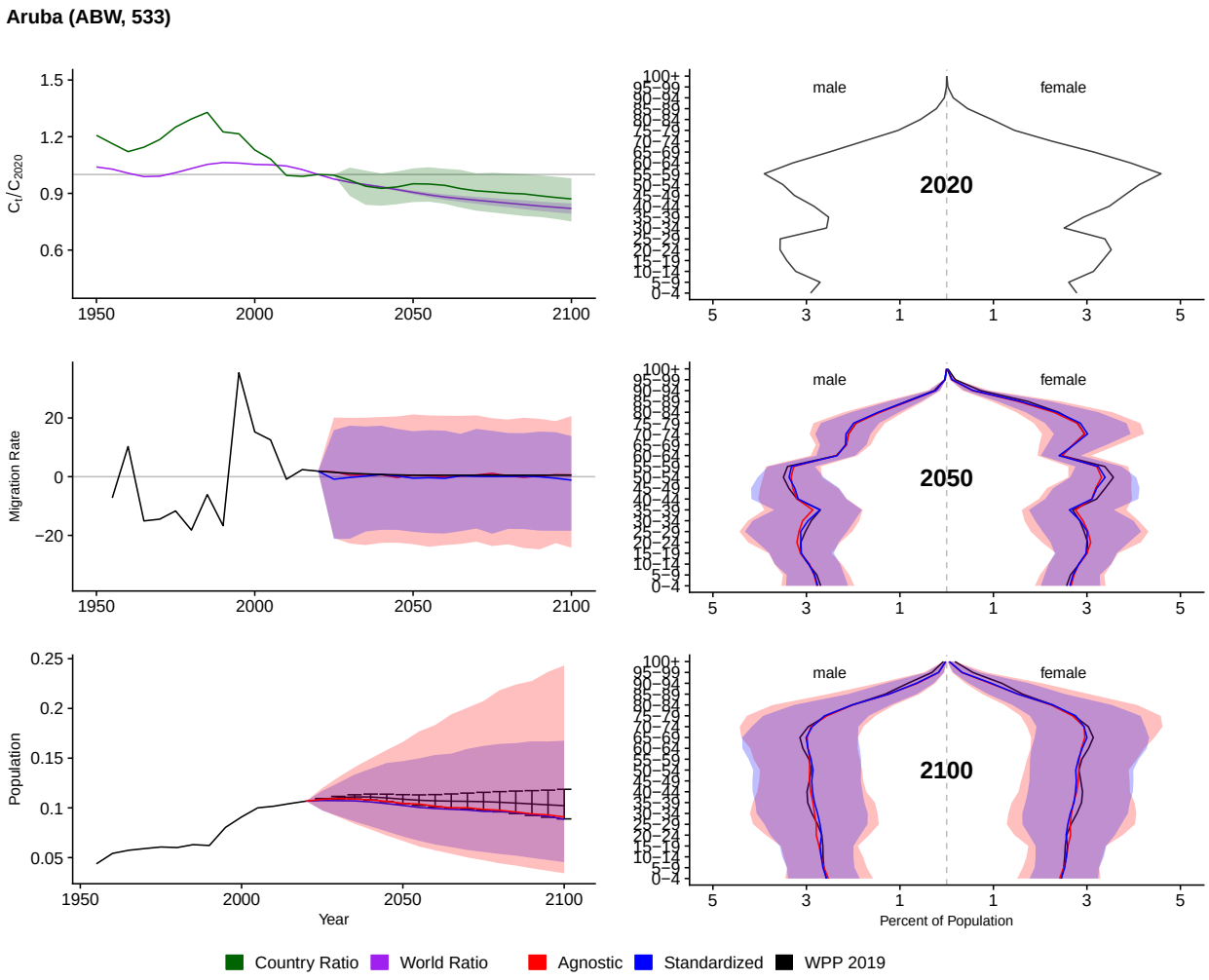


Figure A.56: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

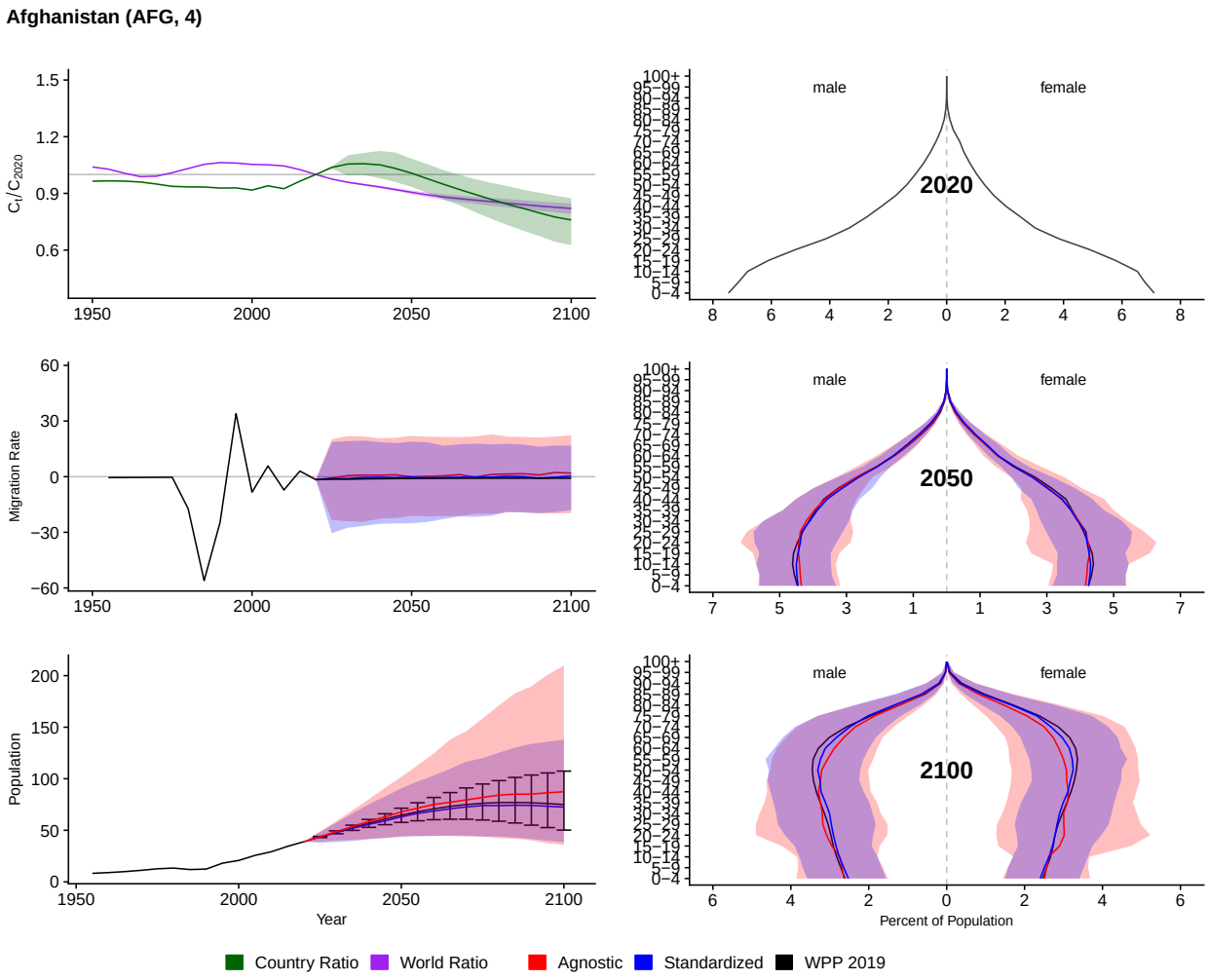


Figure A.57: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

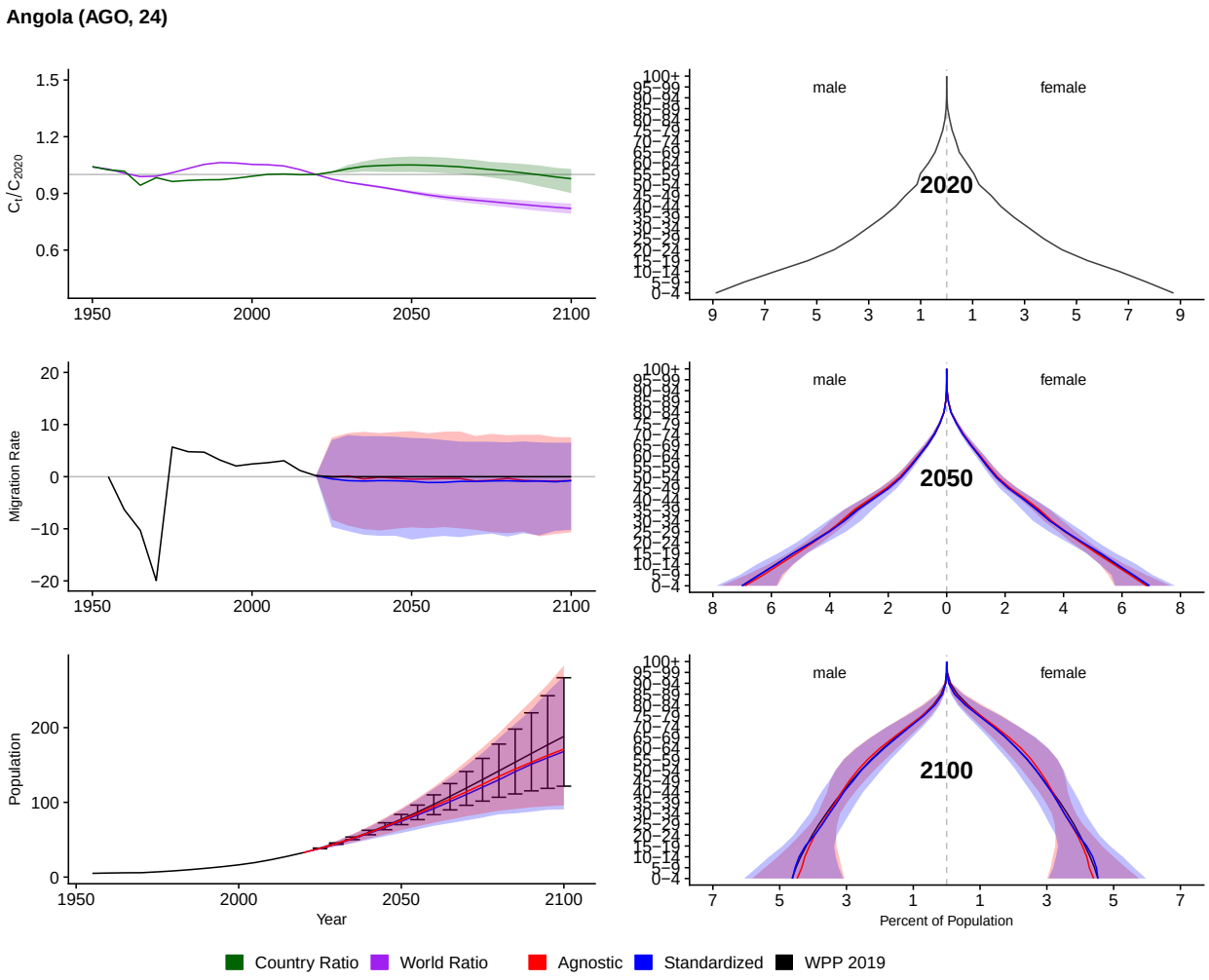


Figure A.58: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

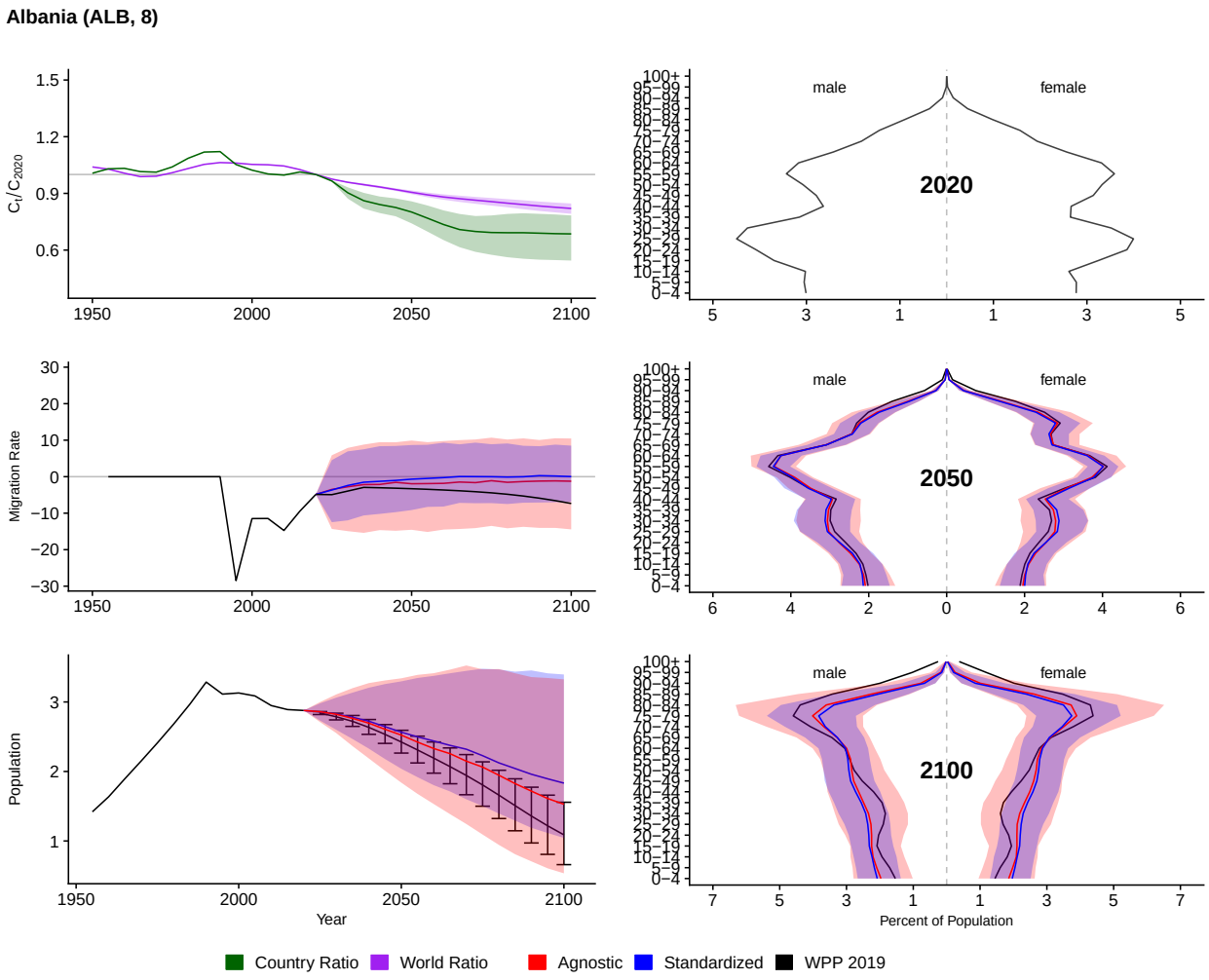


Figure A.59: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

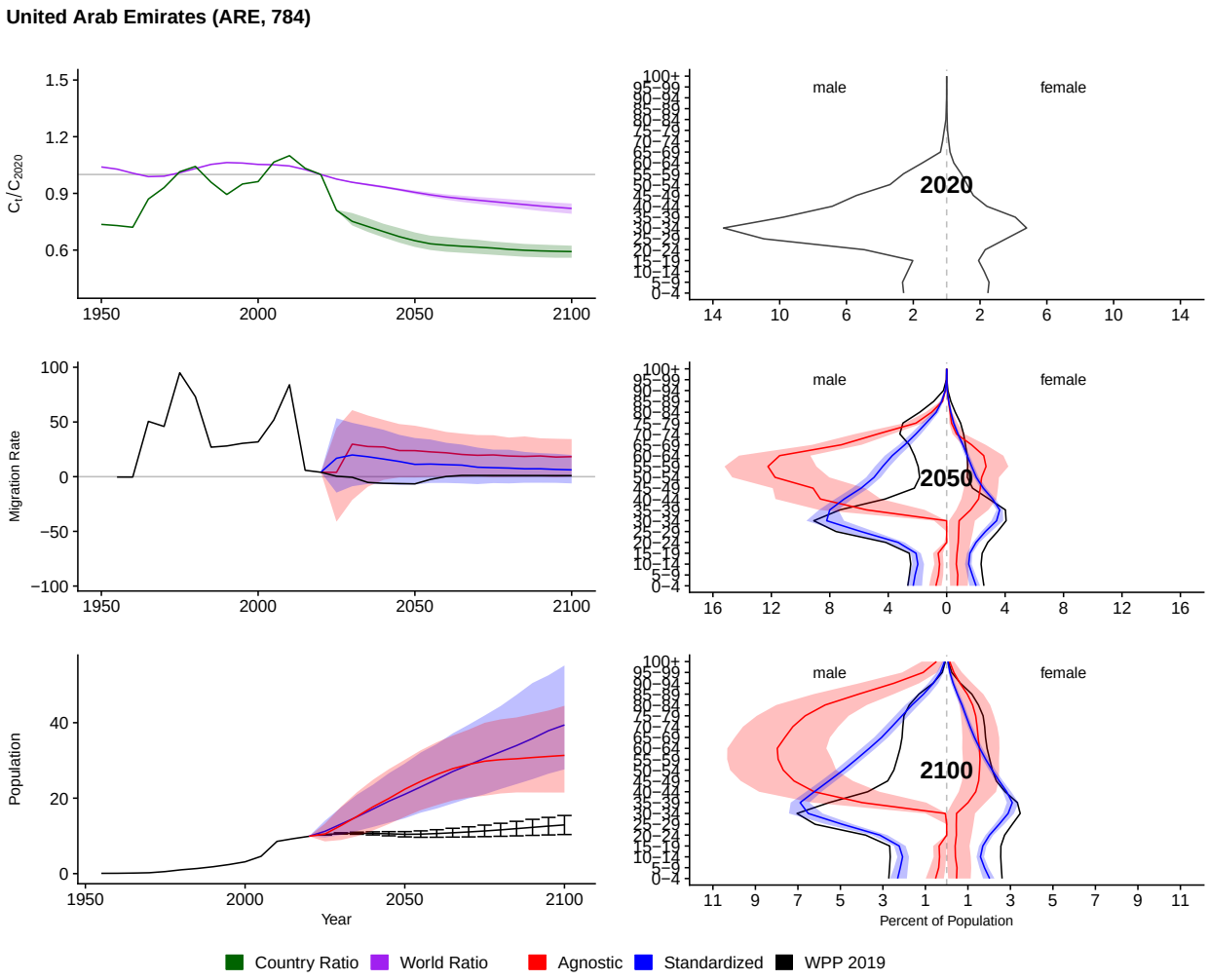


Figure A.60: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

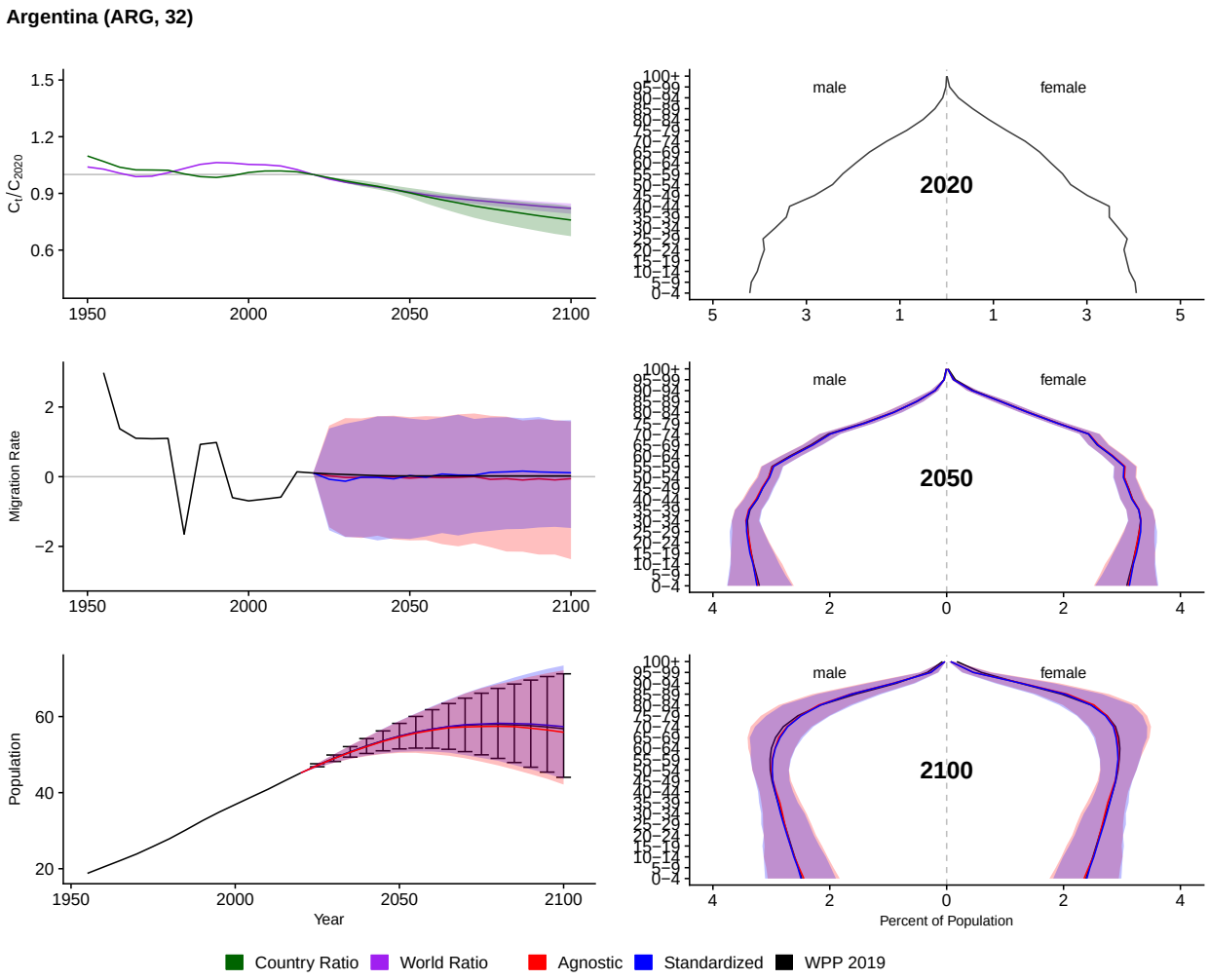


Figure A.61: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

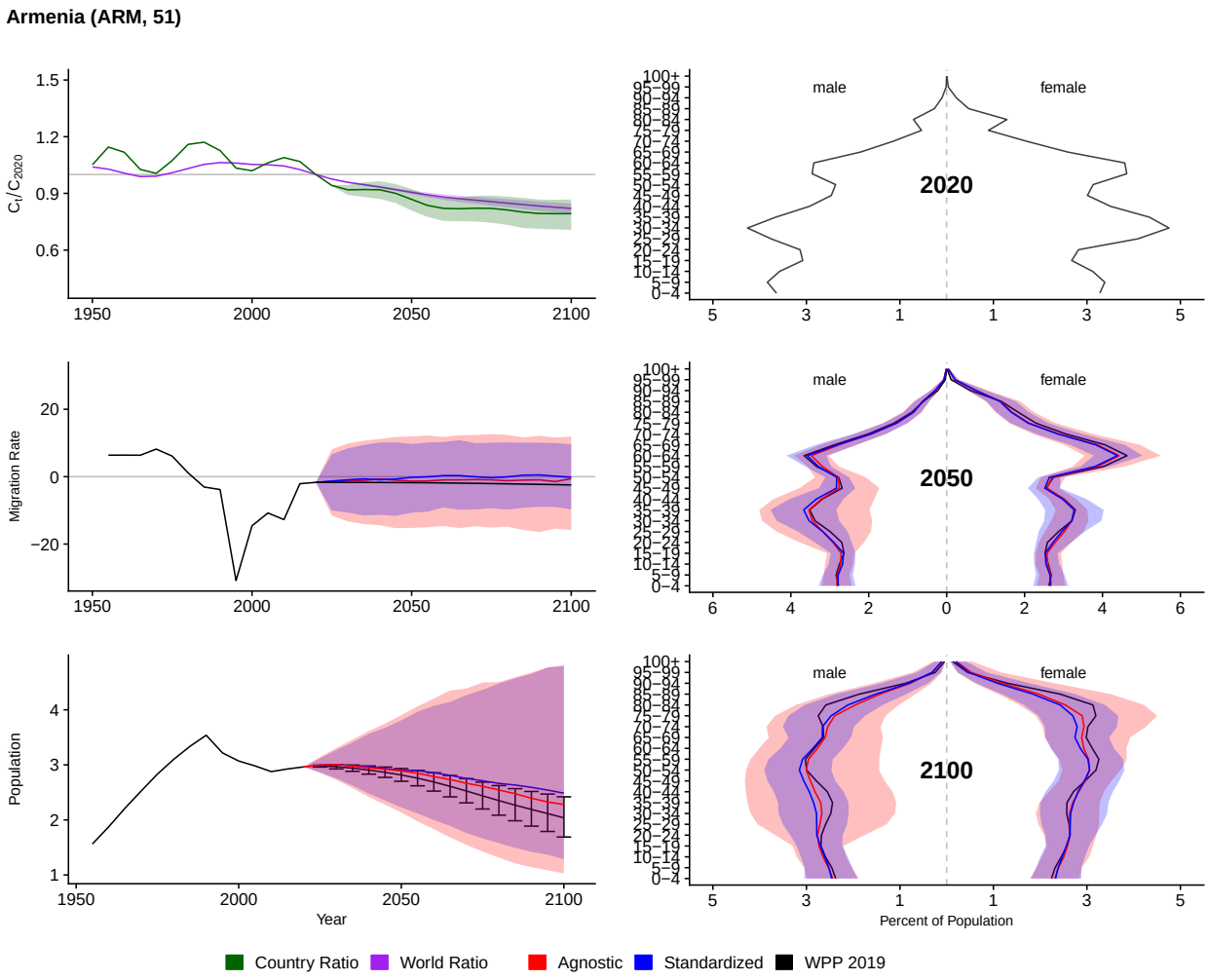


Figure A.62: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

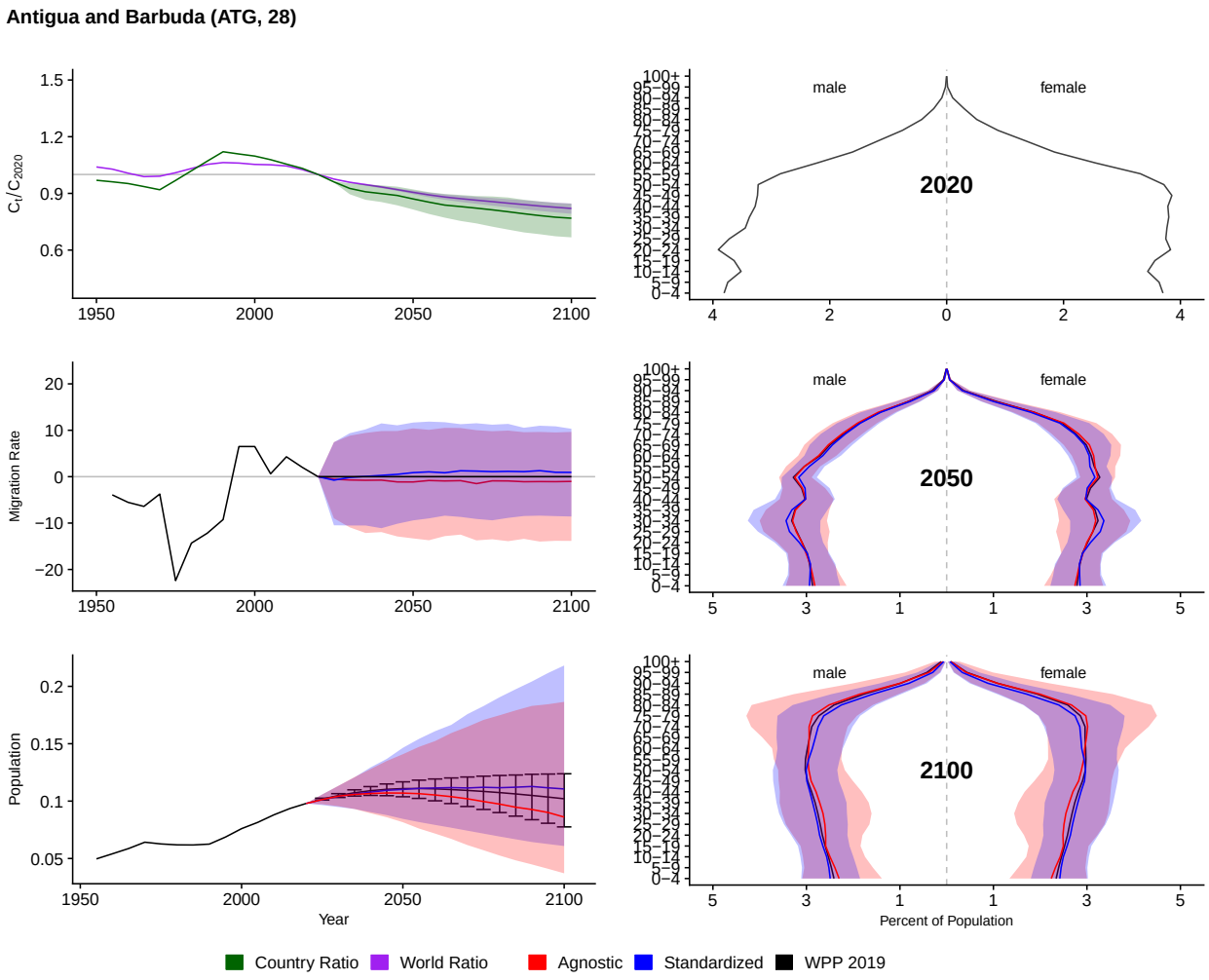


Figure A.63: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

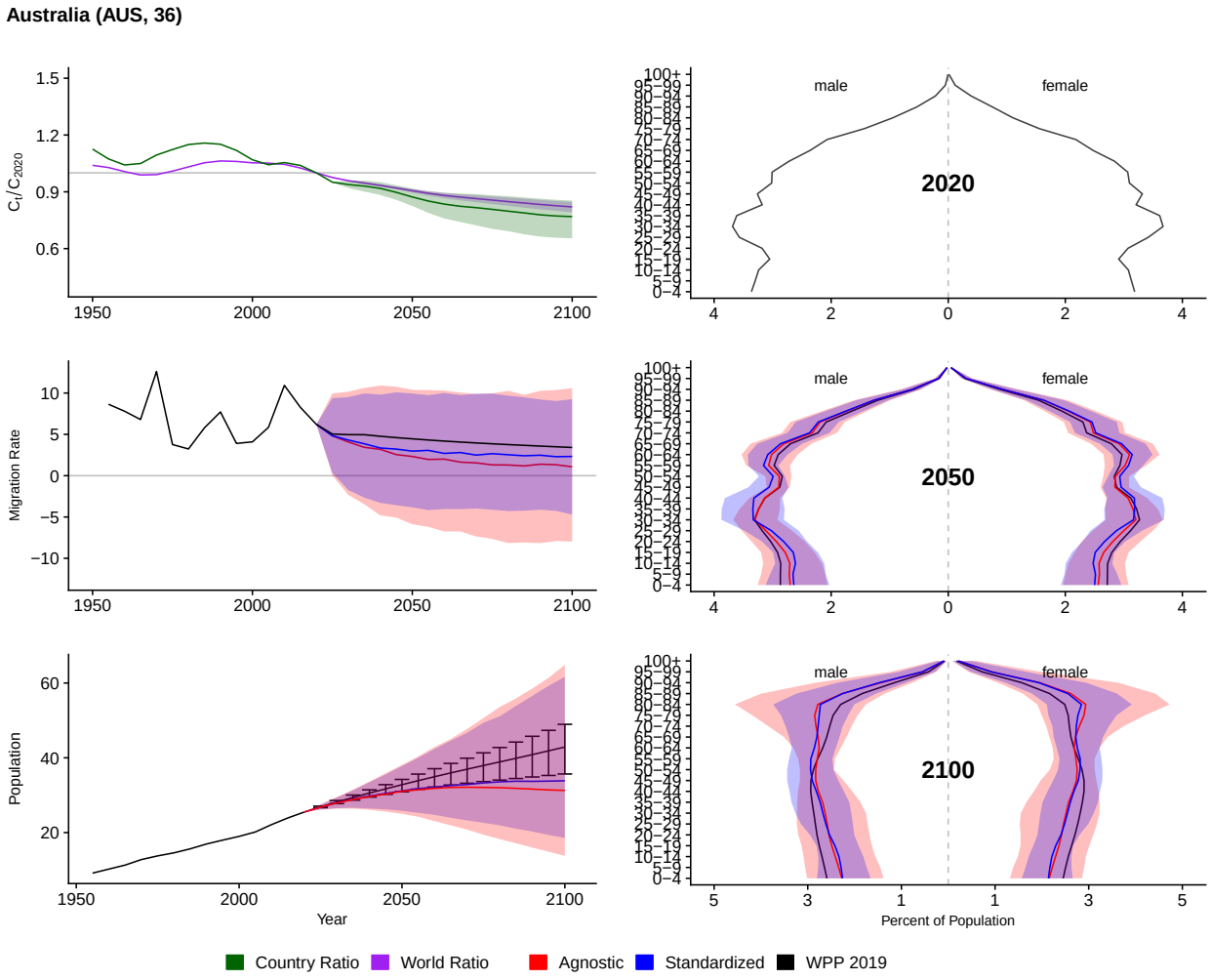


Figure A.64: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

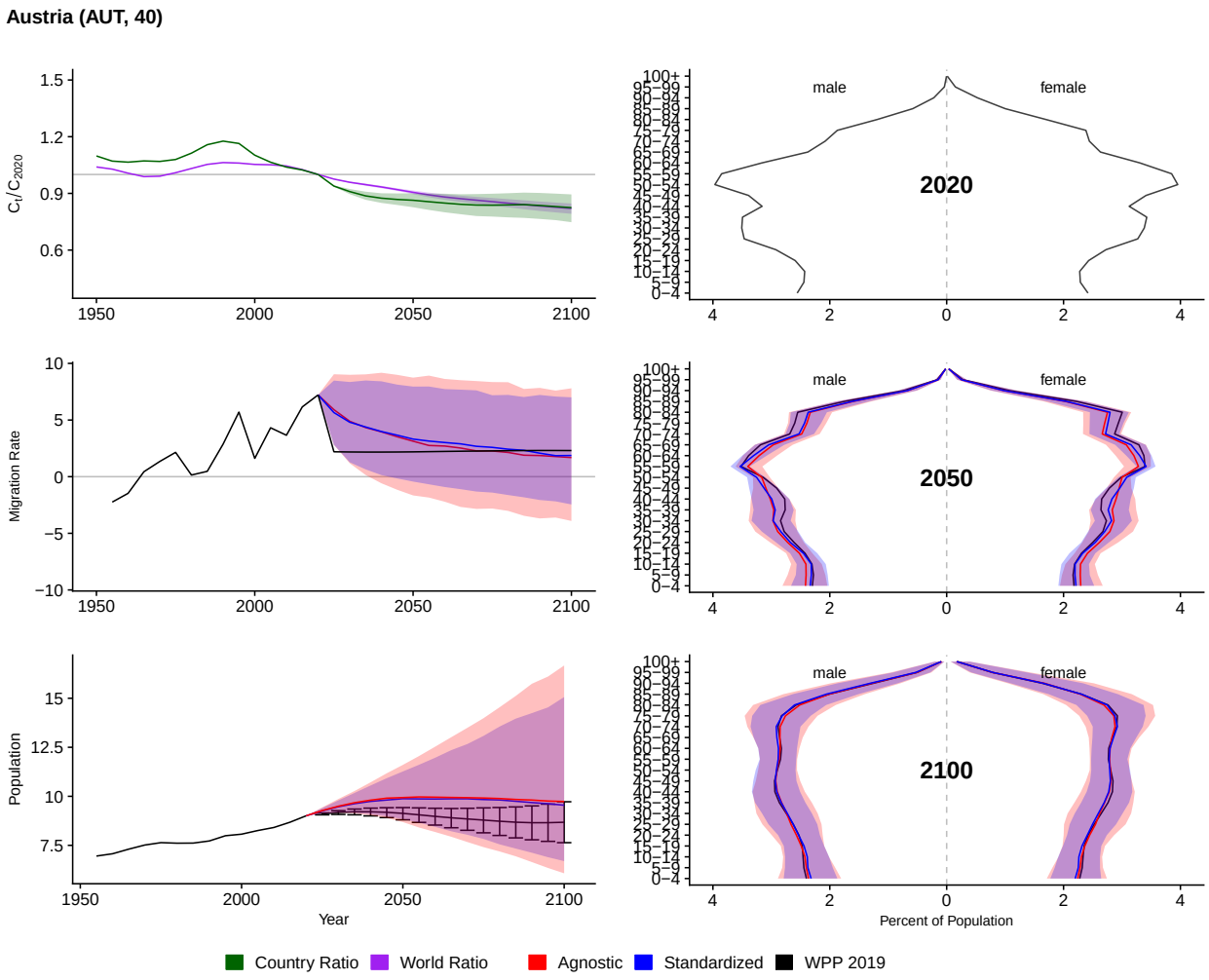


Figure A.65: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

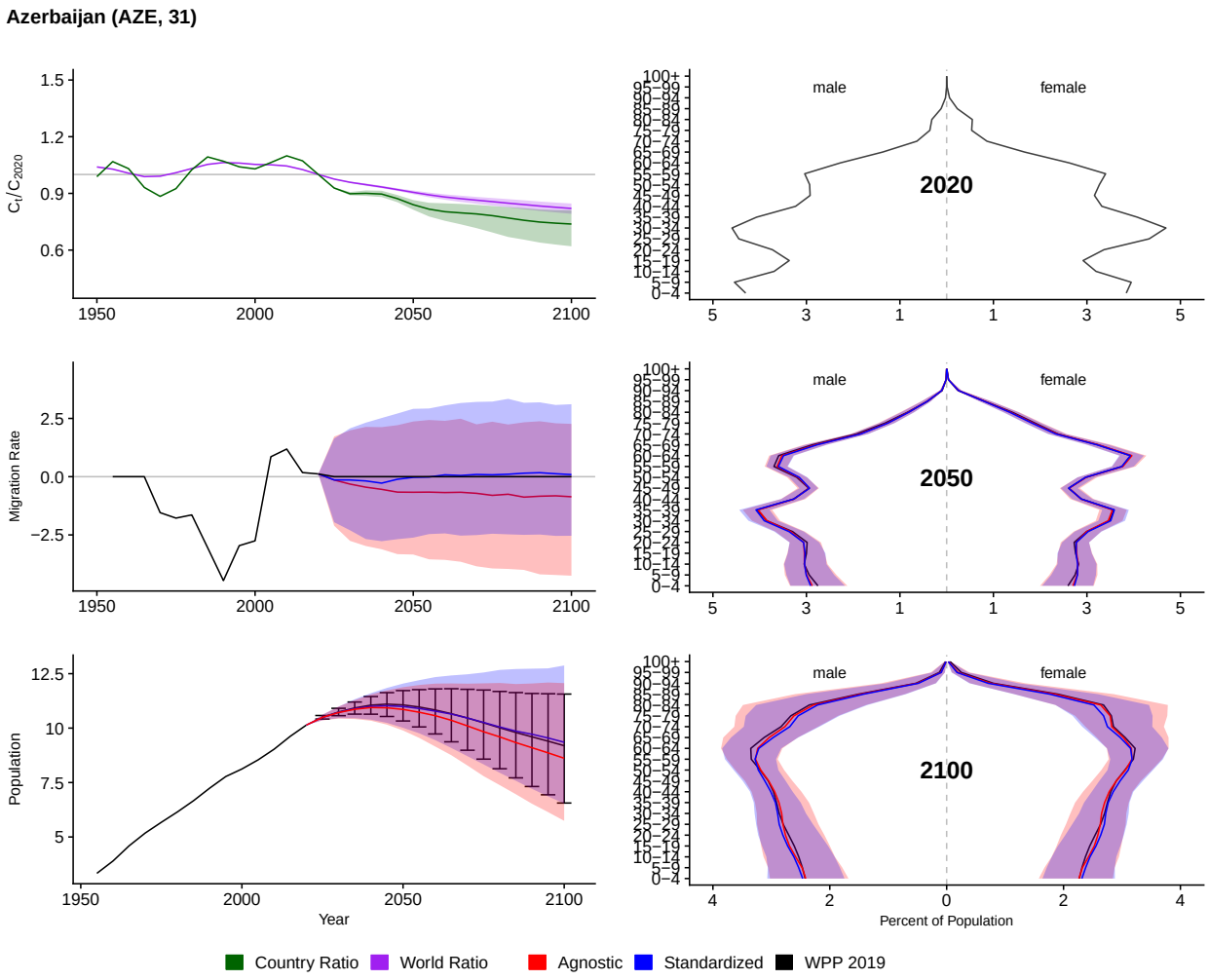


Figure A.66: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

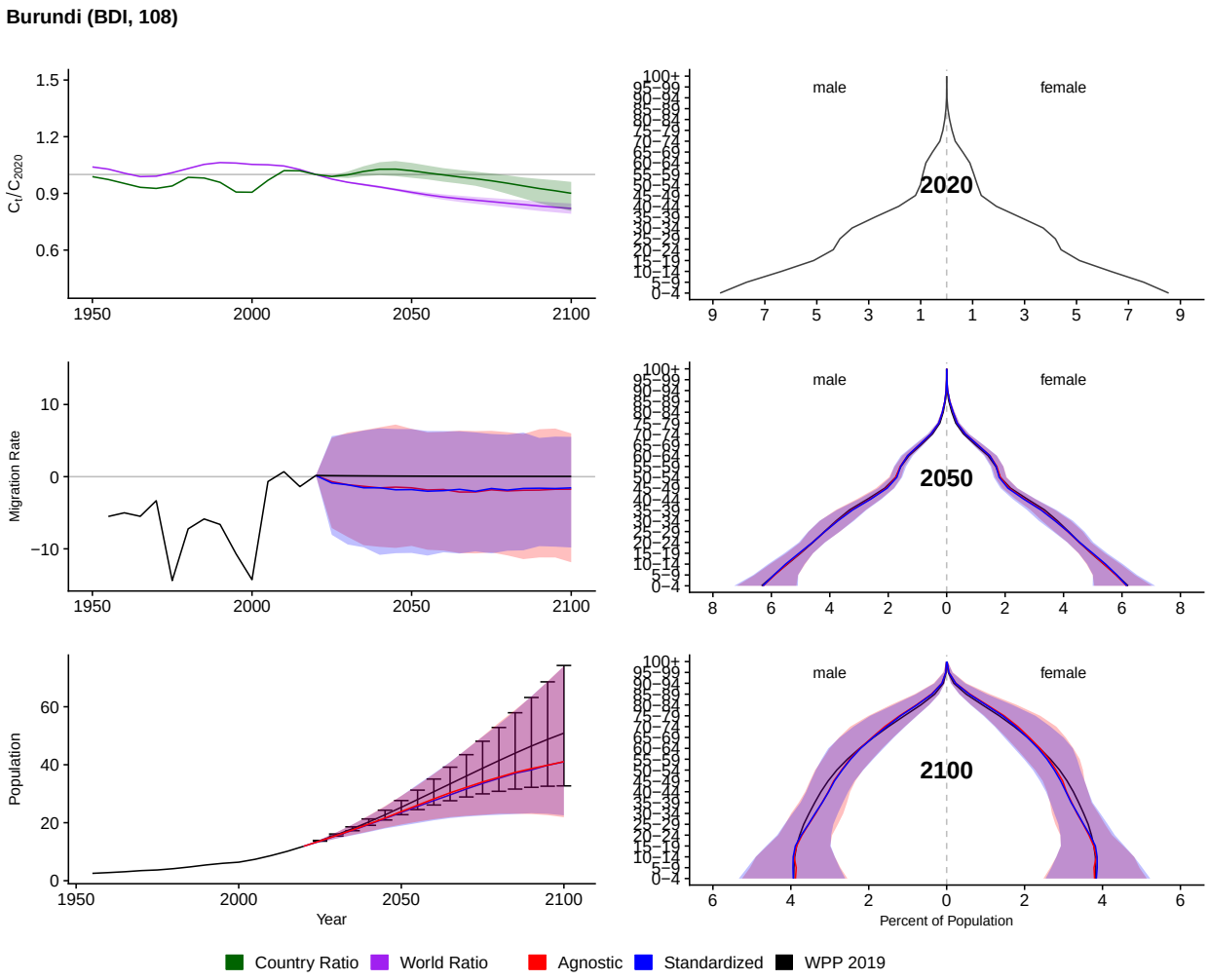


Figure A.67: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

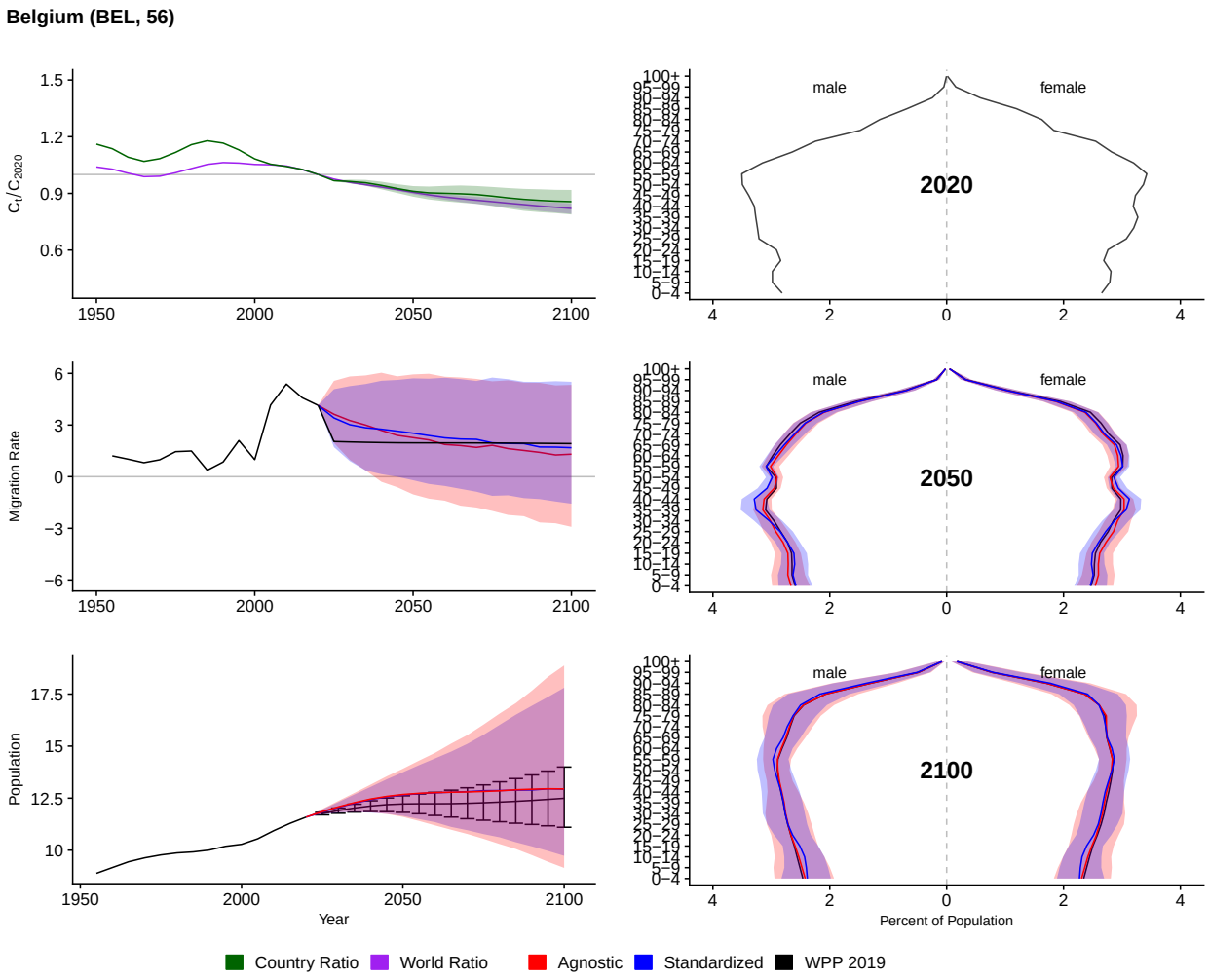


Figure A.68: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

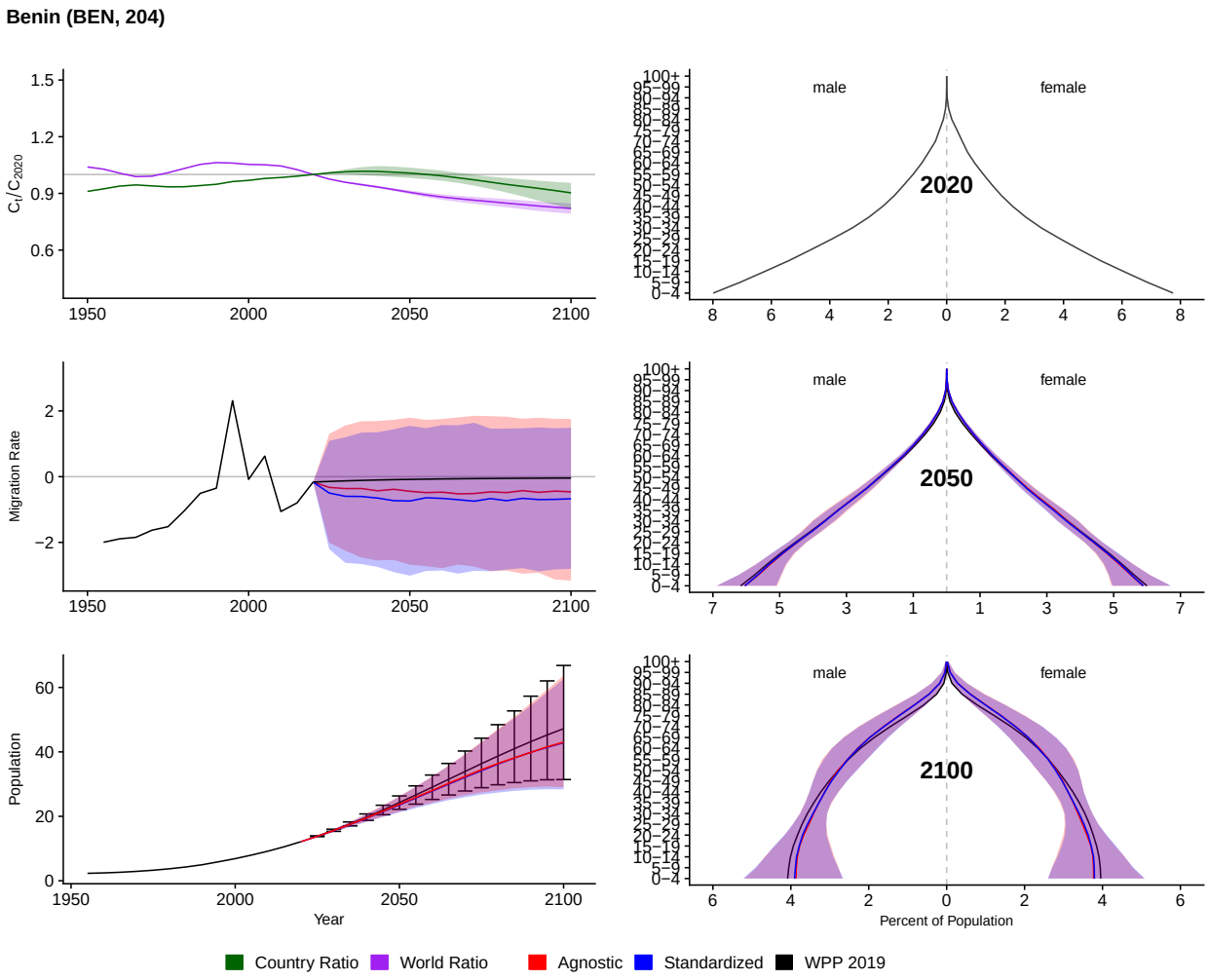


Figure A.69: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

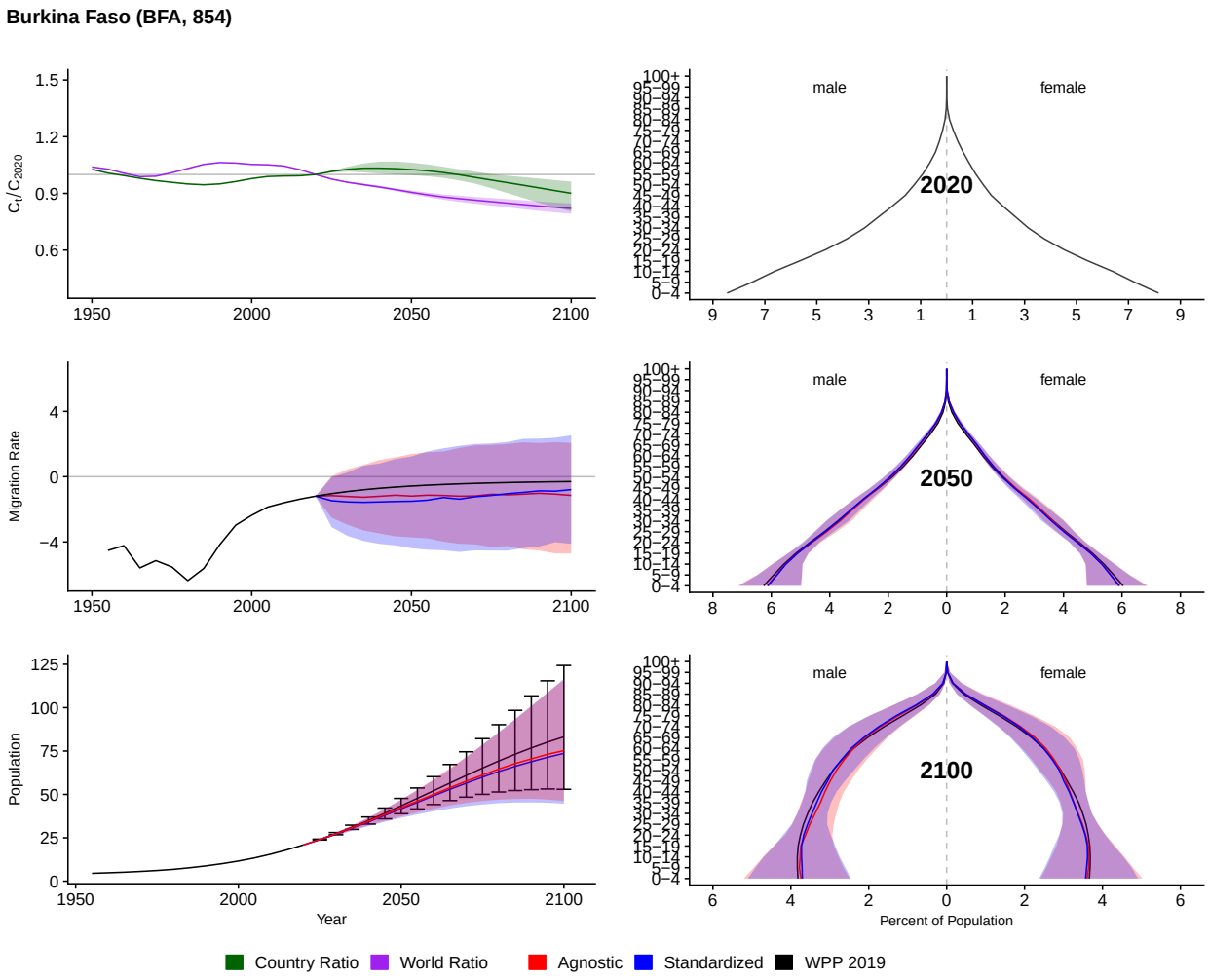


Figure A.70: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

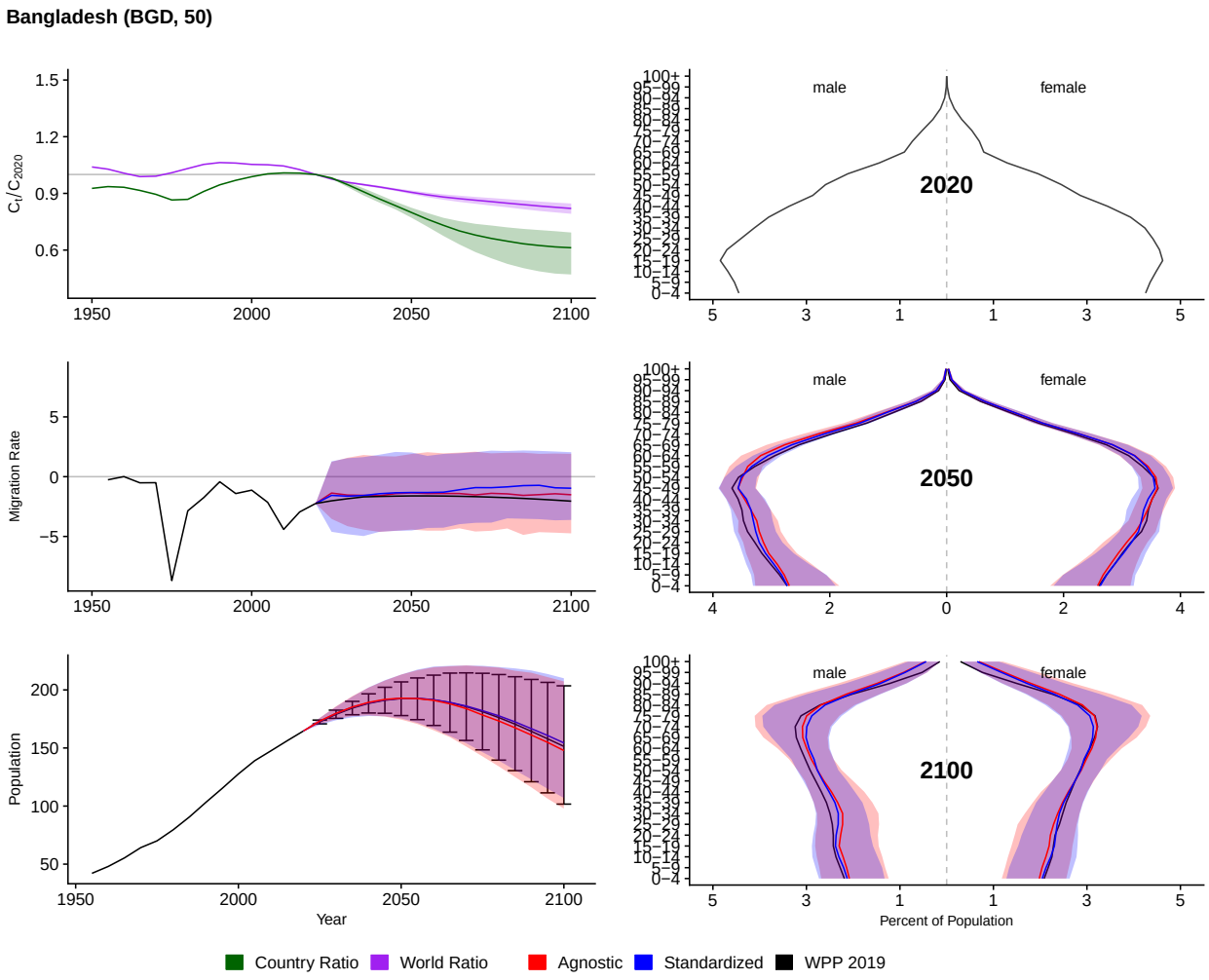


Figure A.71: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

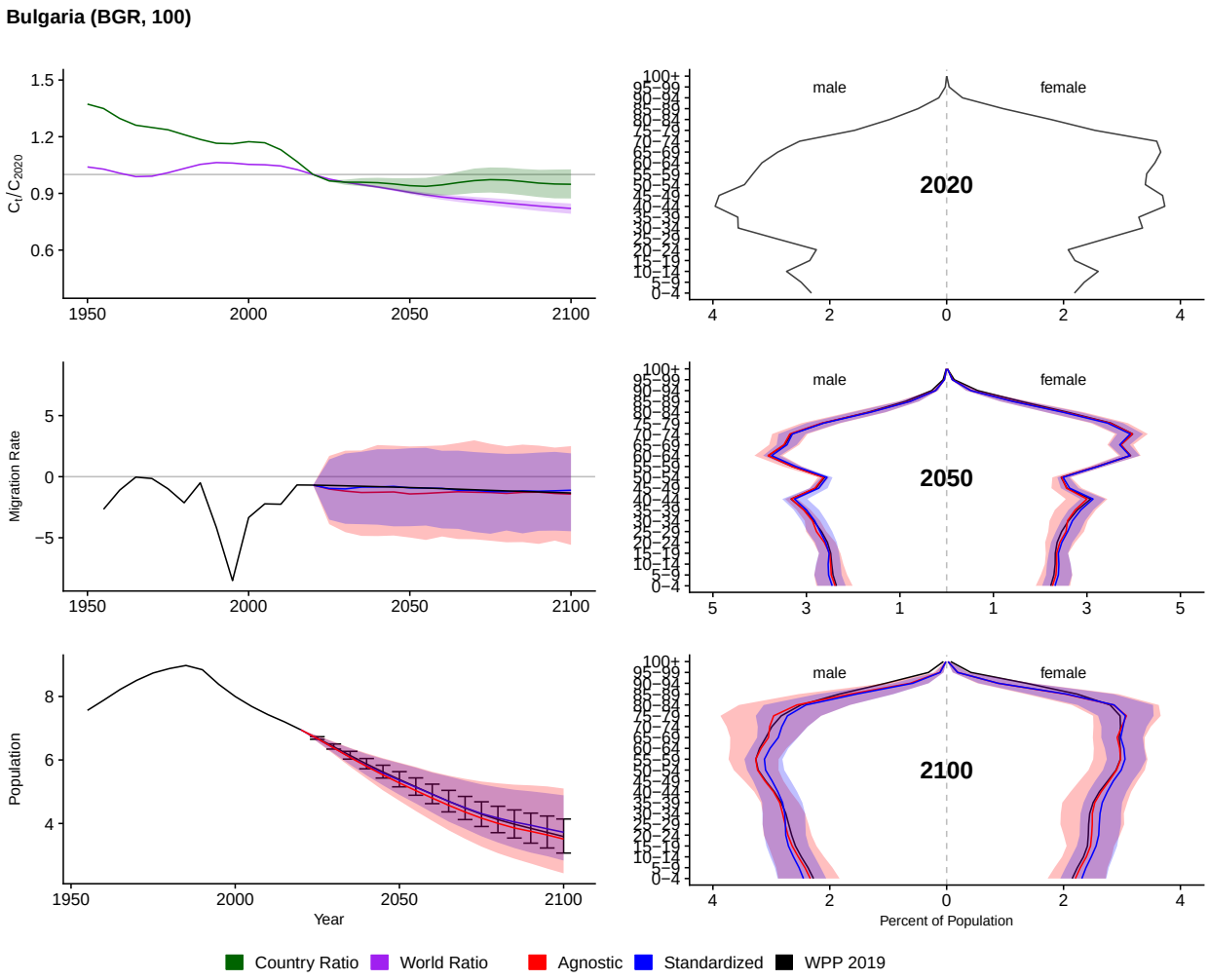


Figure A.72: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

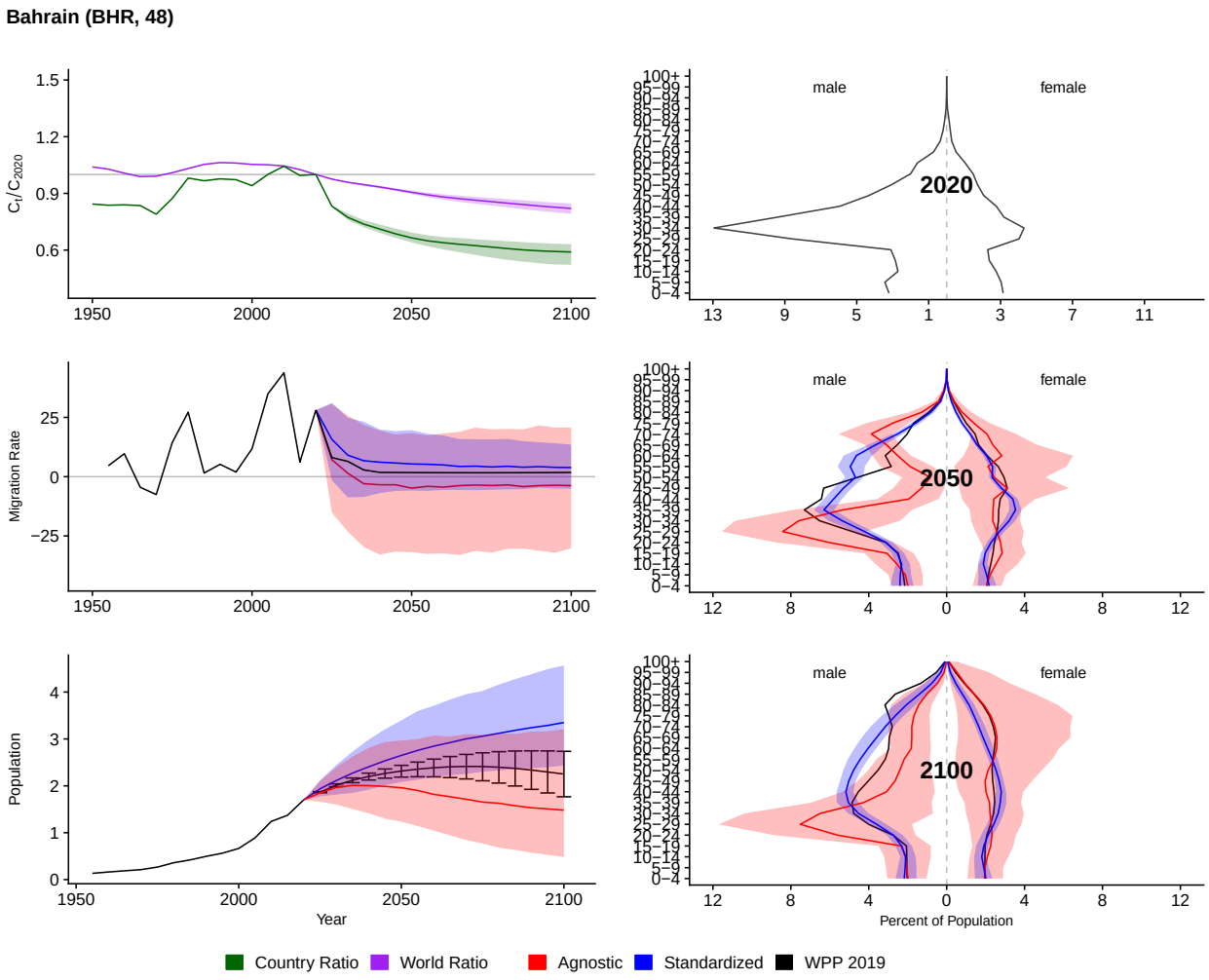


Figure A.73: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

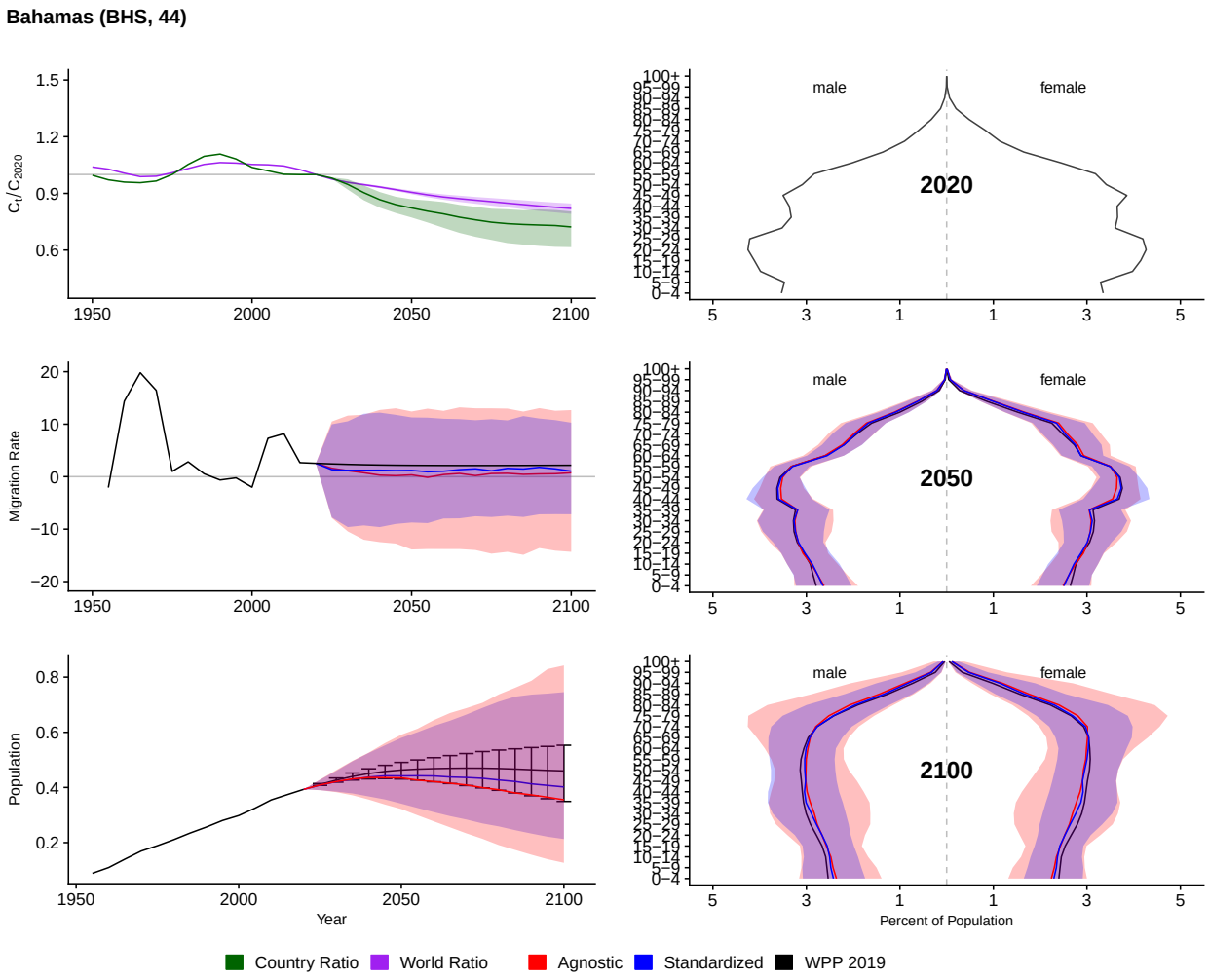


Figure A.74: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

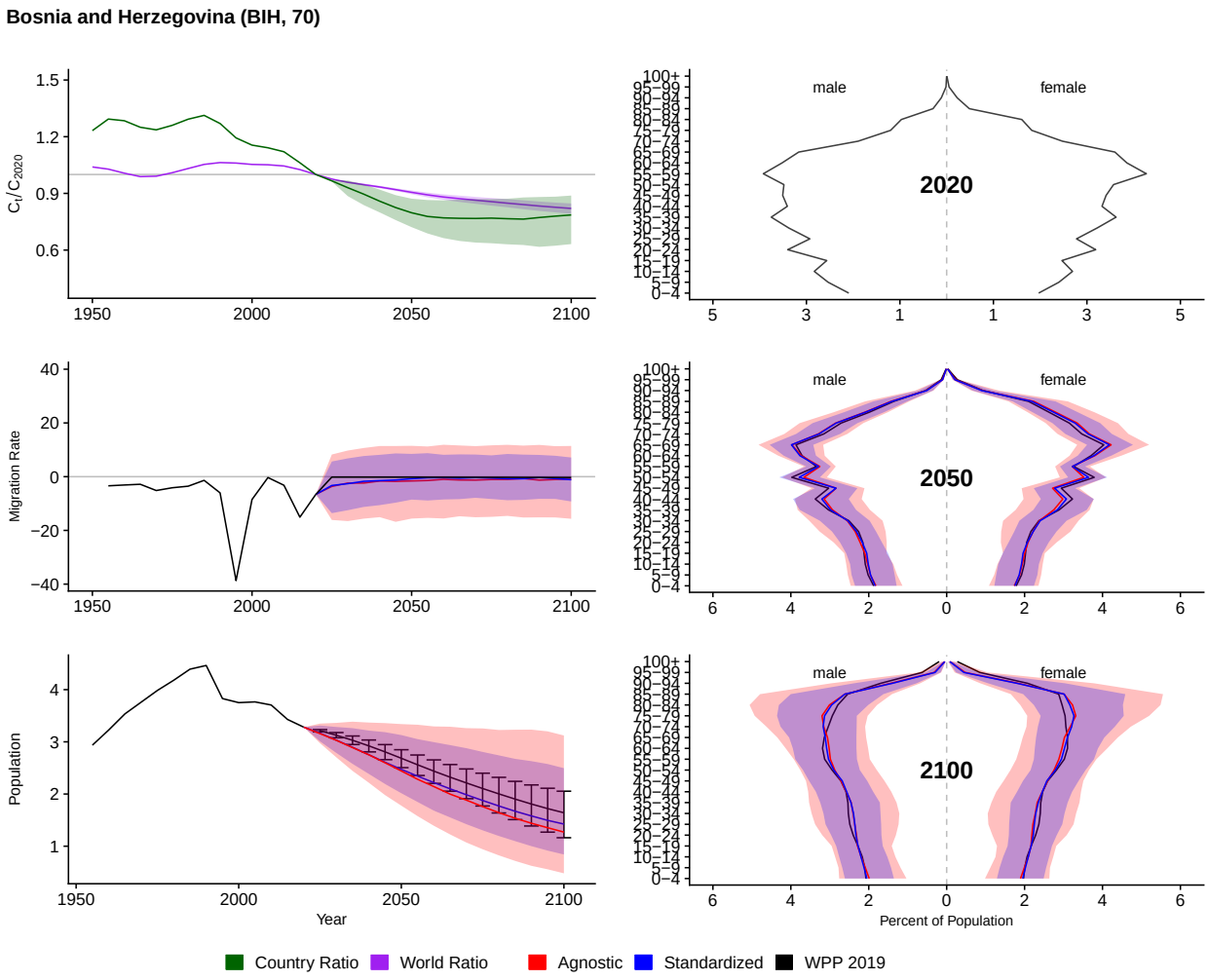


Figure A.75: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

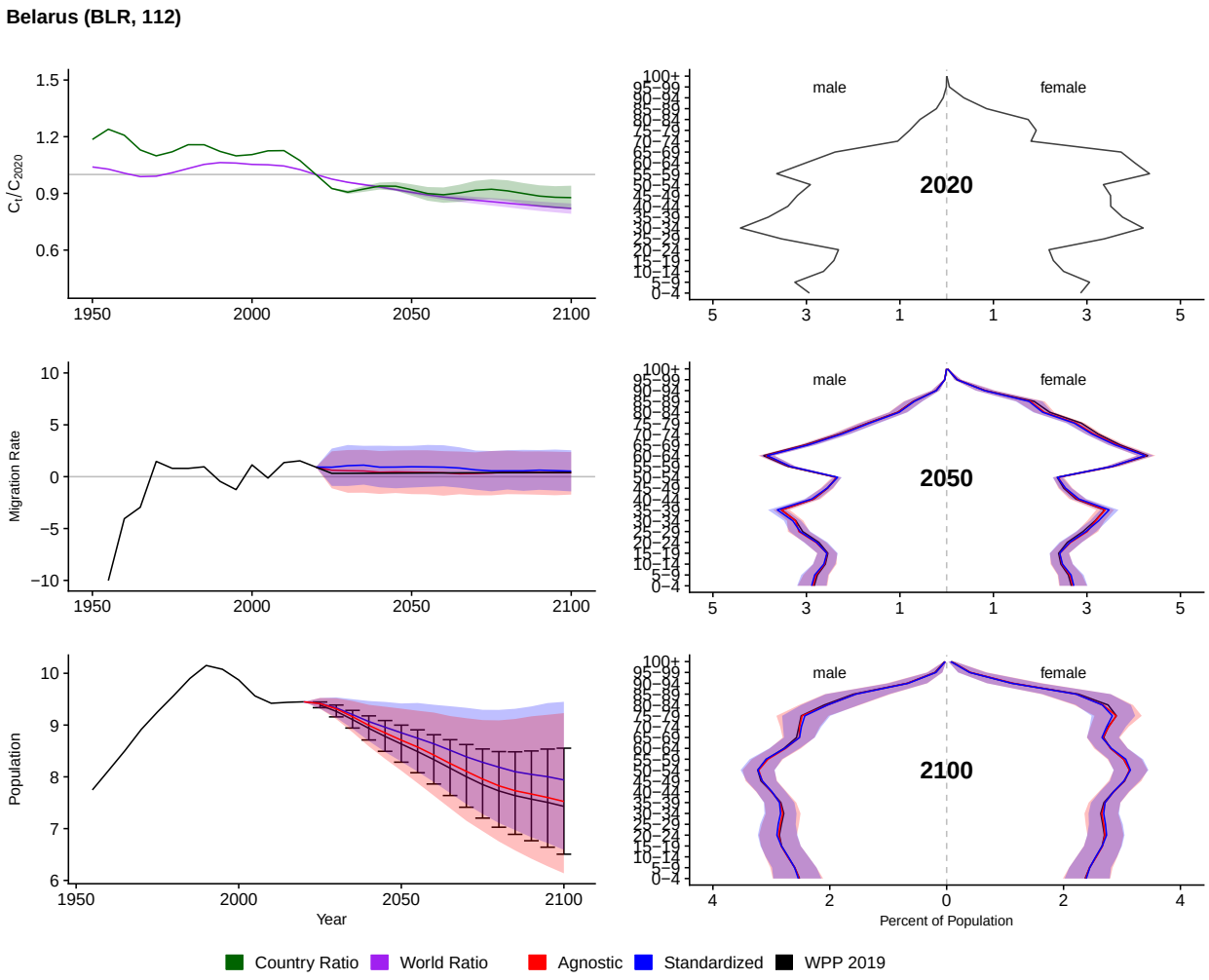


Figure A.76: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

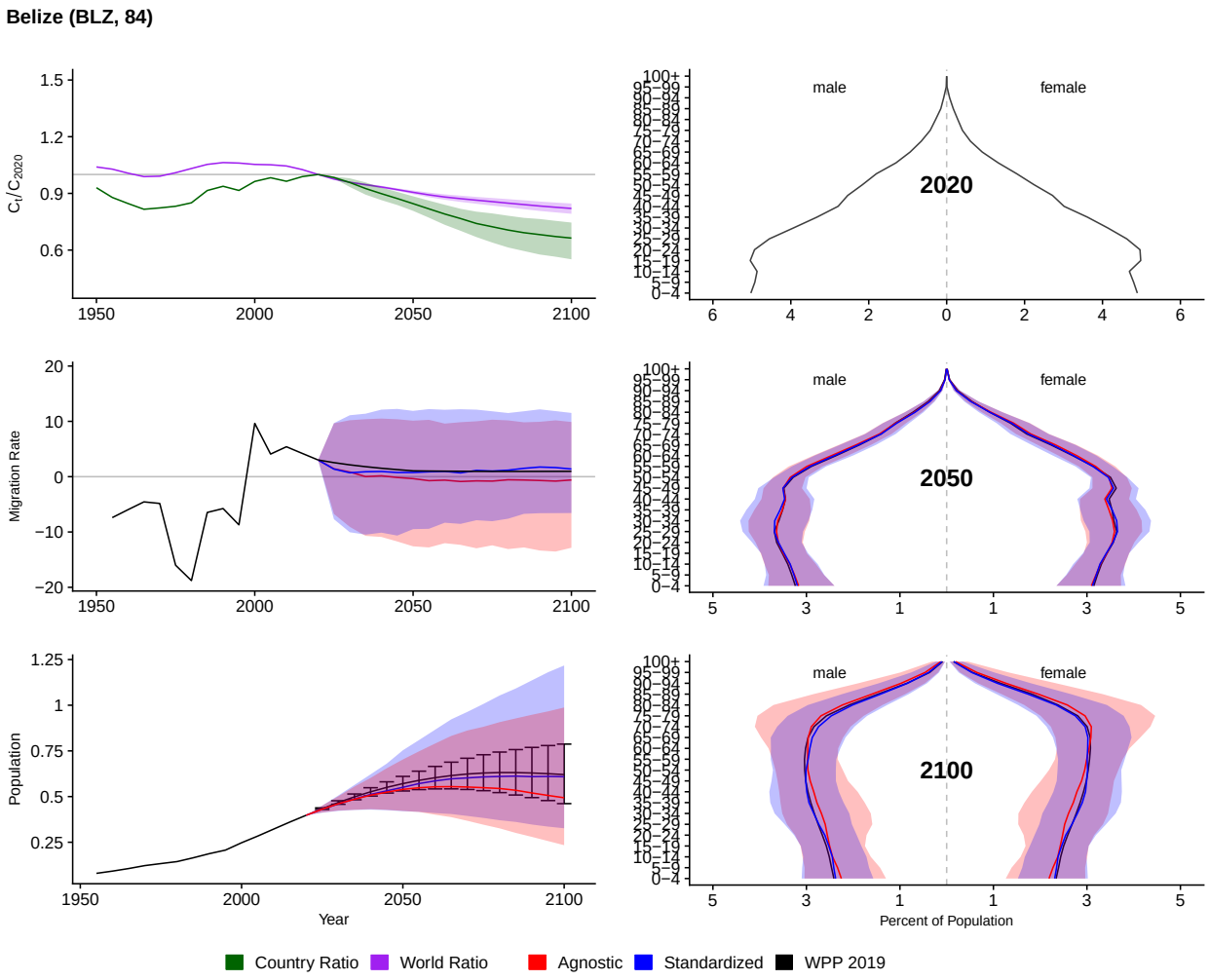


Figure A.77: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

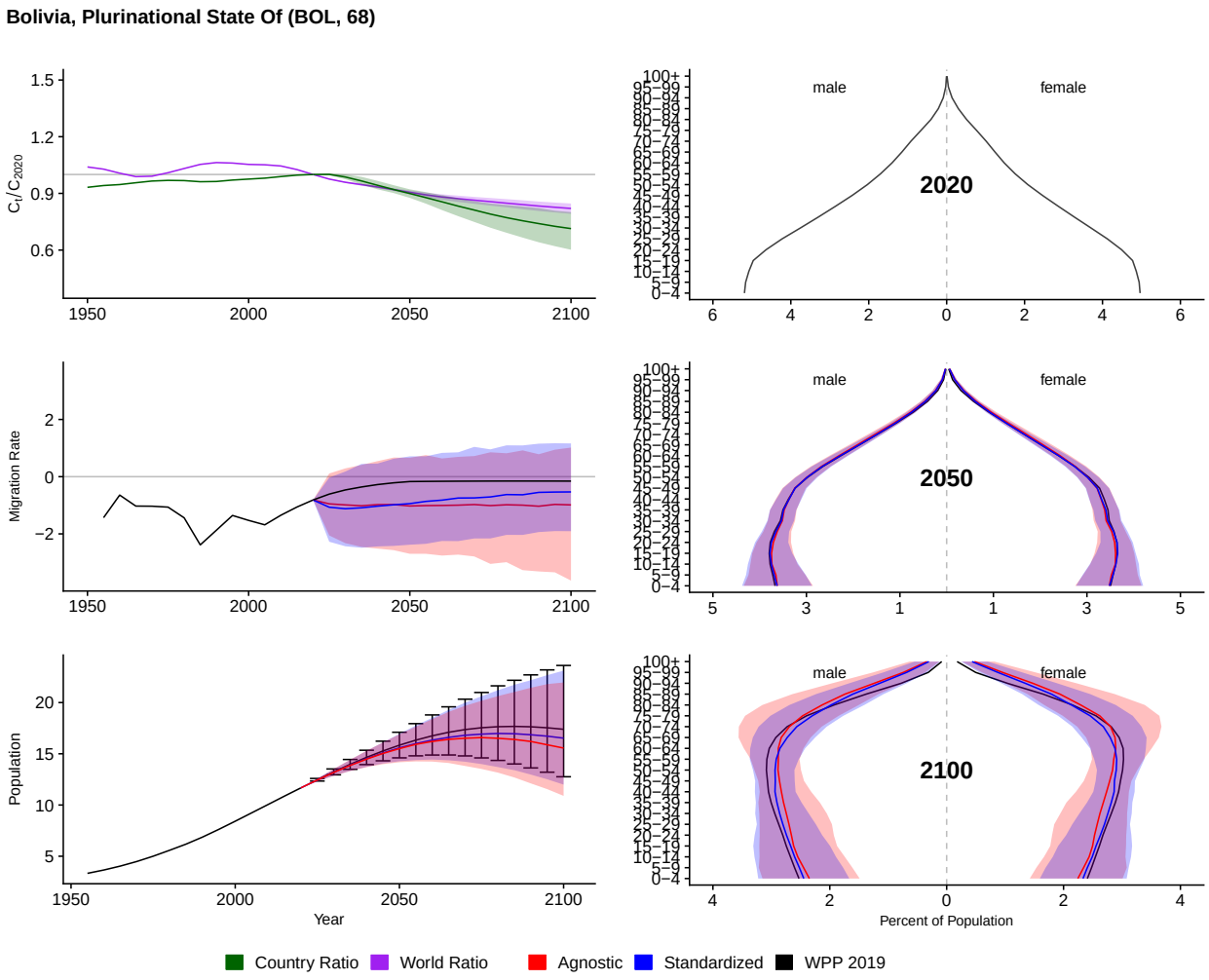


Figure A.78: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

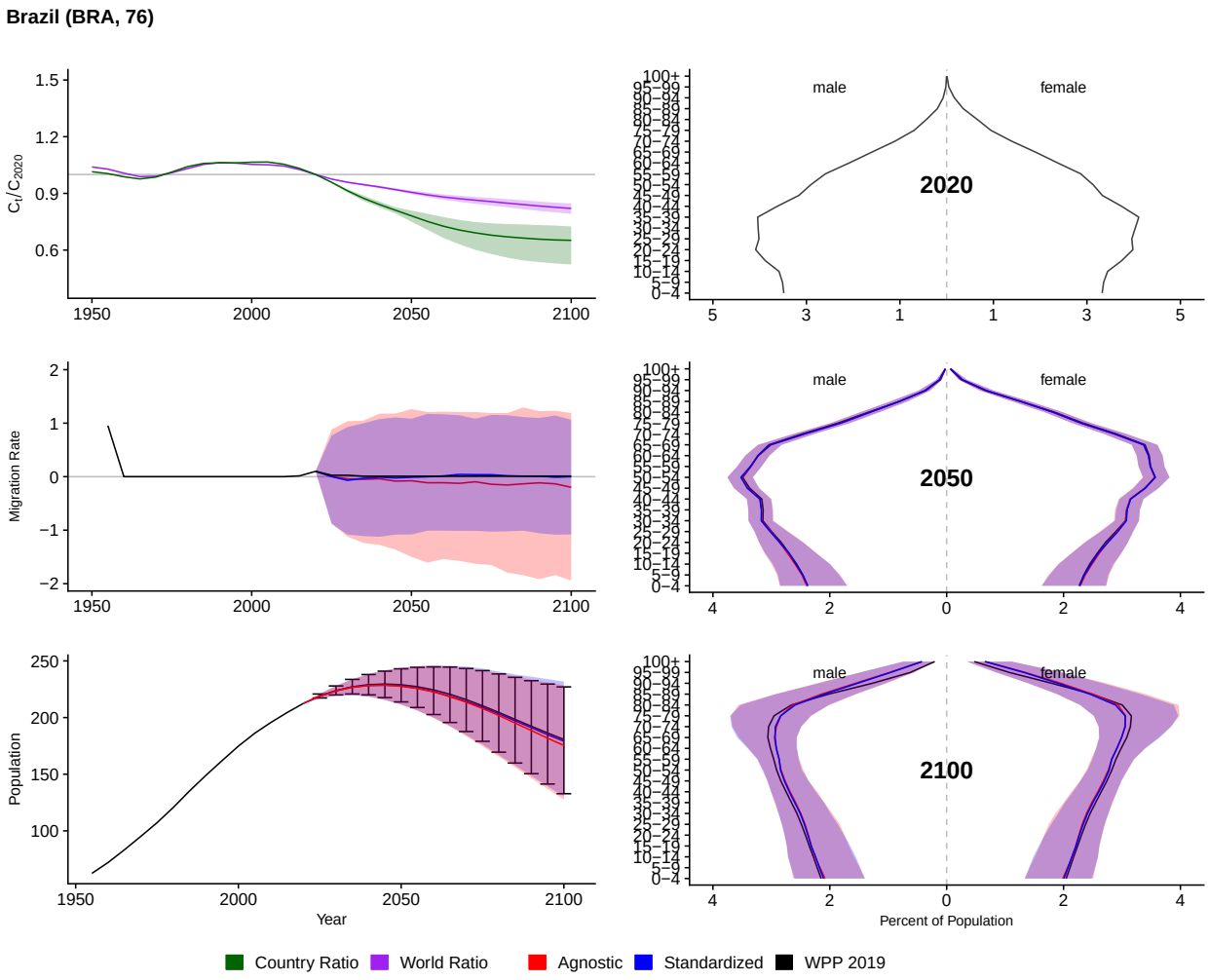


Figure A.79: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

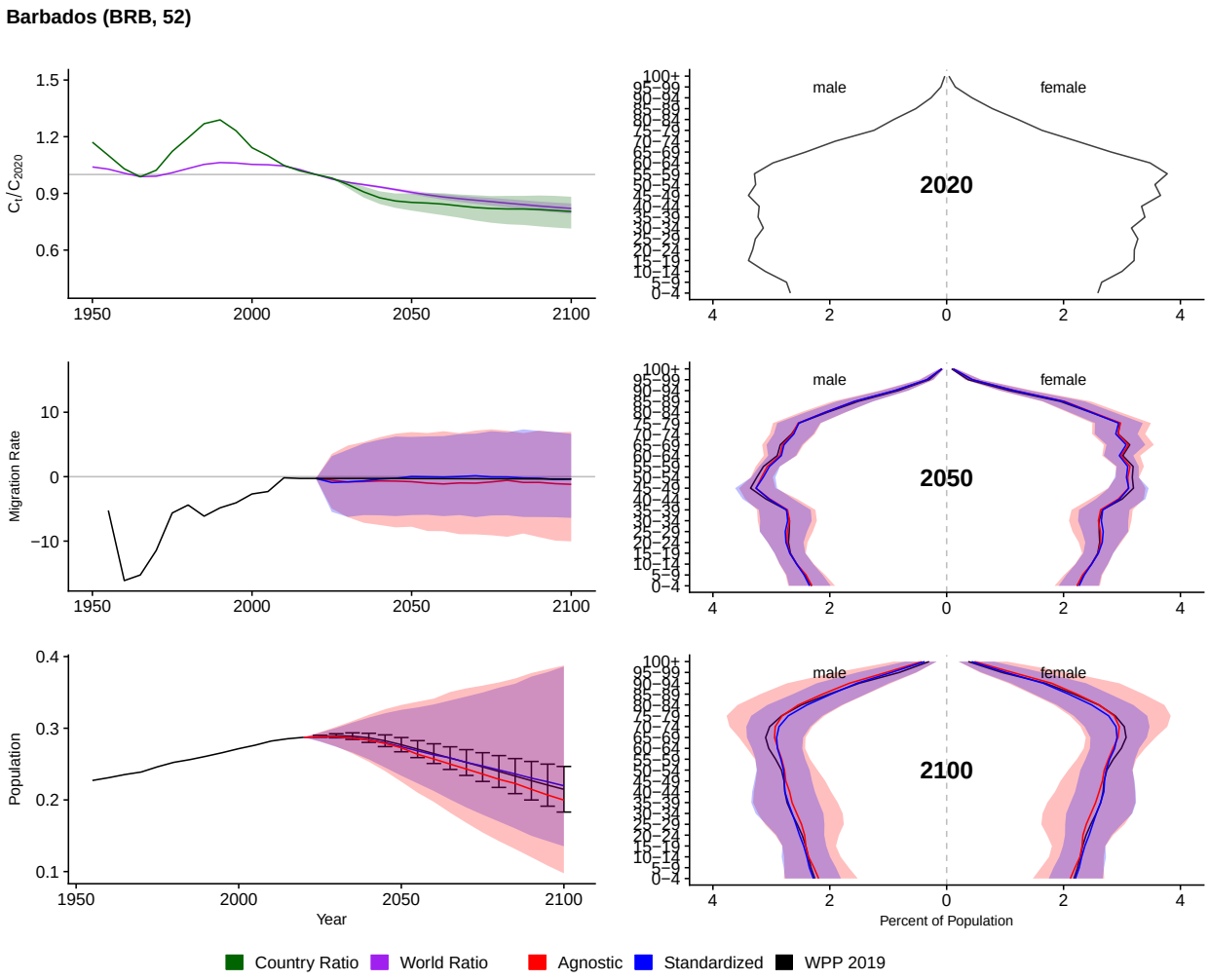


Figure A.80: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

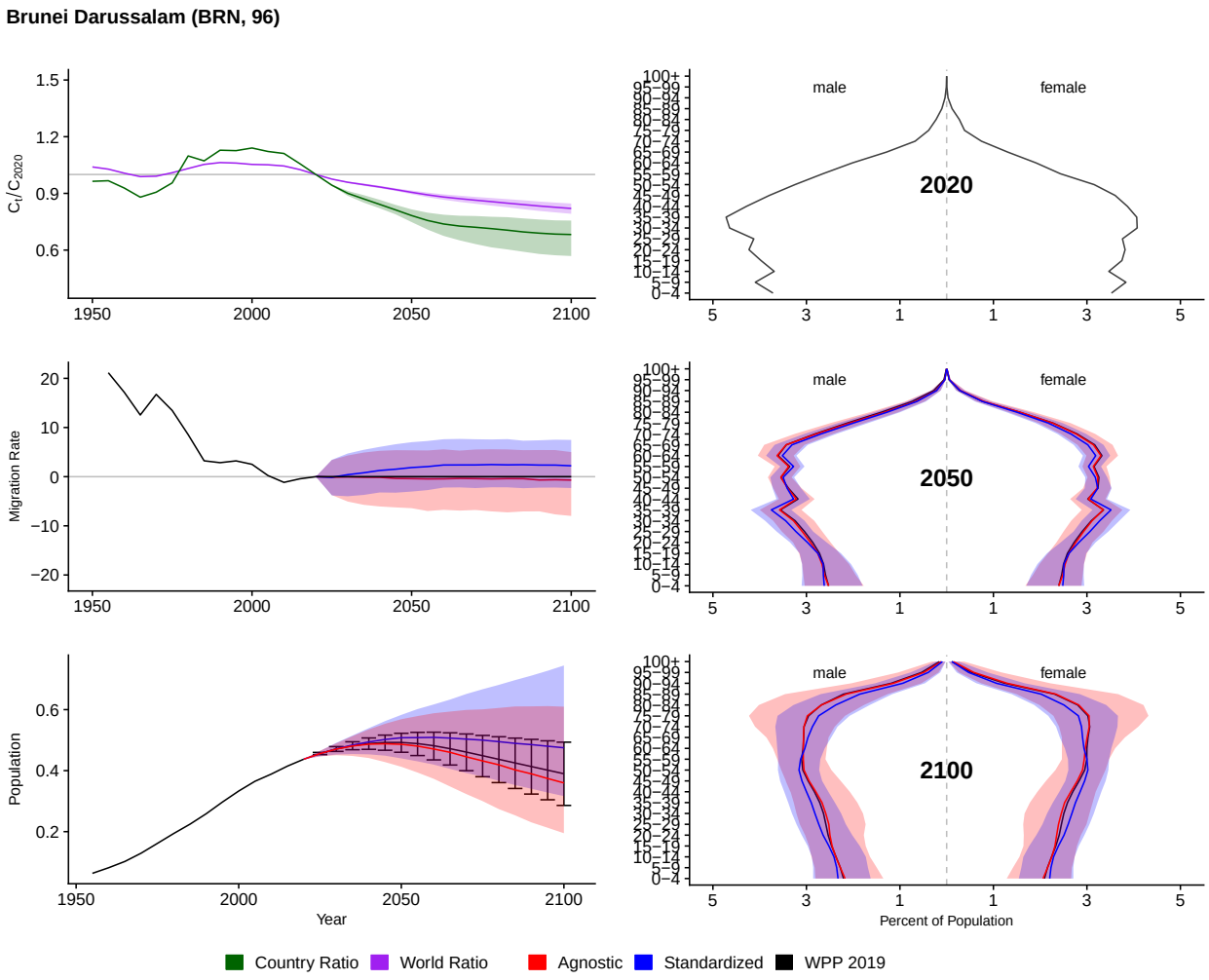


Figure A.81: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

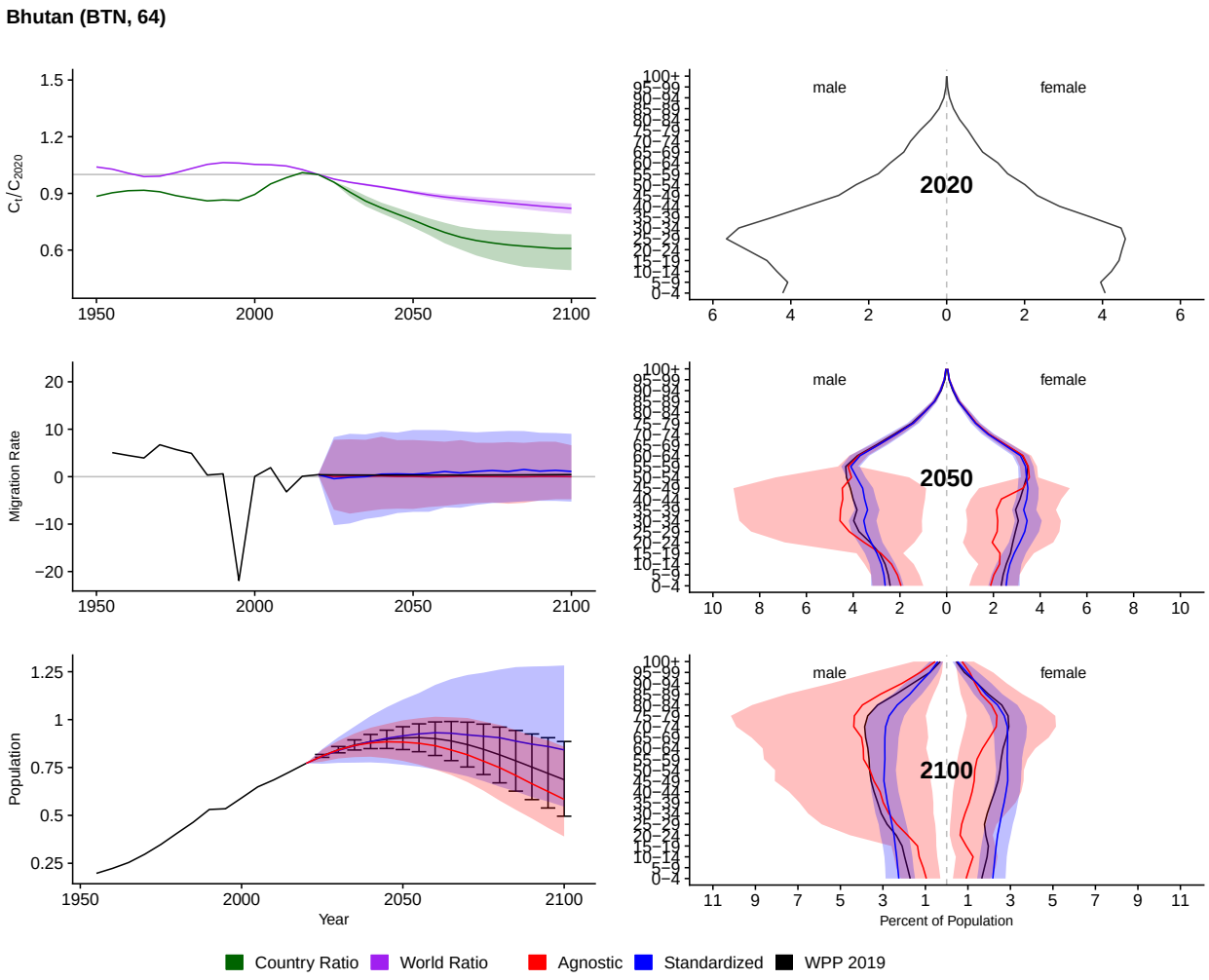


Figure A.82: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

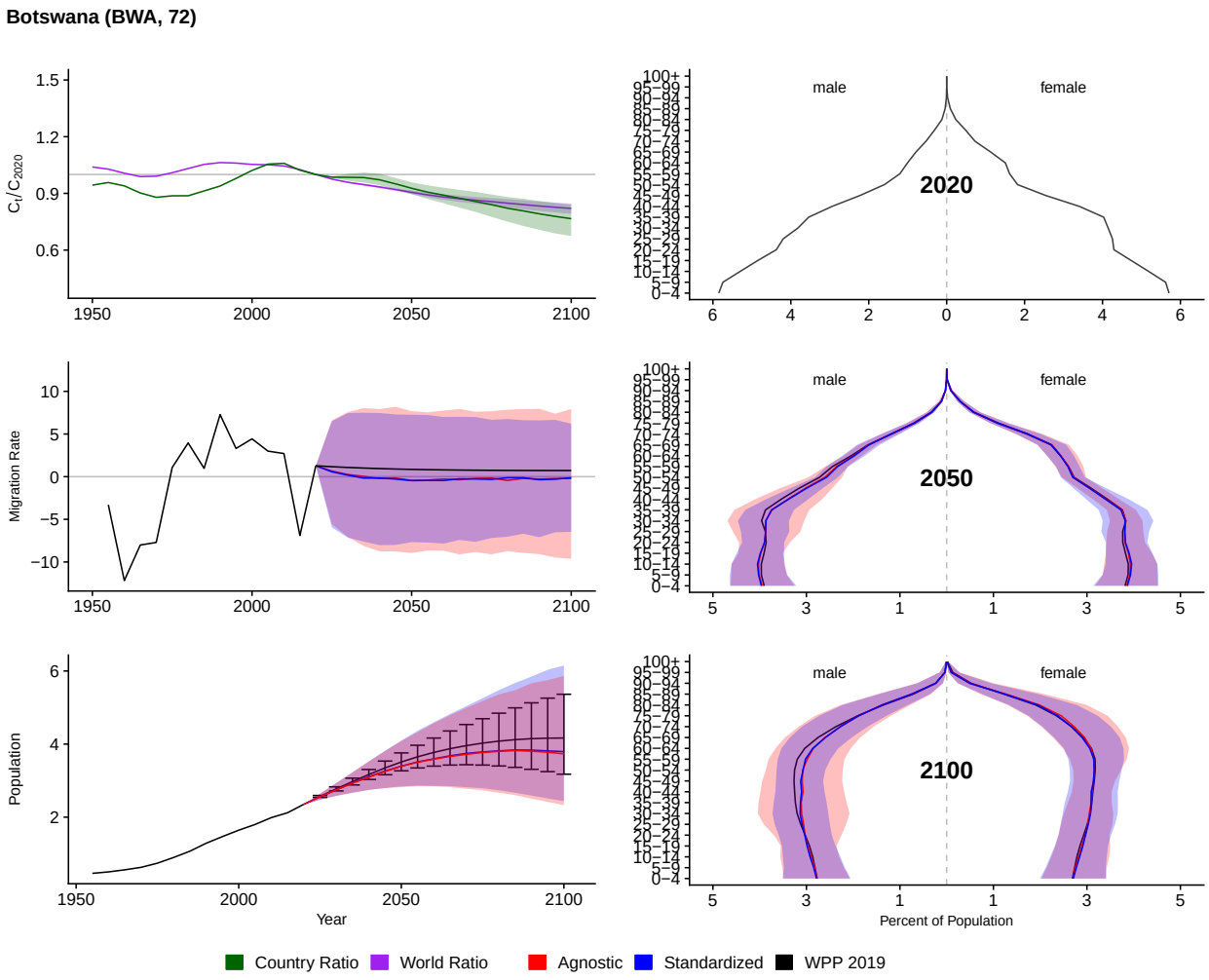


Figure A.83: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

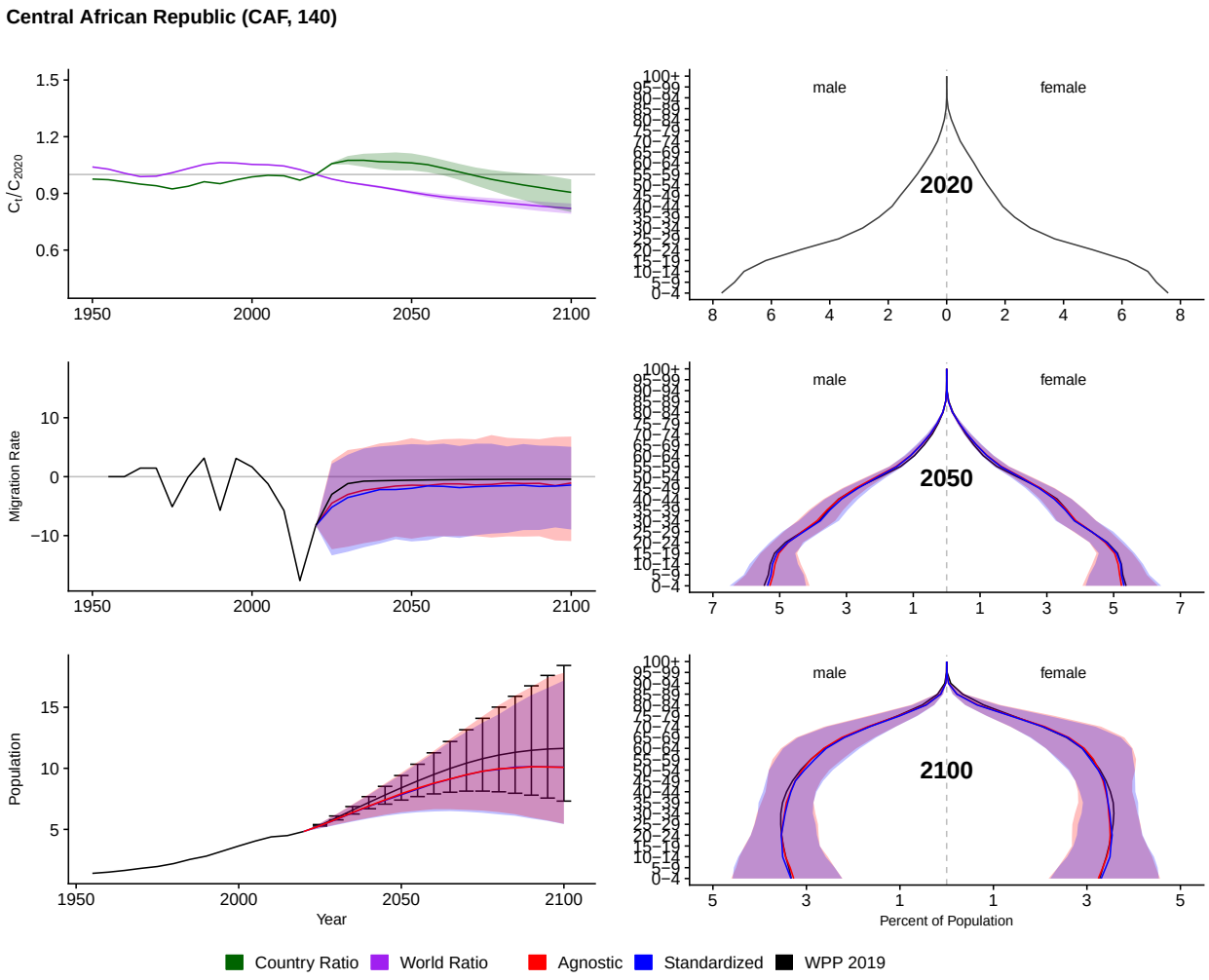


Figure A.84: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

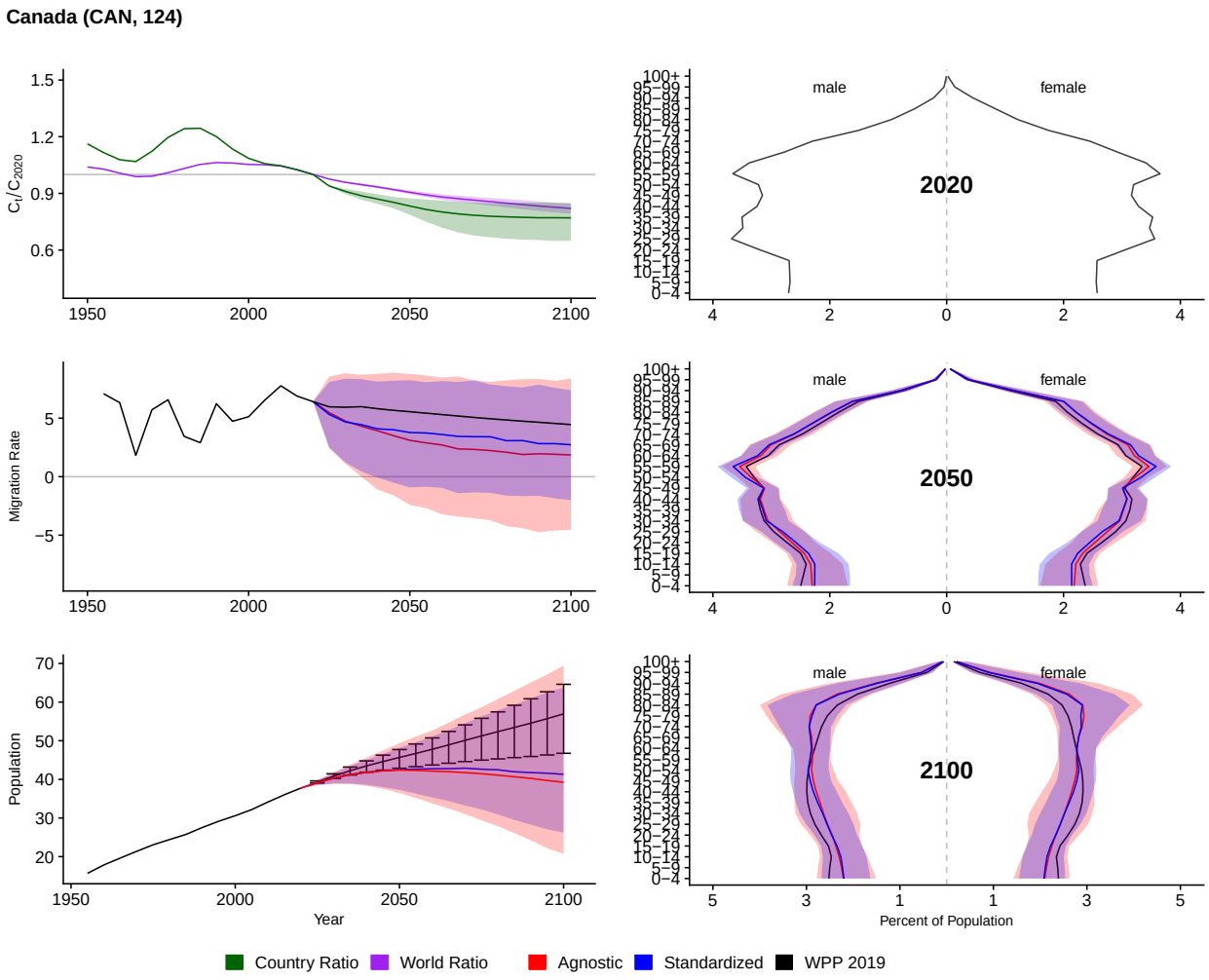


Figure A.85: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

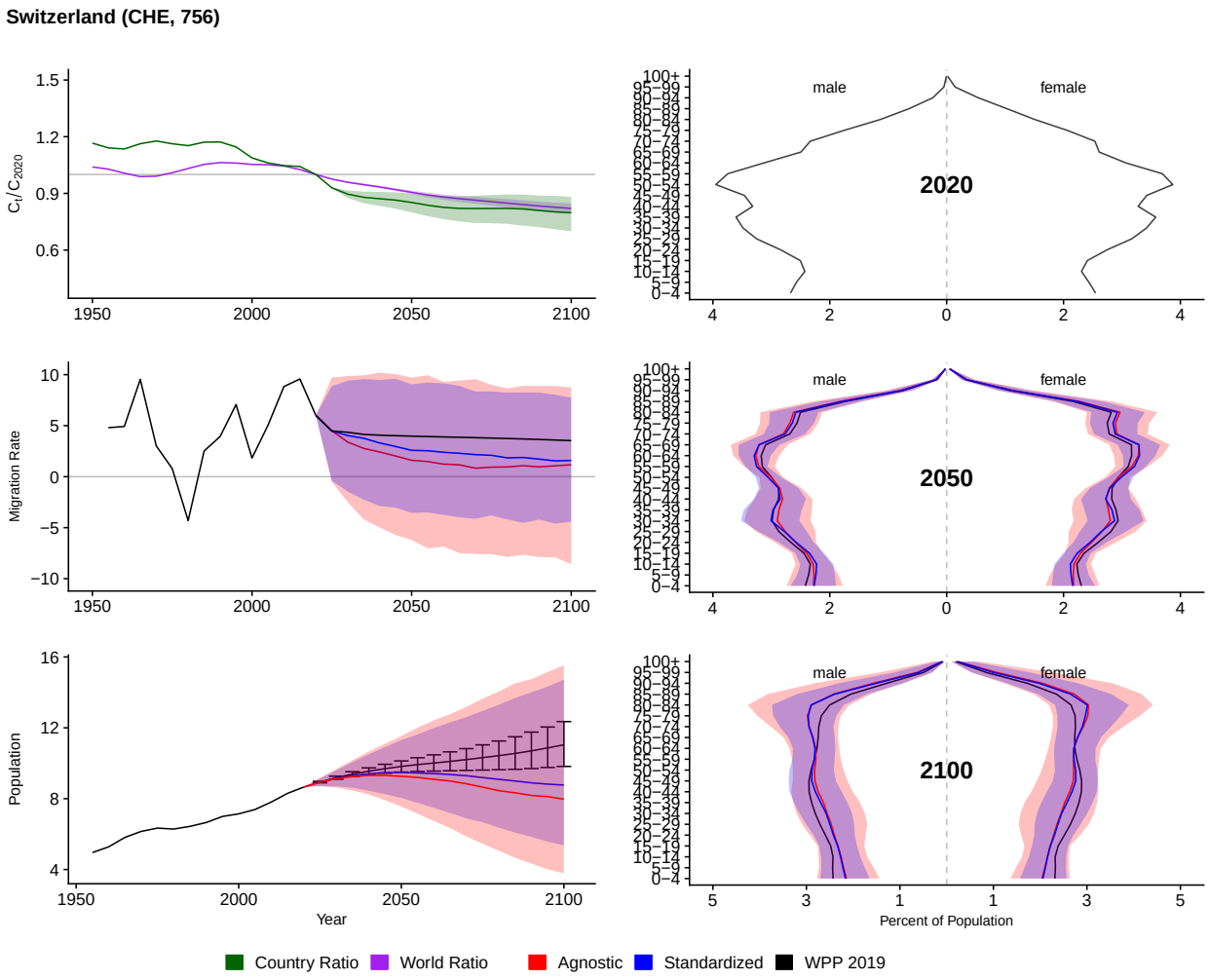


Figure A.86: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

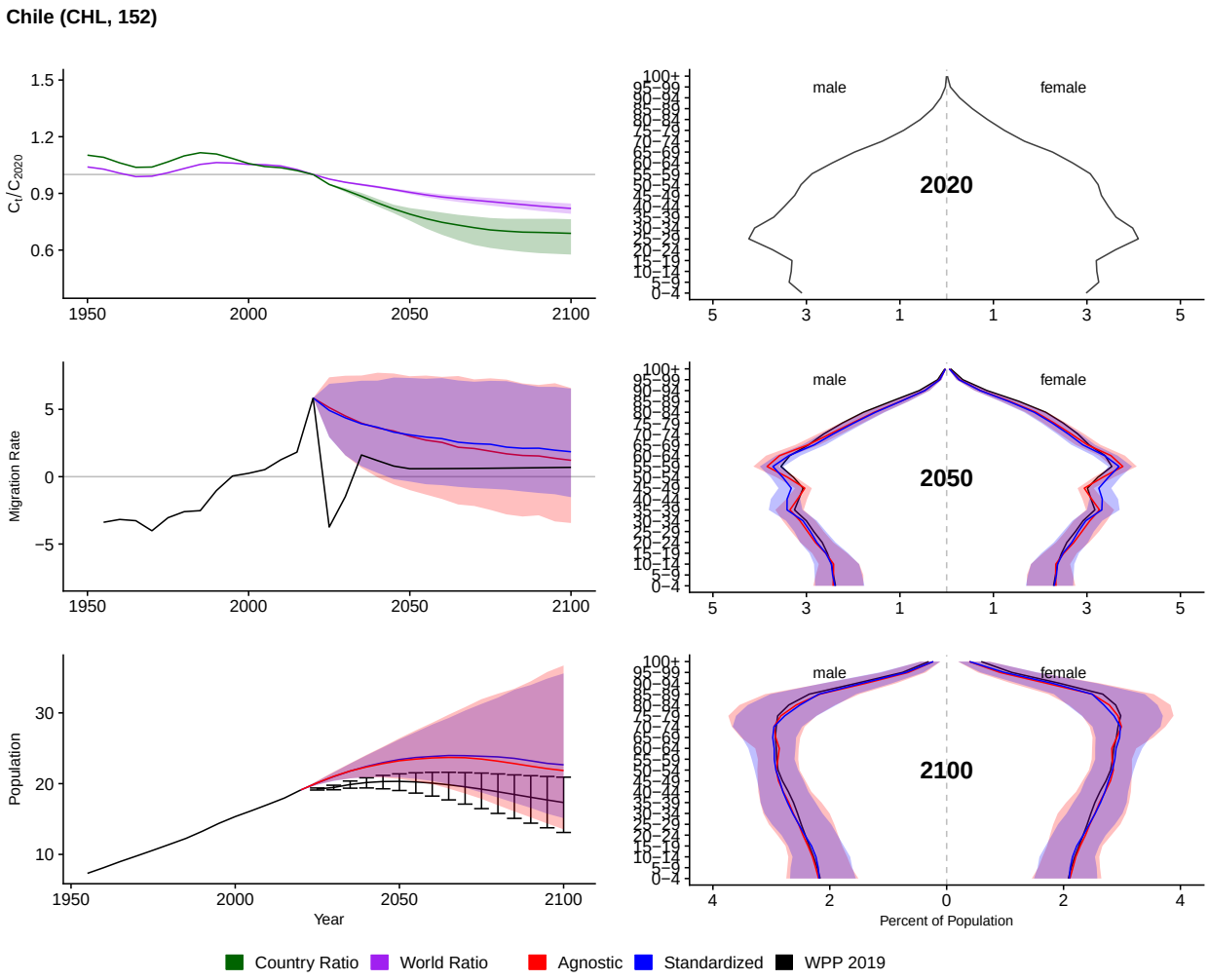


Figure A.87: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

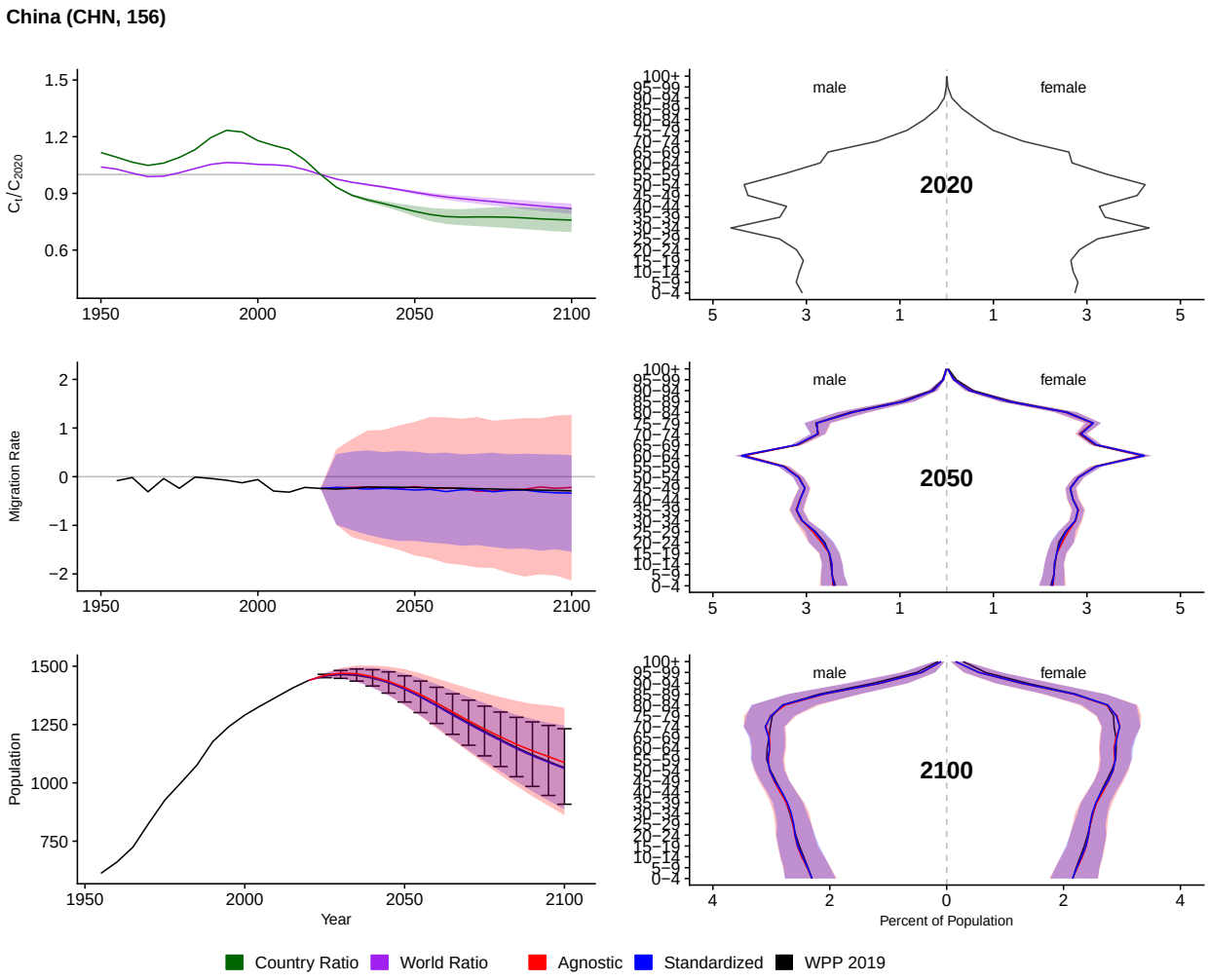


Figure A.88: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

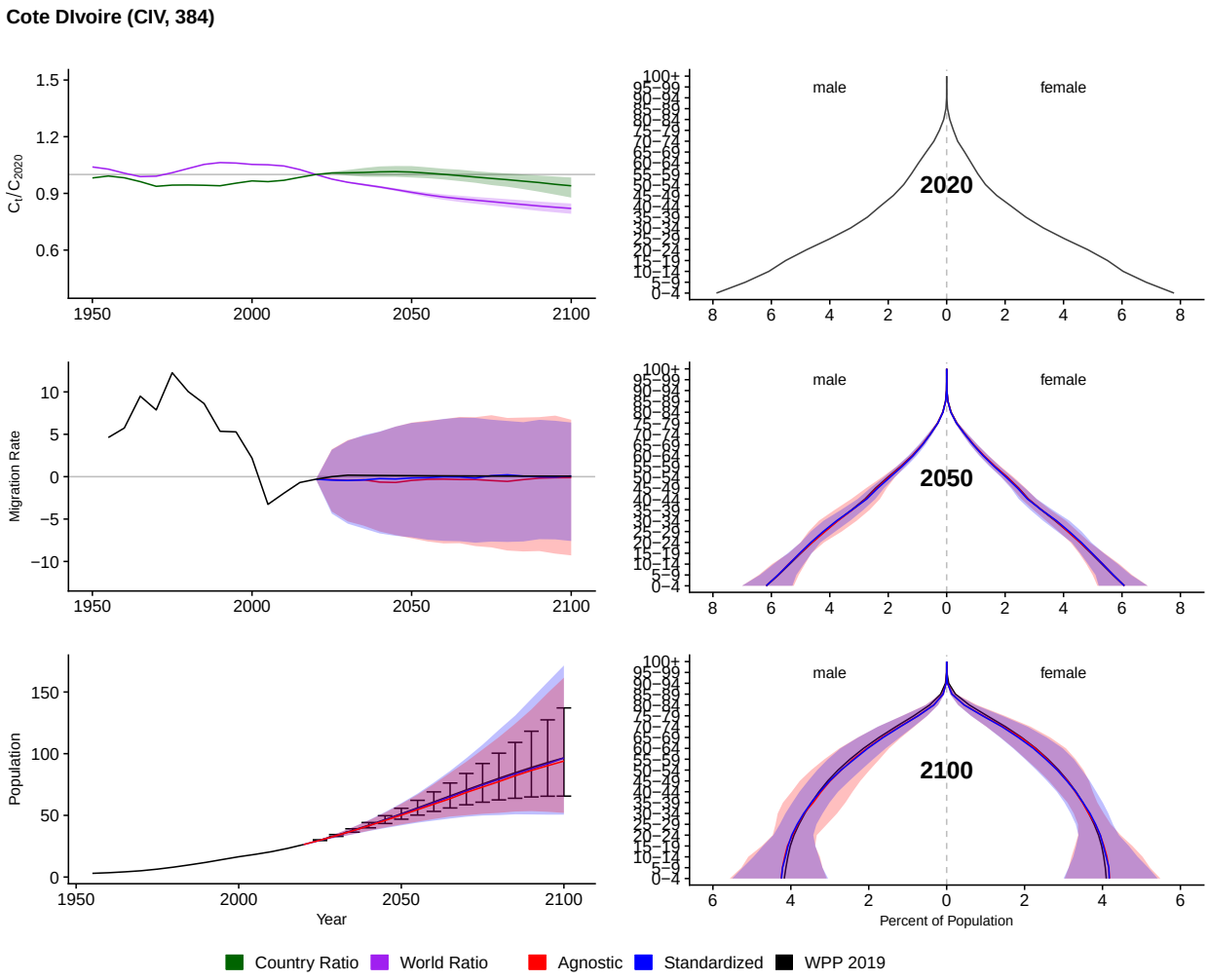


Figure A.89: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

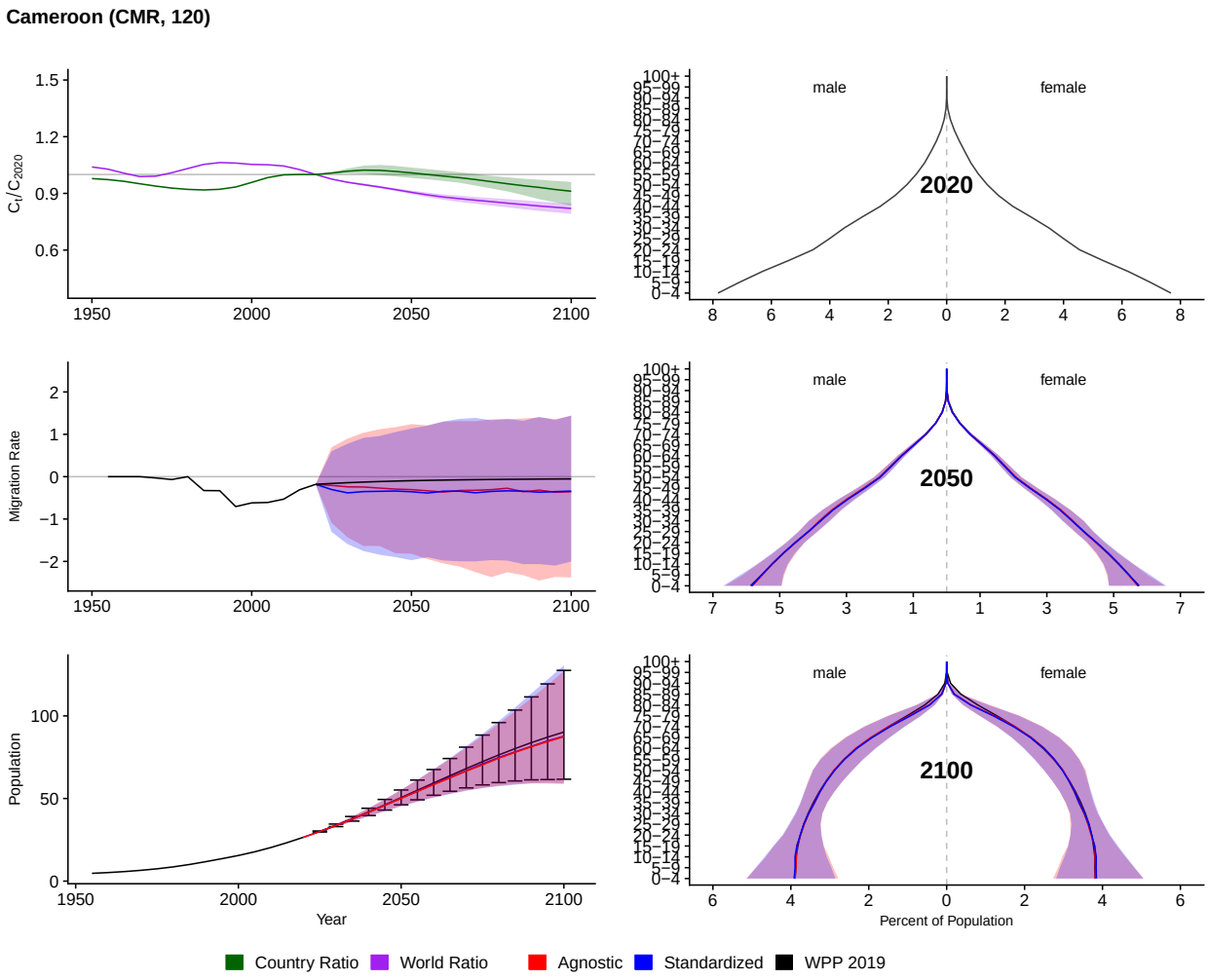


Figure A.90: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

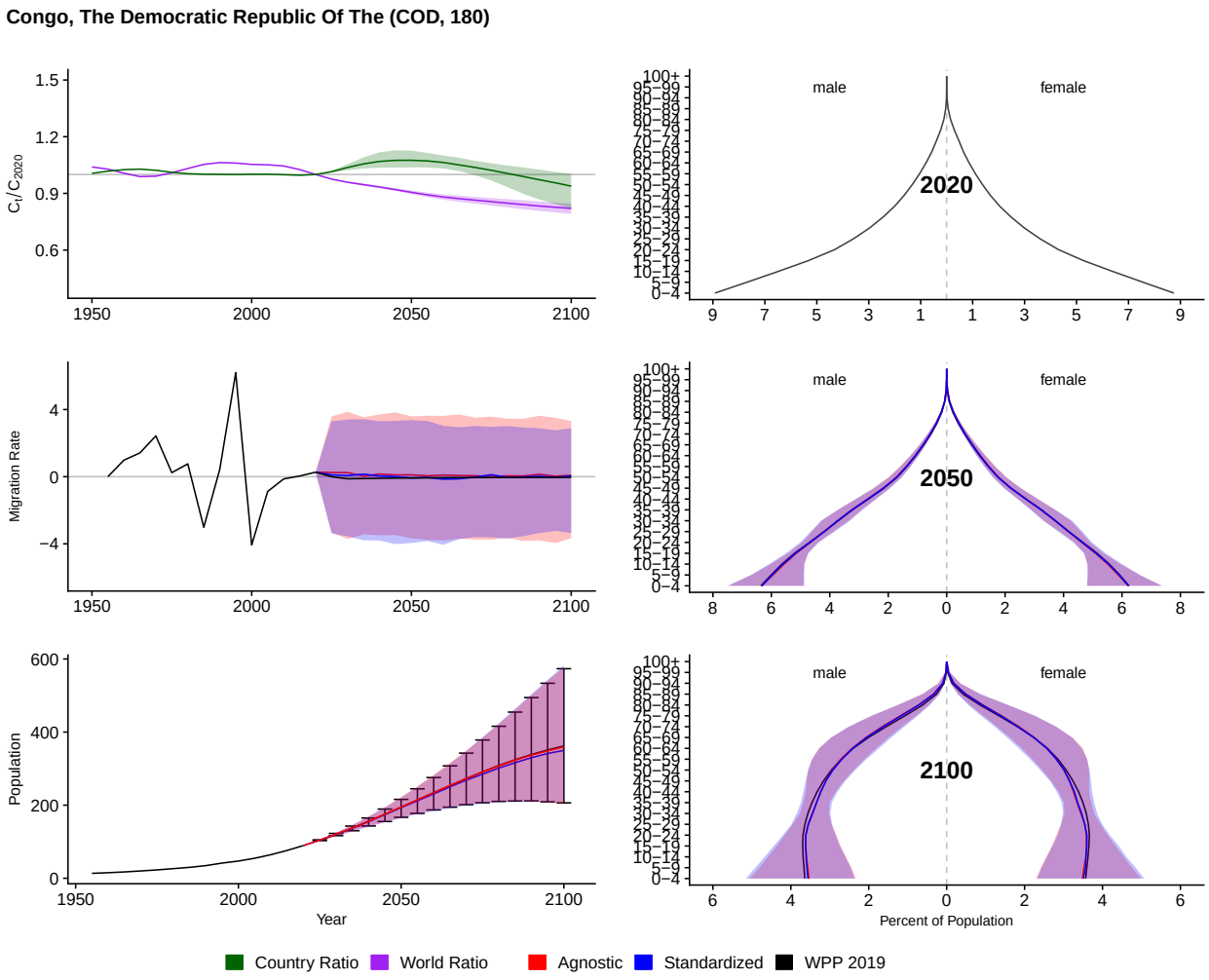


Figure A.91: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

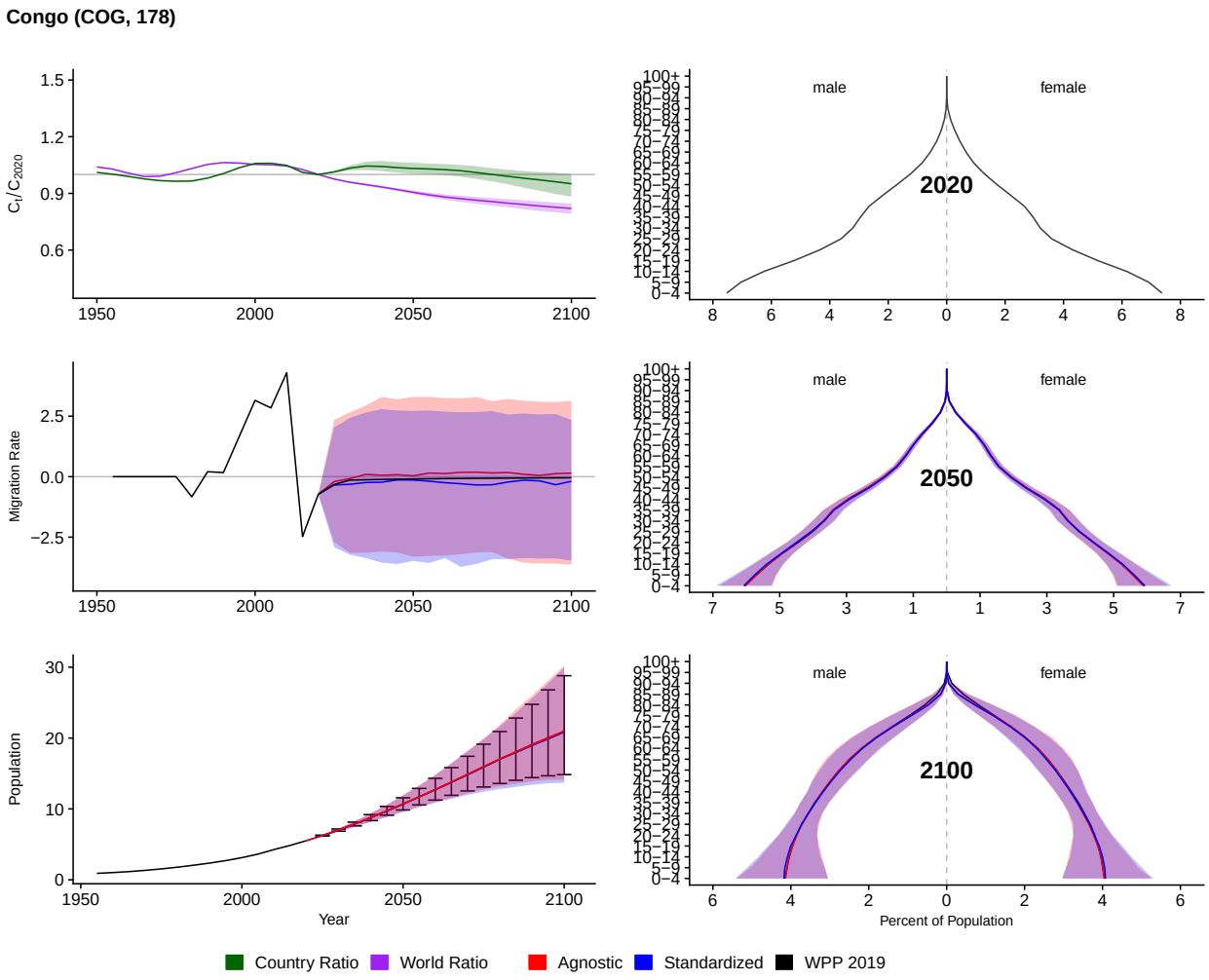


Figure A.92: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

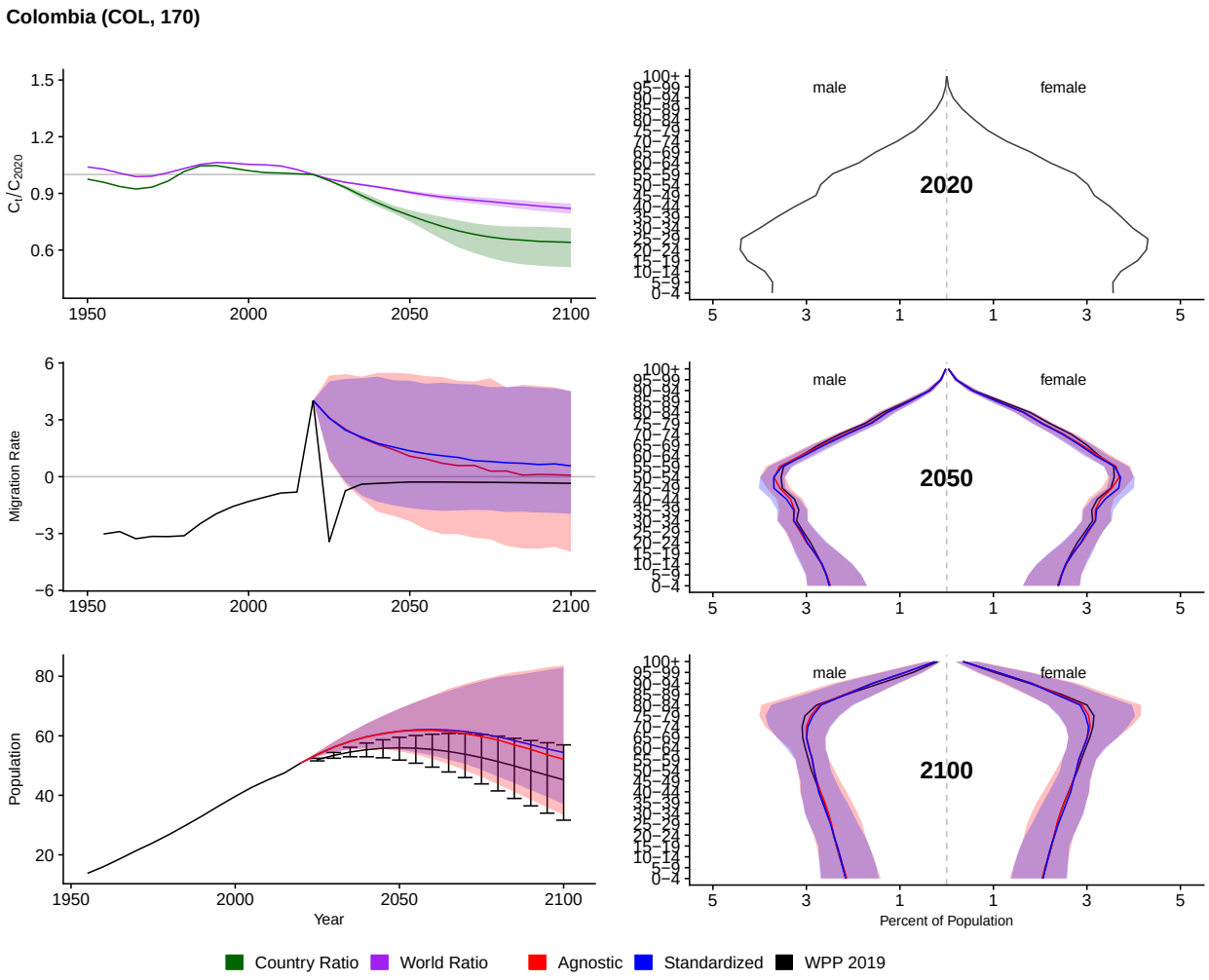


Figure A.93: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

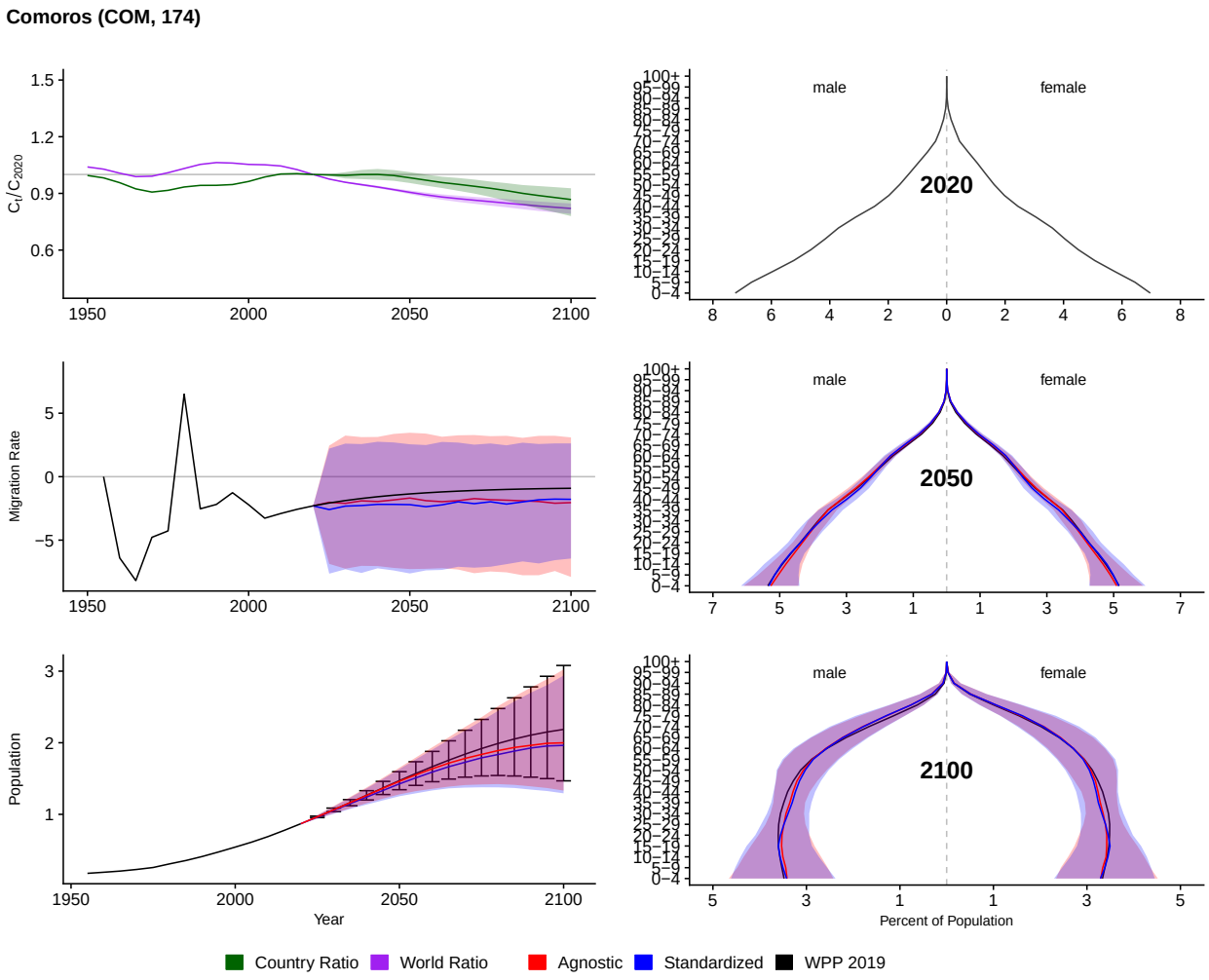


Figure A.94: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

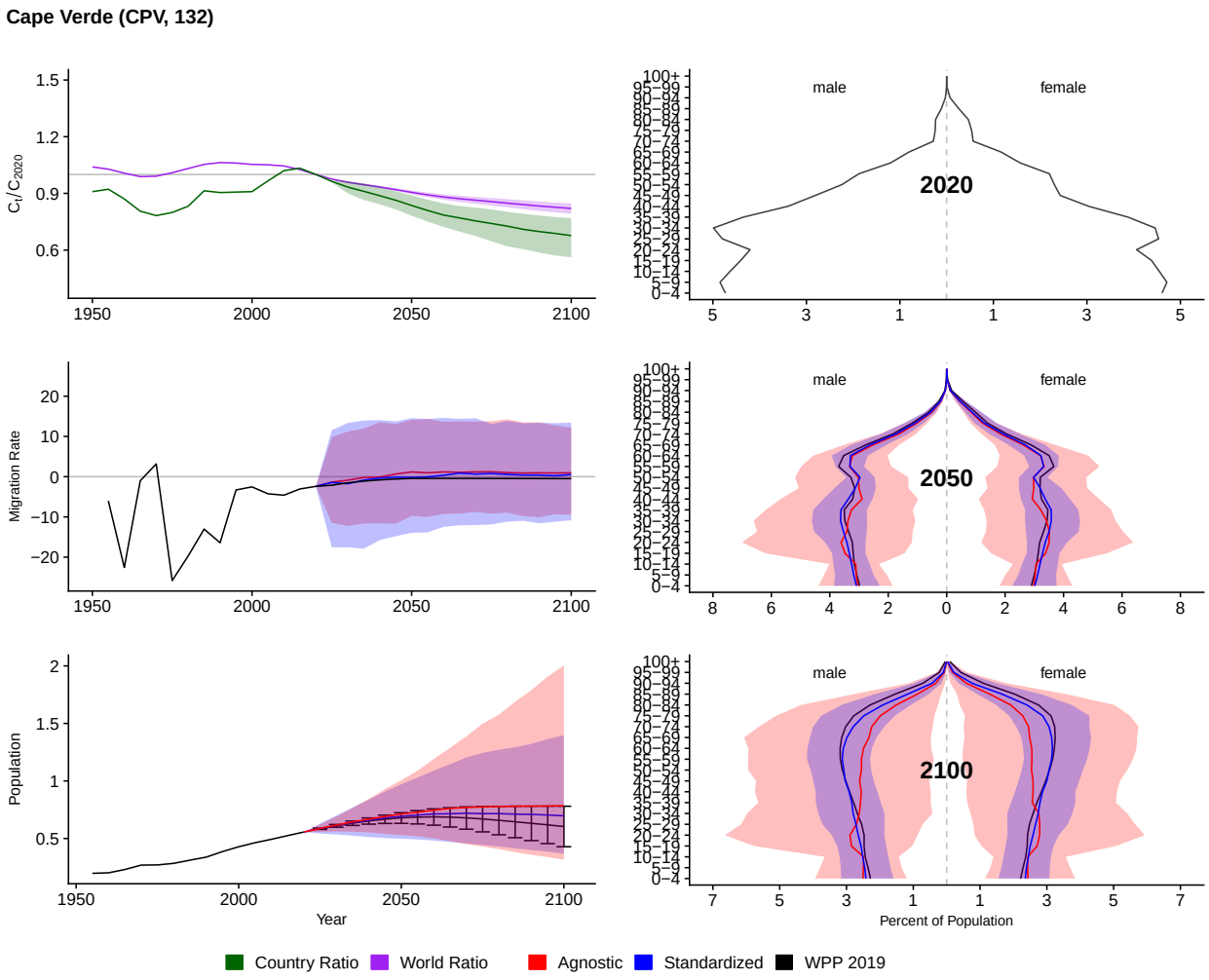


Figure A.95: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

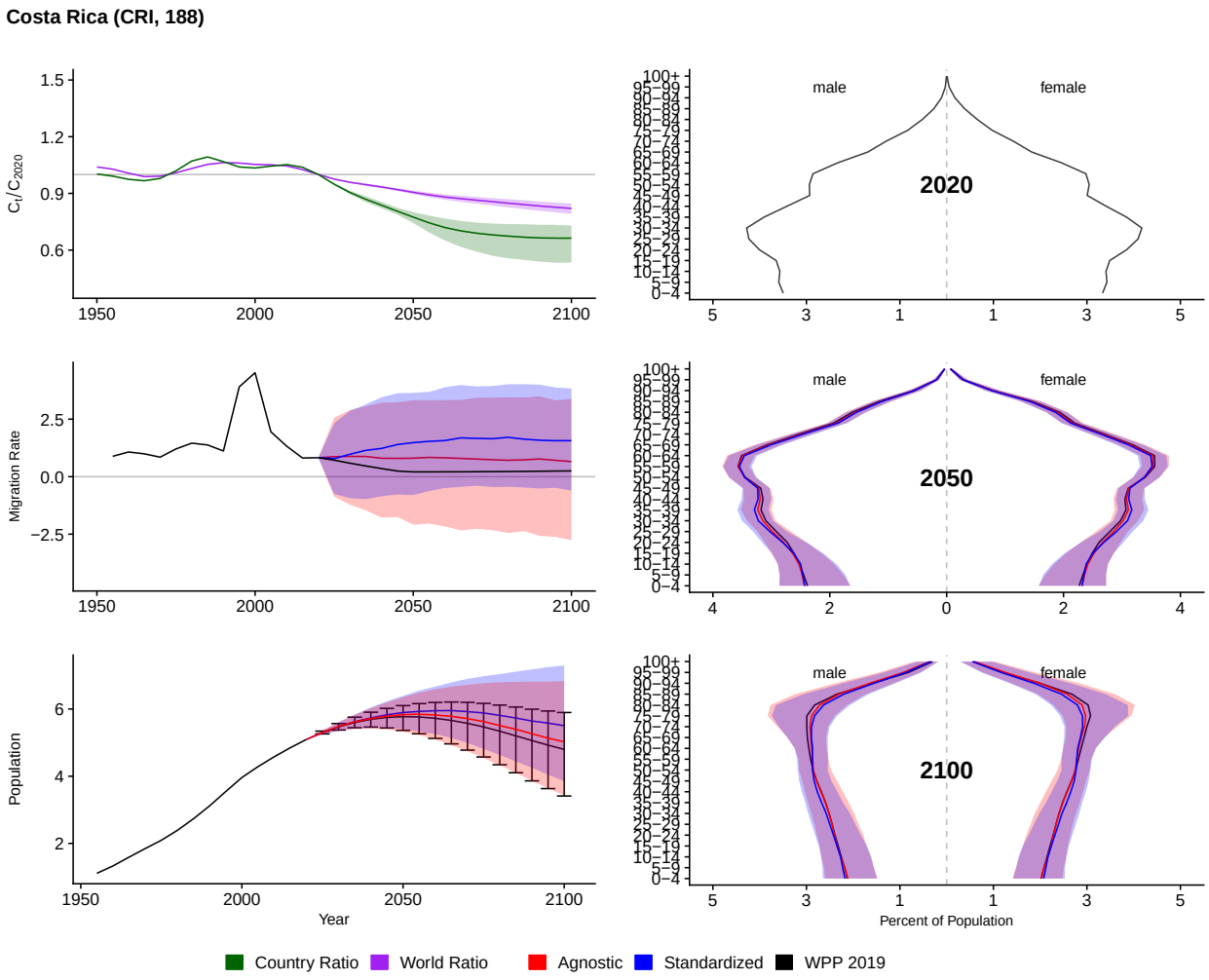


Figure A.96: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

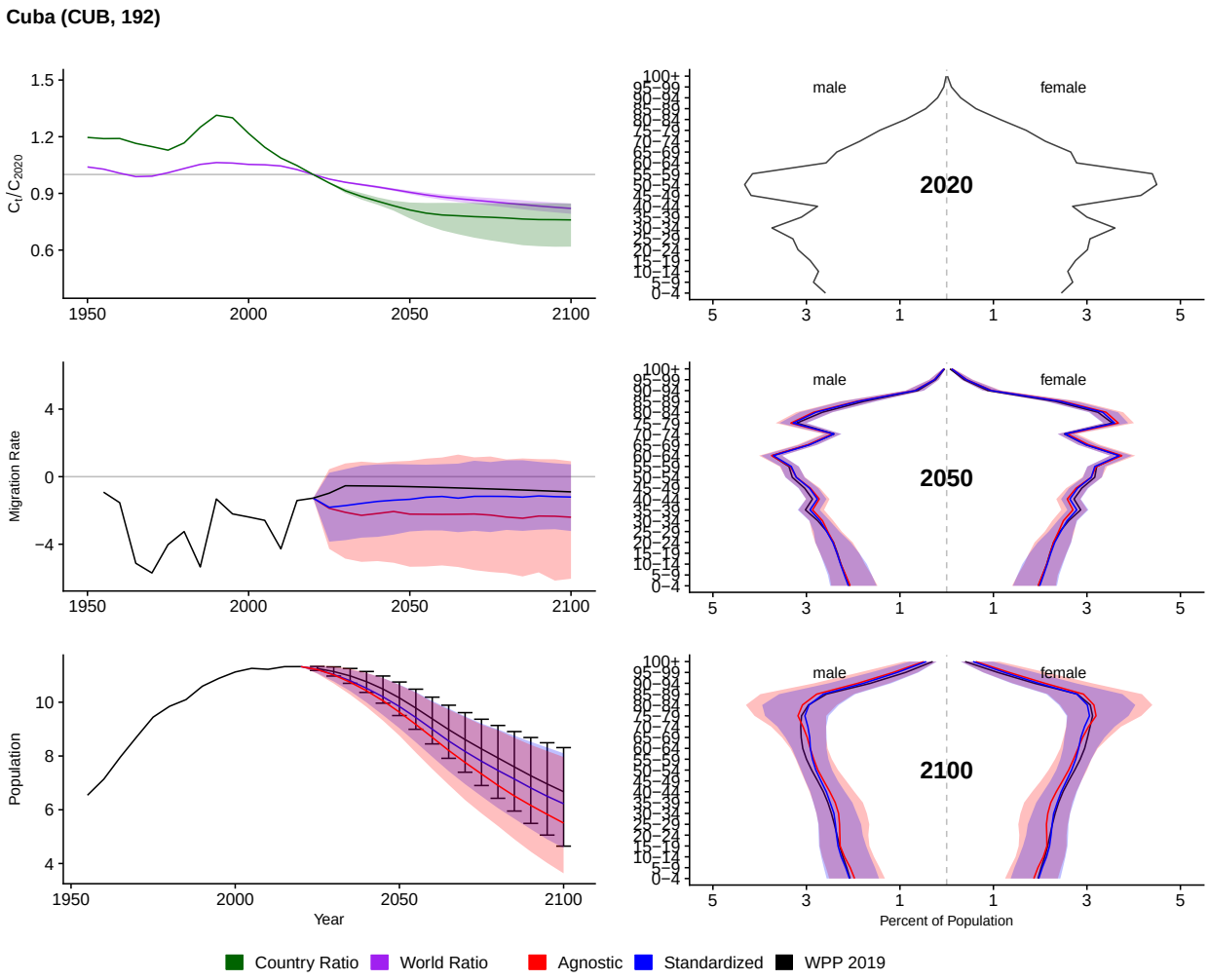


Figure A.97: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

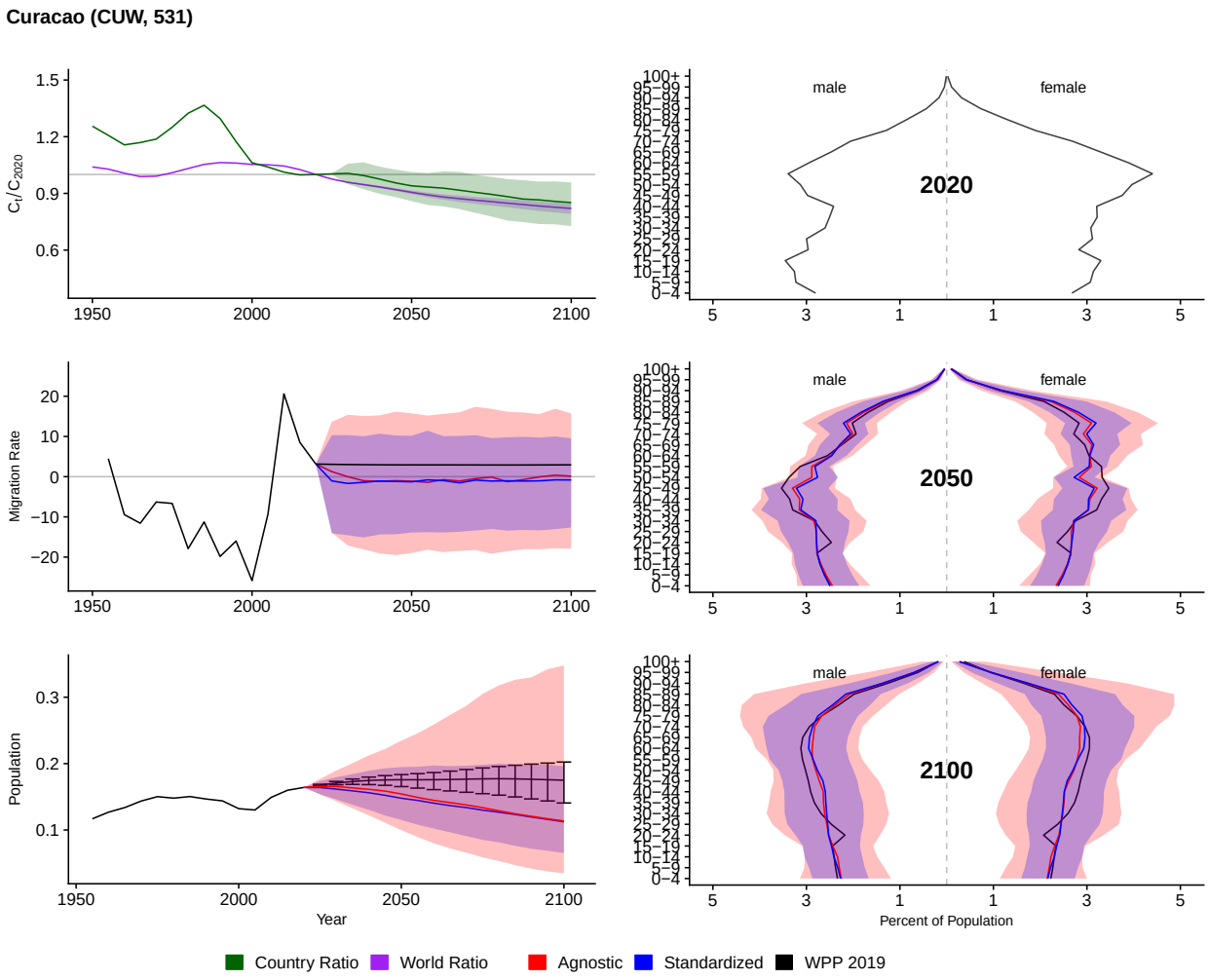


Figure A.98: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

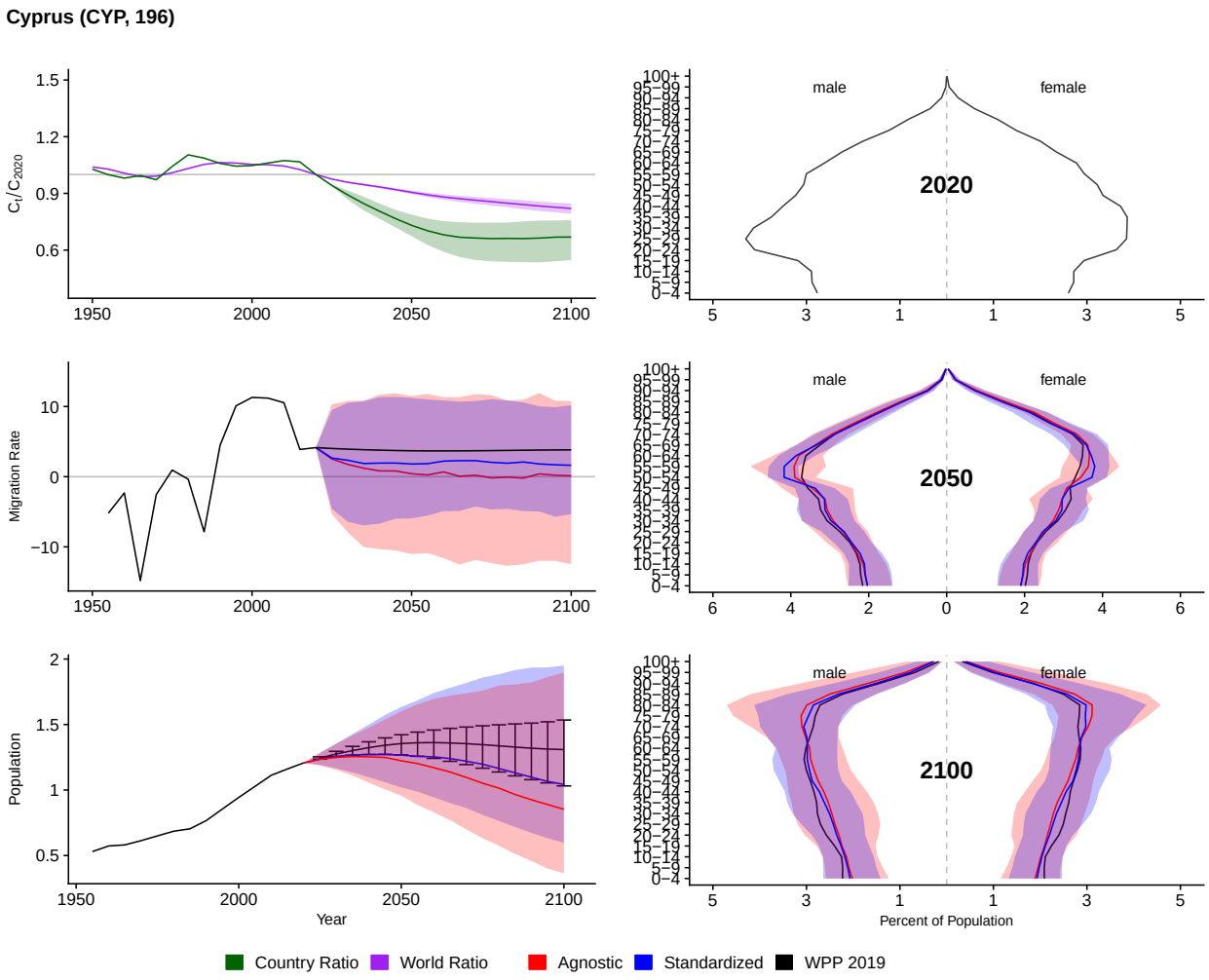


Figure A.99: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

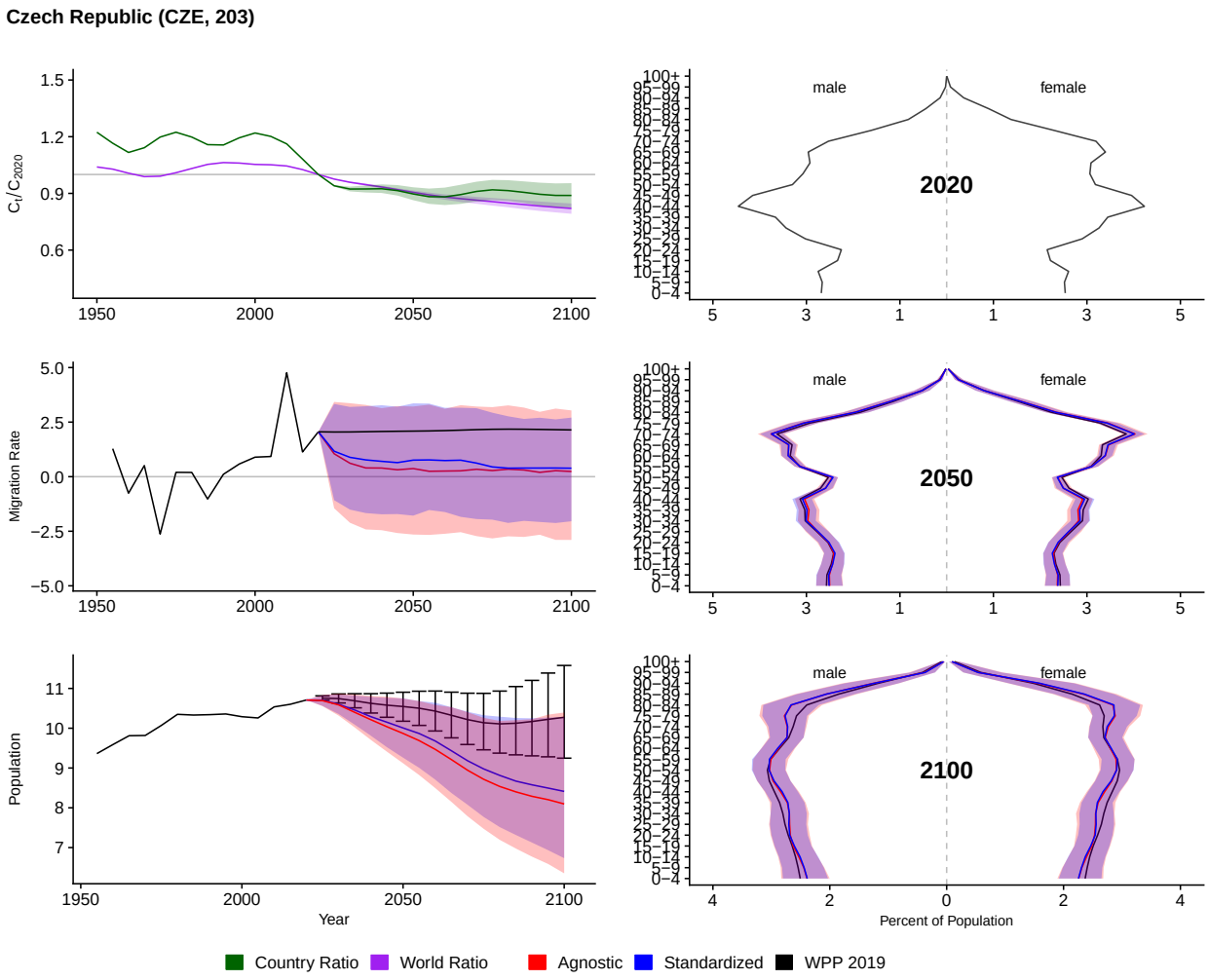


Figure A.100: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

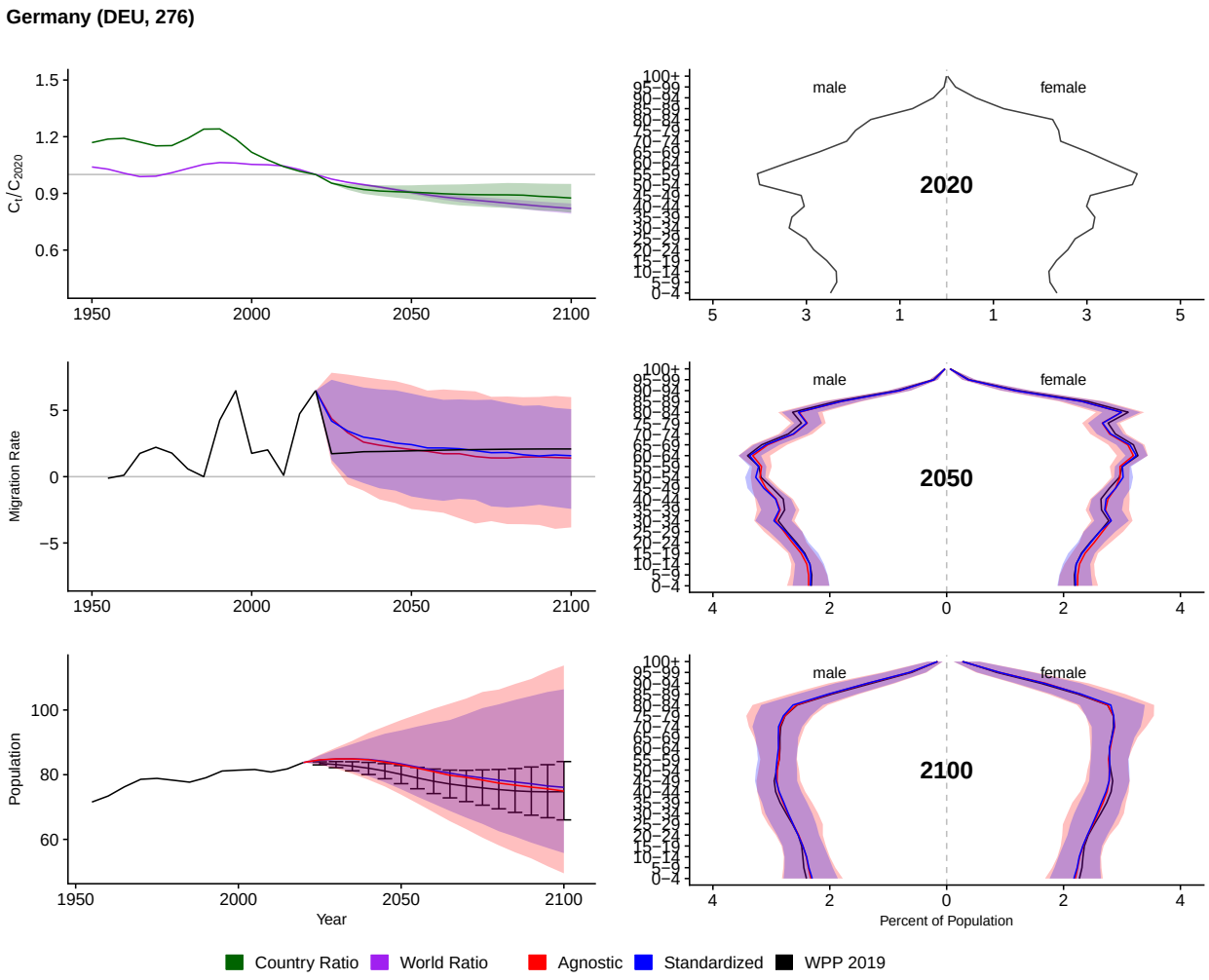


Figure A.101: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

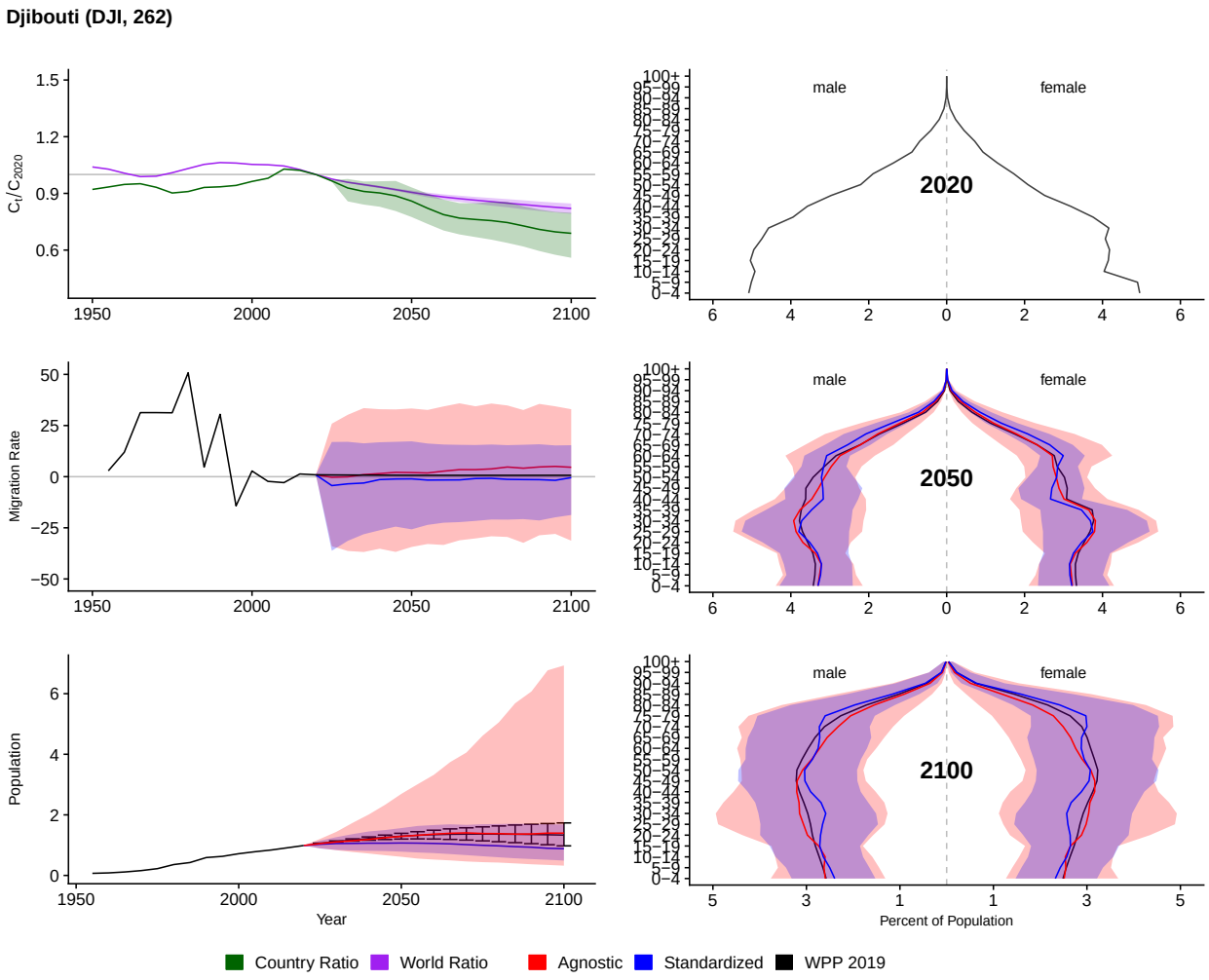


Figure A.102: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

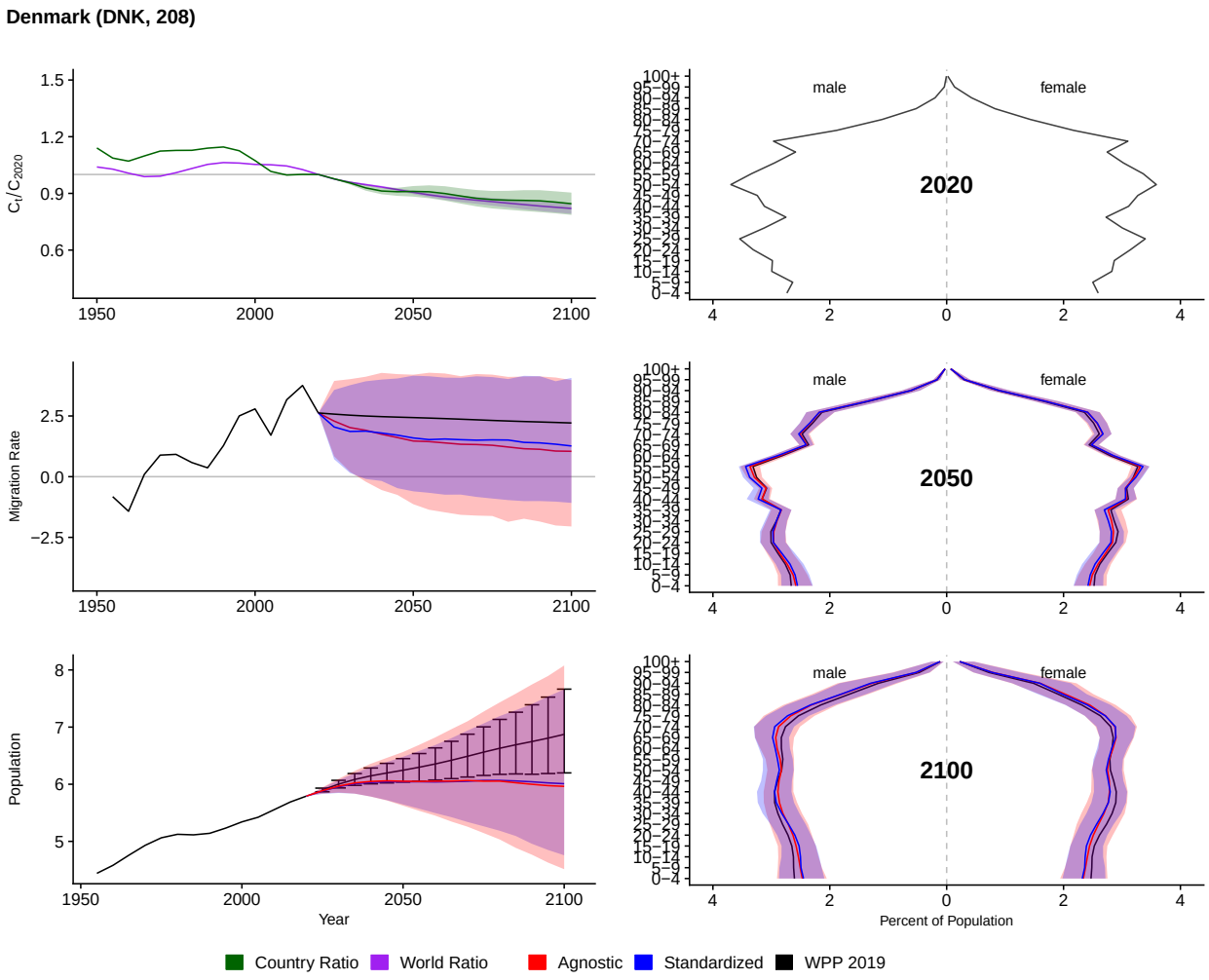


Figure A.103: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

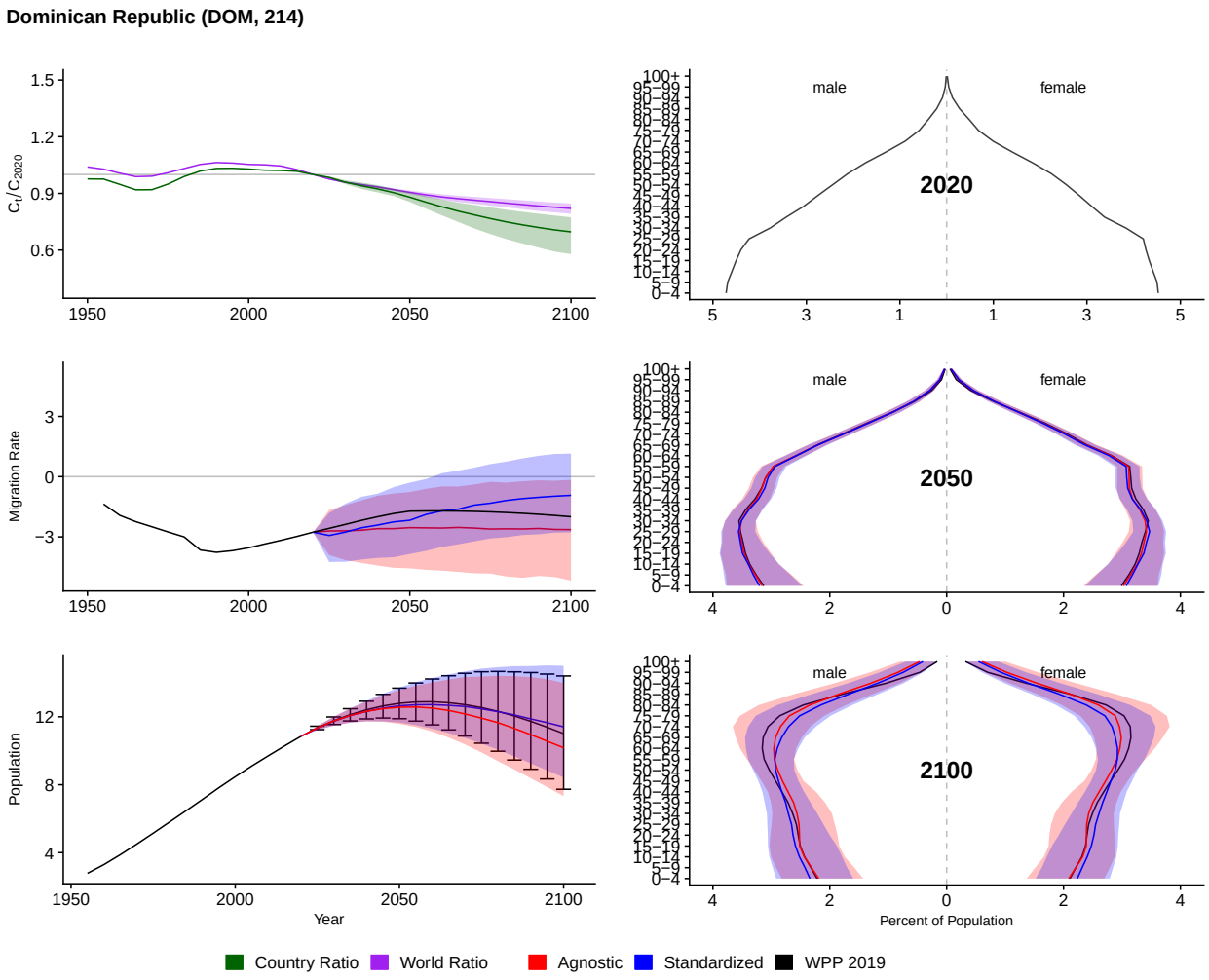


Figure A.104: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

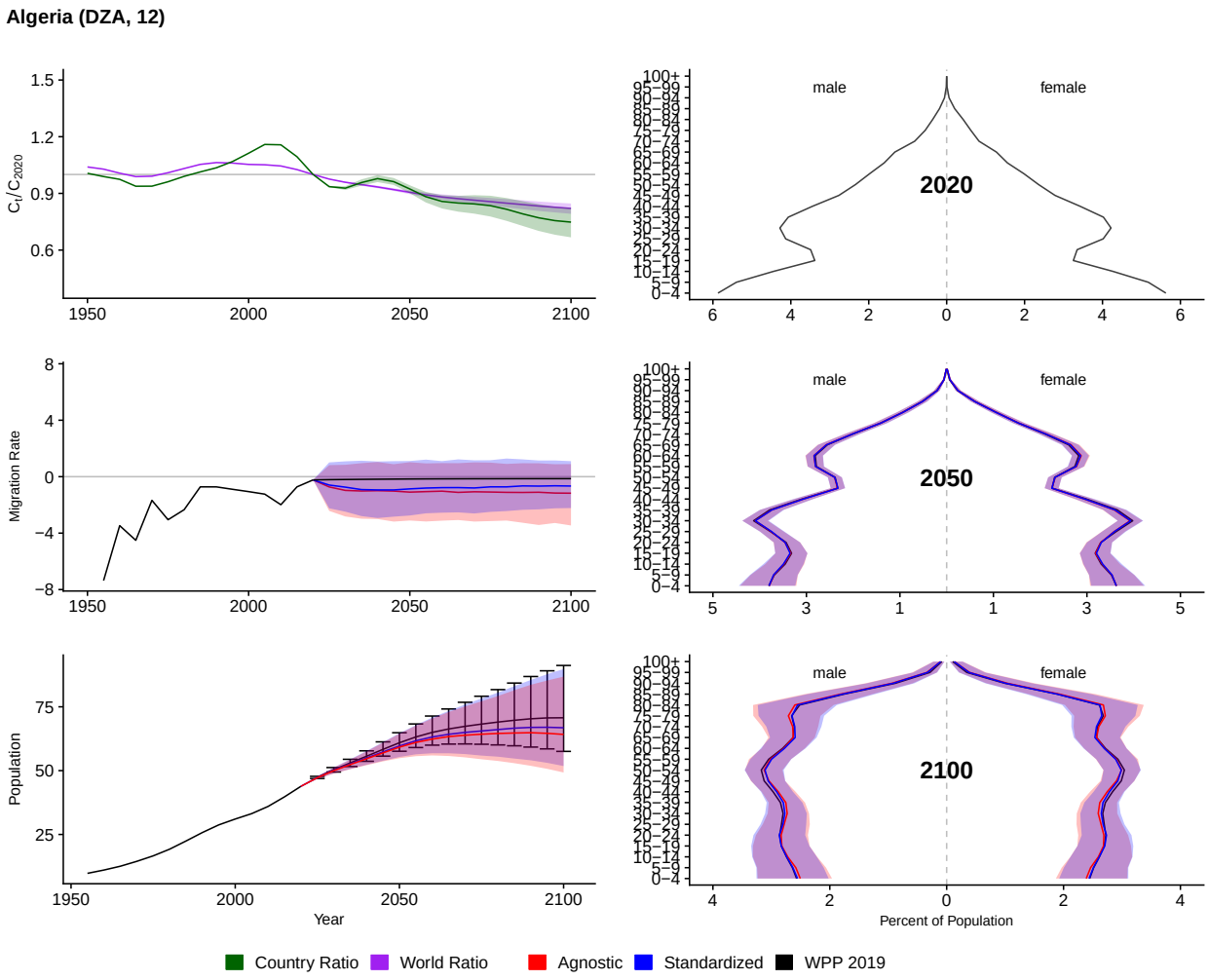


Figure A.105: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

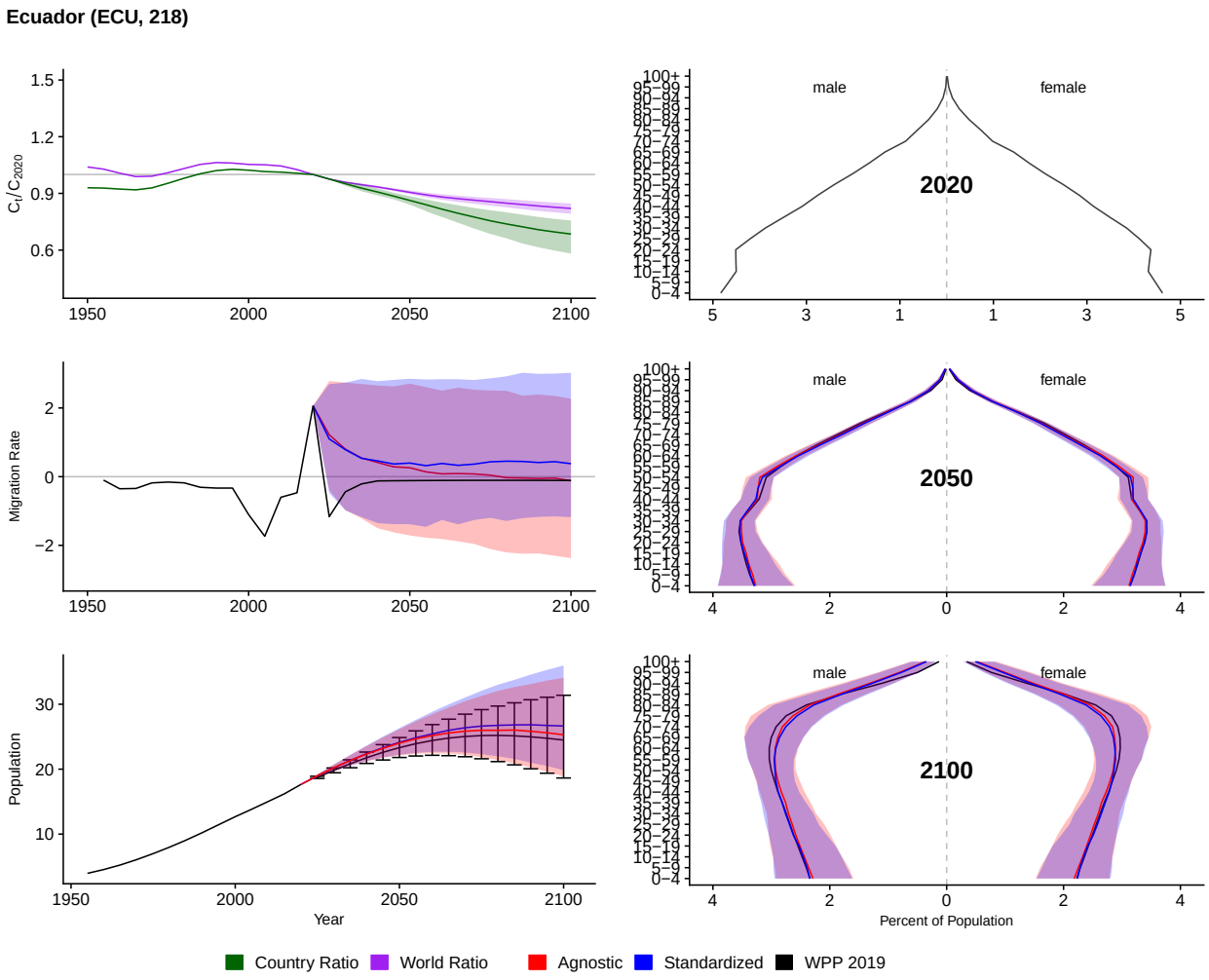


Figure A.106: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

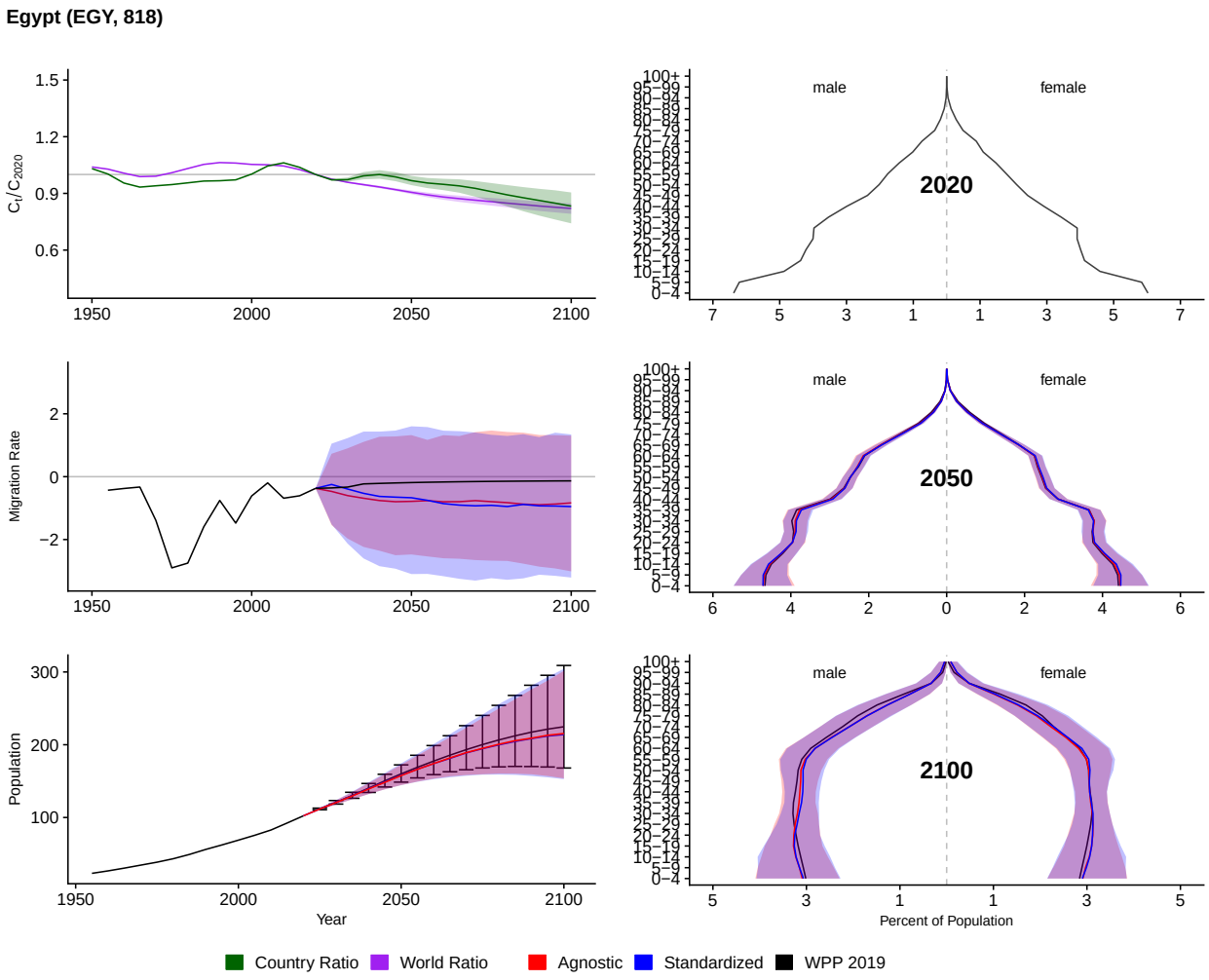


Figure A.107: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

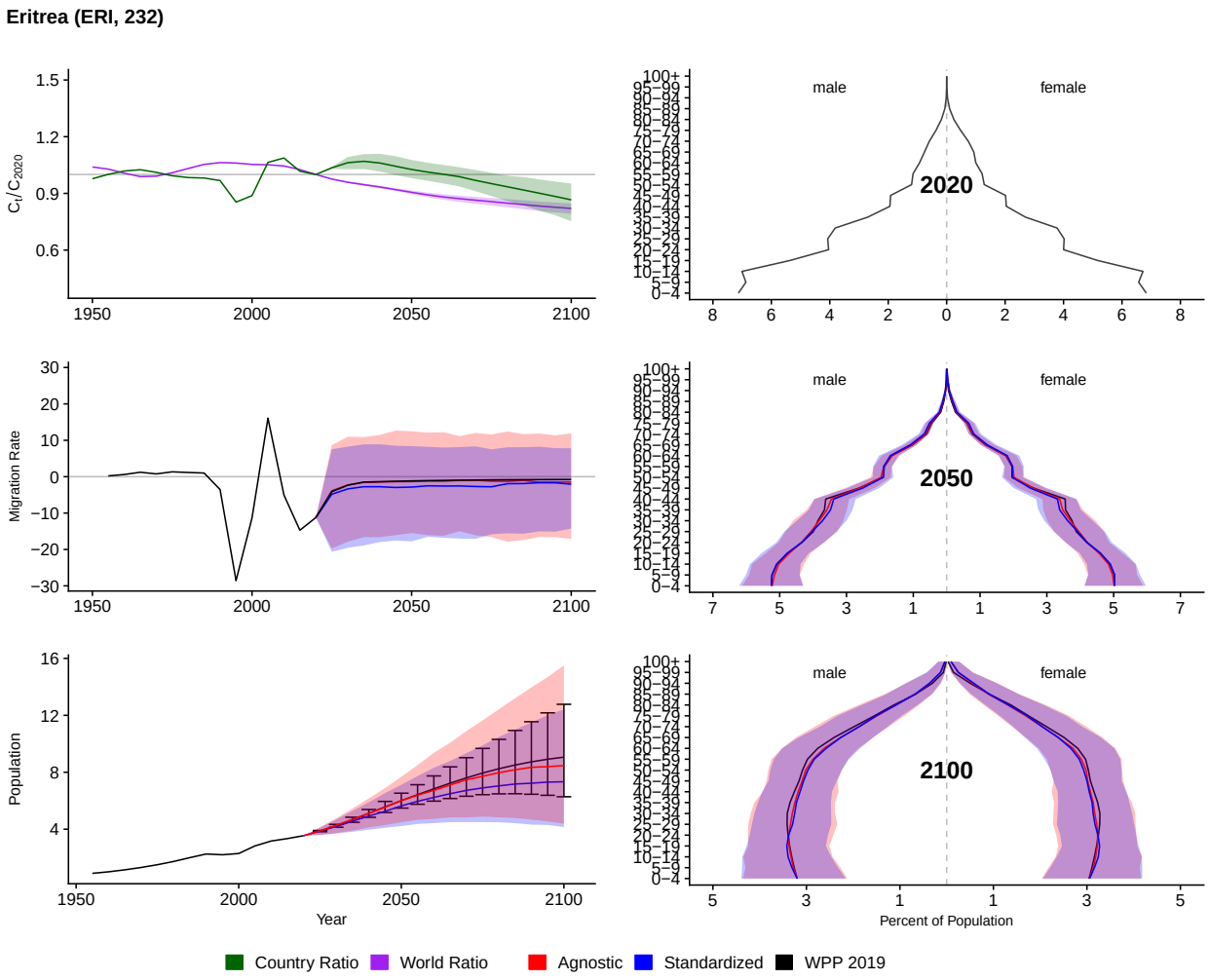


Figure A.108: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

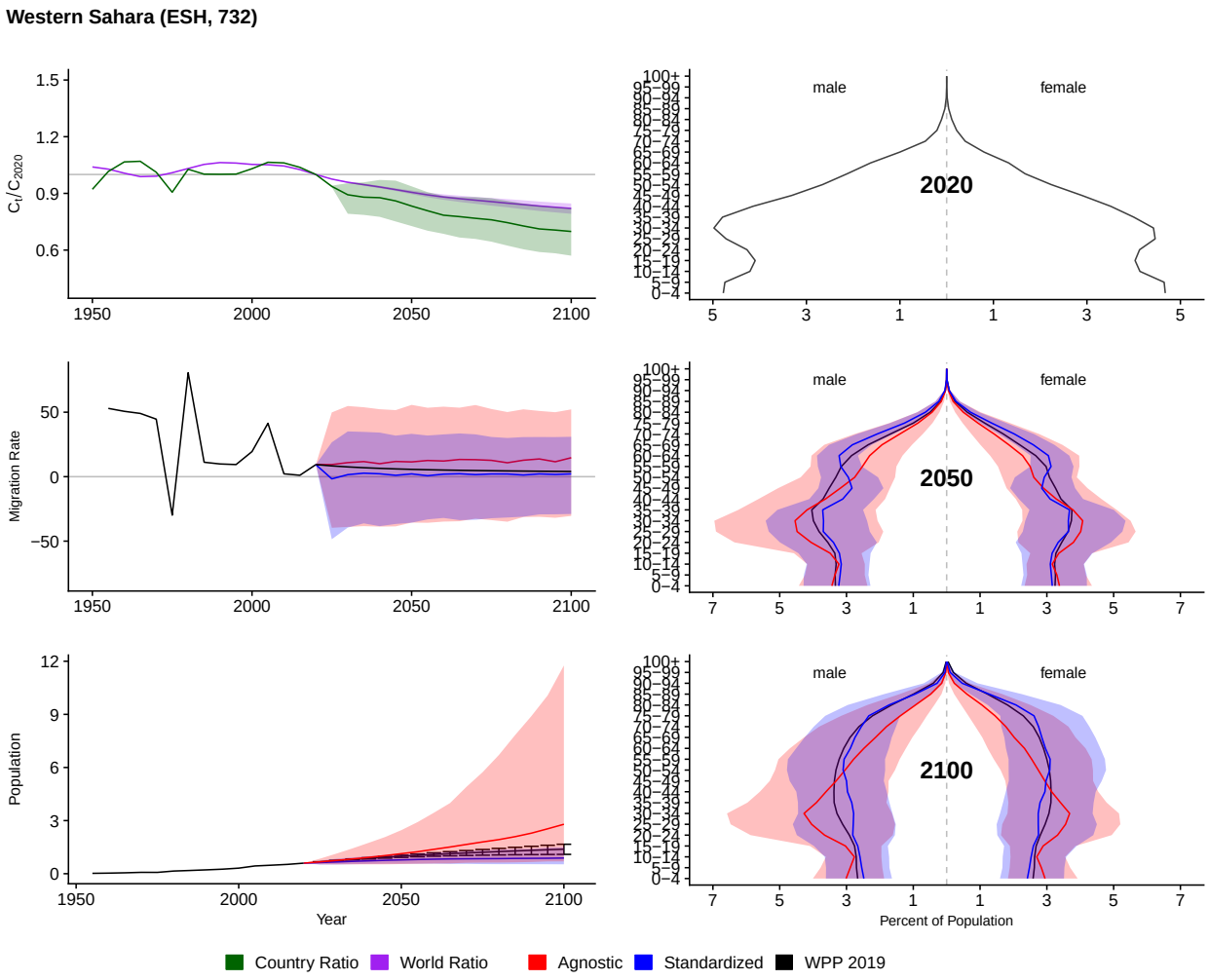


Figure A.109: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

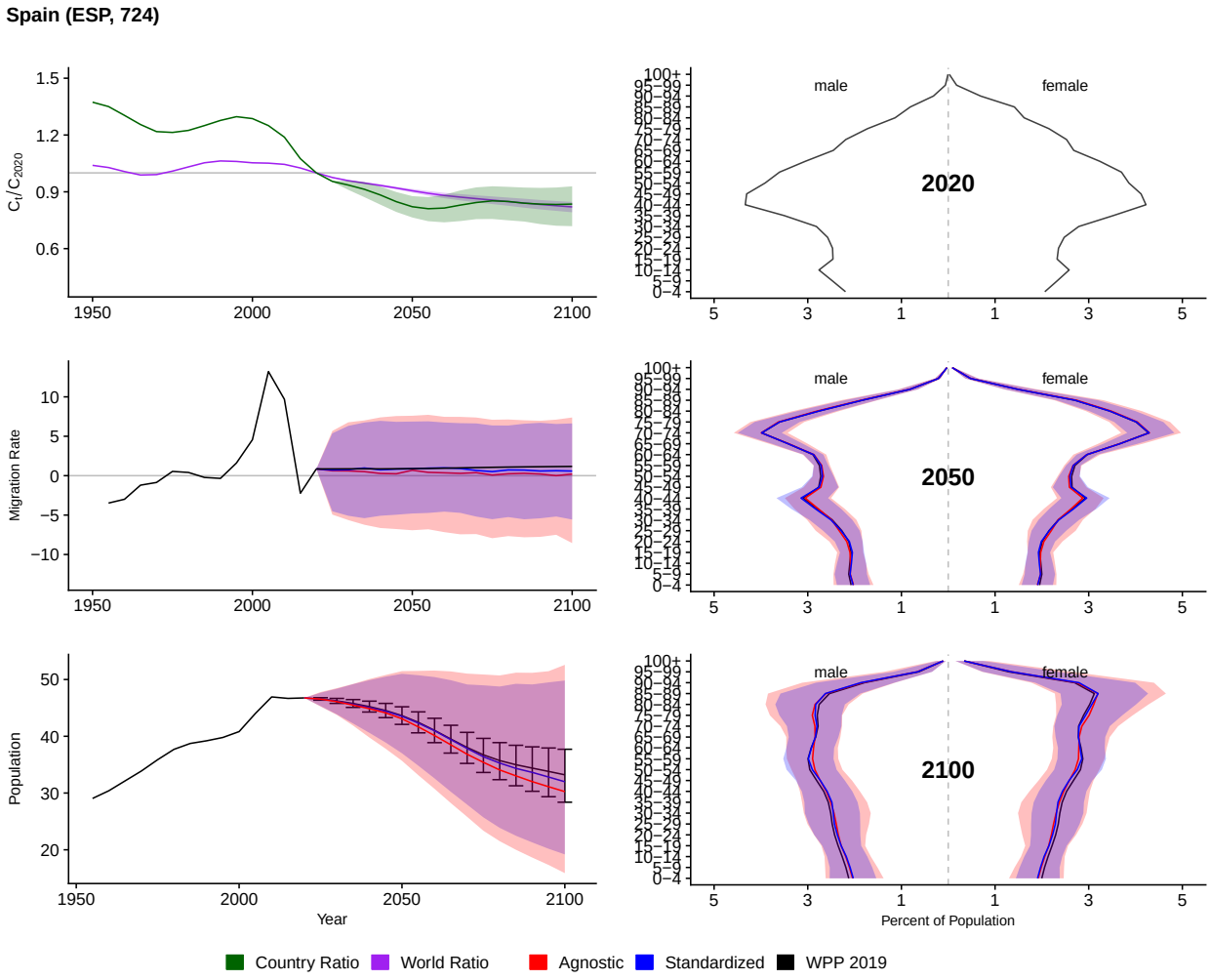


Figure A.110: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

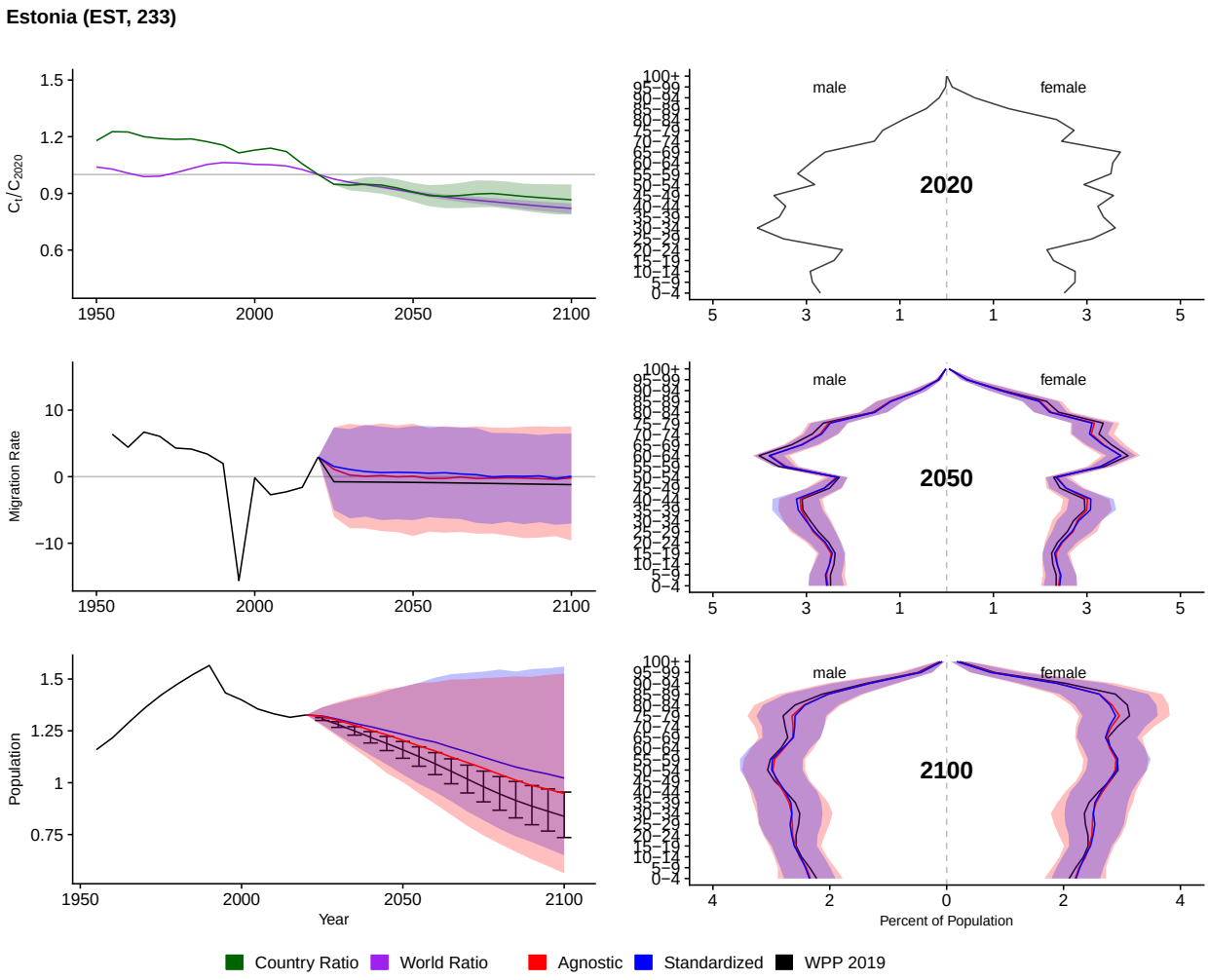


Figure A.111: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

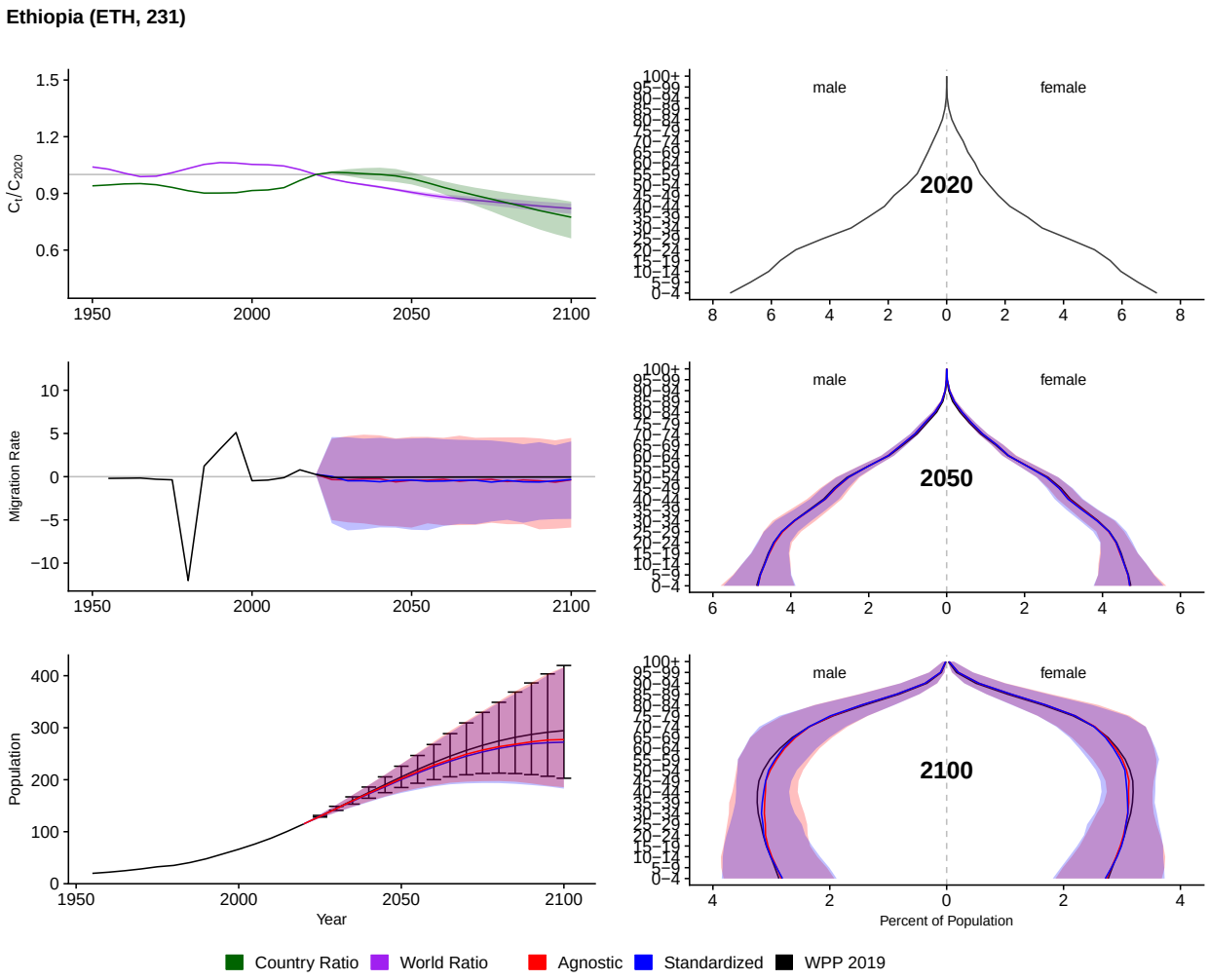


Figure A.112: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

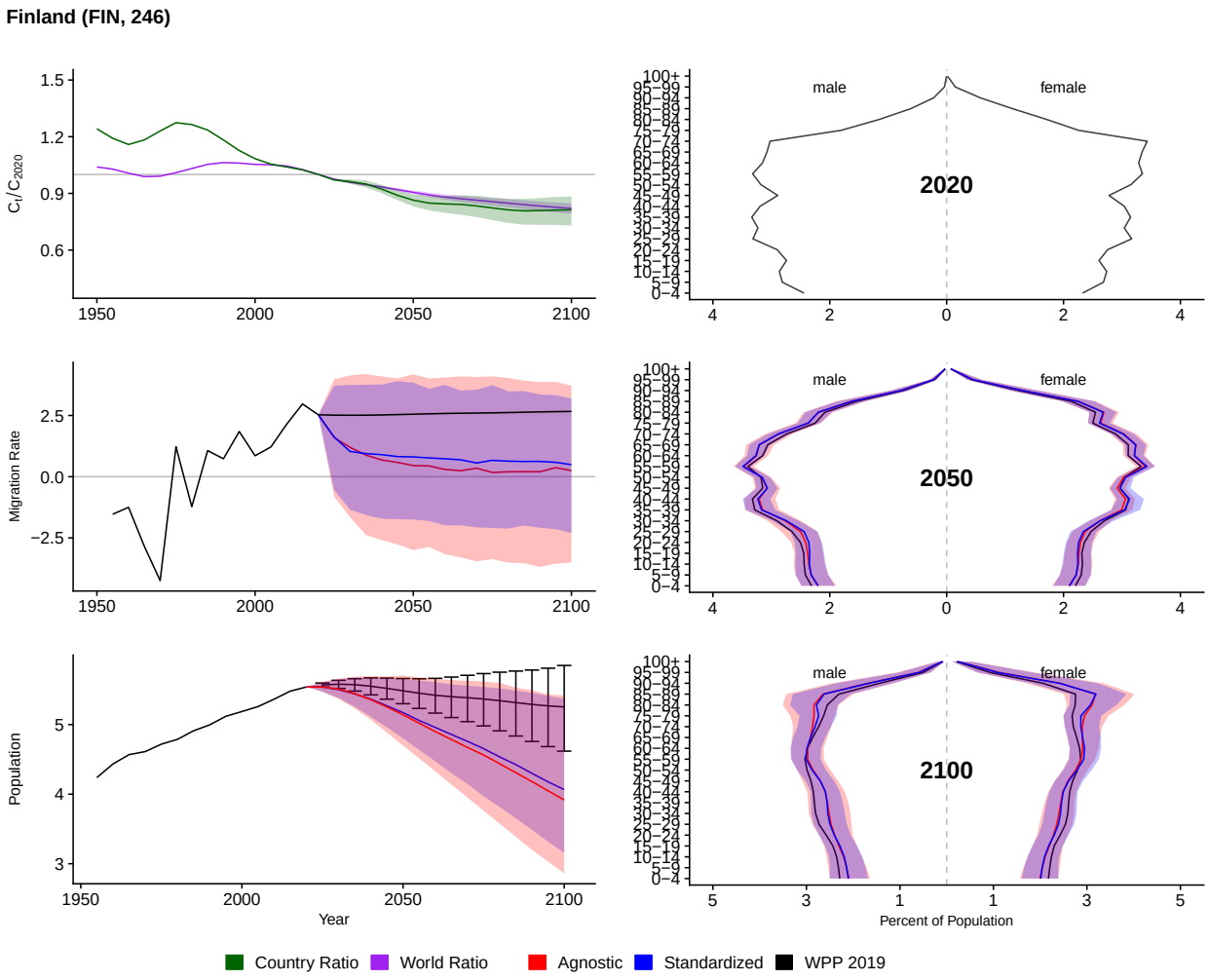


Figure A.113: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

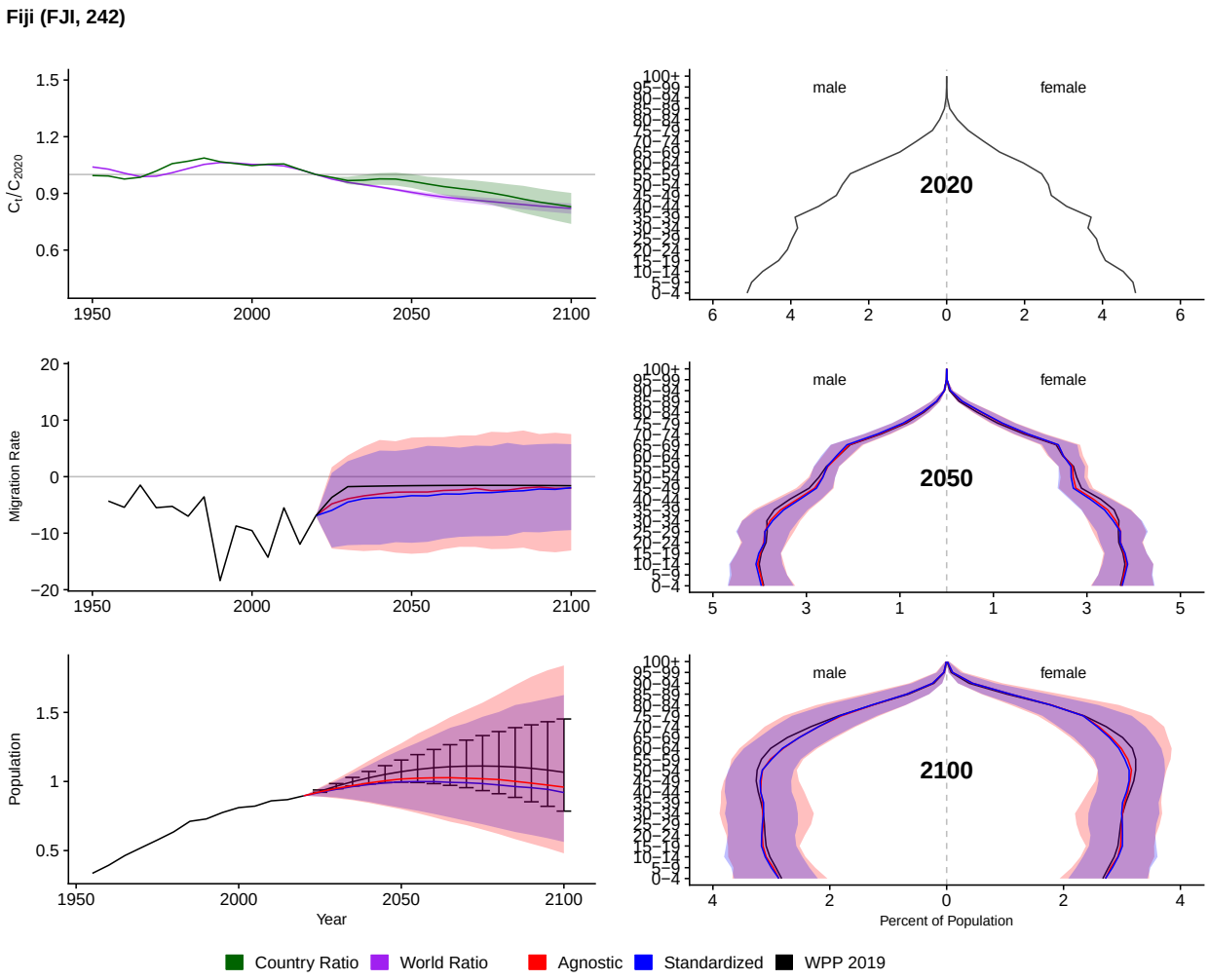


Figure A.114: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

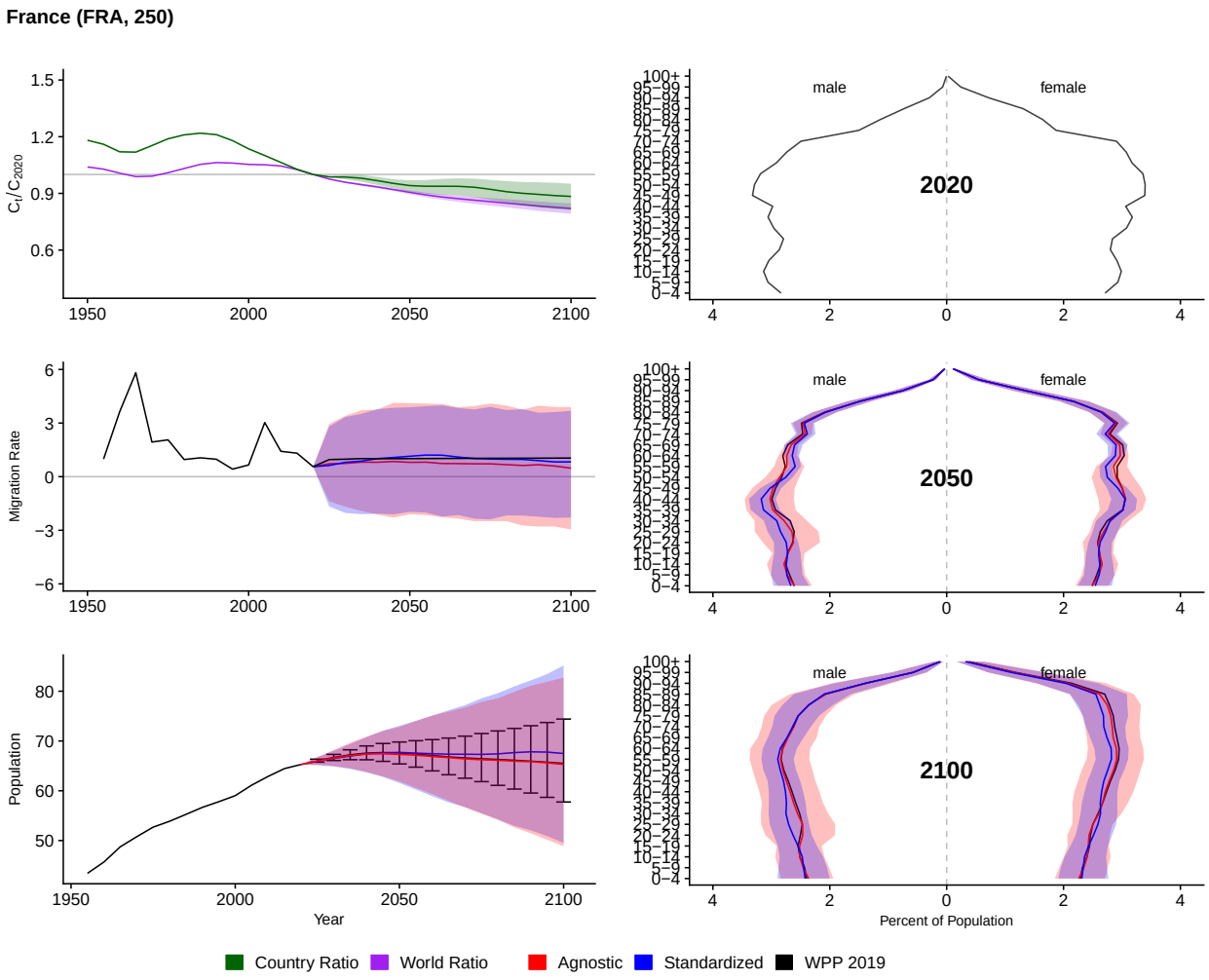


Figure A.115: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

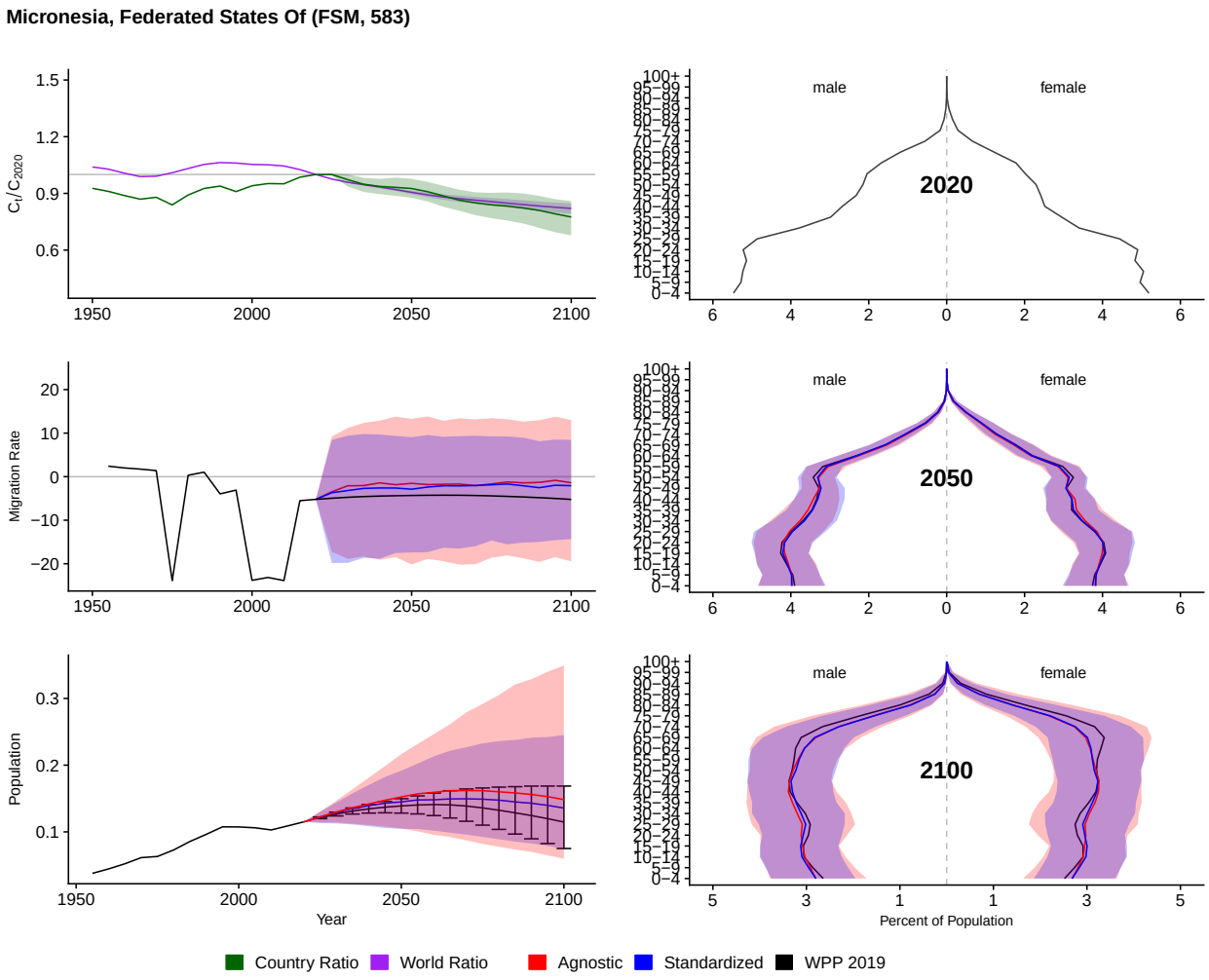


Figure A.116: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

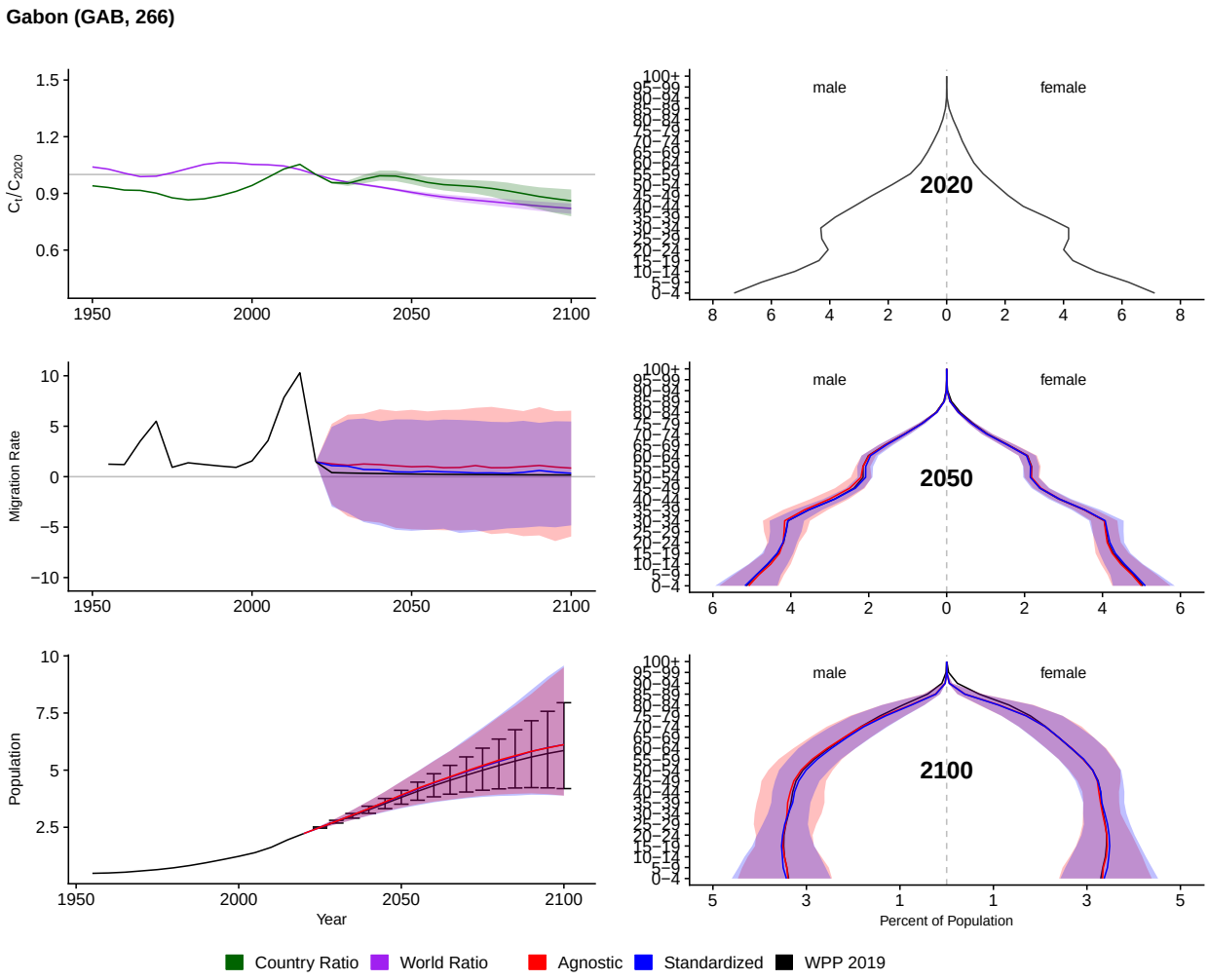


Figure A.117: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

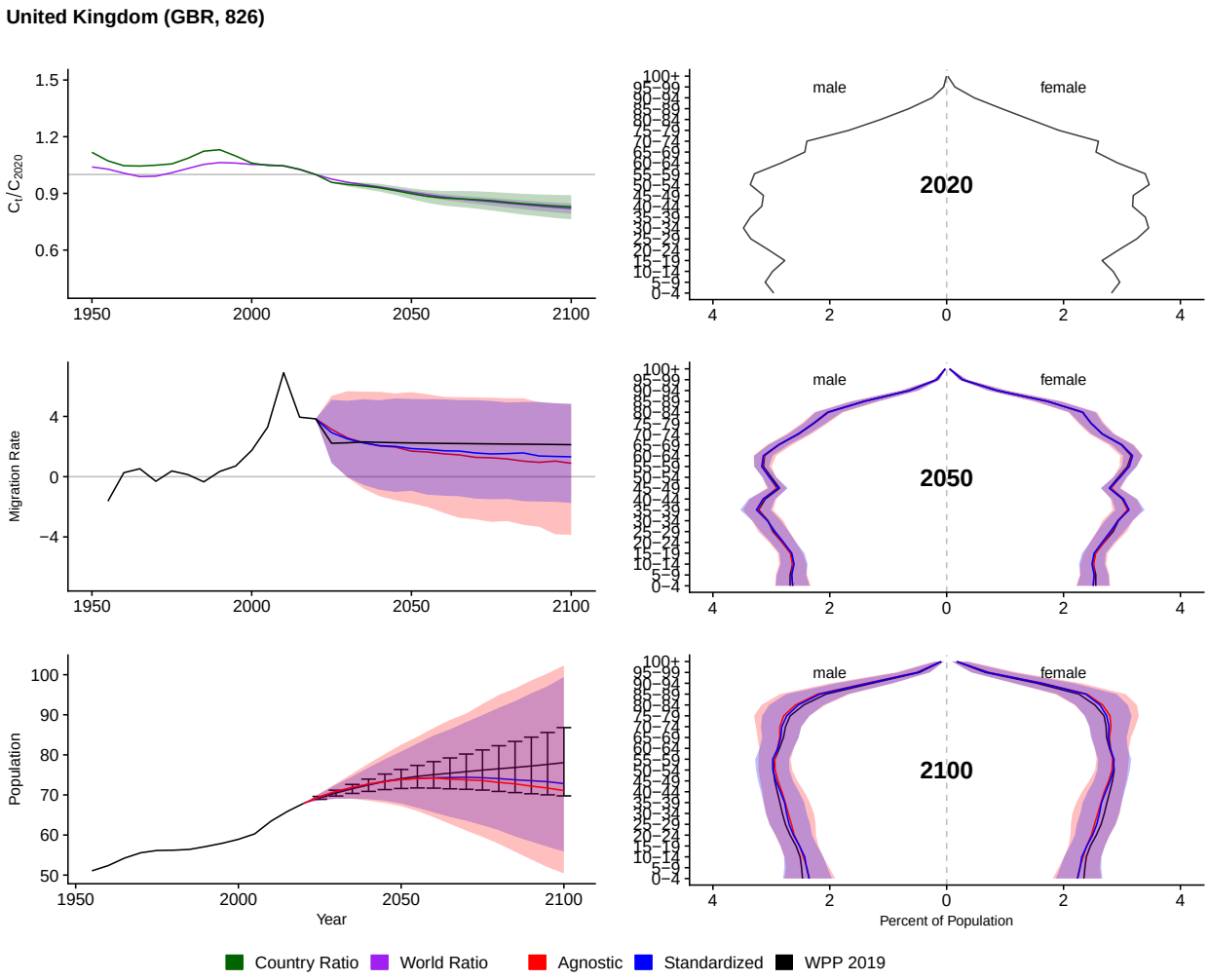


Figure A.118: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

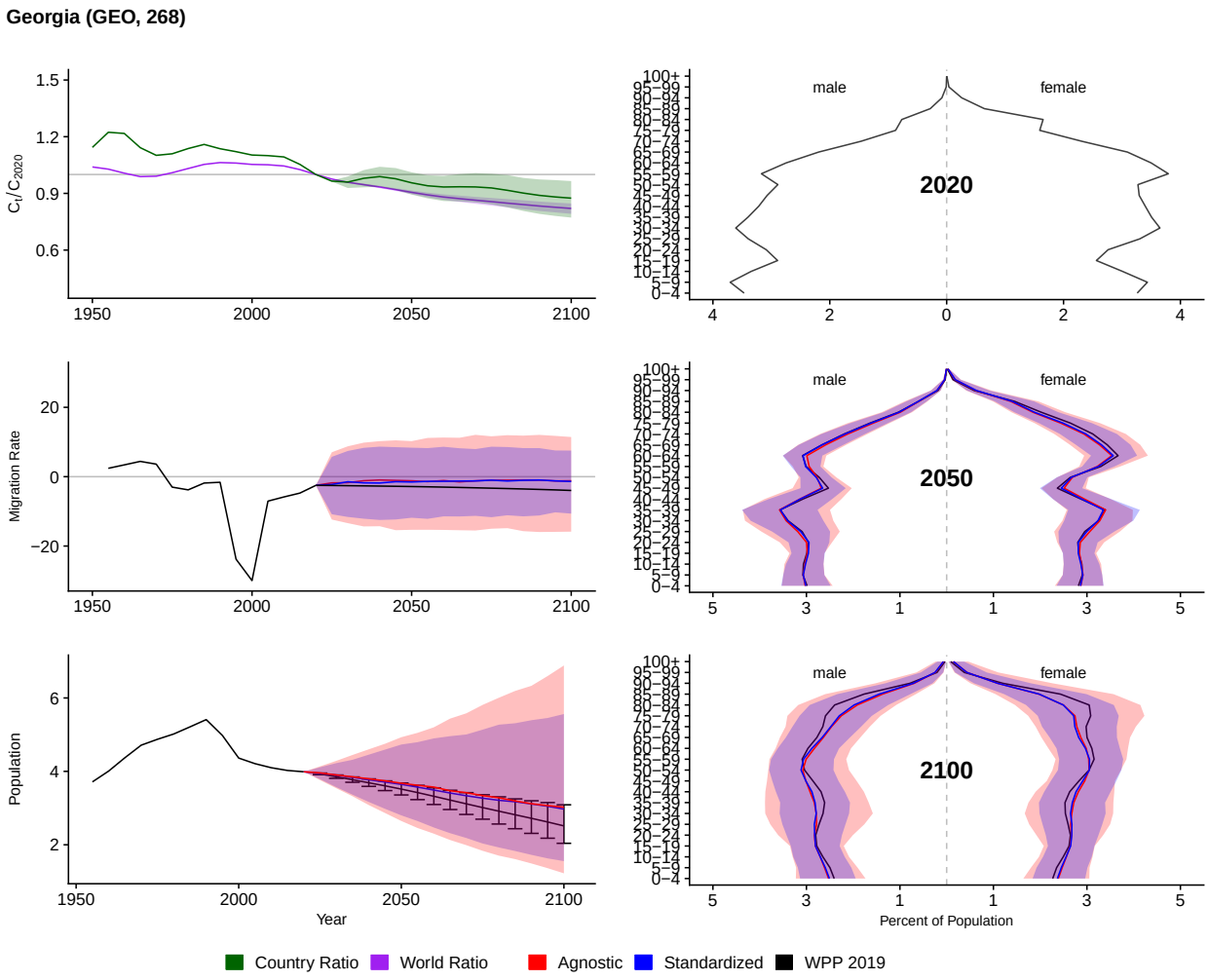


Figure A.119: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

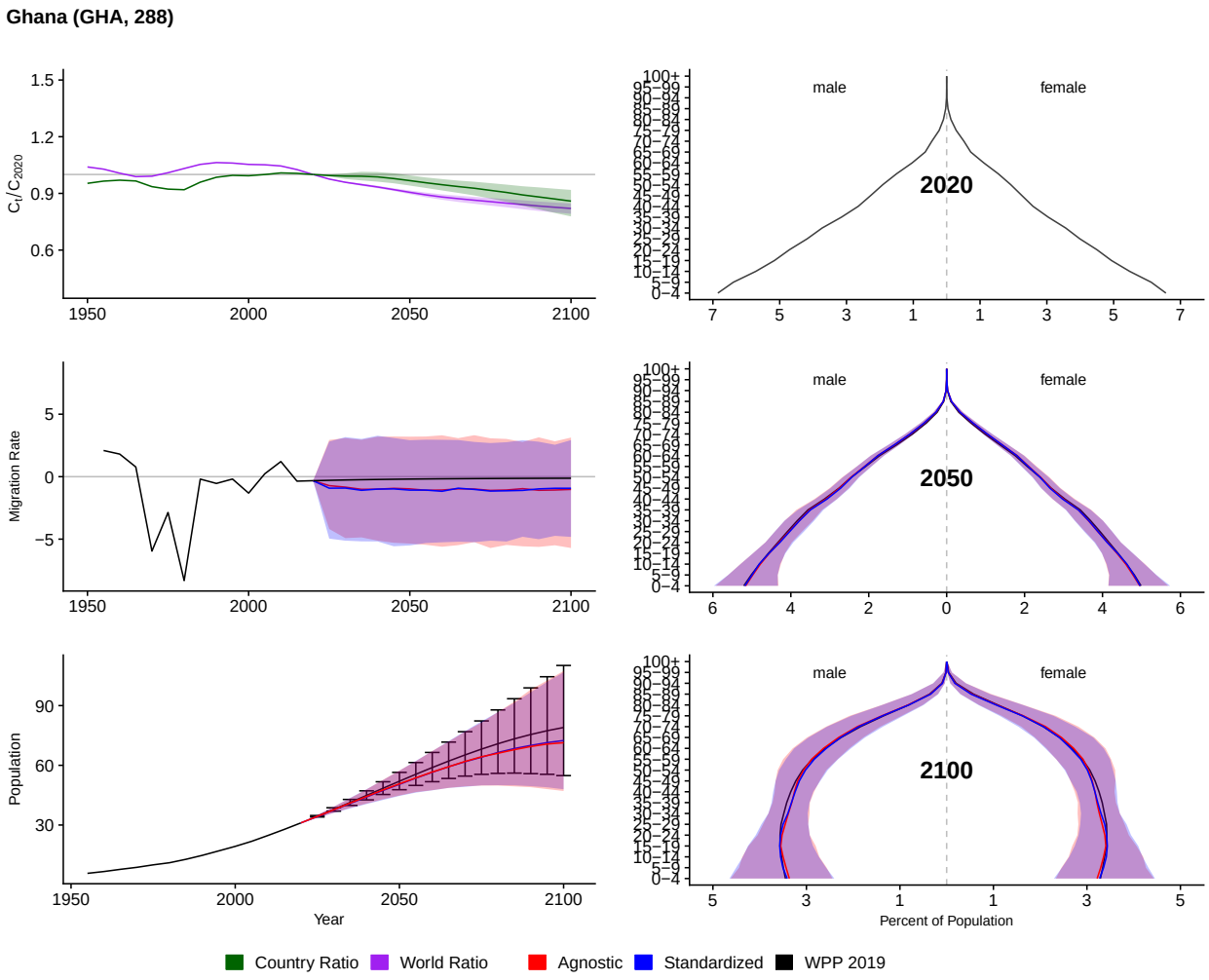


Figure A.120: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

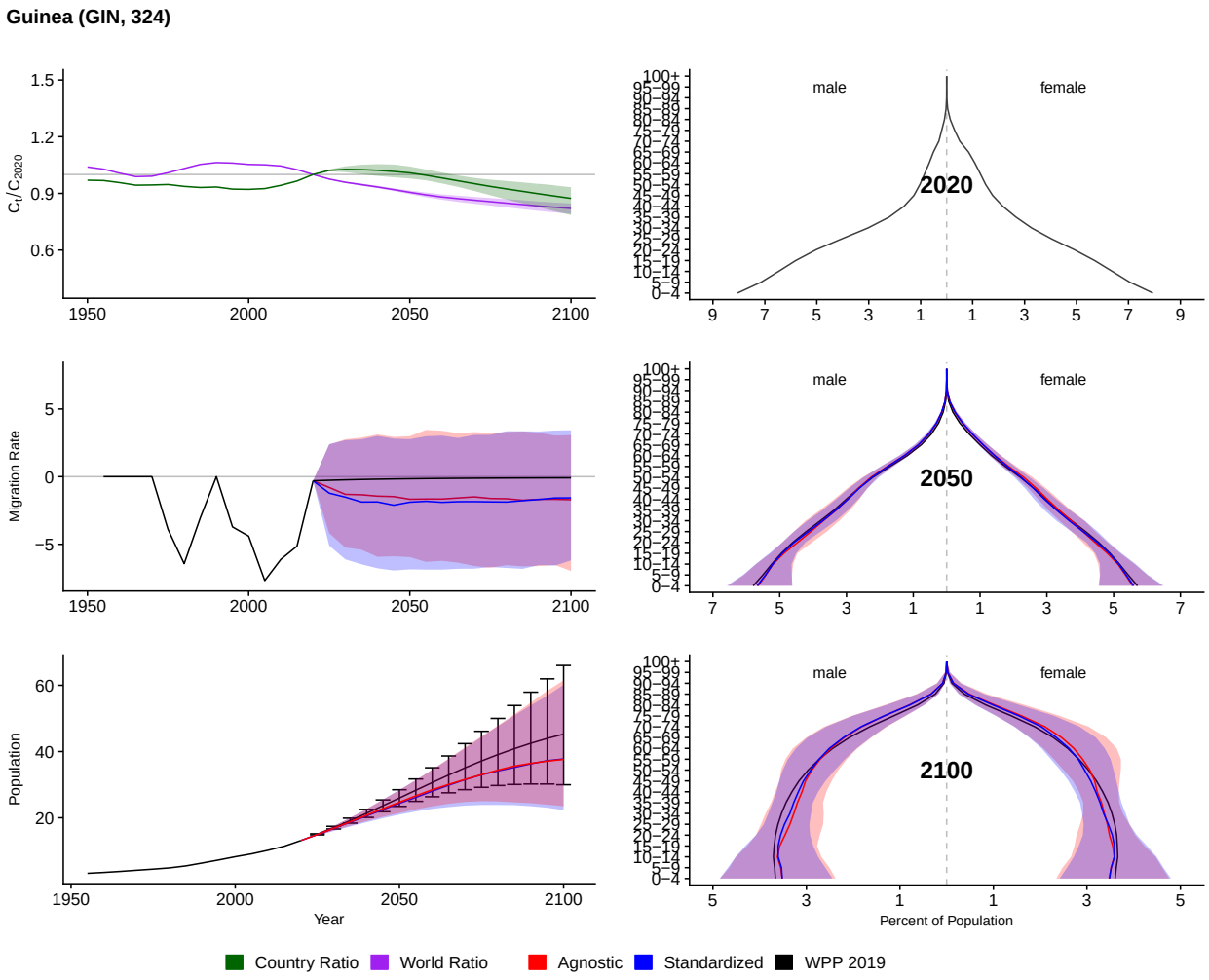


Figure A.121: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

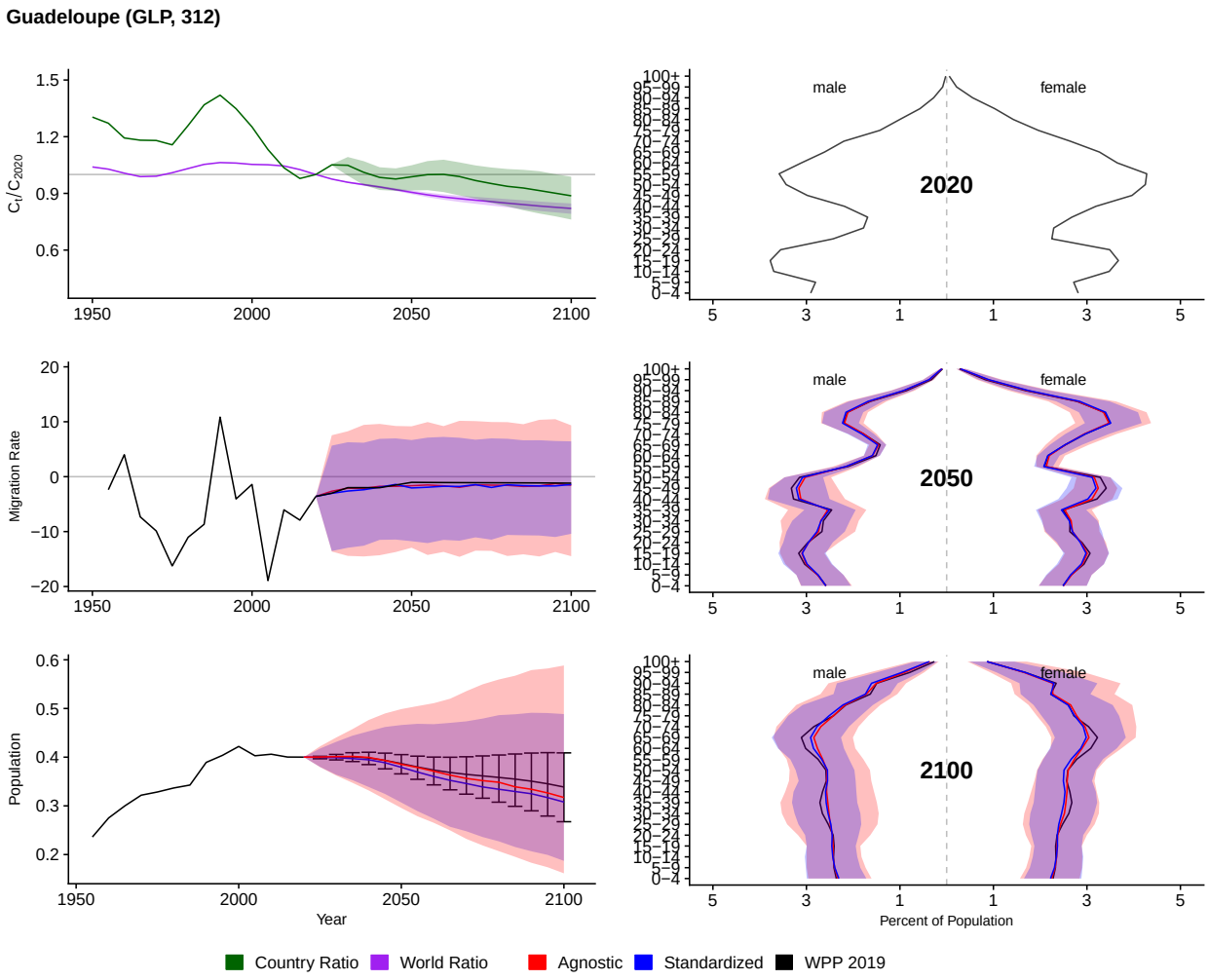


Figure A.122: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

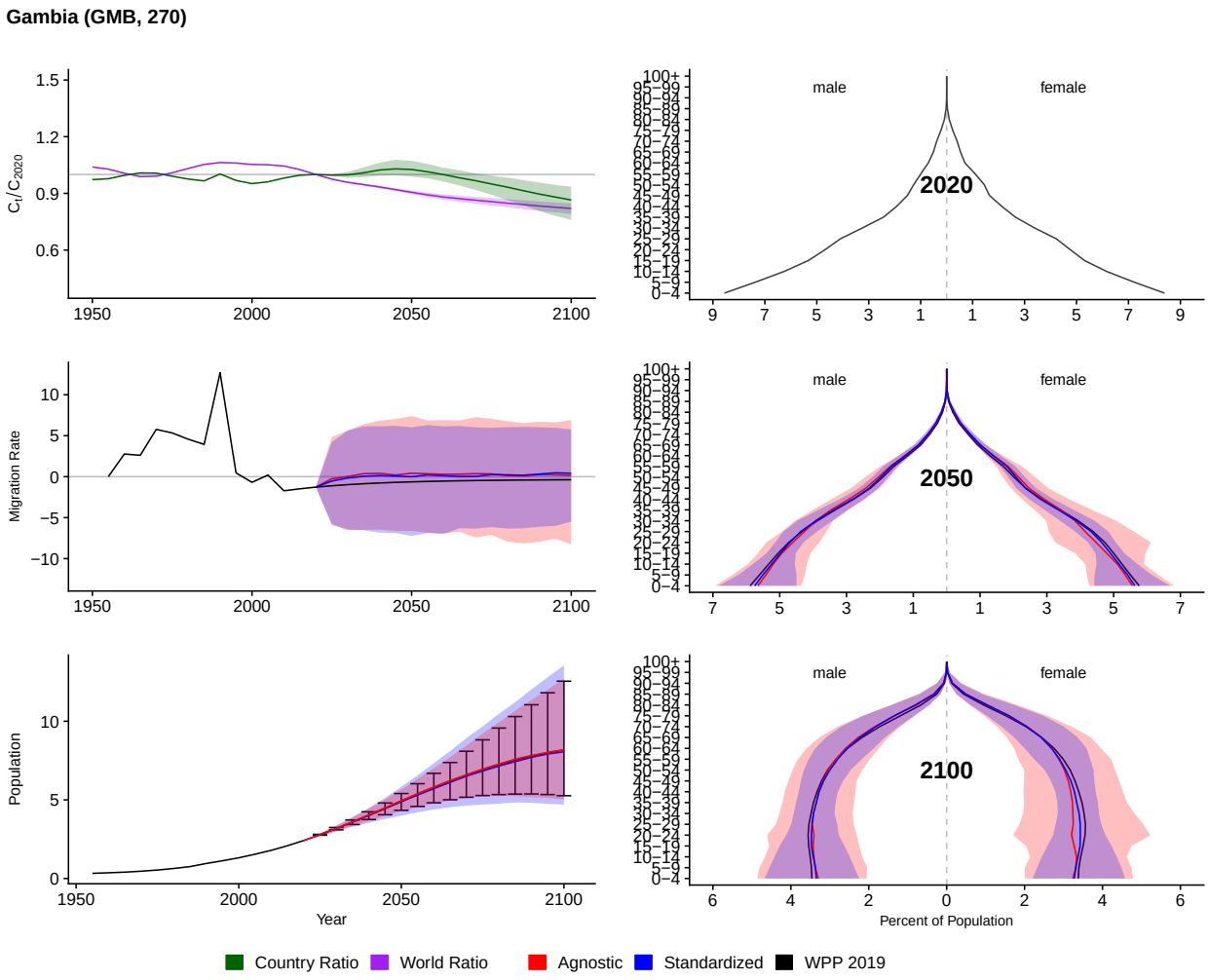


Figure A.123: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

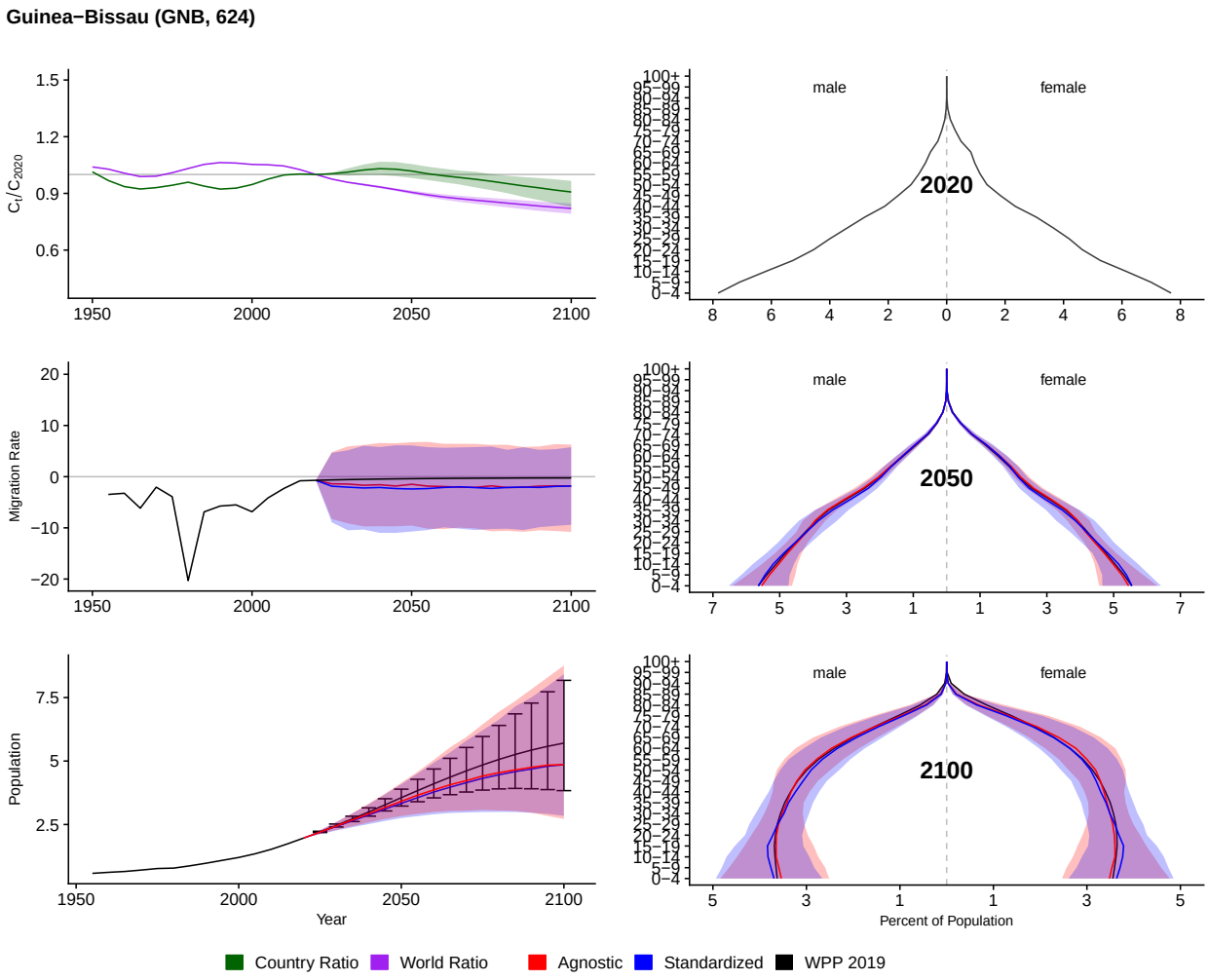


Figure A.124: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

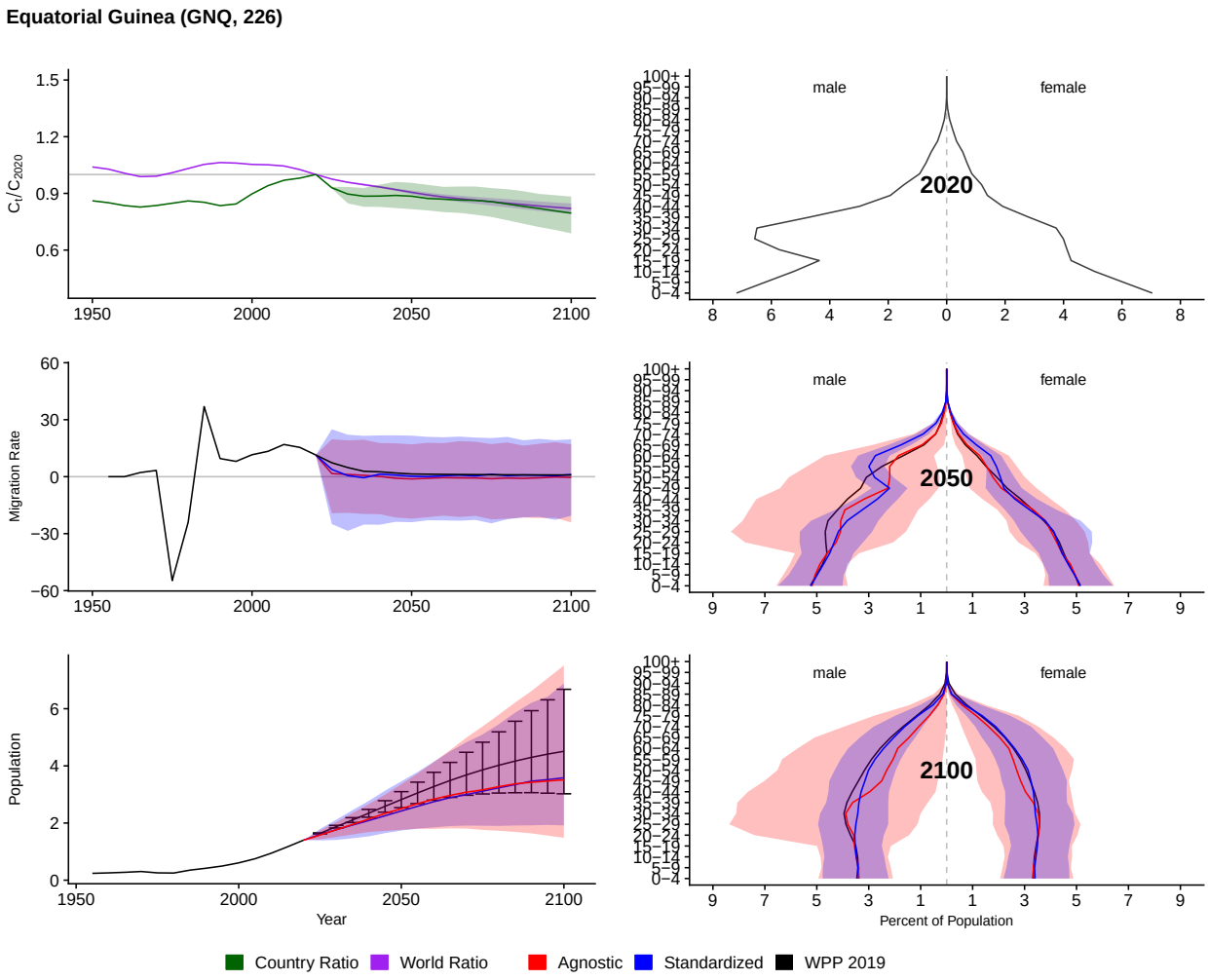


Figure A.125: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

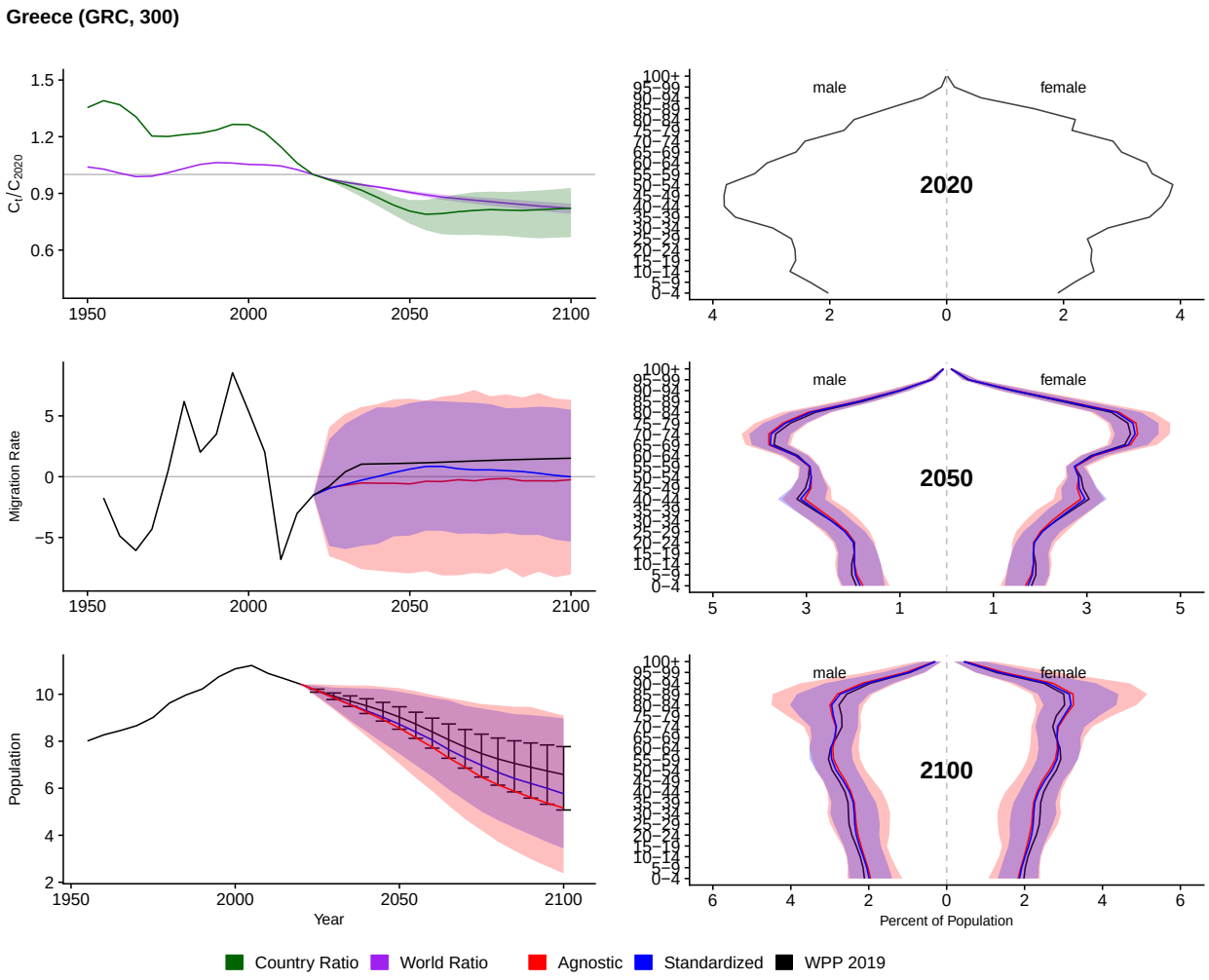


Figure A.126: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

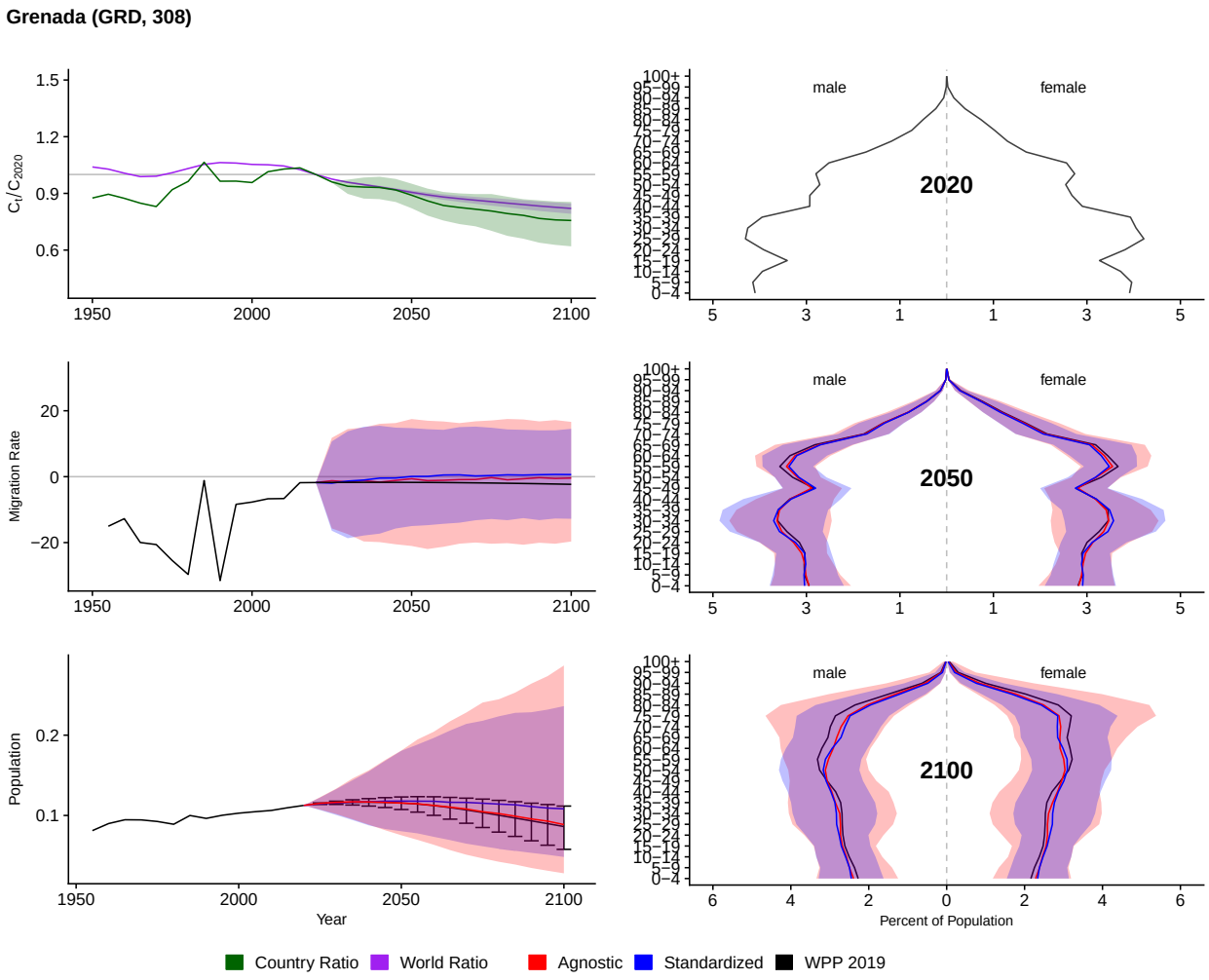


Figure A.127: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

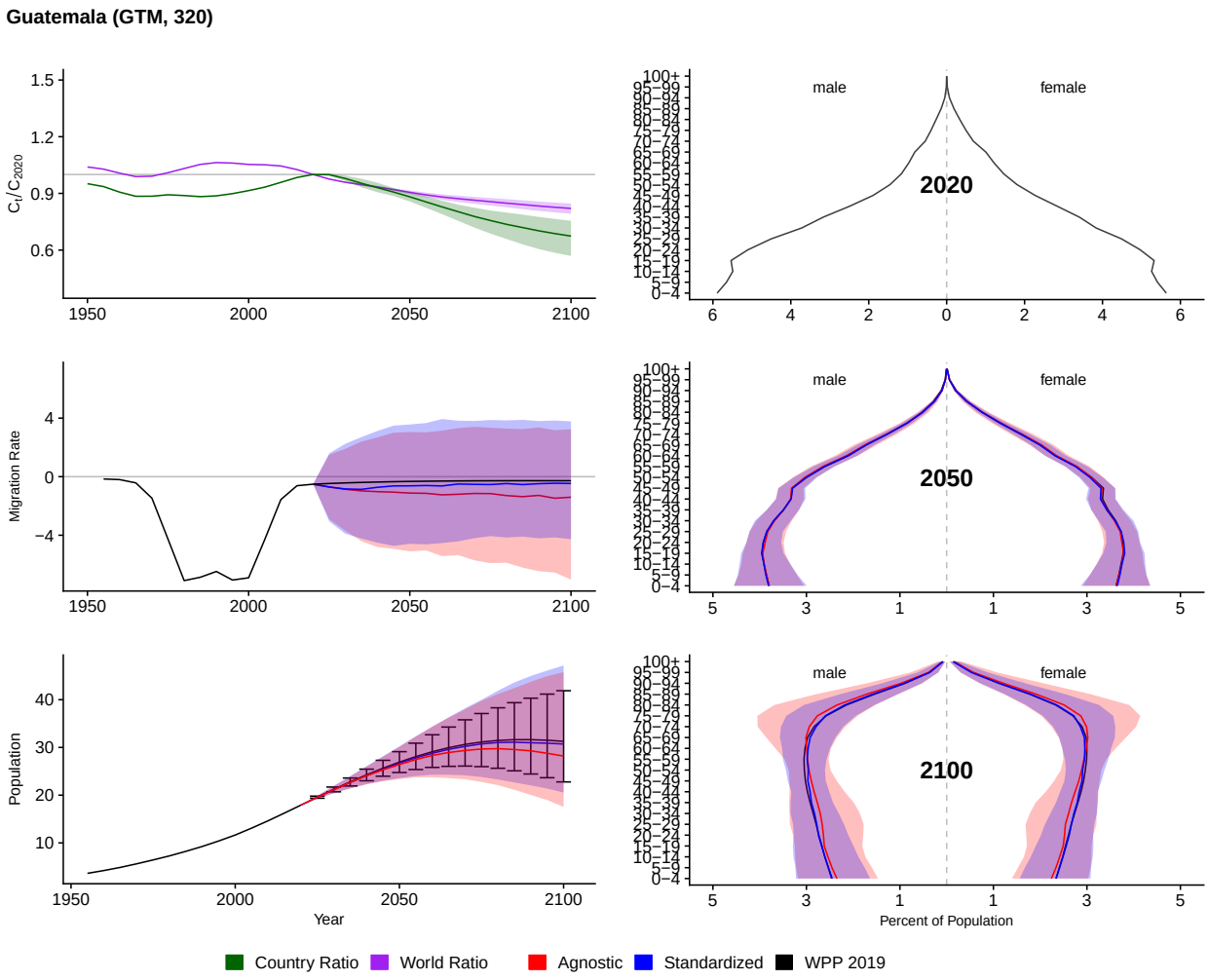


Figure A.128: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

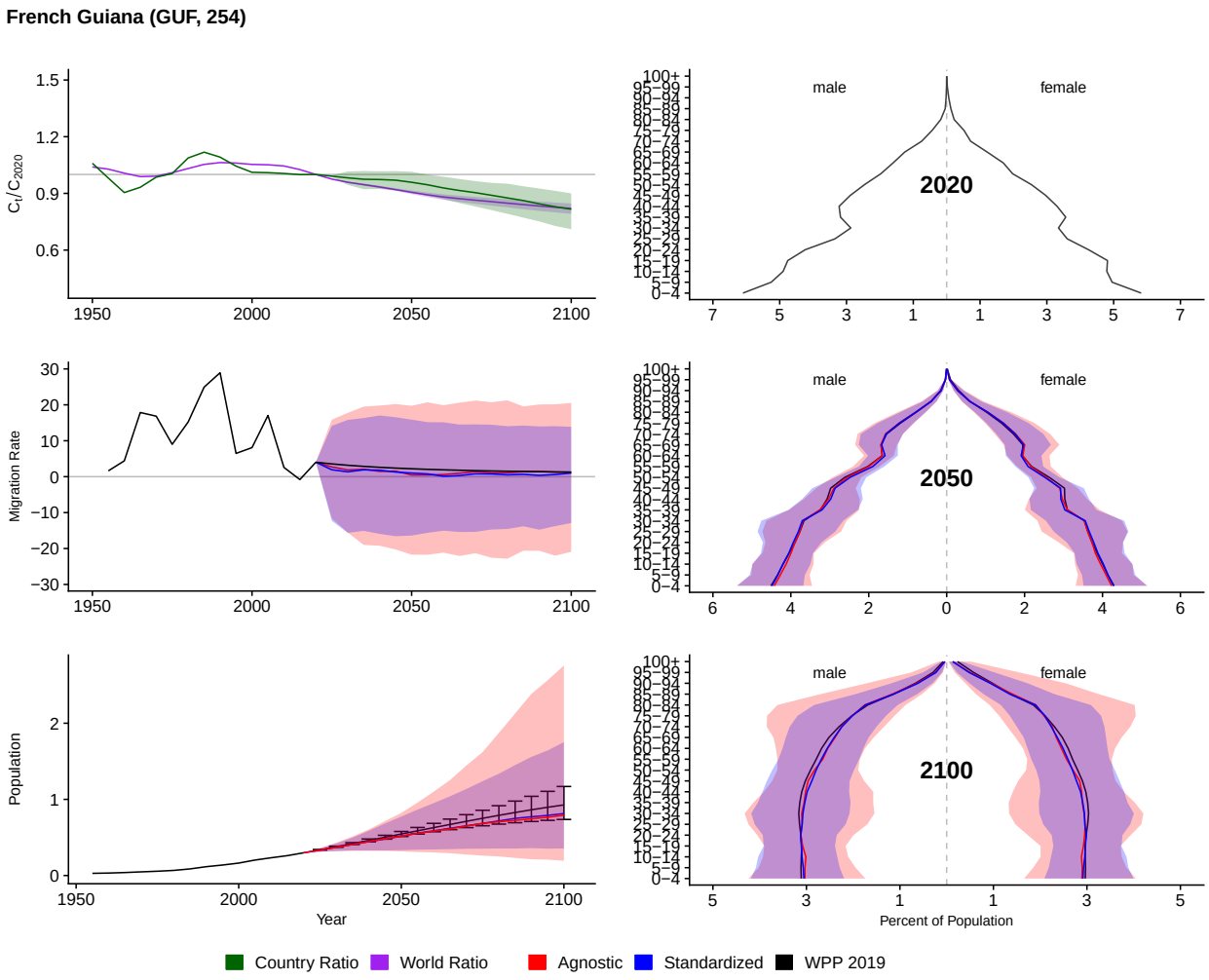


Figure A.129: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

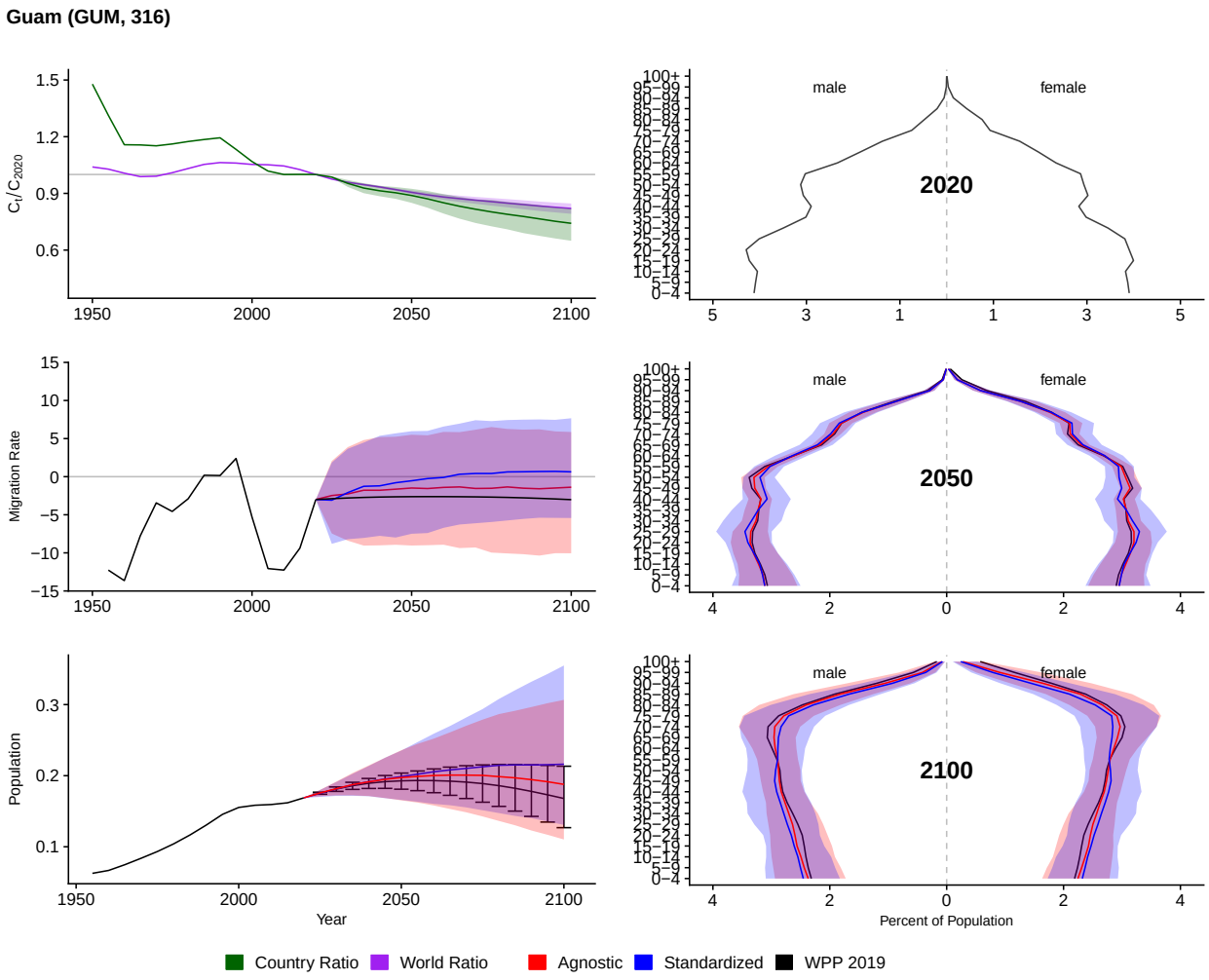


Figure A.130: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

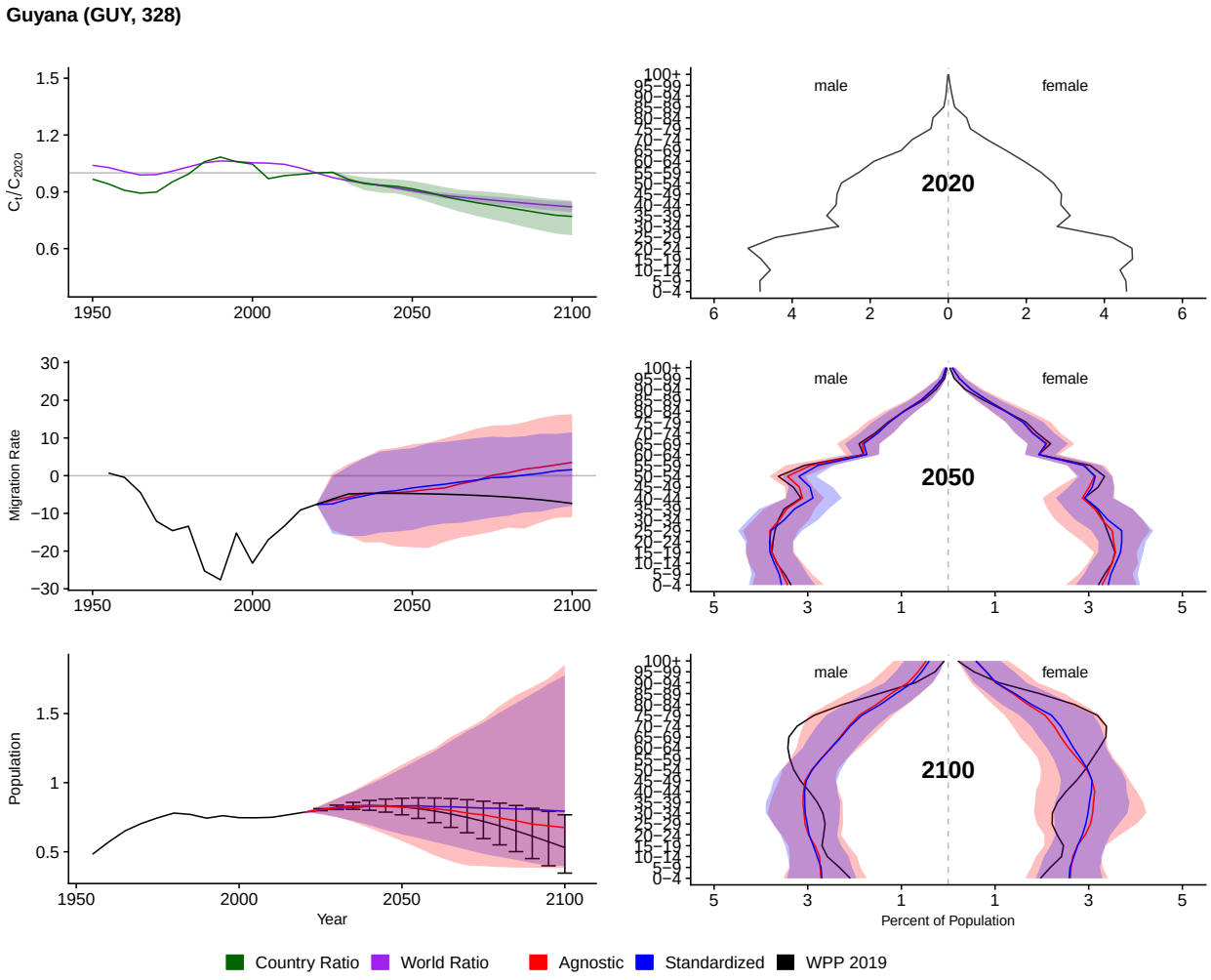


Figure A.131: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

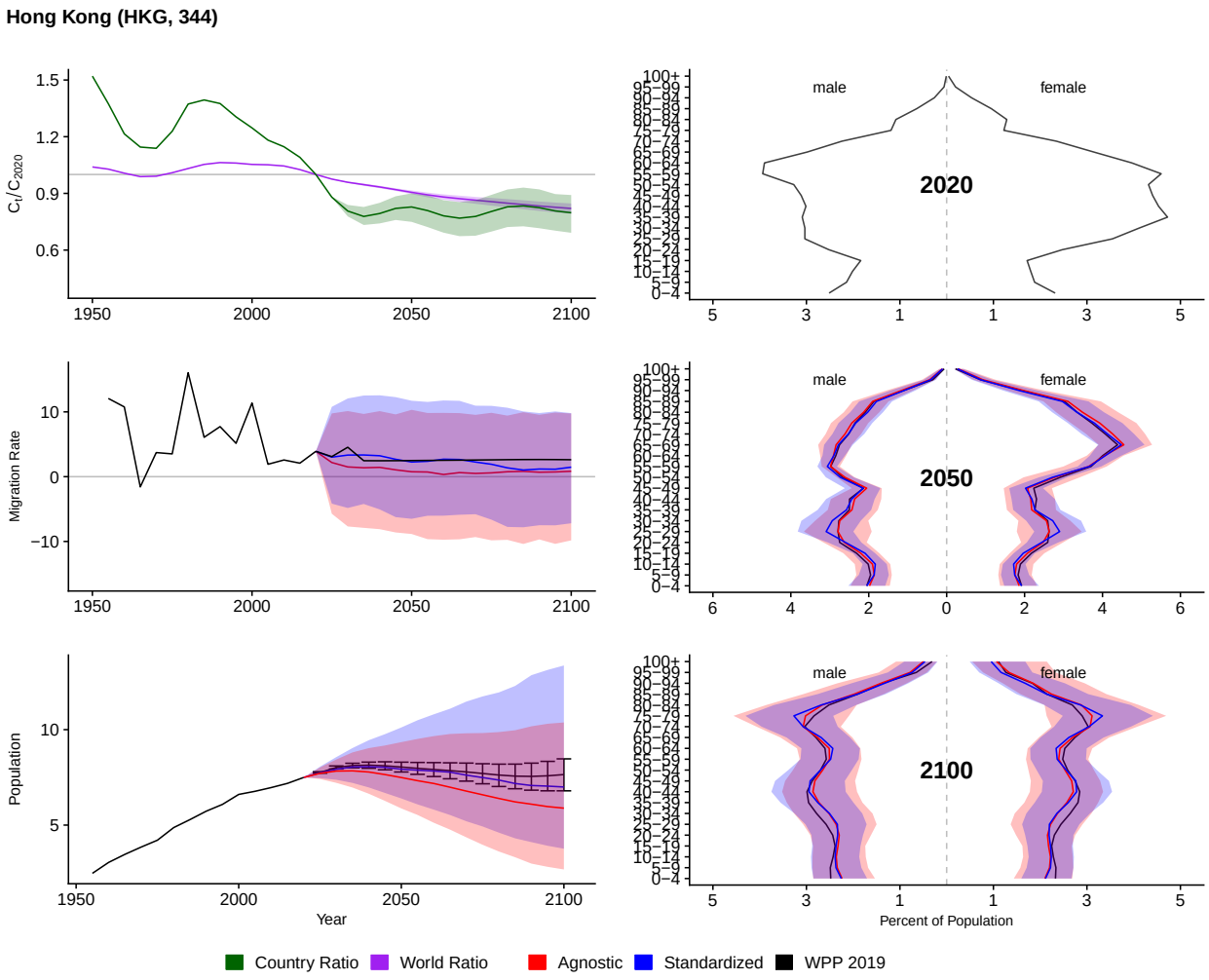


Figure A.132: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

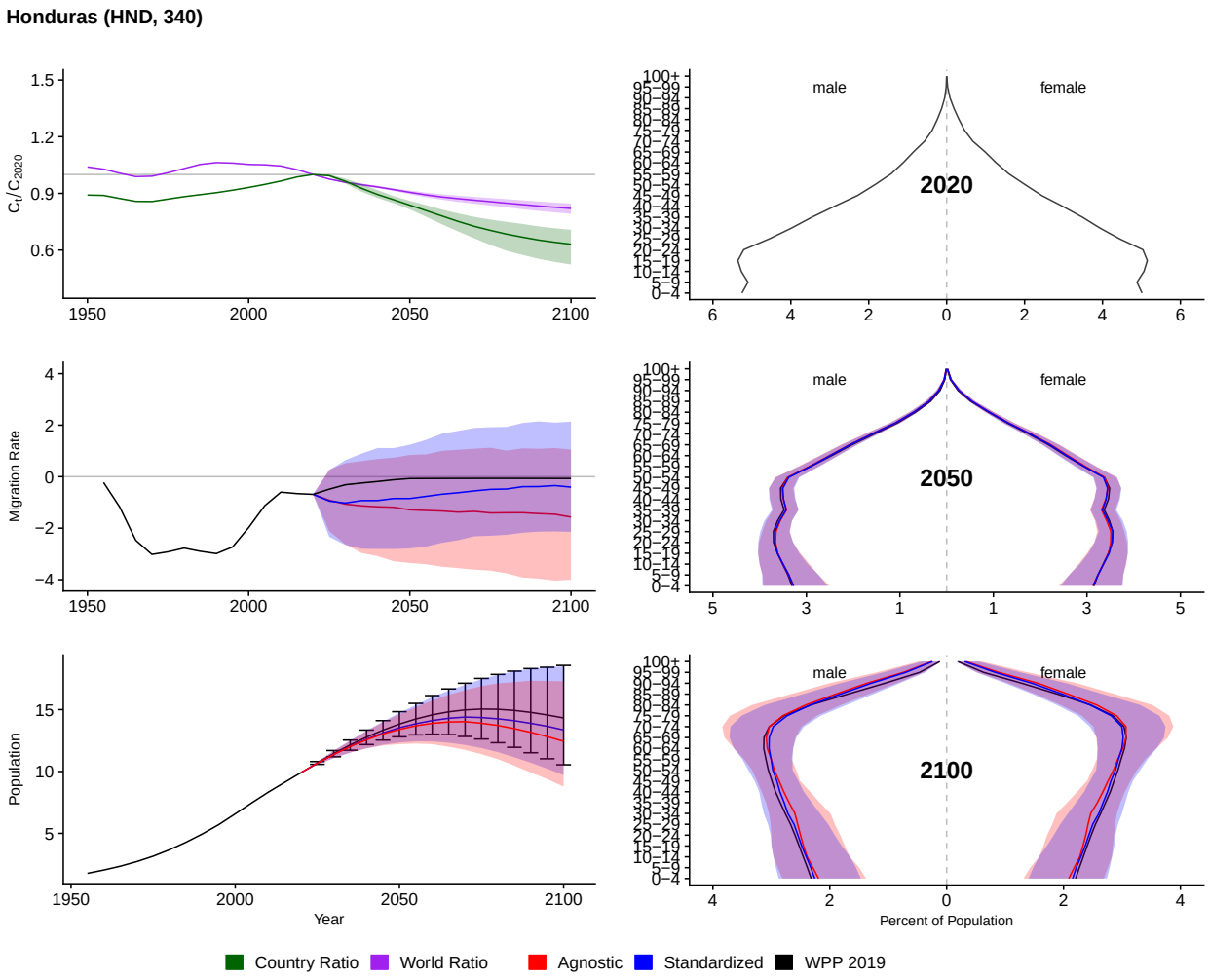


Figure A.133: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

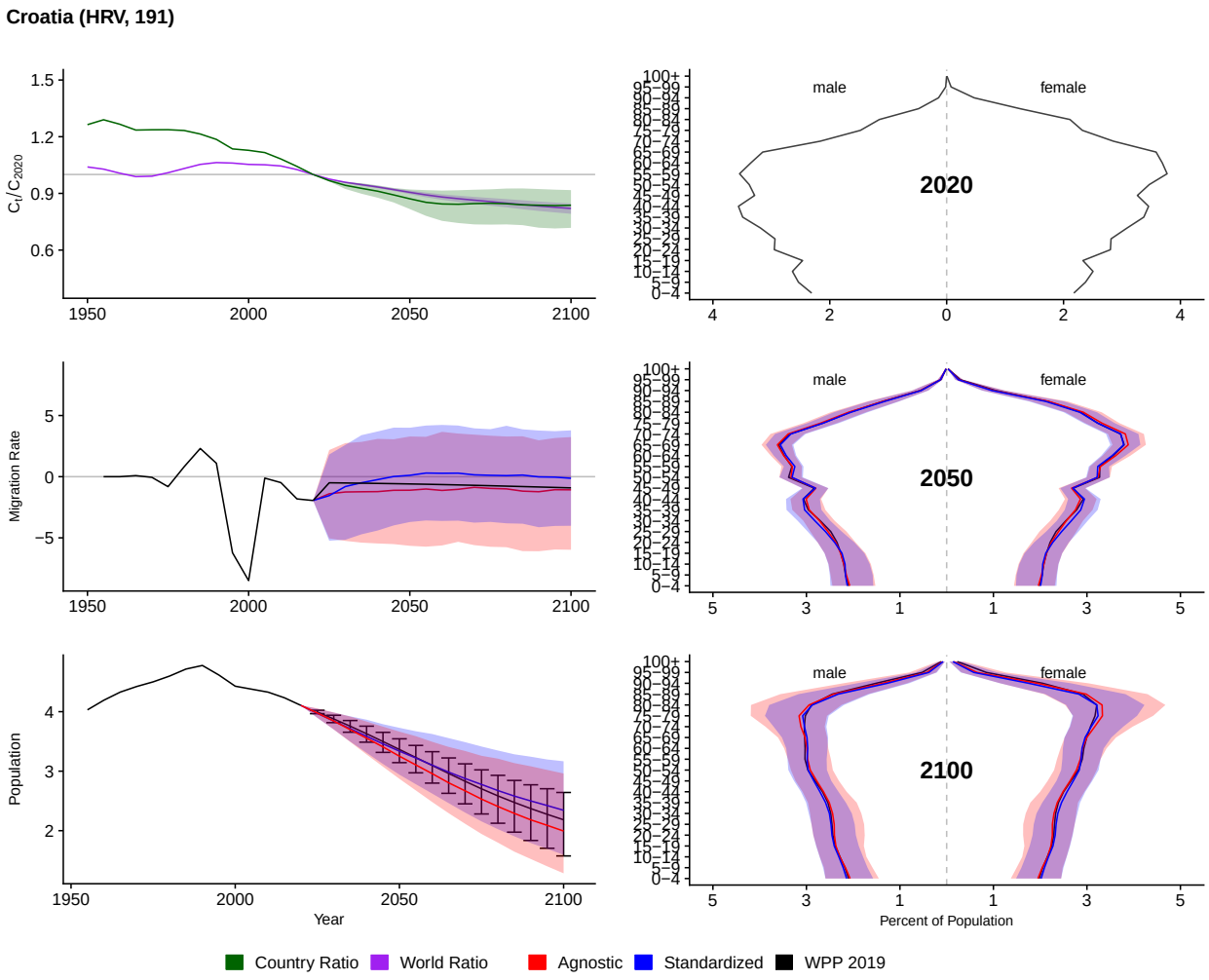


Figure A.134: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

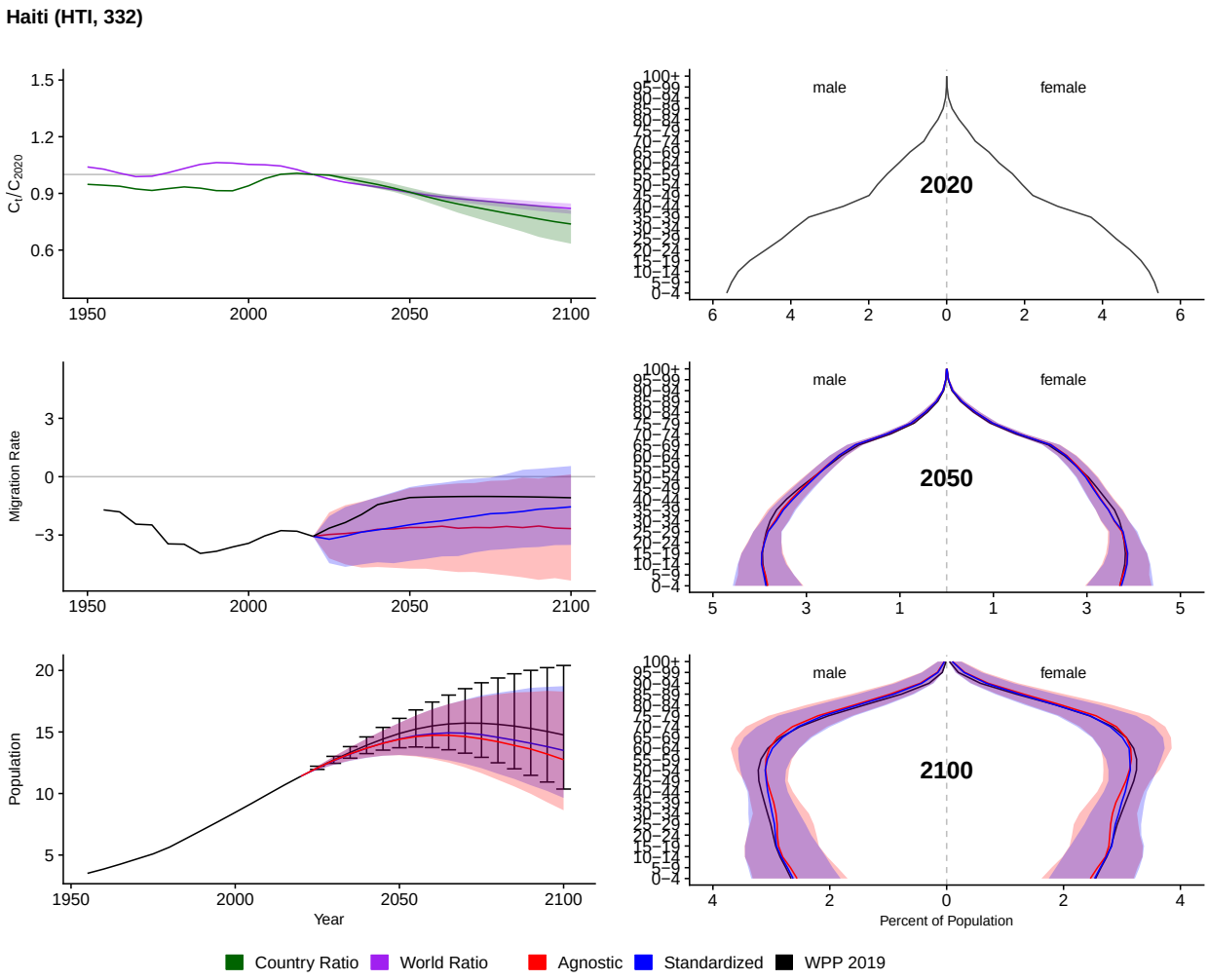


Figure A.135: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

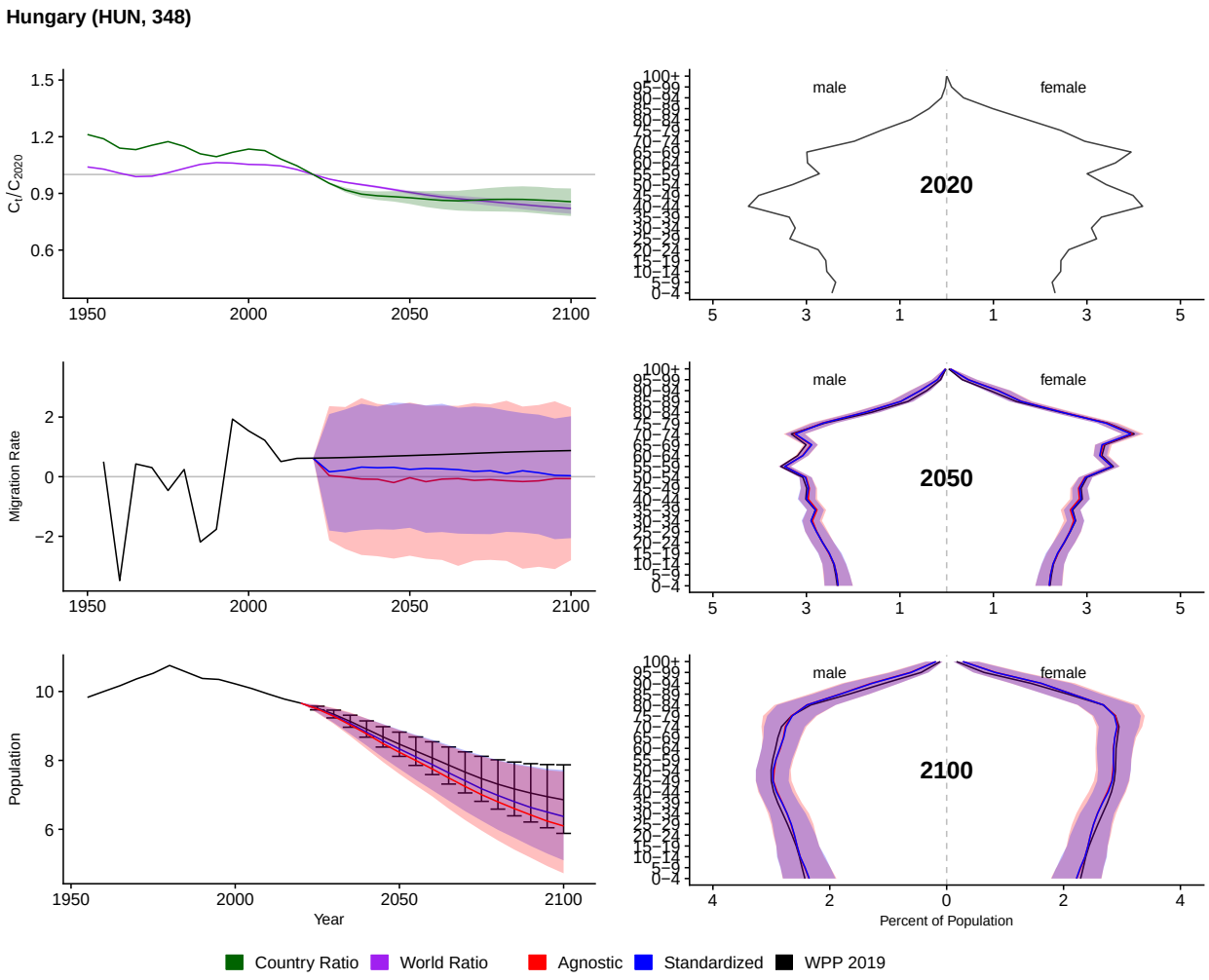


Figure A.136: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

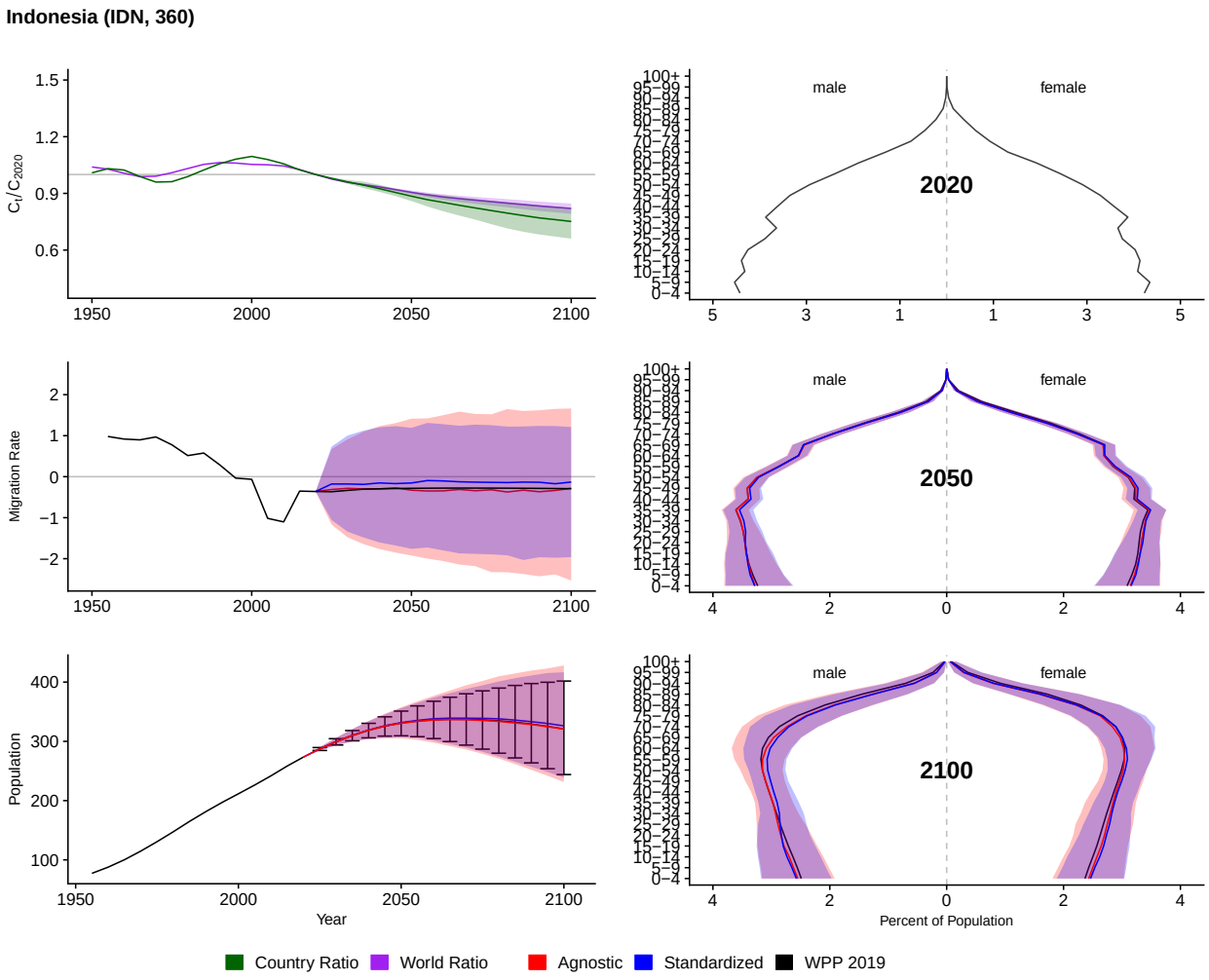


Figure A.137: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

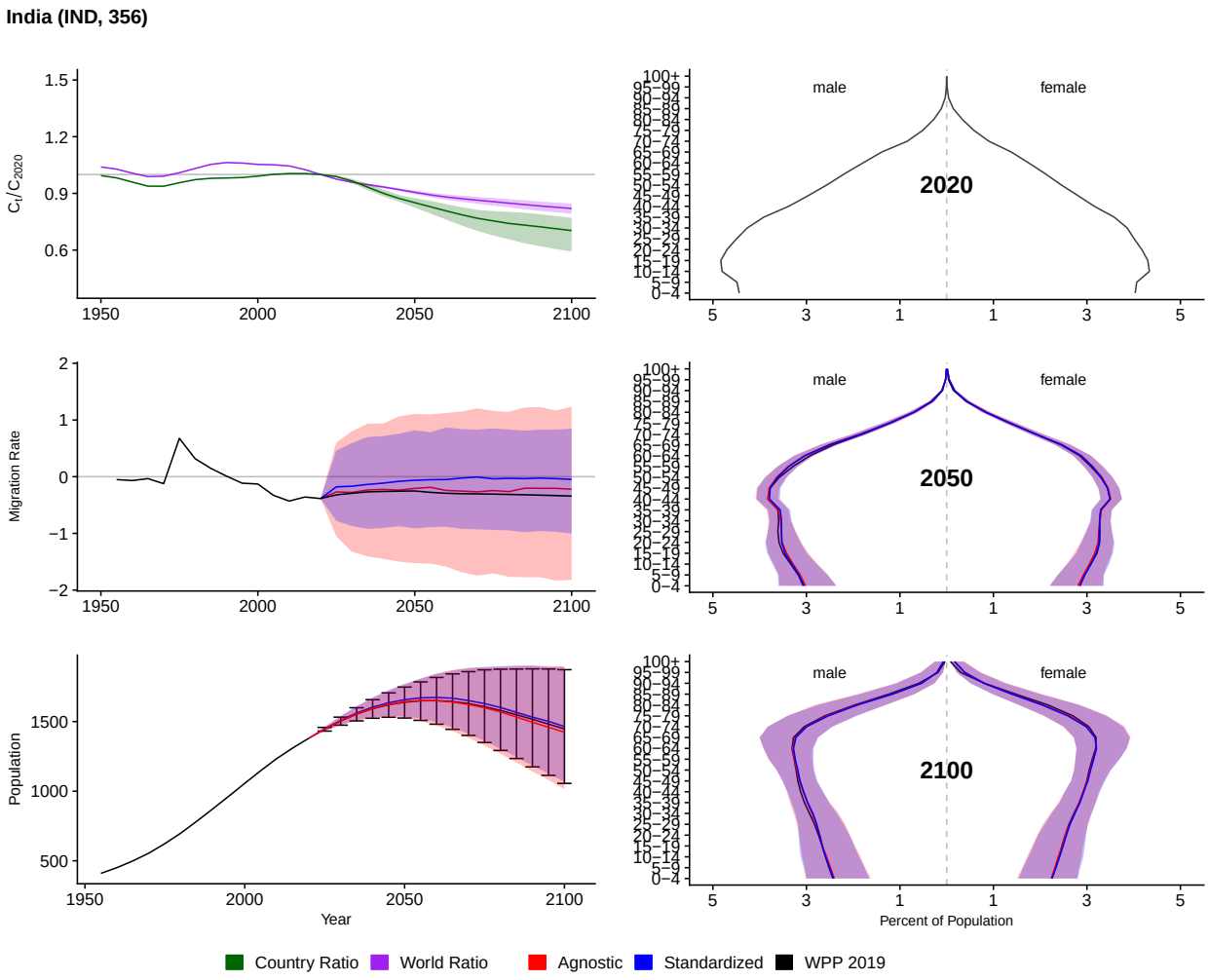


Figure A.138: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

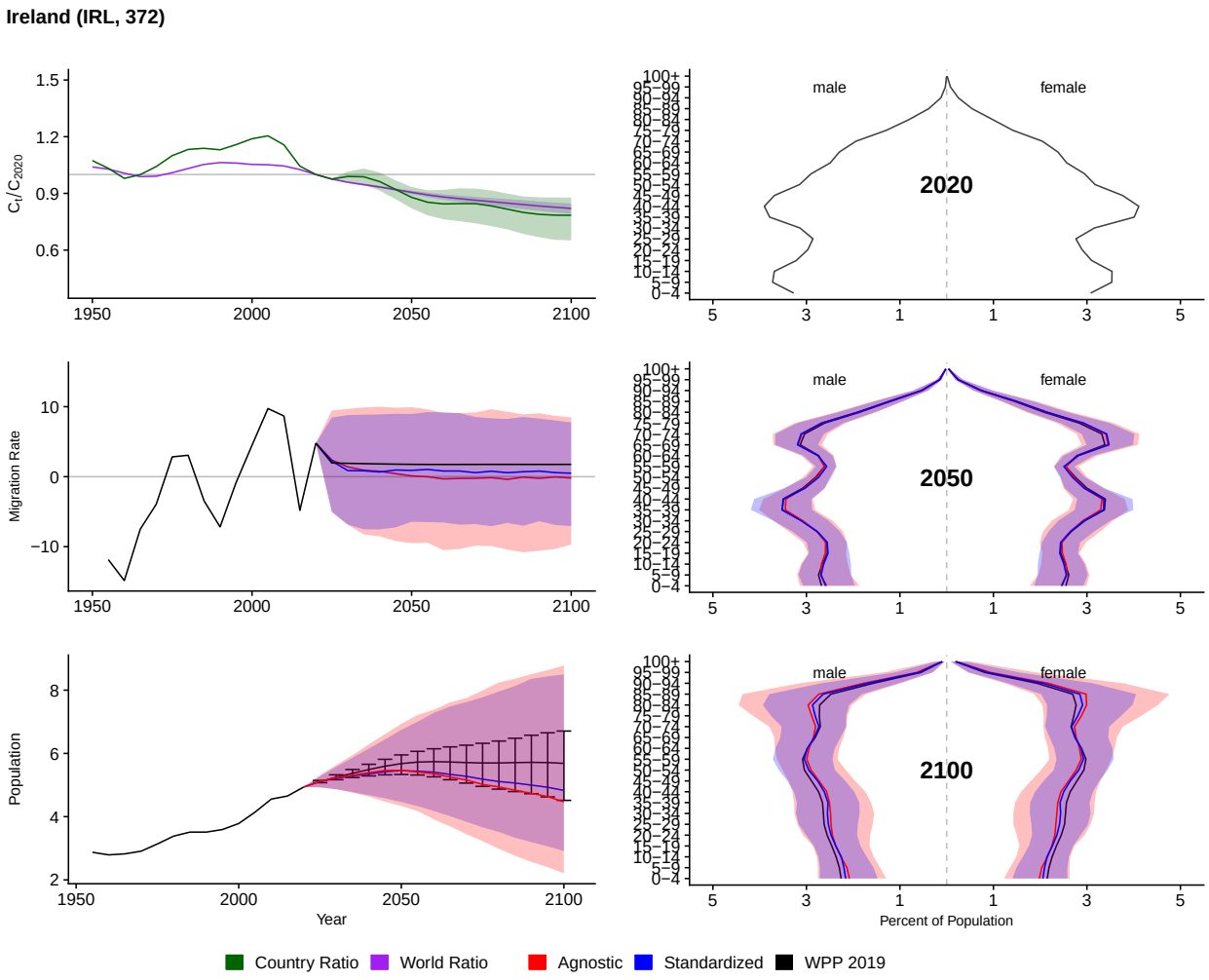


Figure A.139: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

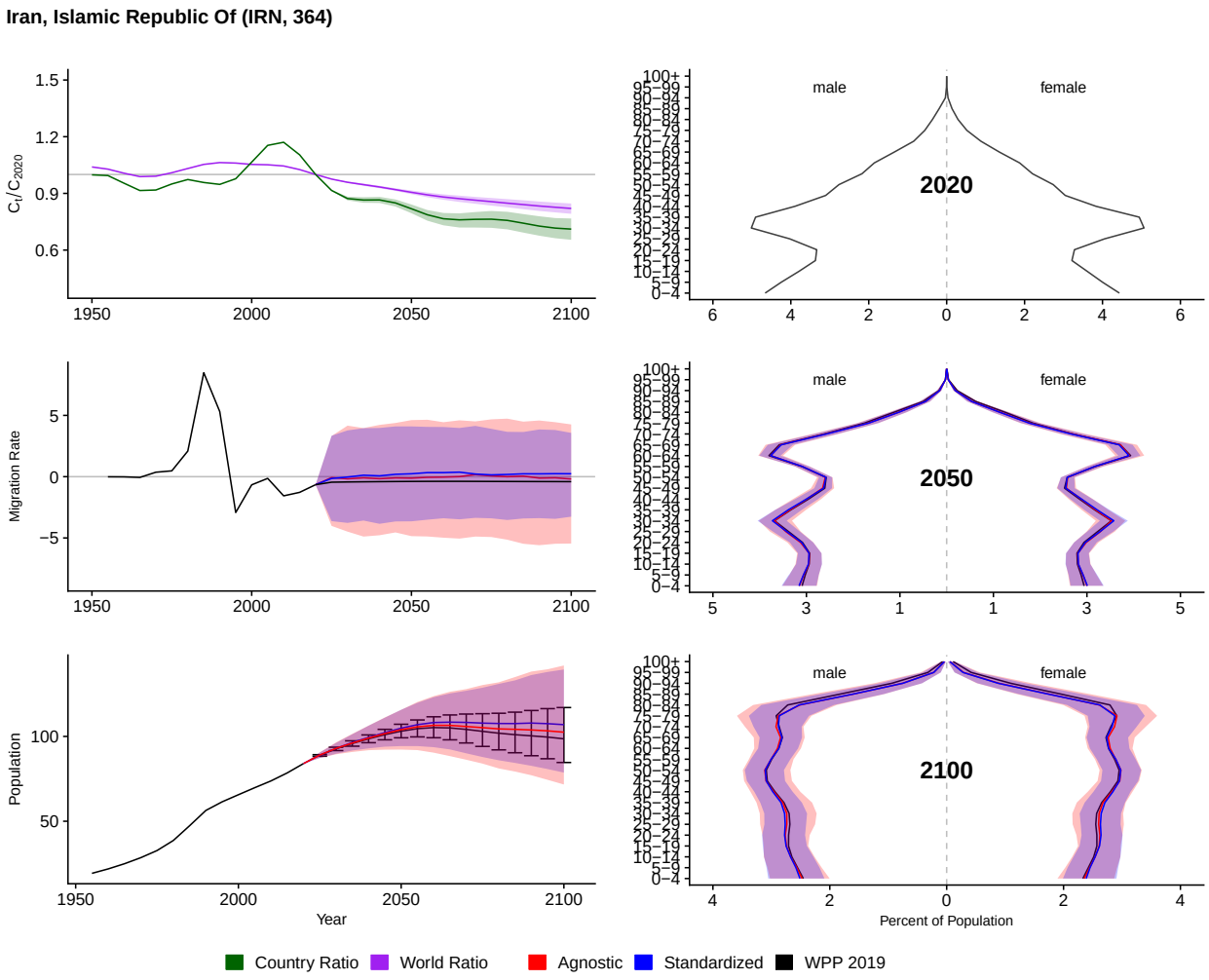


Figure A.140: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

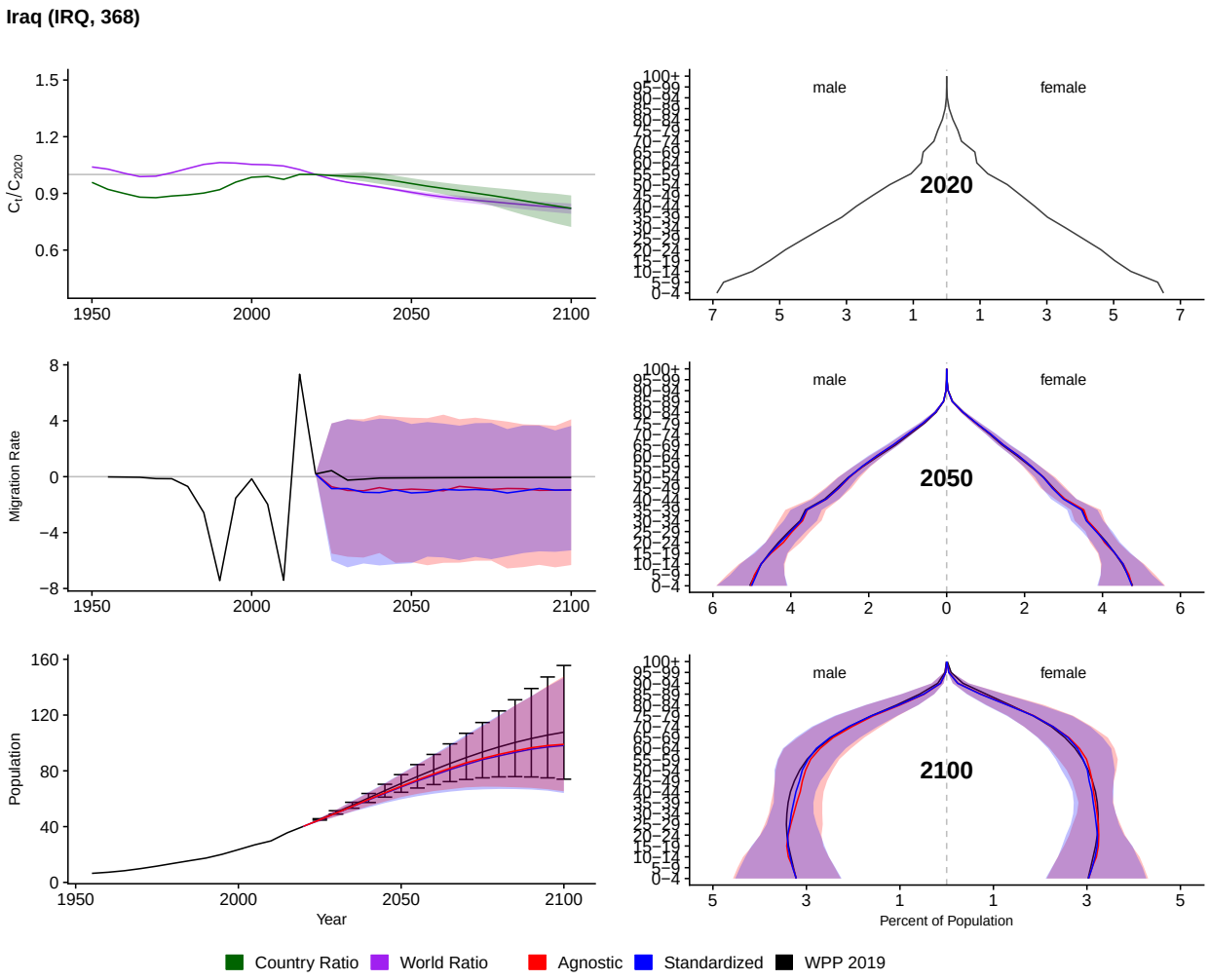


Figure A.141: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

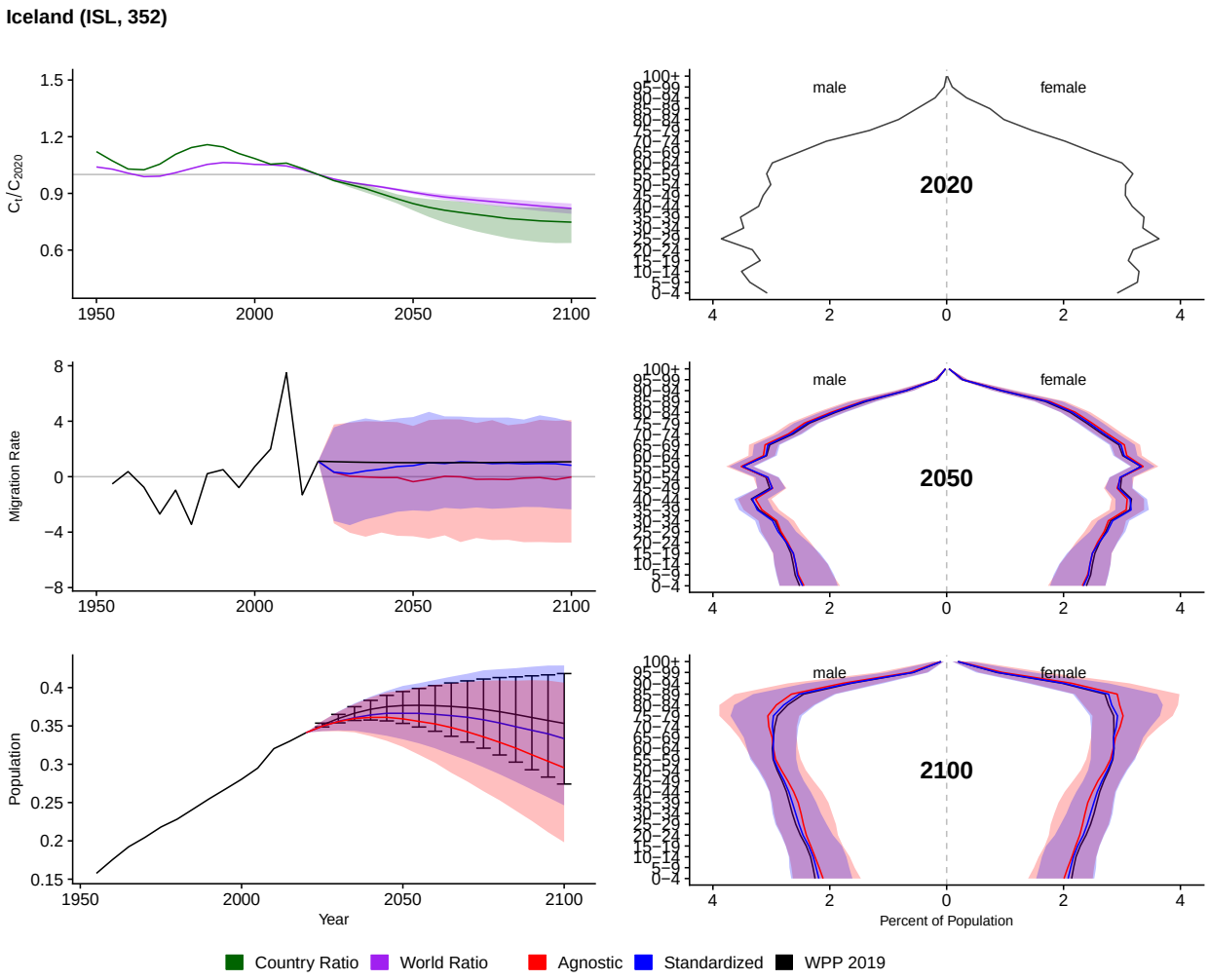


Figure A.142: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

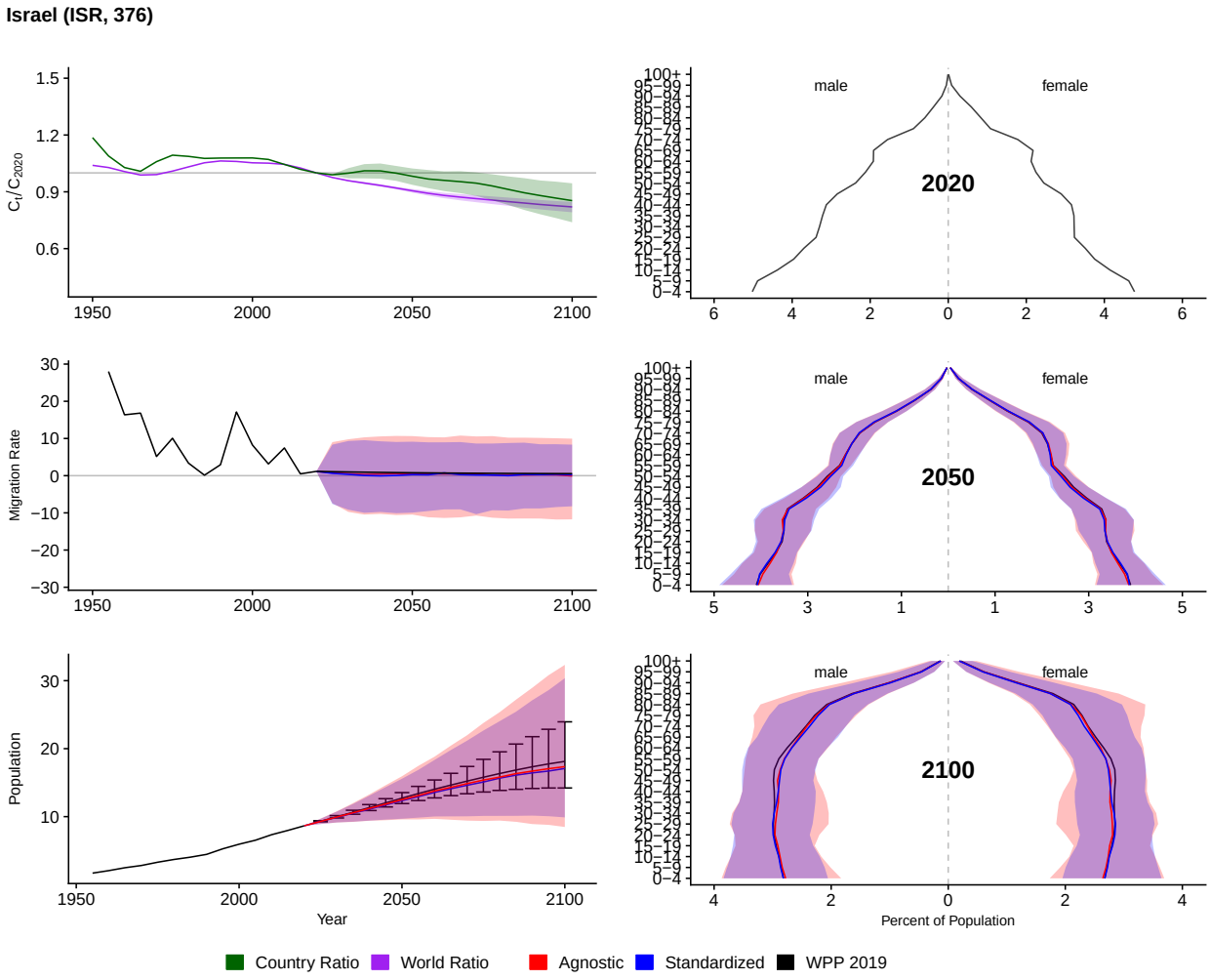


Figure A.143: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

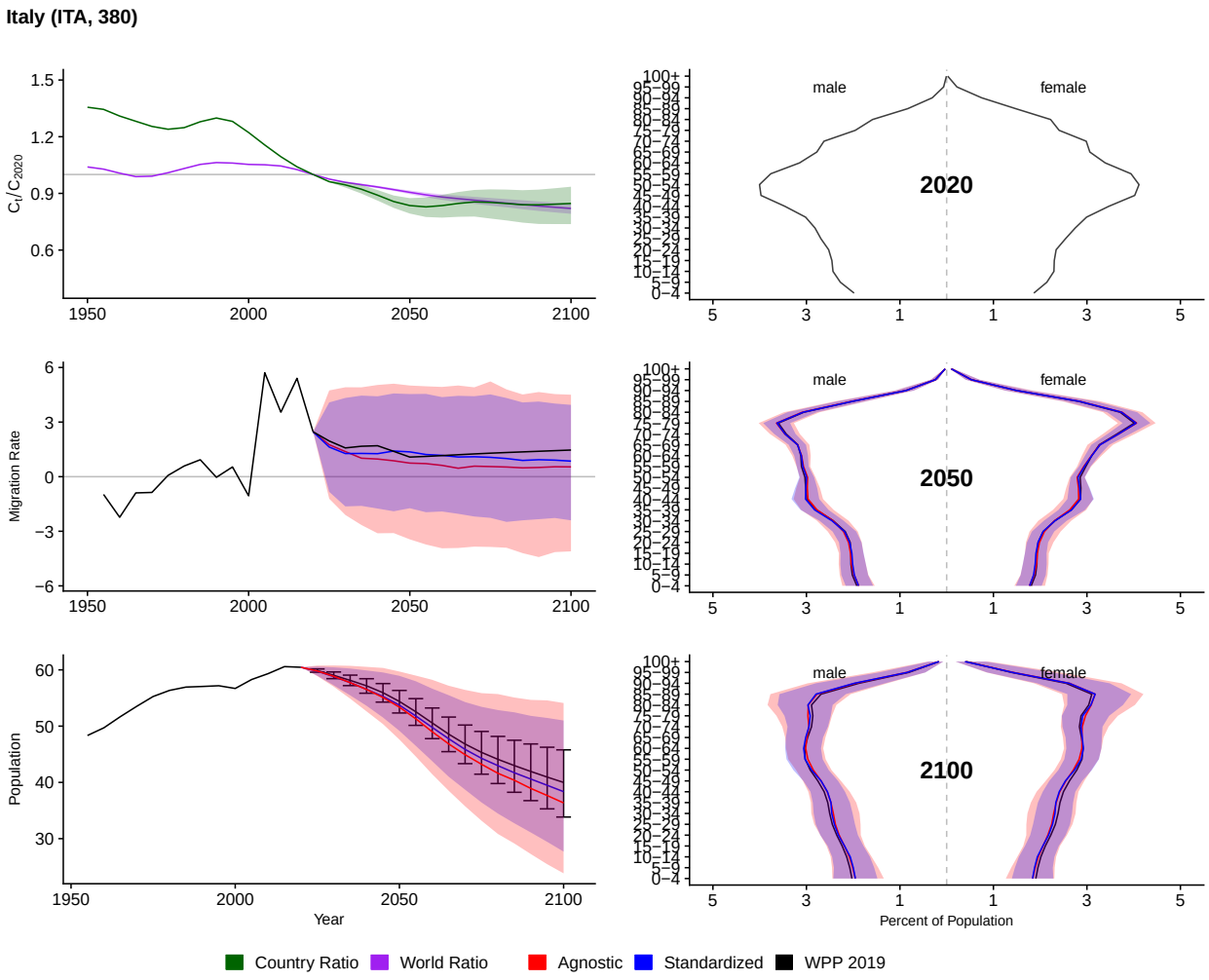


Figure A.144: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

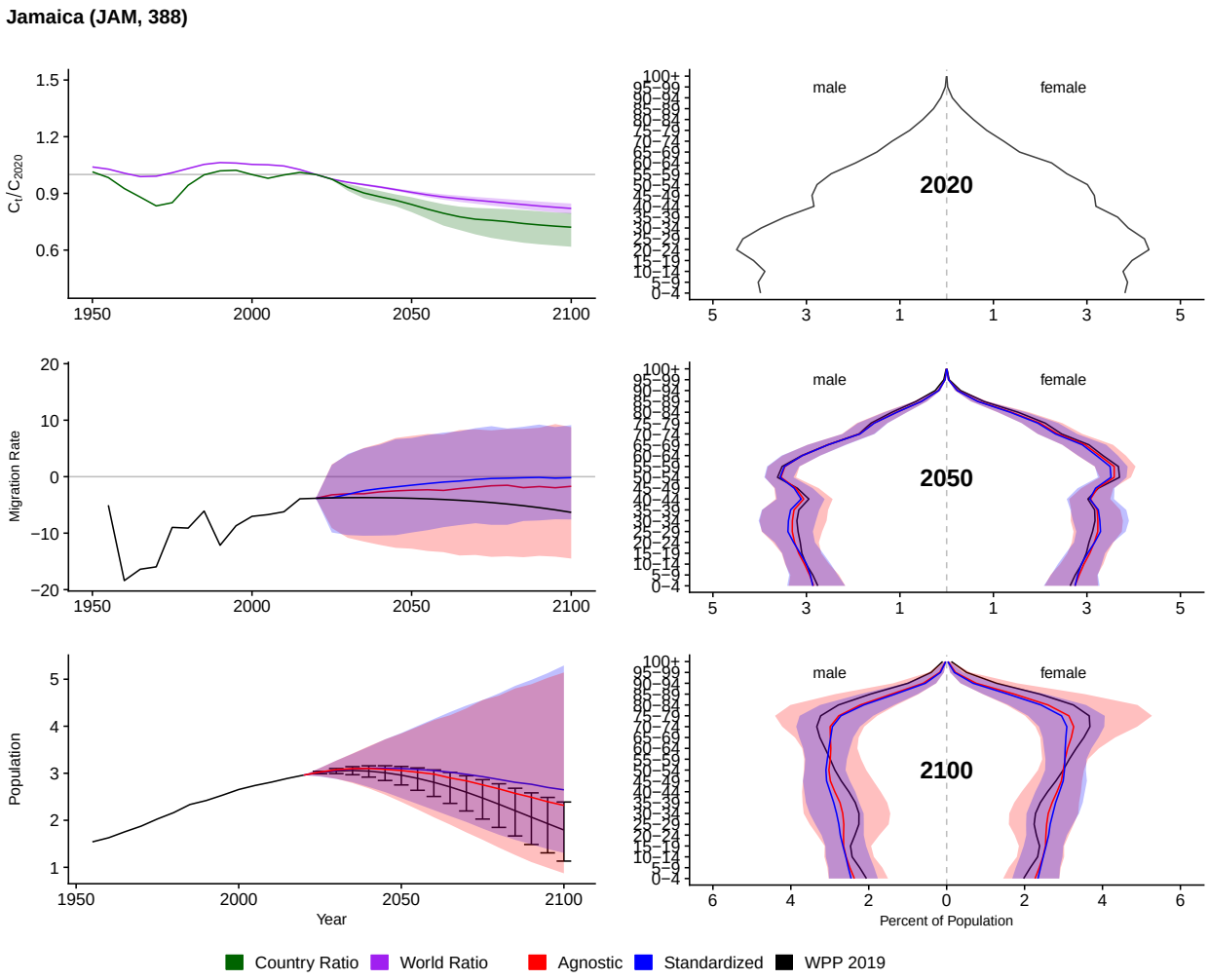


Figure A.145: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

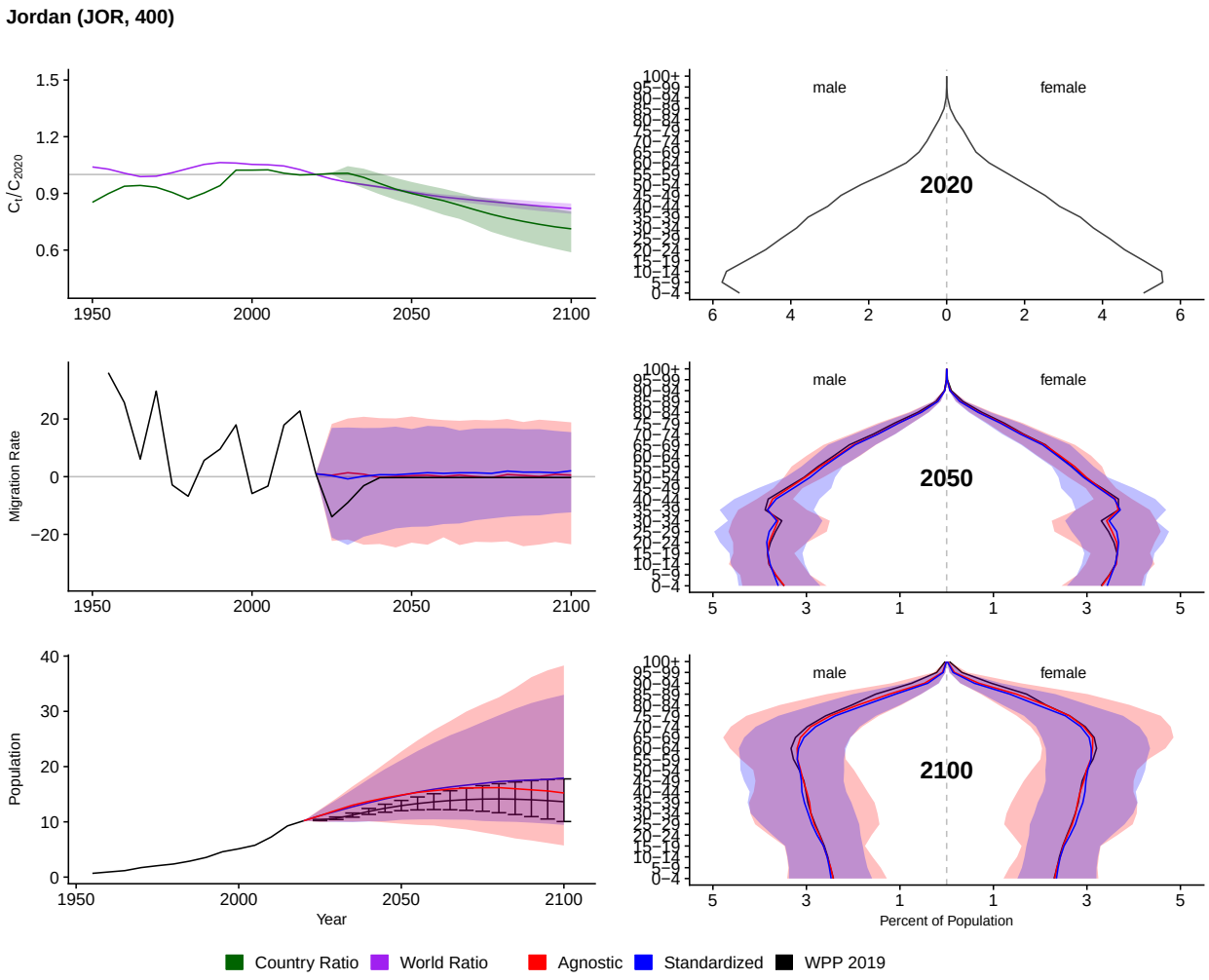


Figure A.146: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

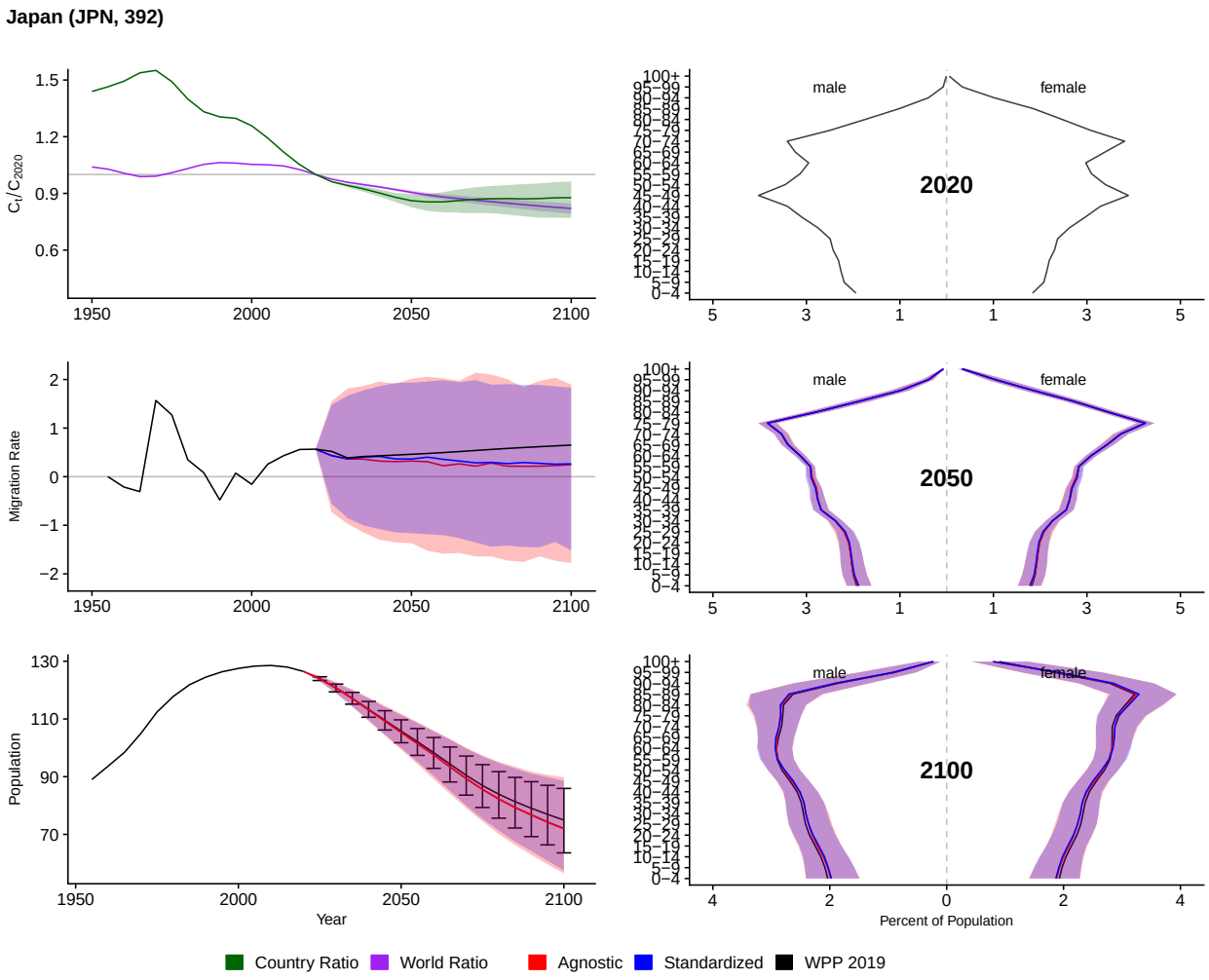


Figure A.147: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

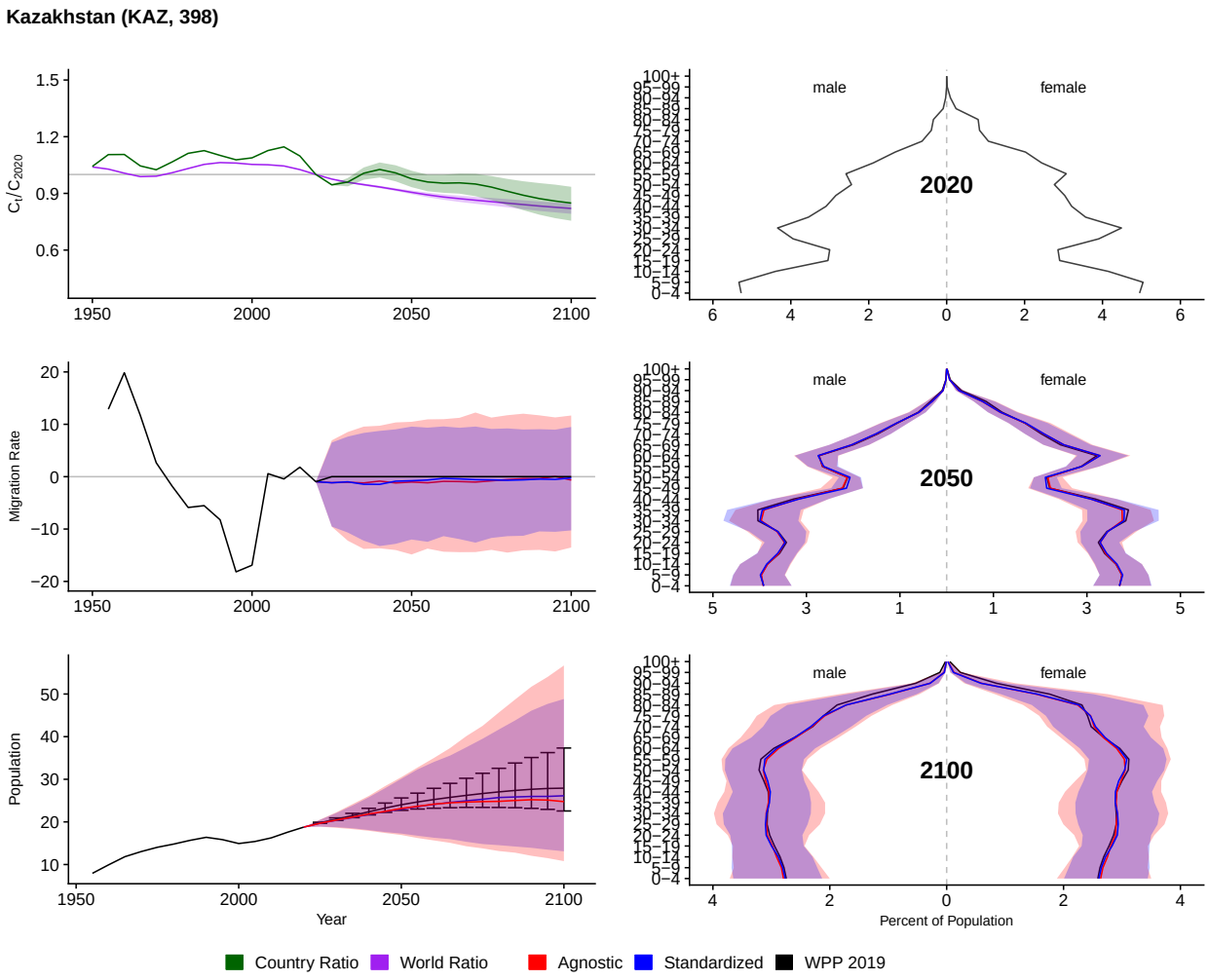


Figure A.148: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

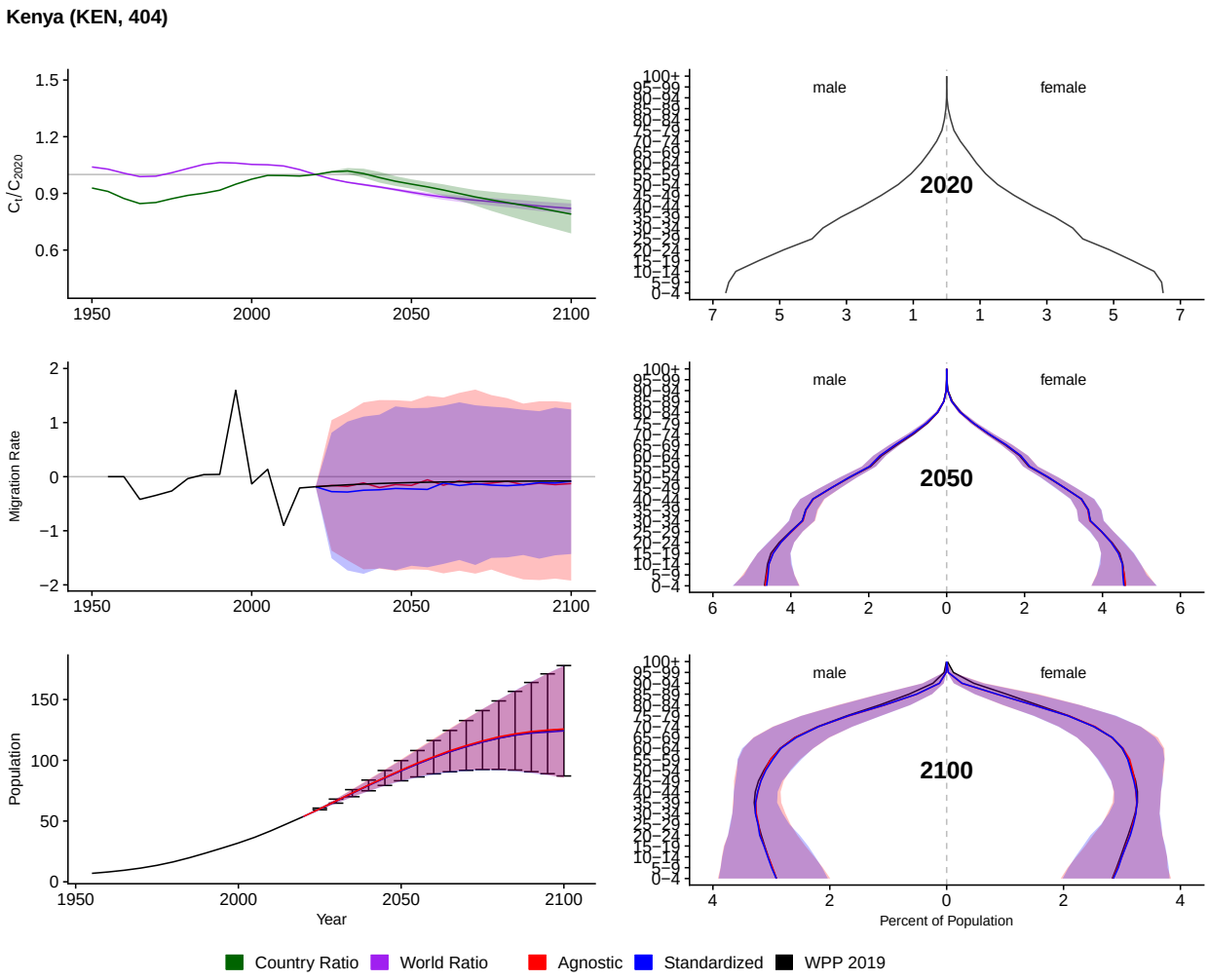


Figure A.149: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

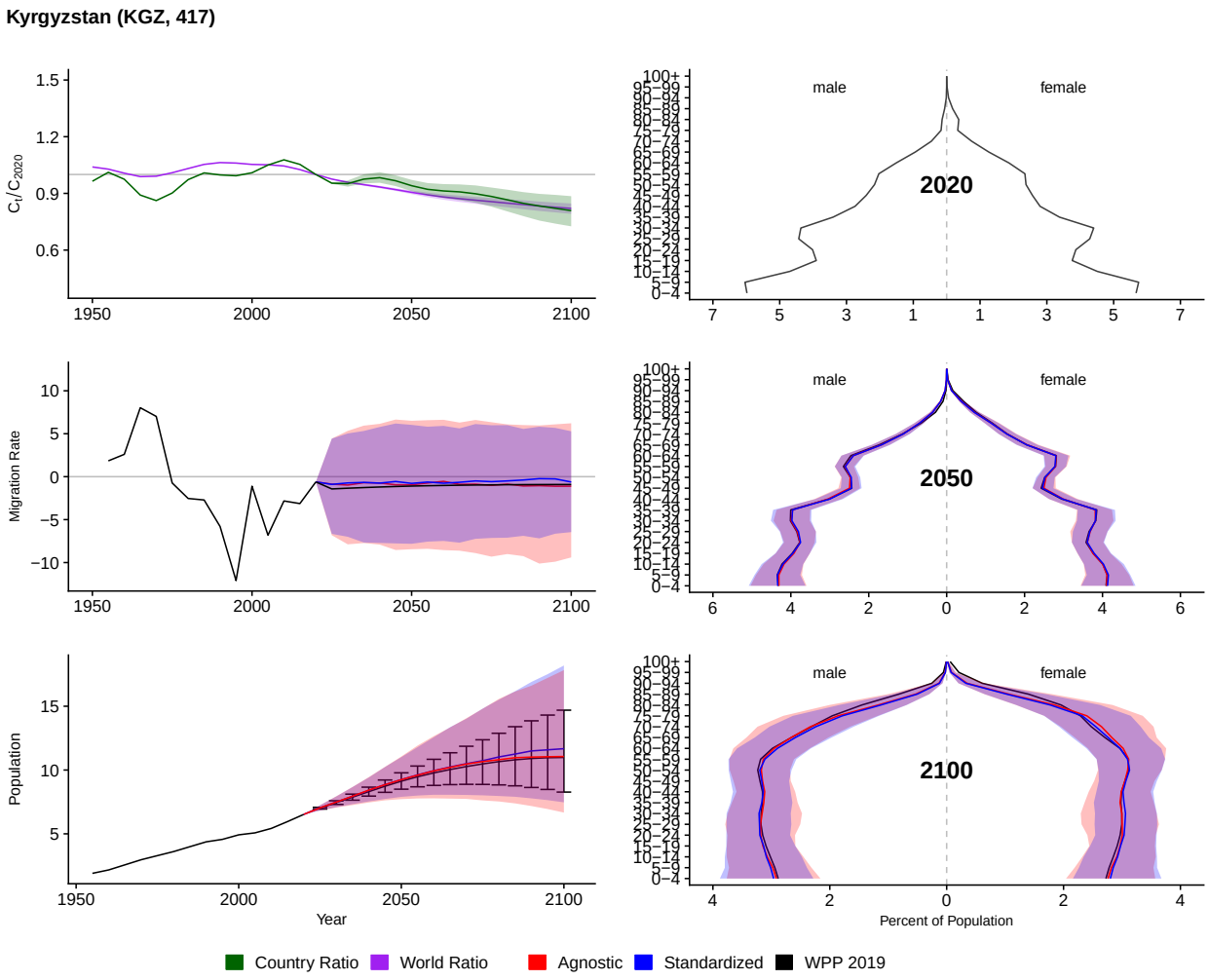


Figure A.150: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

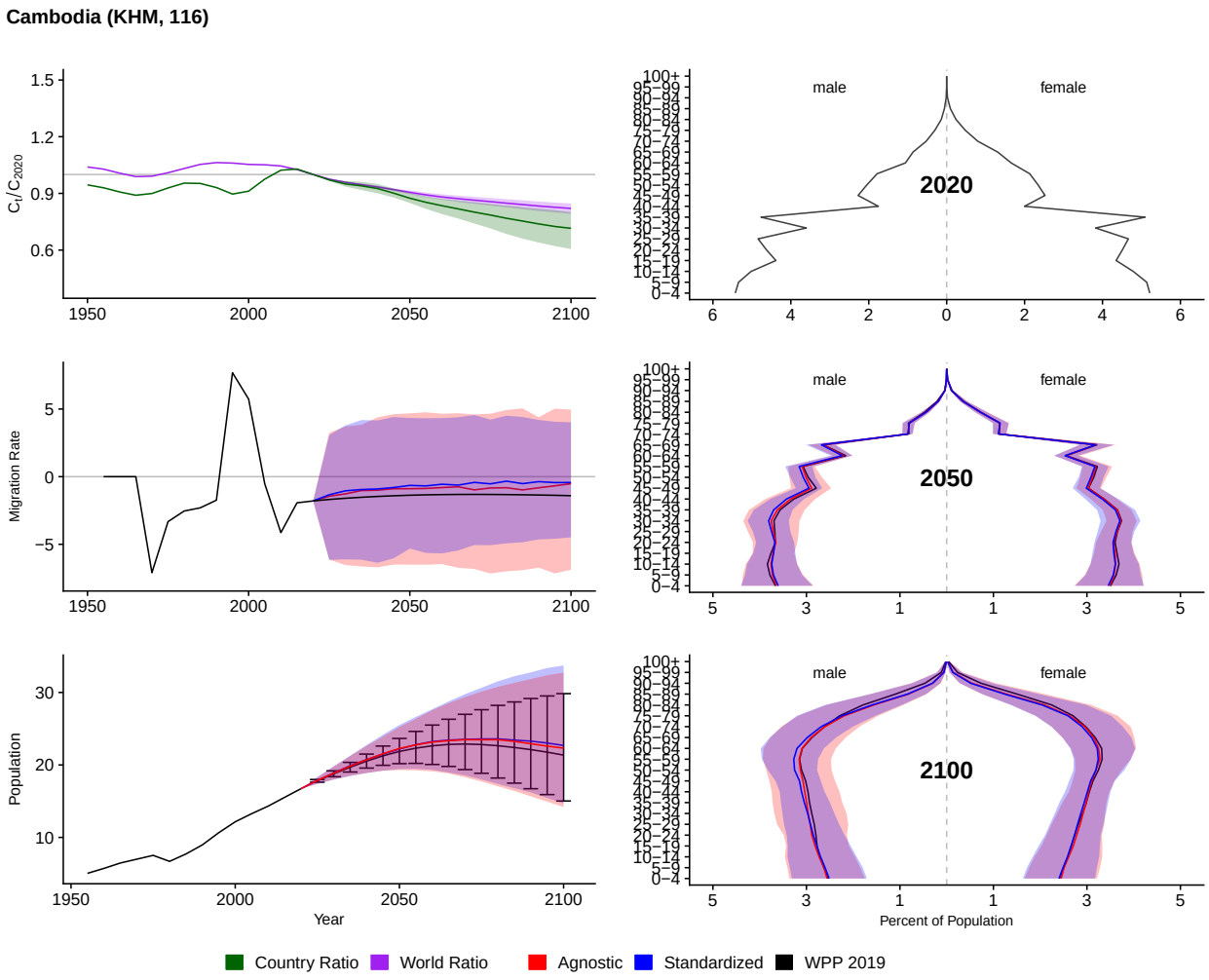


Figure A.151: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

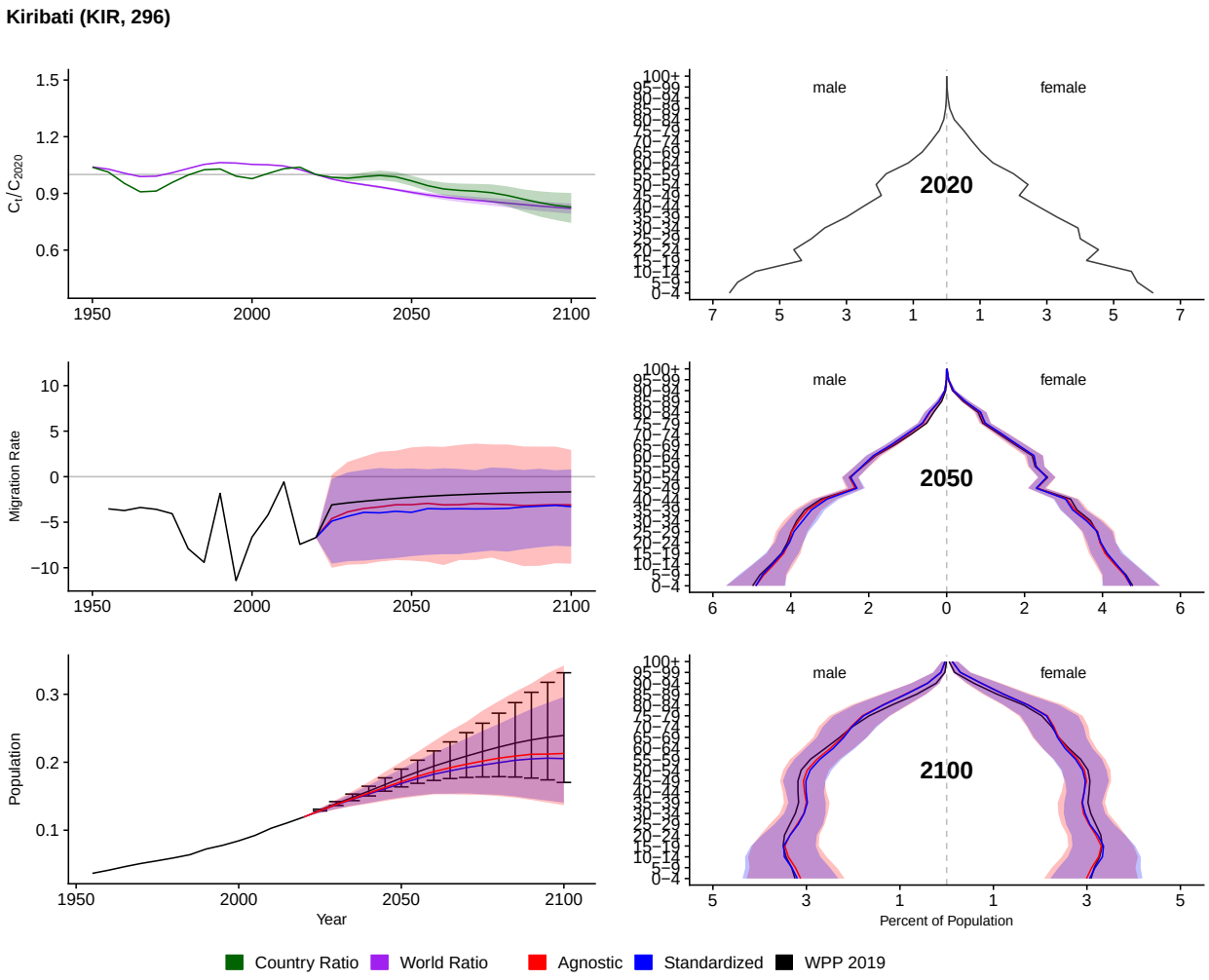


Figure A.152: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

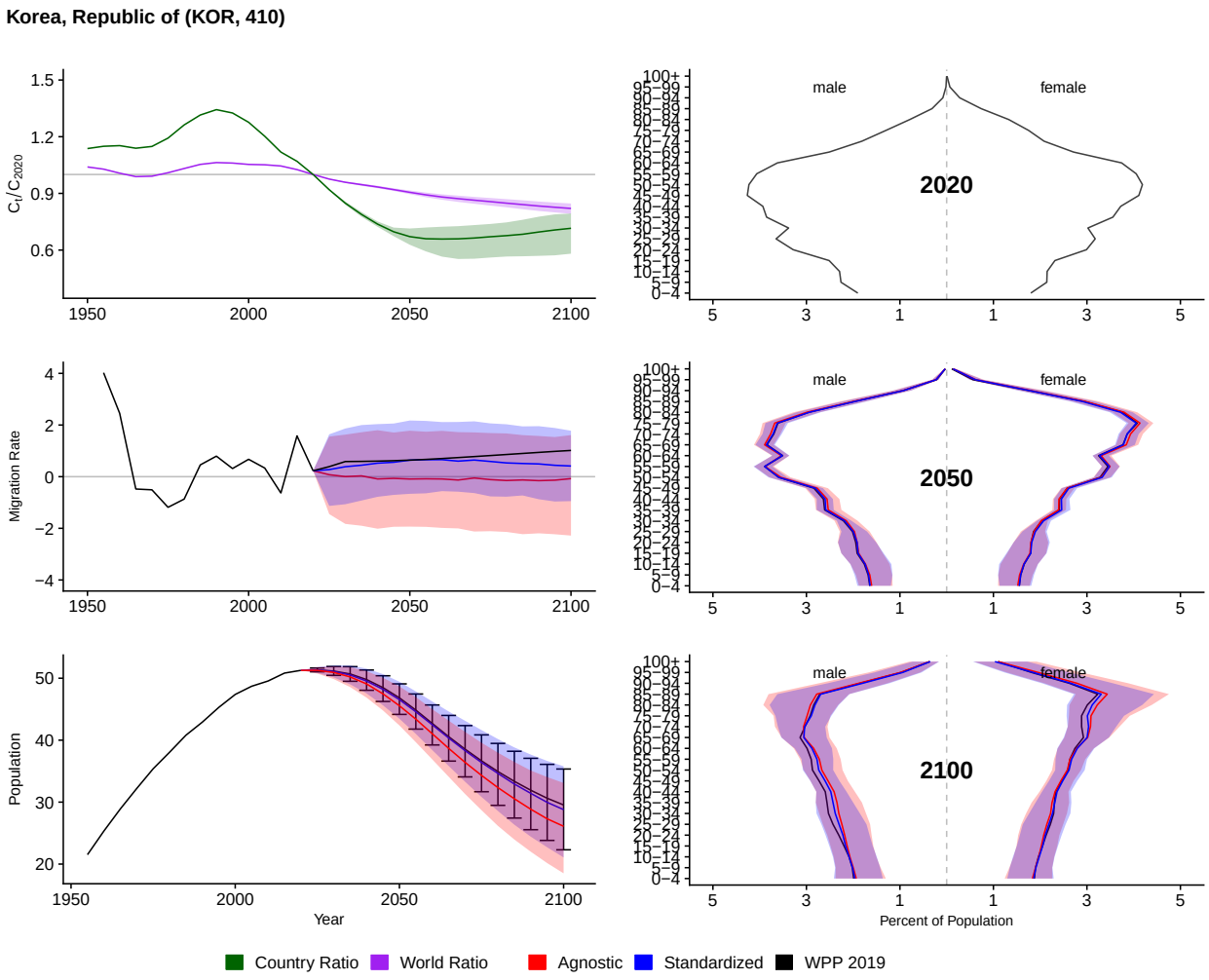


Figure A.153: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

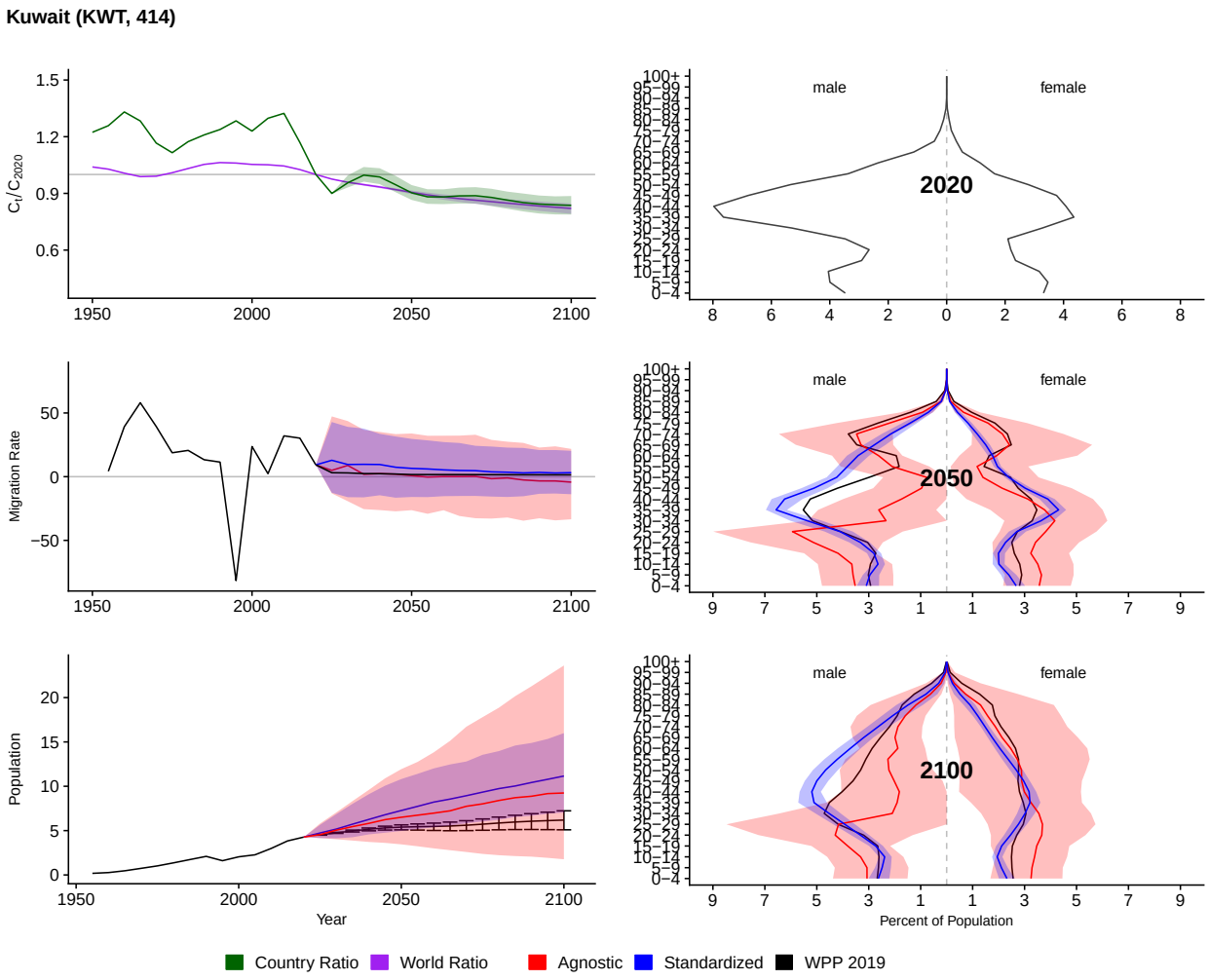


Figure A.154: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

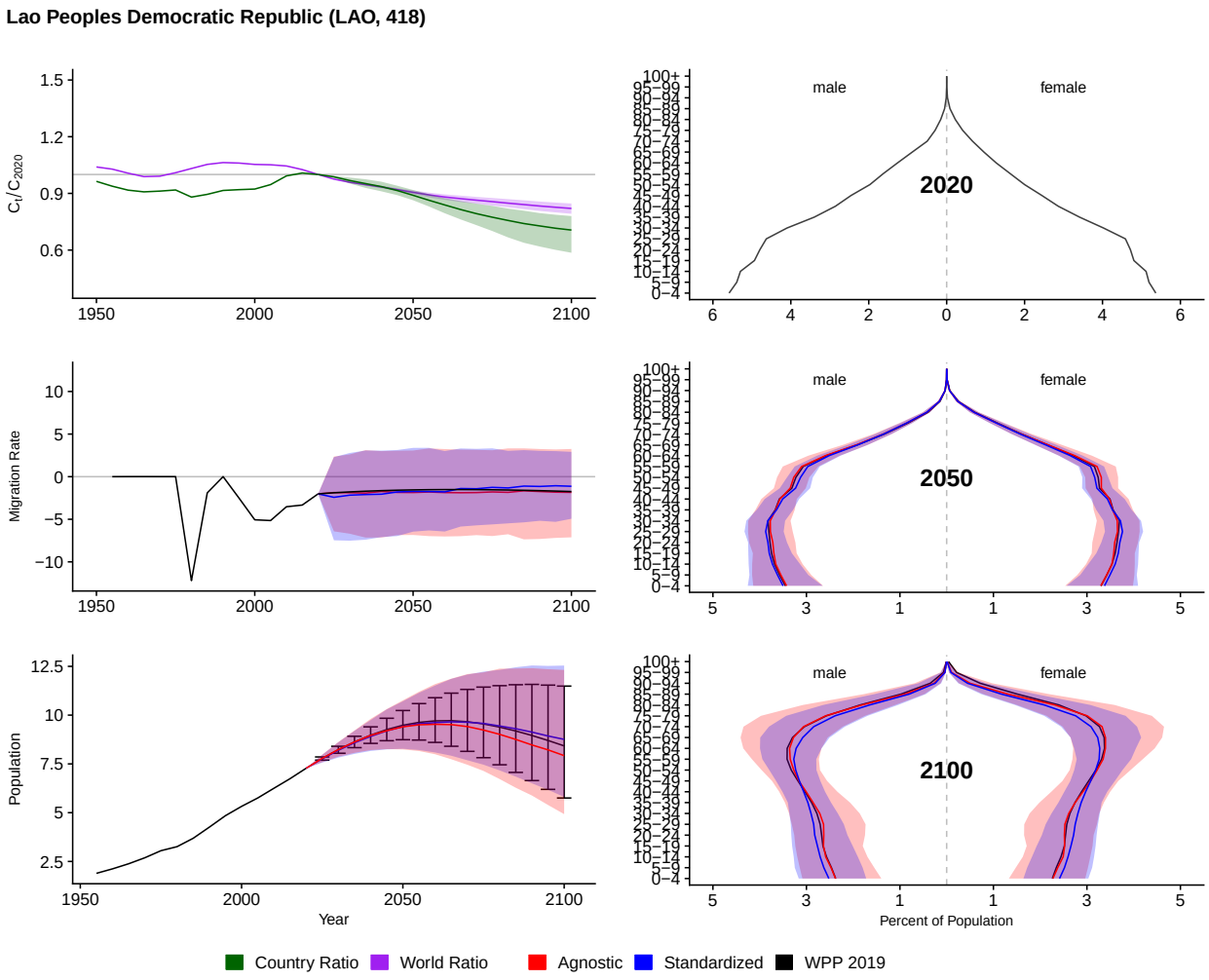


Figure A.155: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

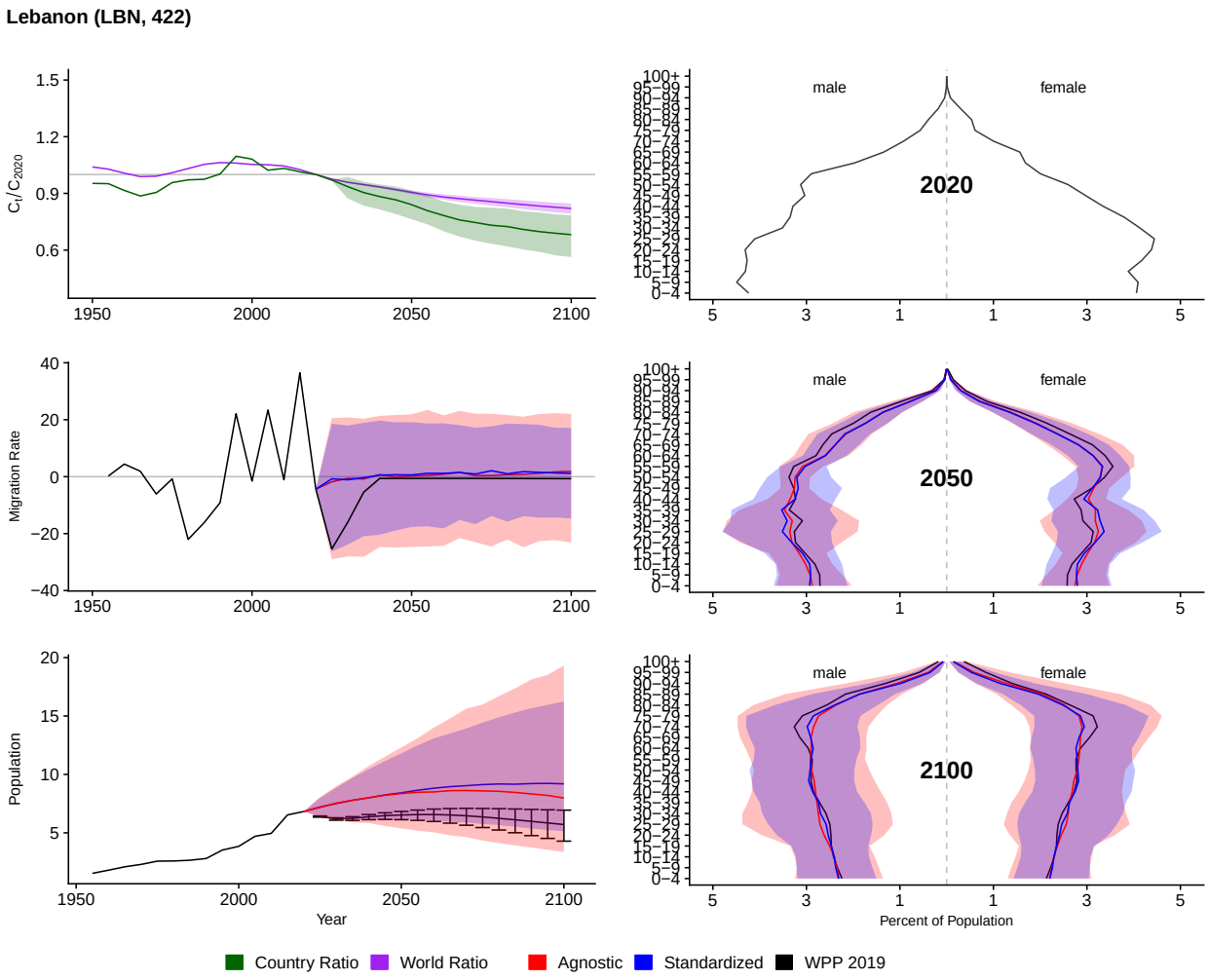


Figure A.156: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

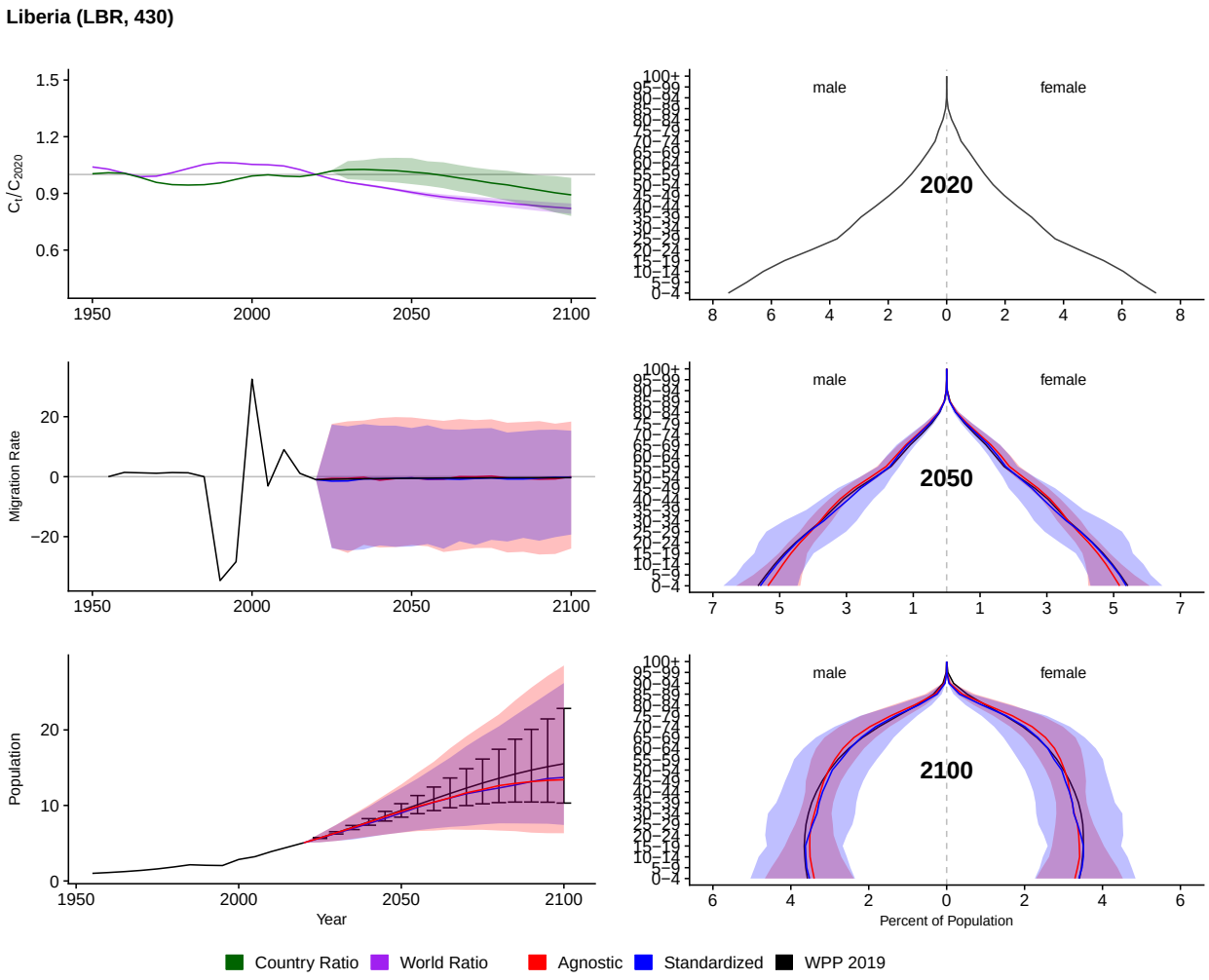


Figure A.157: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

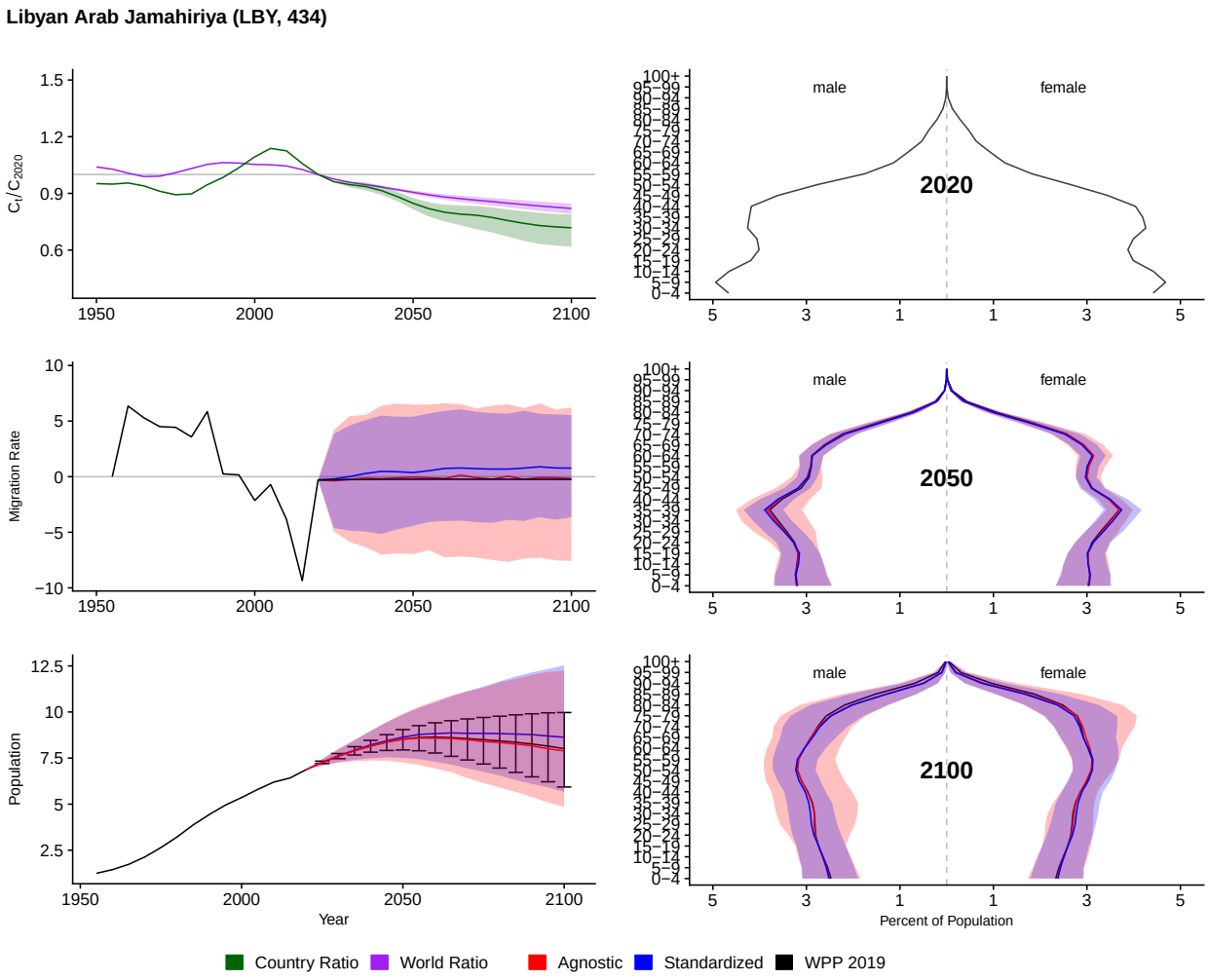


Figure A.158: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

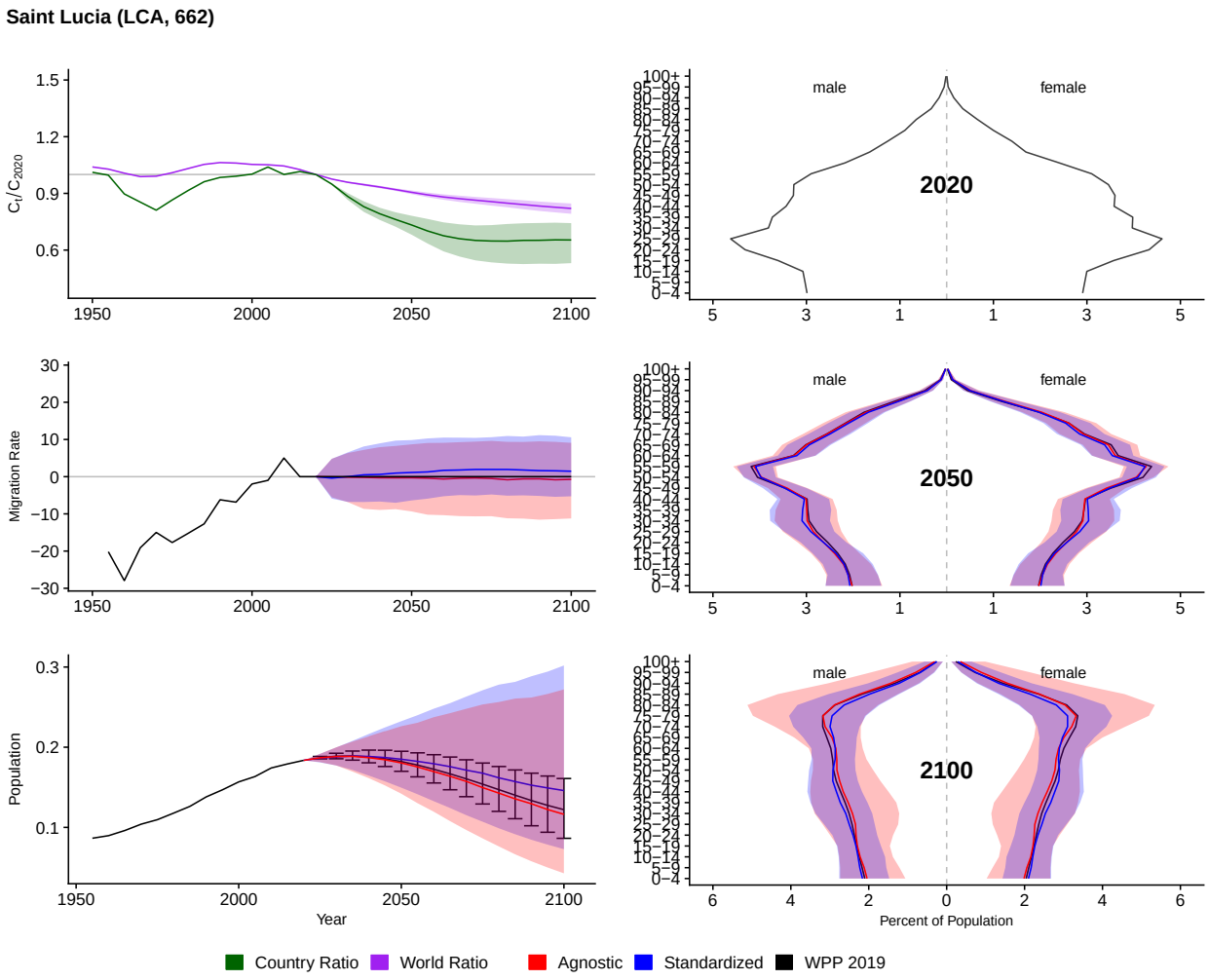


Figure A.159: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

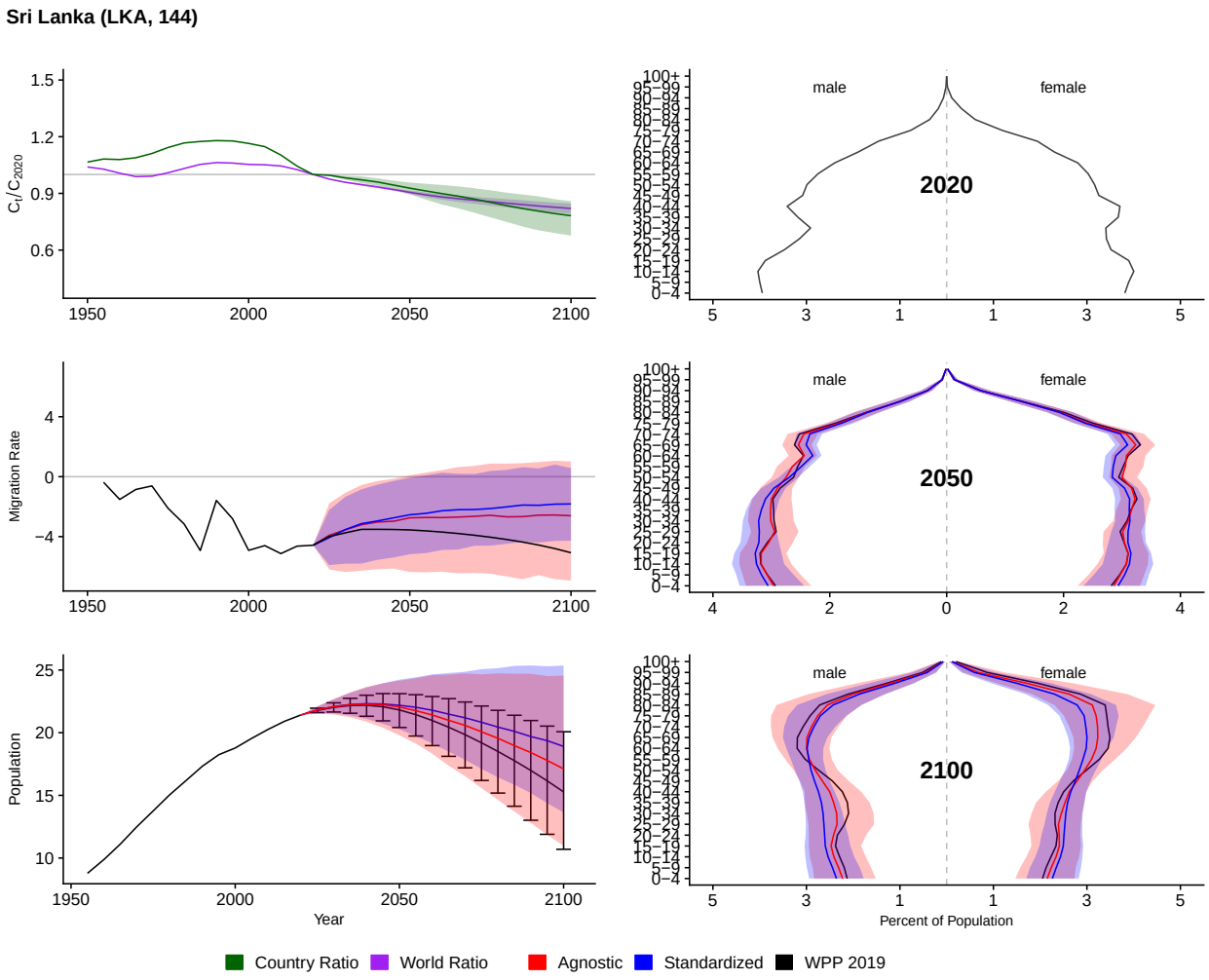


Figure A.160: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

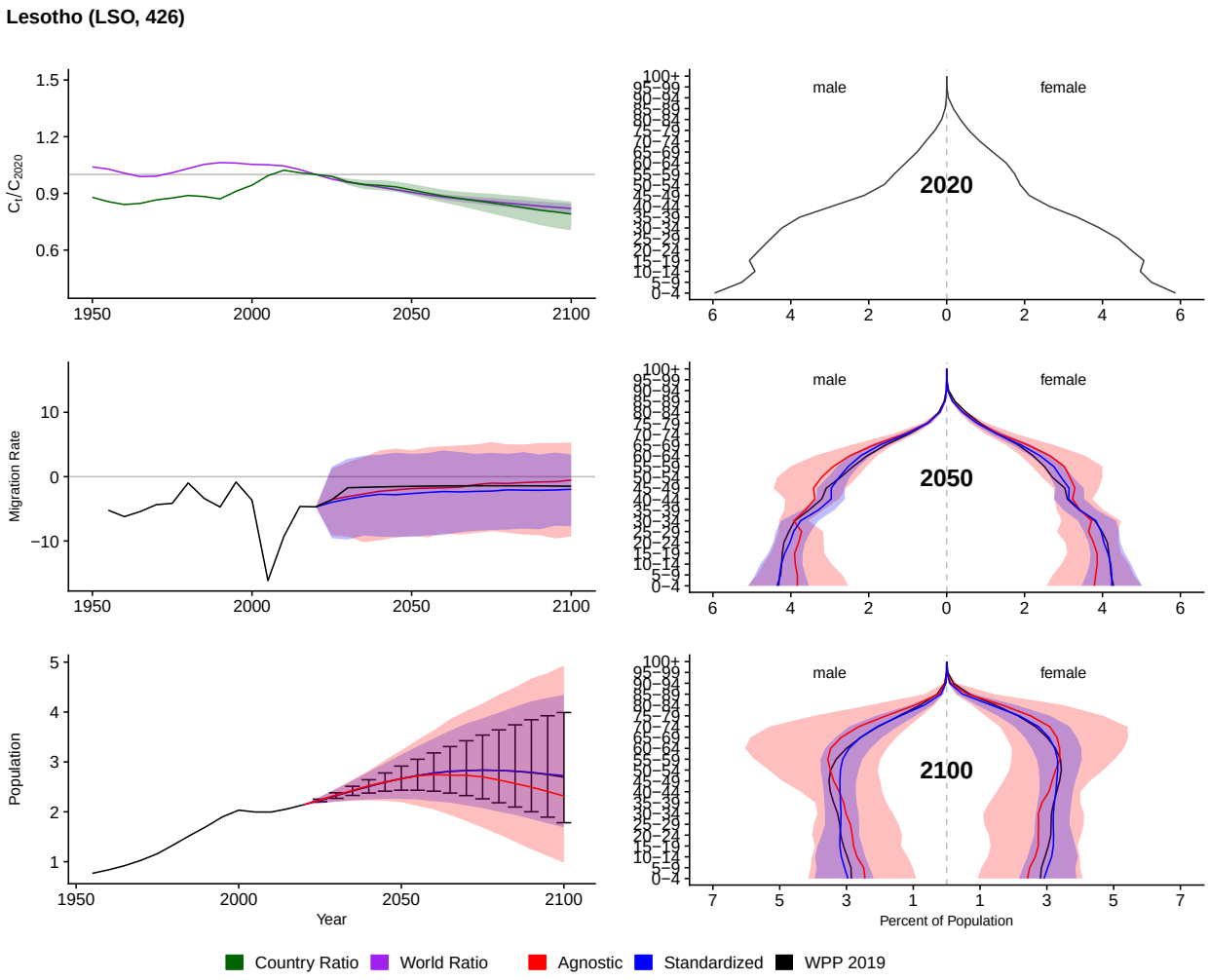


Figure A.161: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

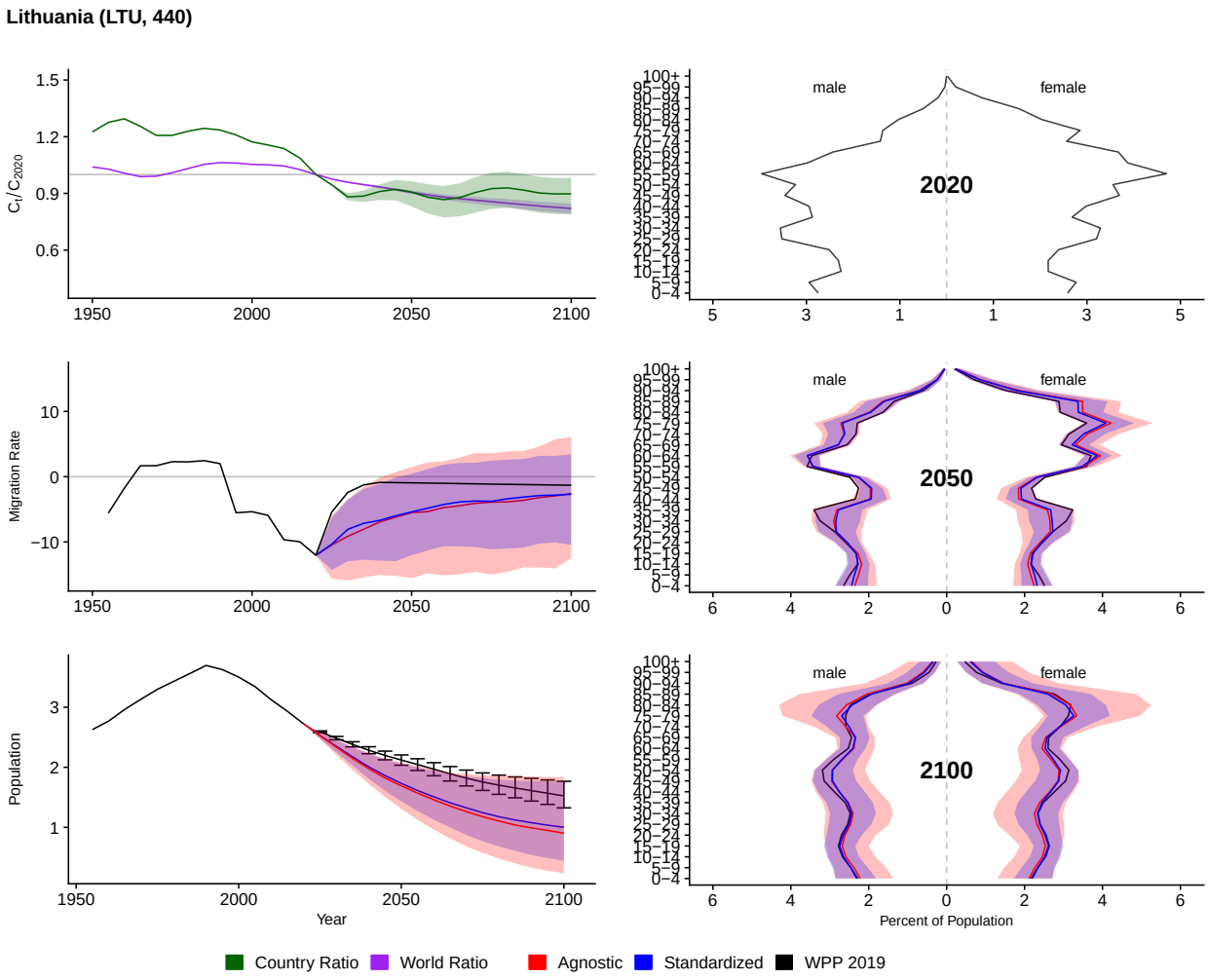


Figure A.162: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

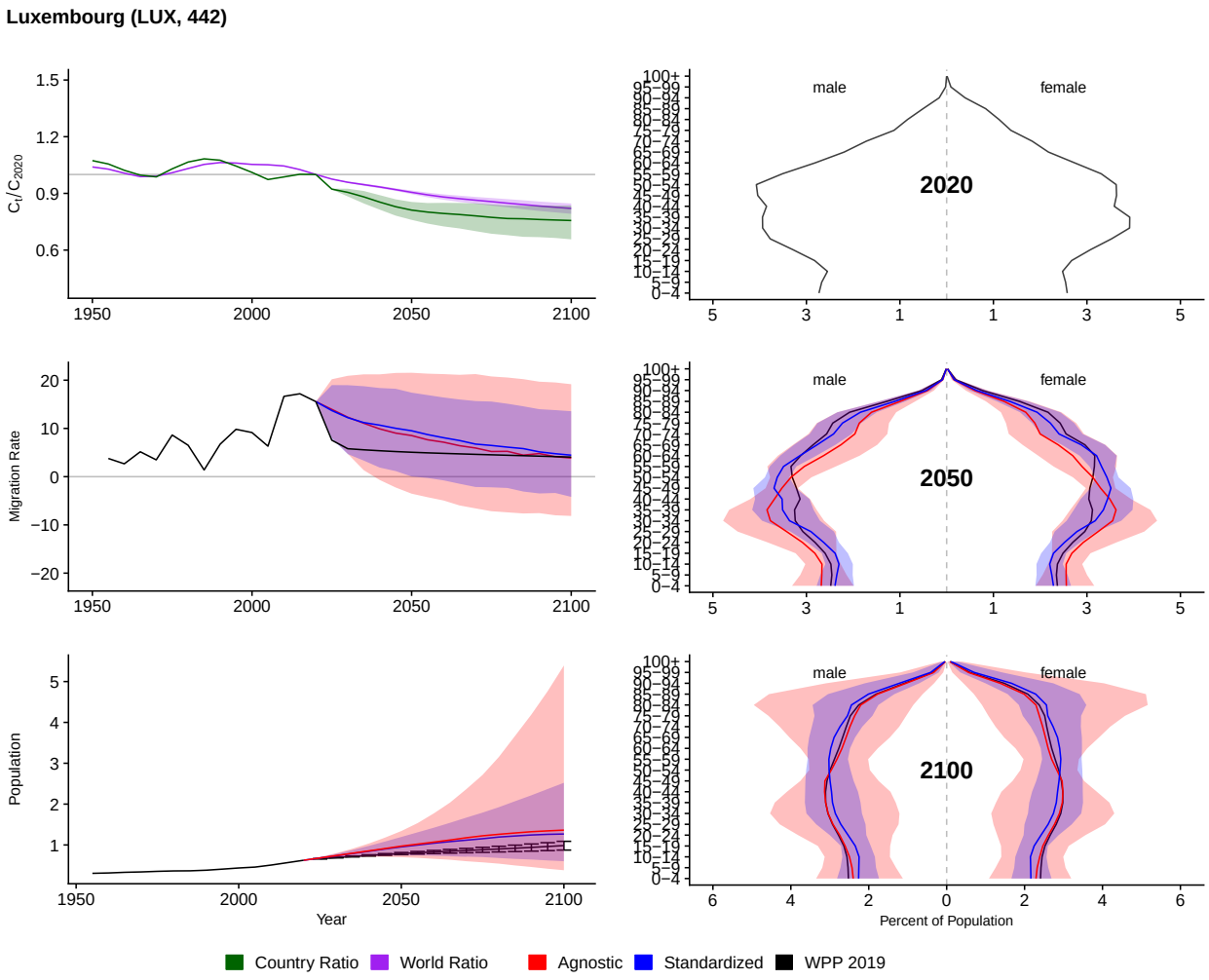


Figure A.163: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

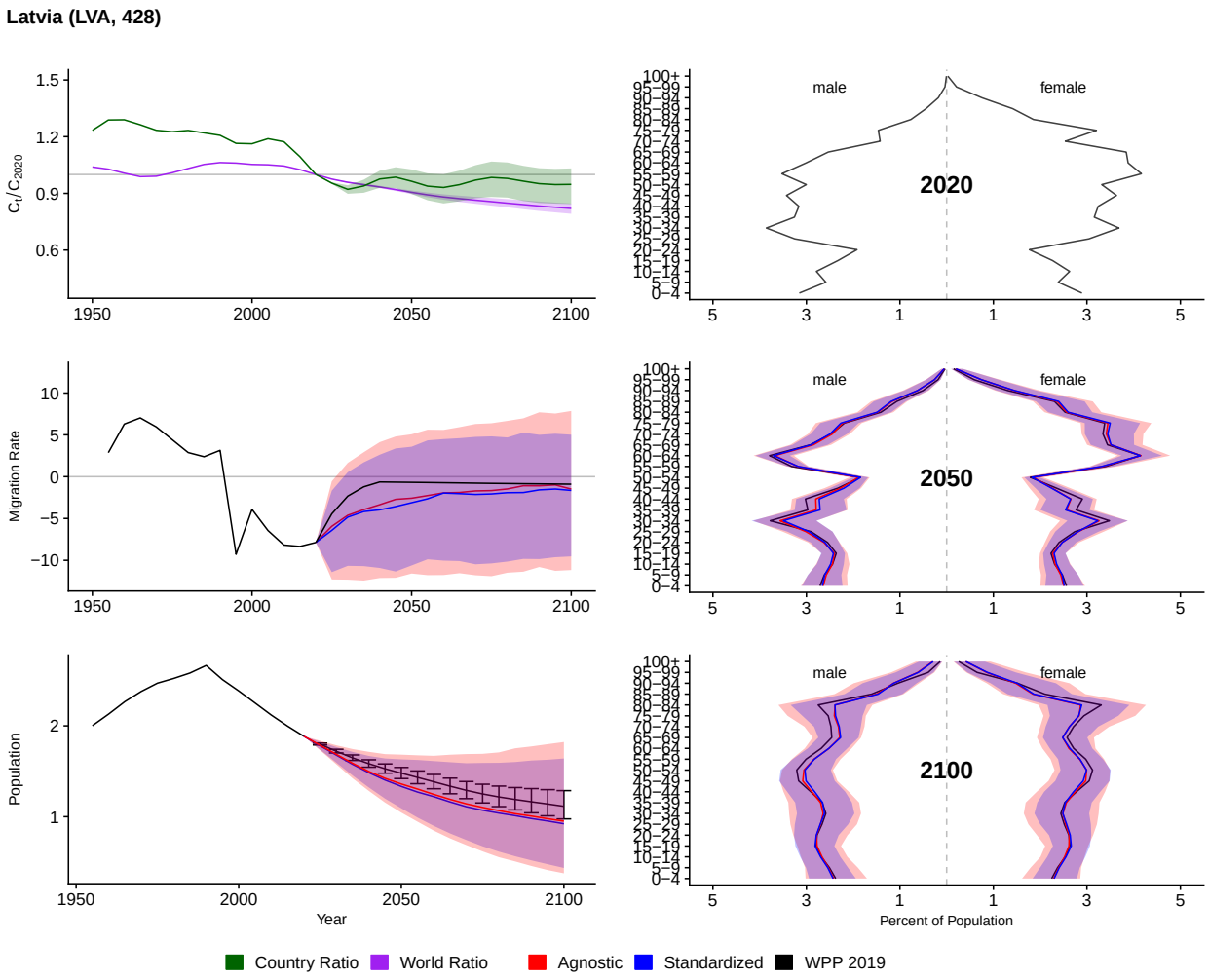


Figure A.164: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

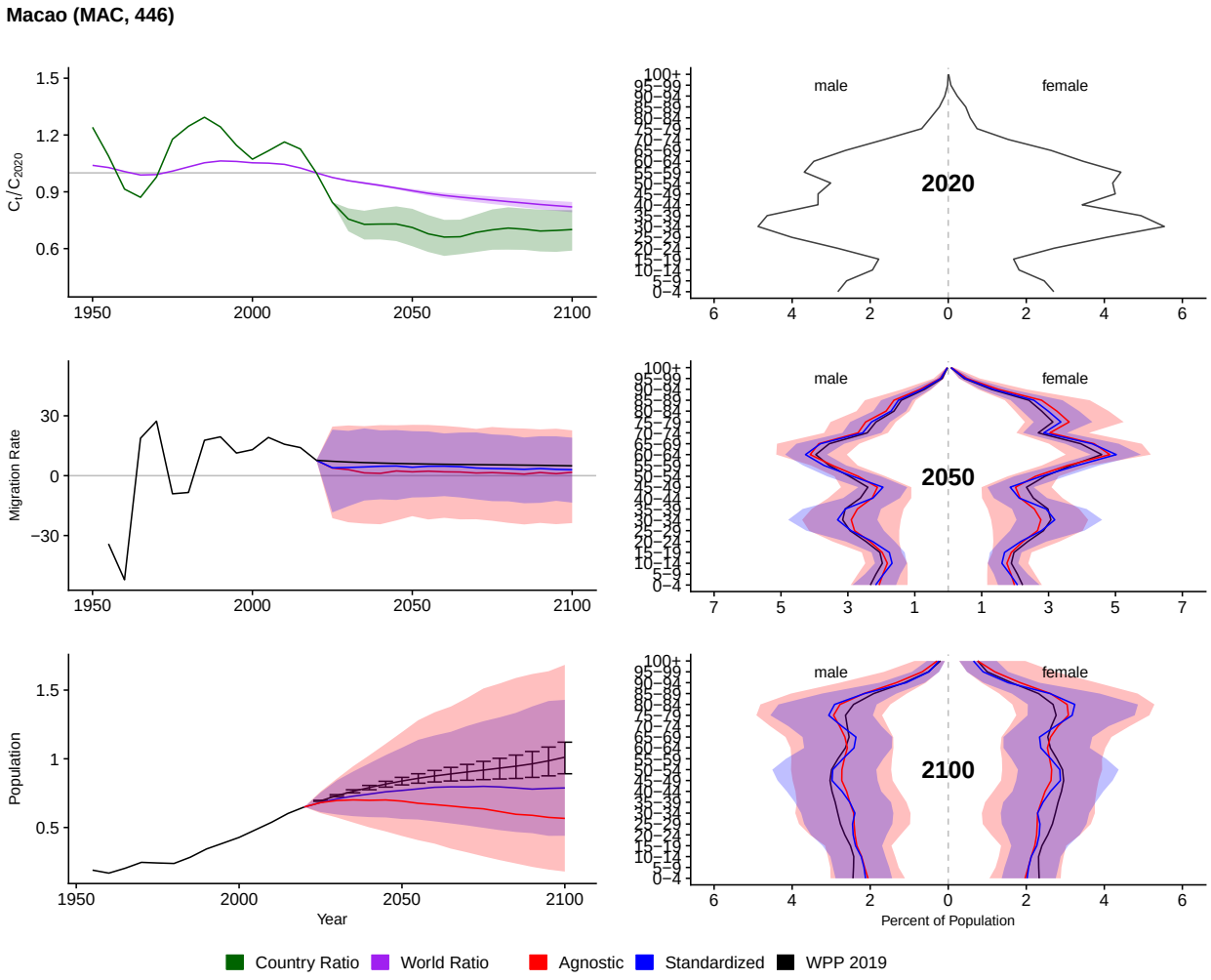


Figure A.165: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

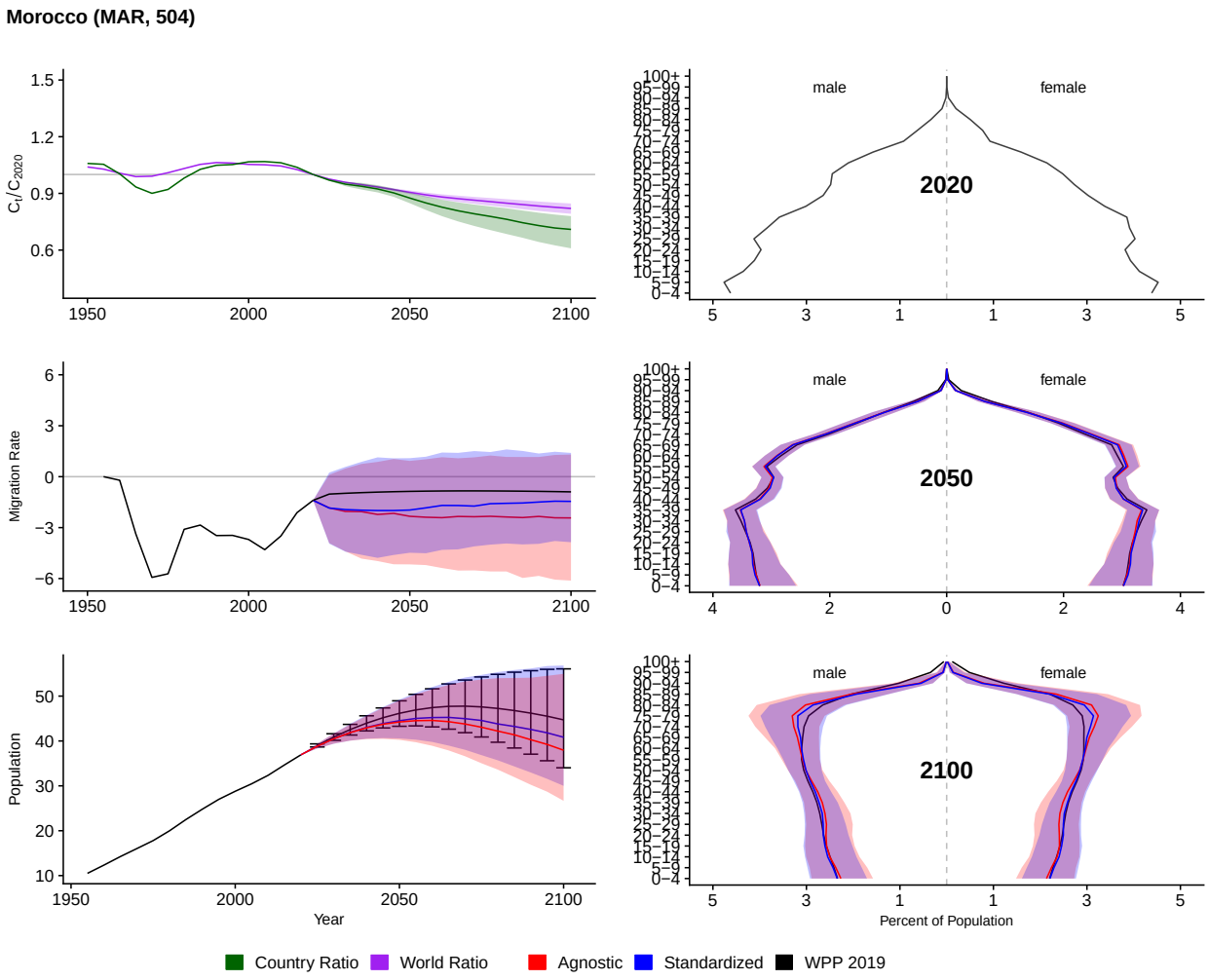


Figure A.166: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

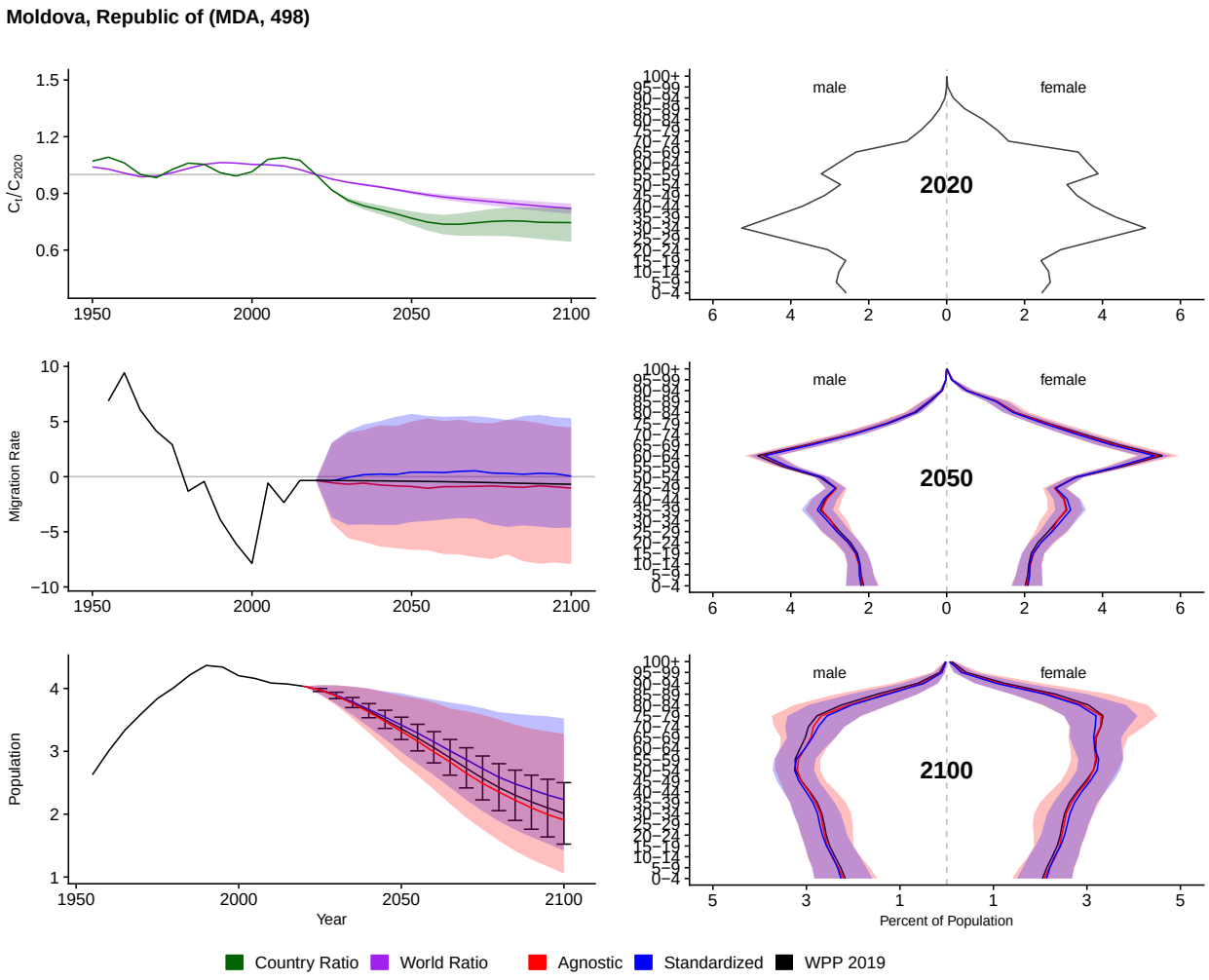


Figure A.167: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

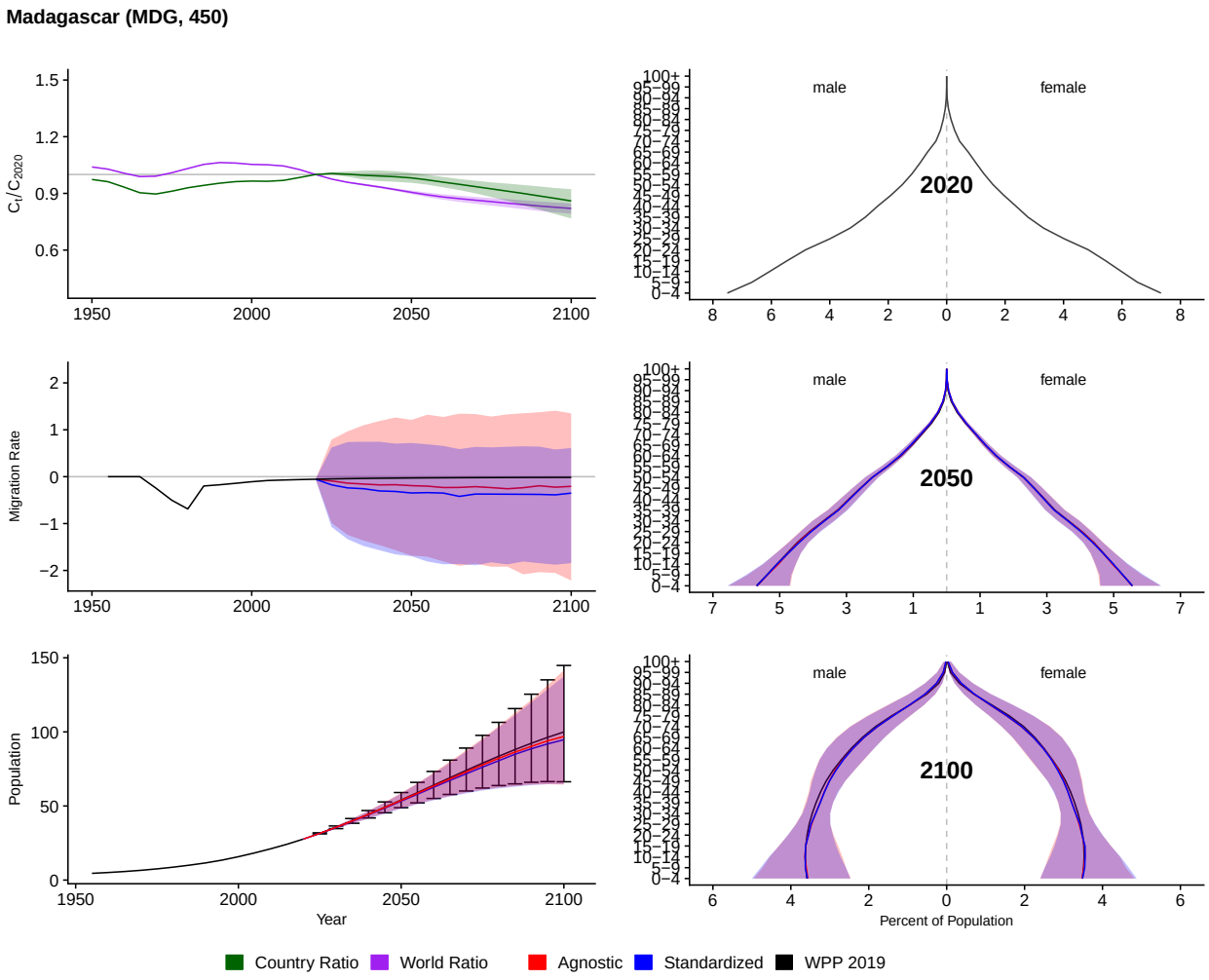


Figure A.168: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

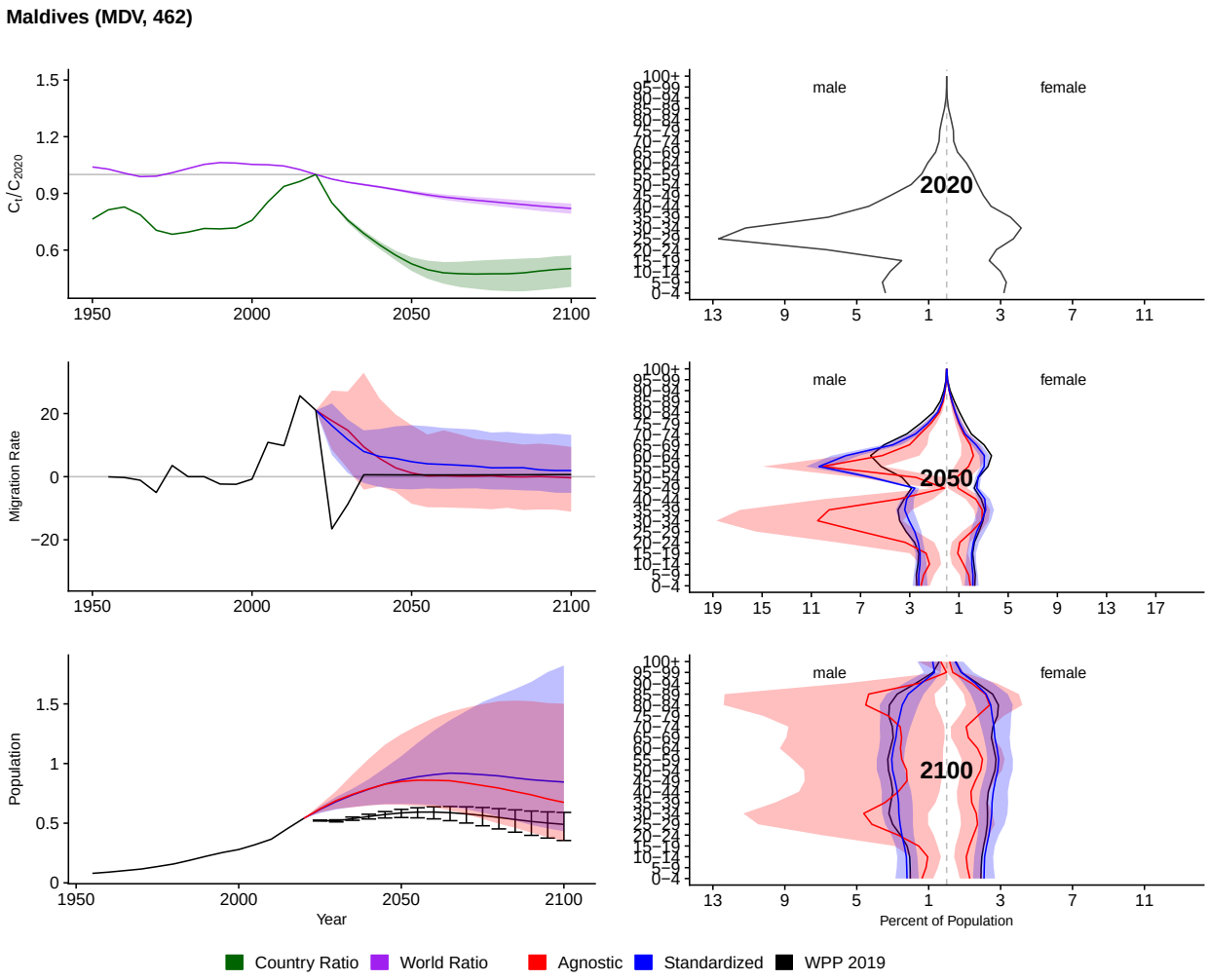


Figure A.169: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

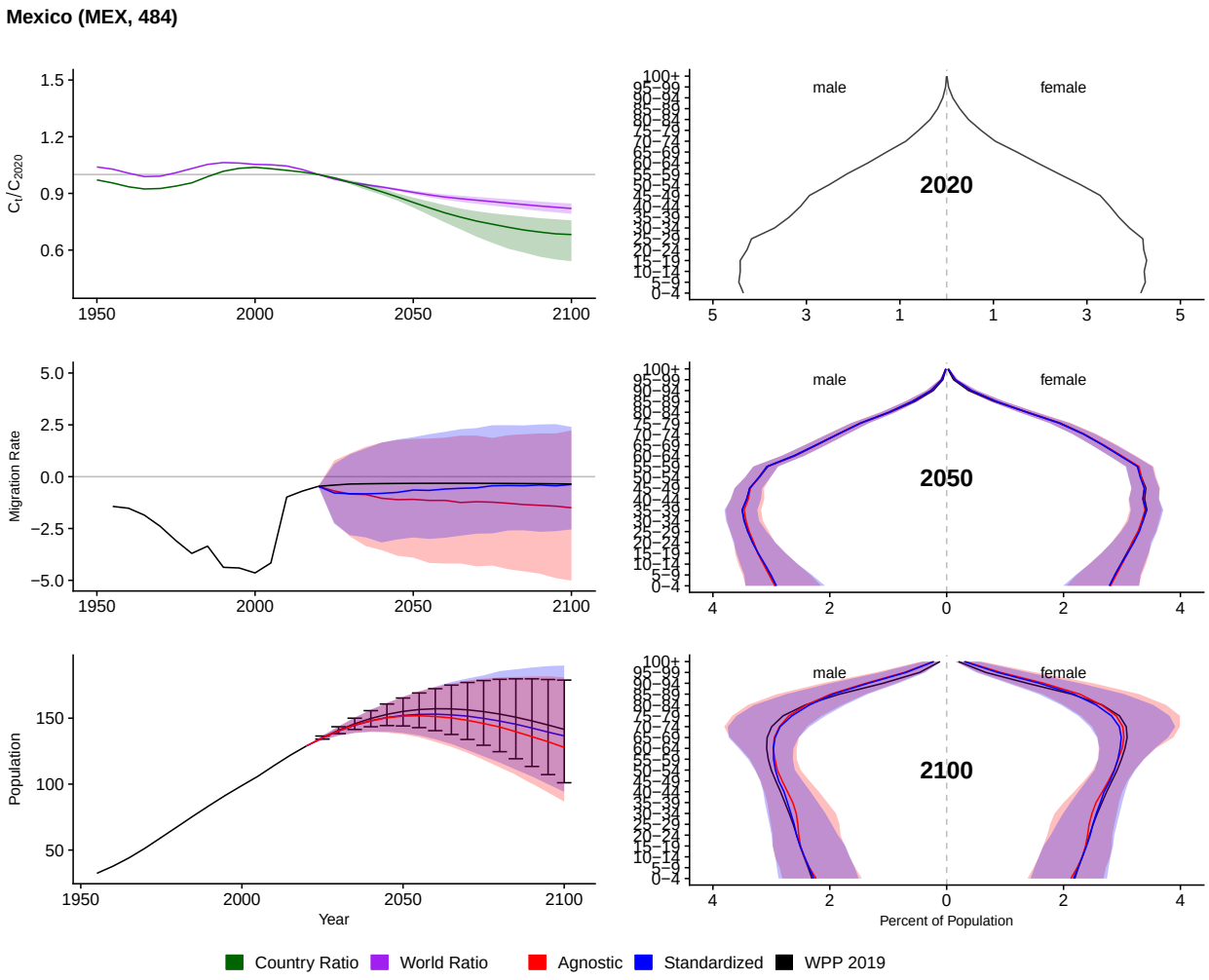


Figure A.170: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

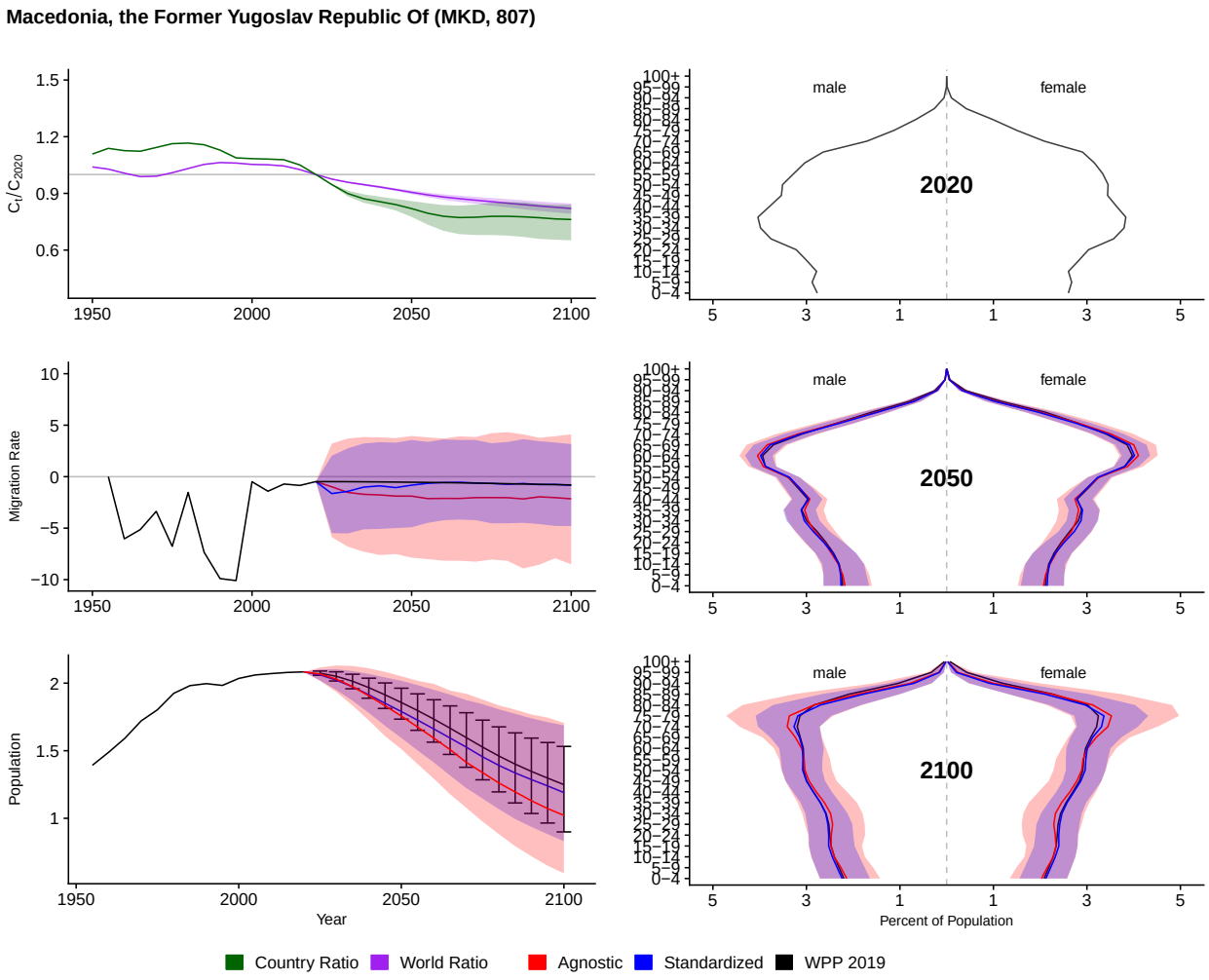


Figure A.171: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

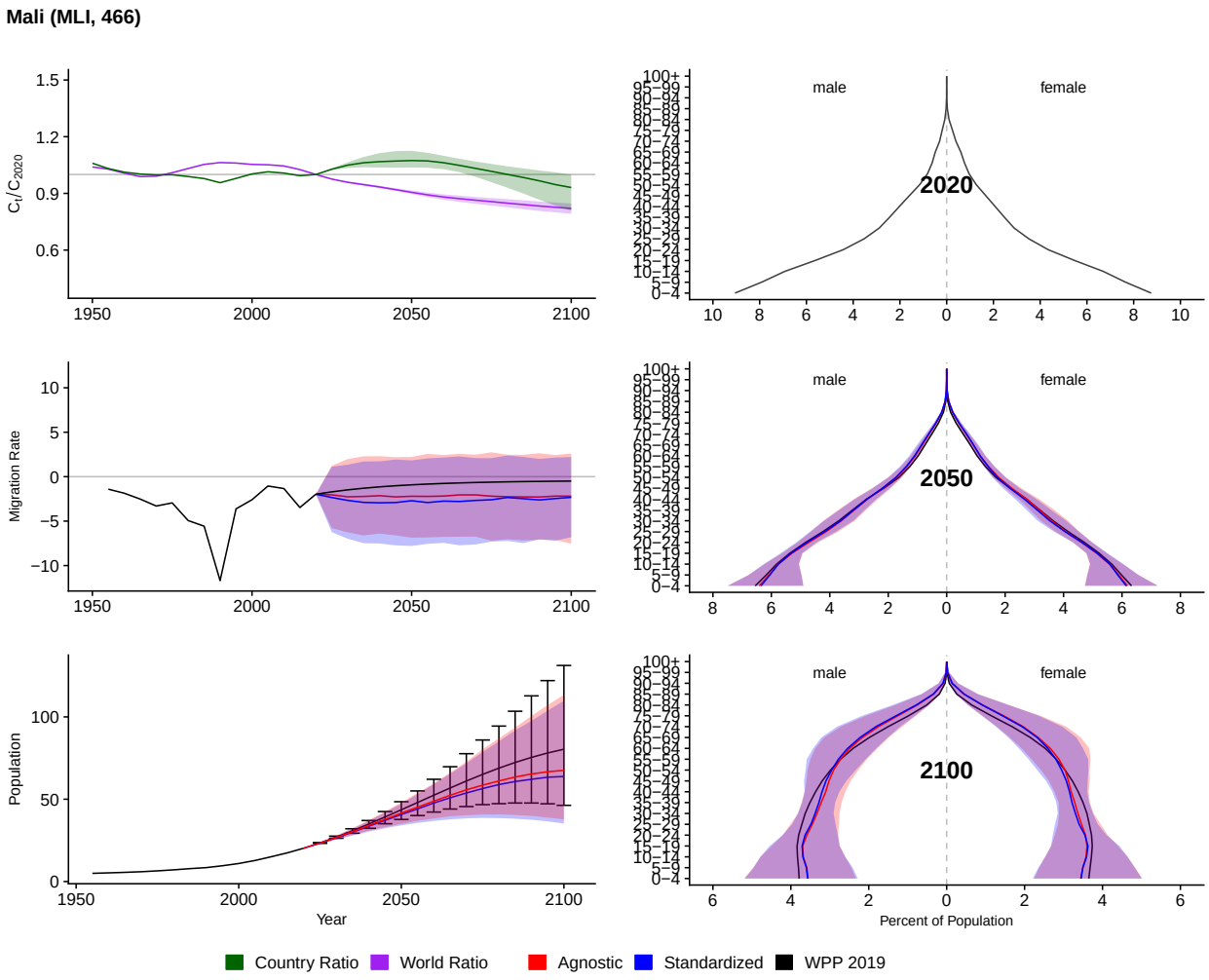


Figure A.172: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

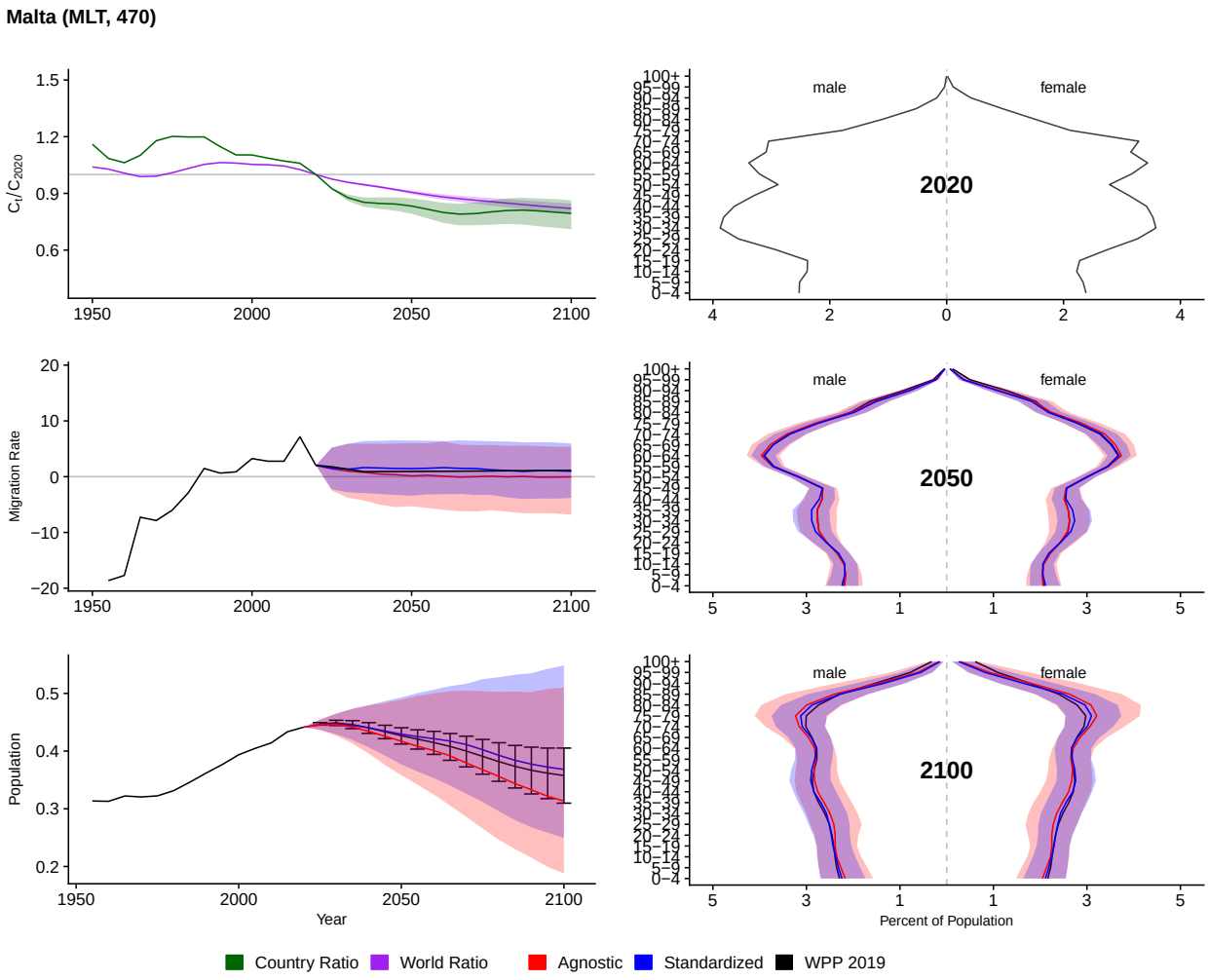


Figure A.173: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

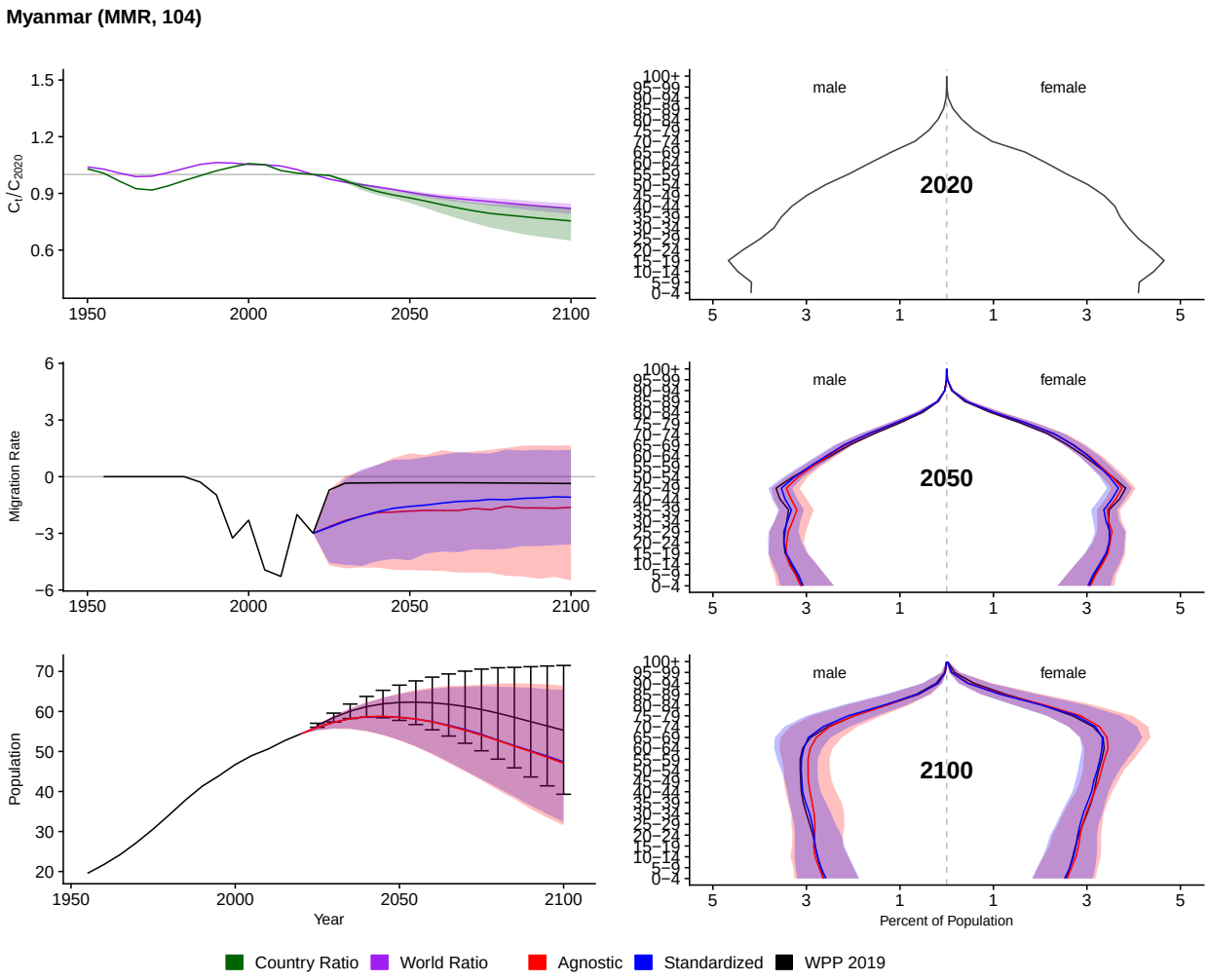


Figure A.174: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

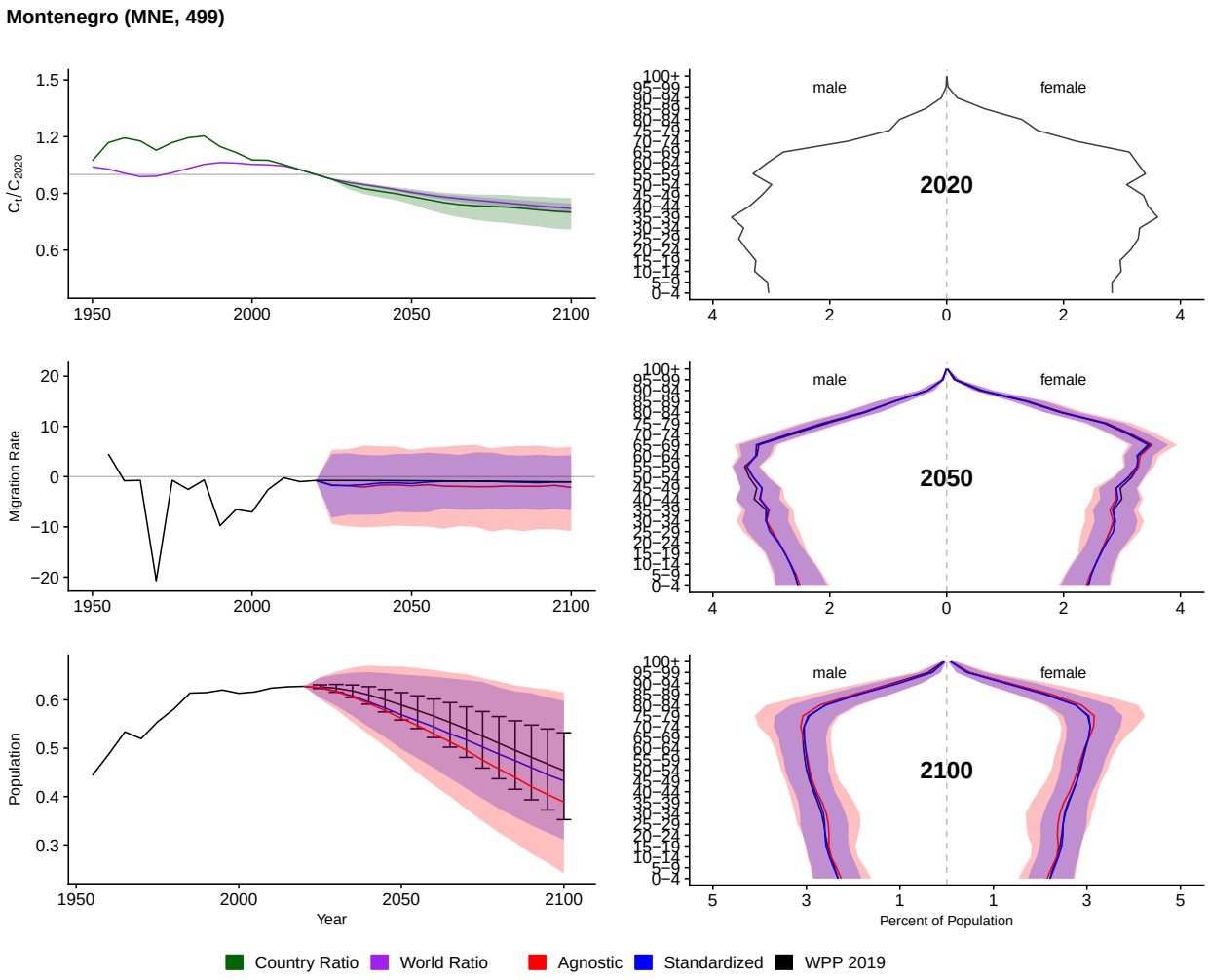


Figure A.175: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

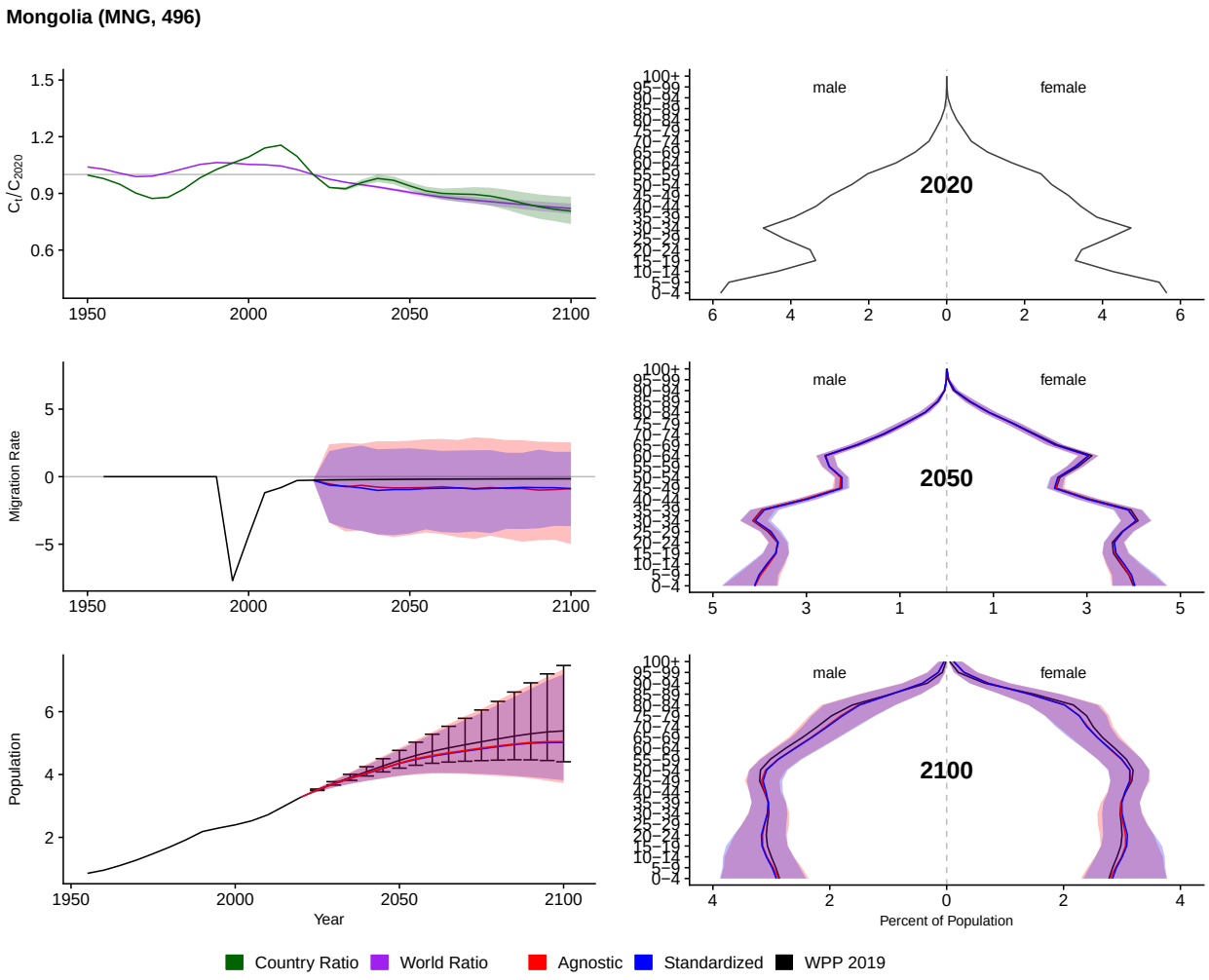


Figure A.176: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

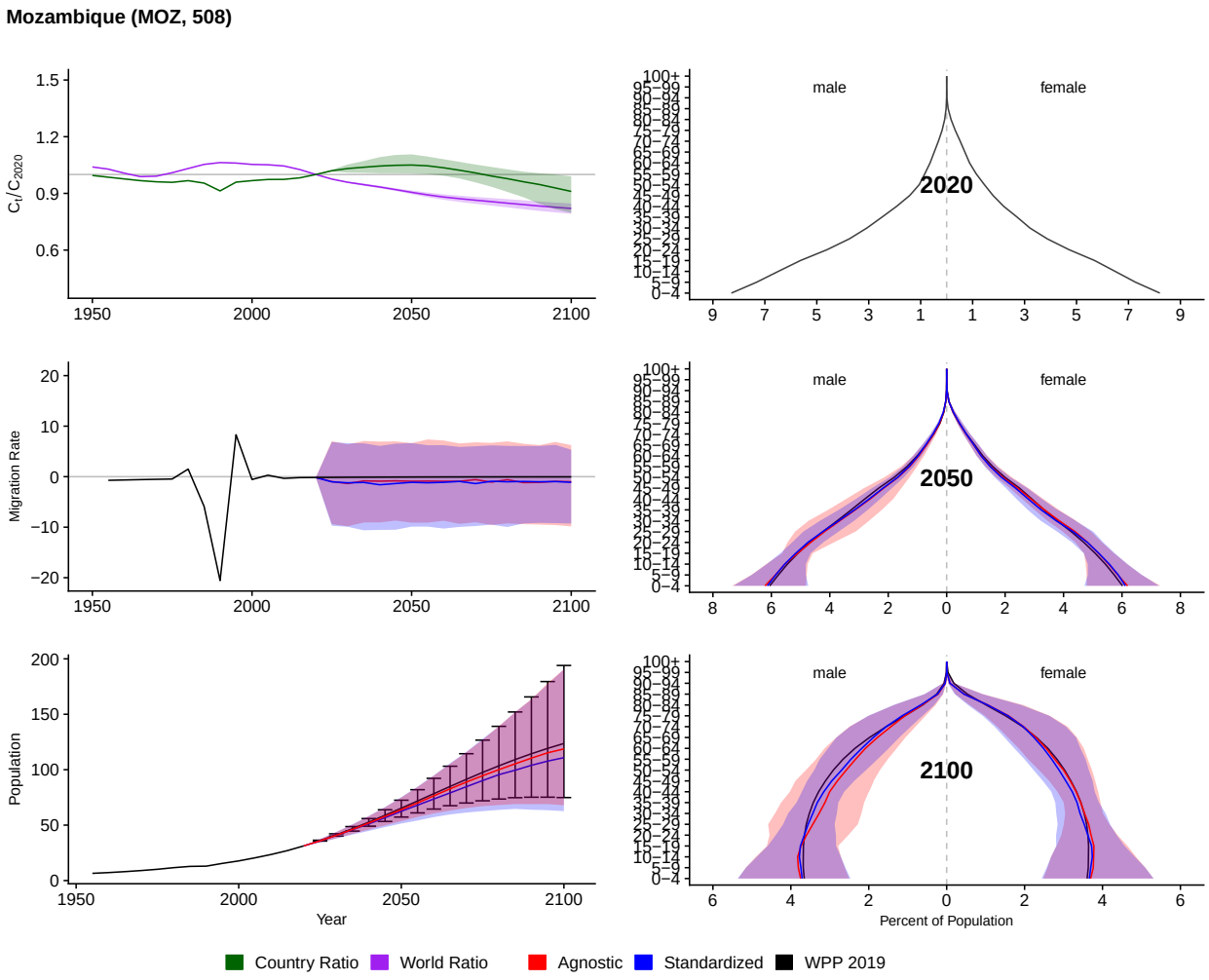


Figure A.177: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

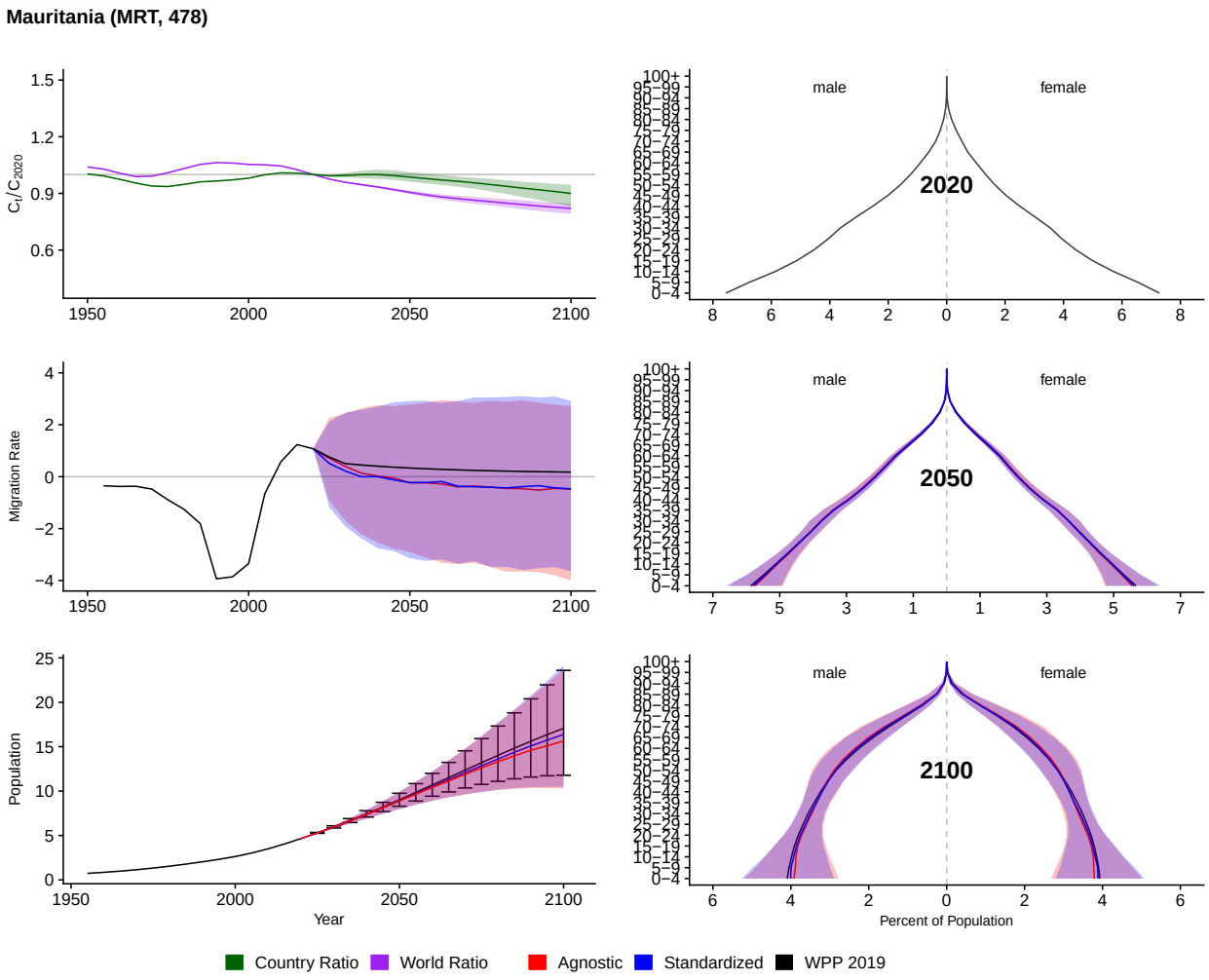


Figure A.178: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

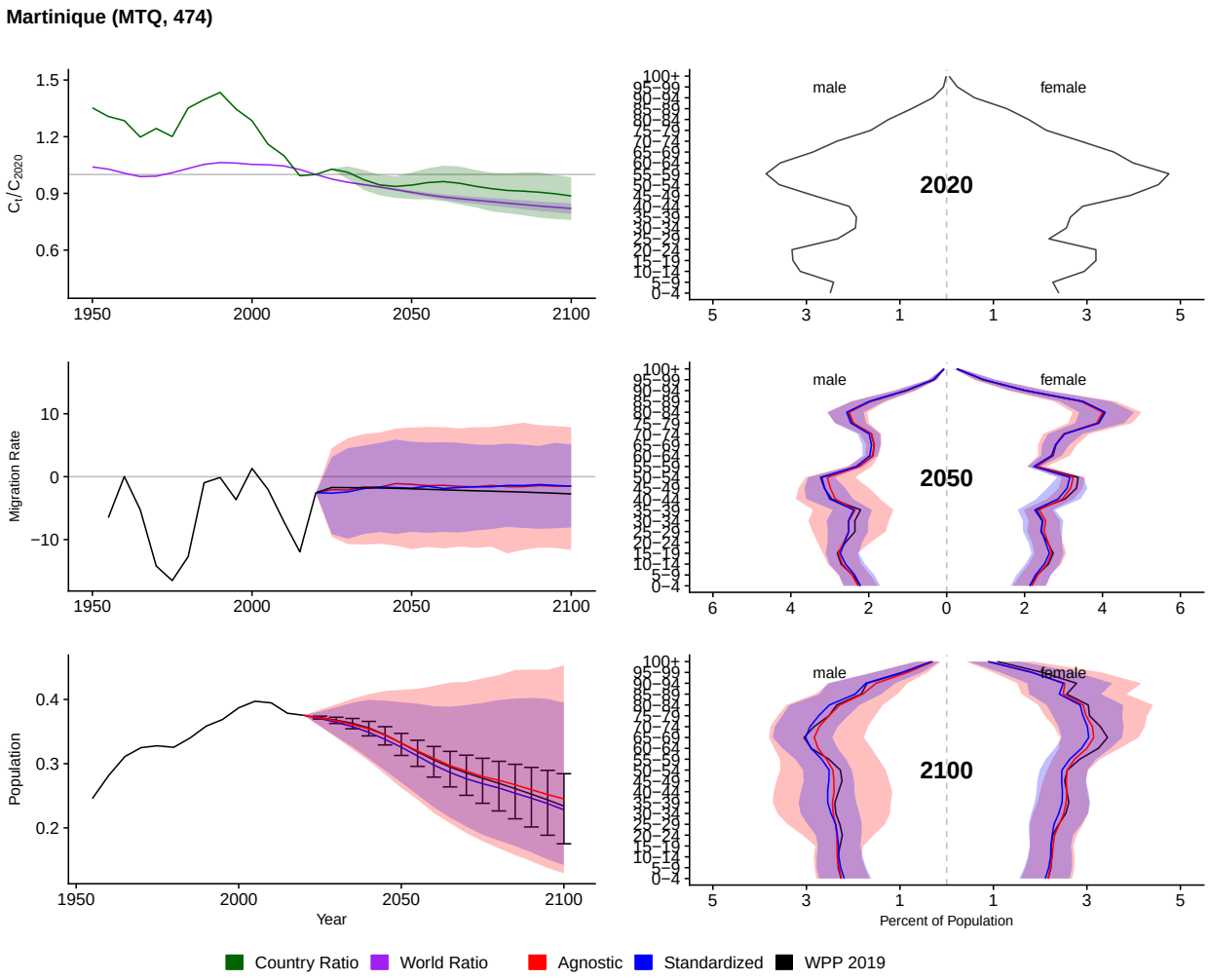


Figure A.179: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

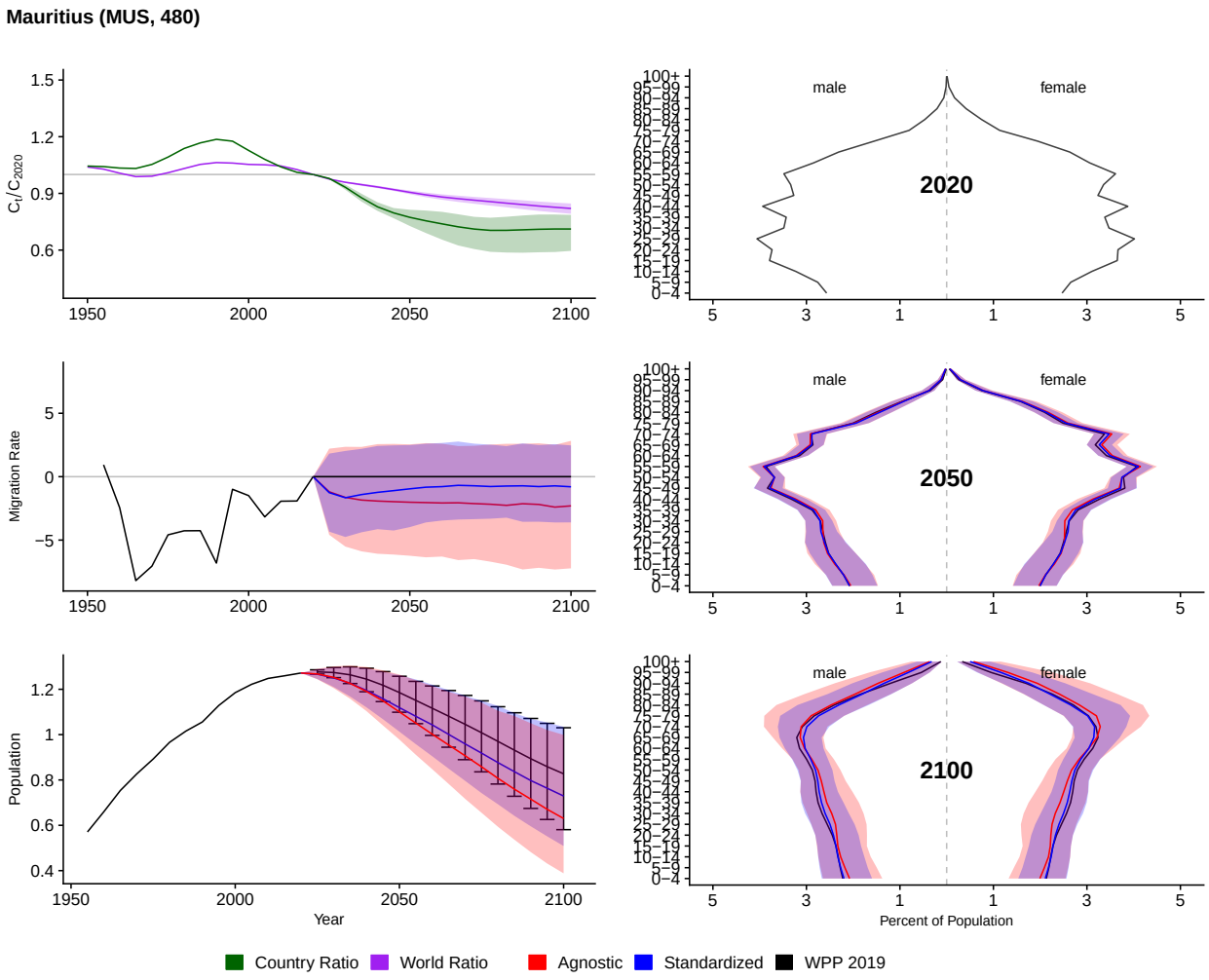


Figure A.180: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

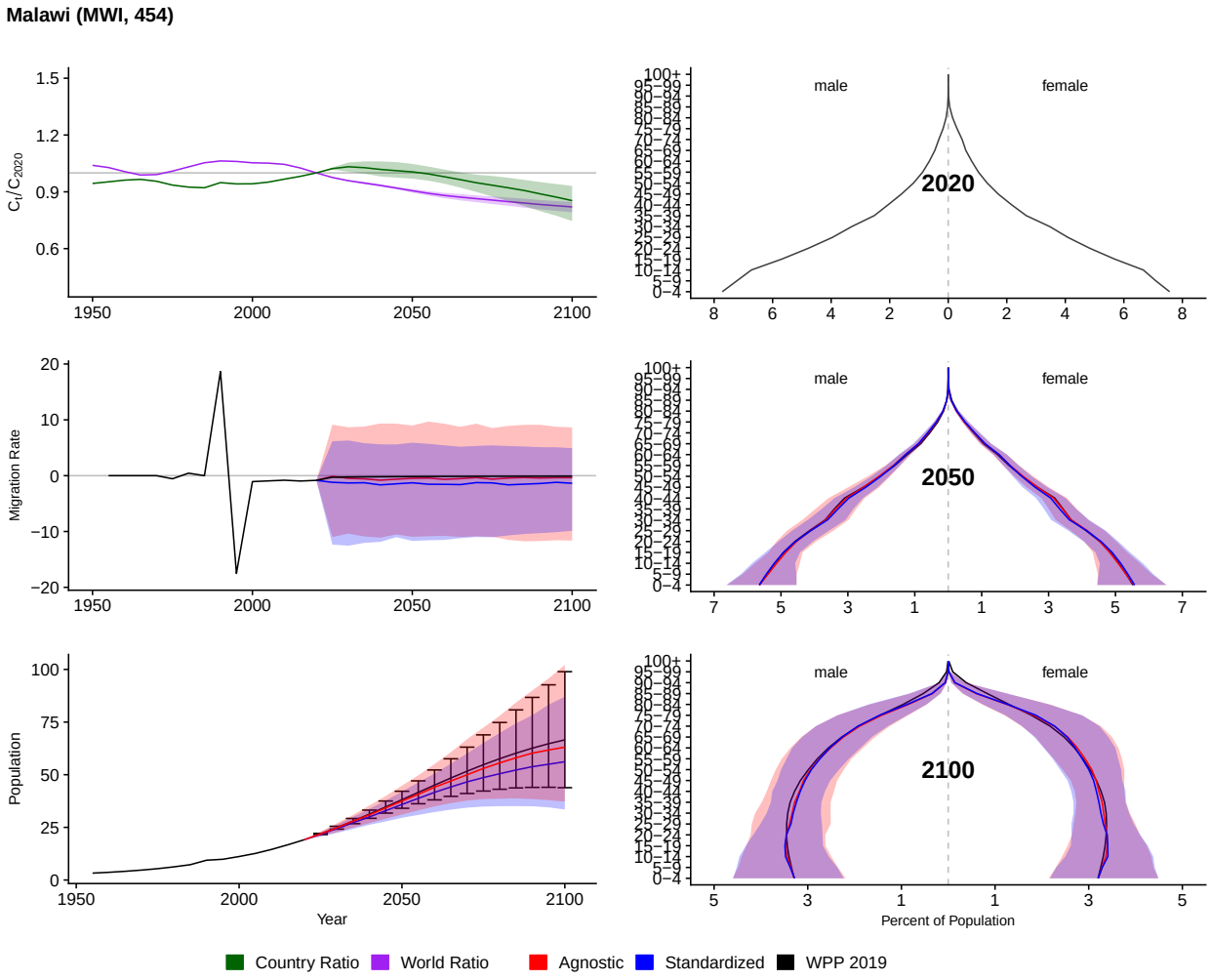


Figure A.181: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

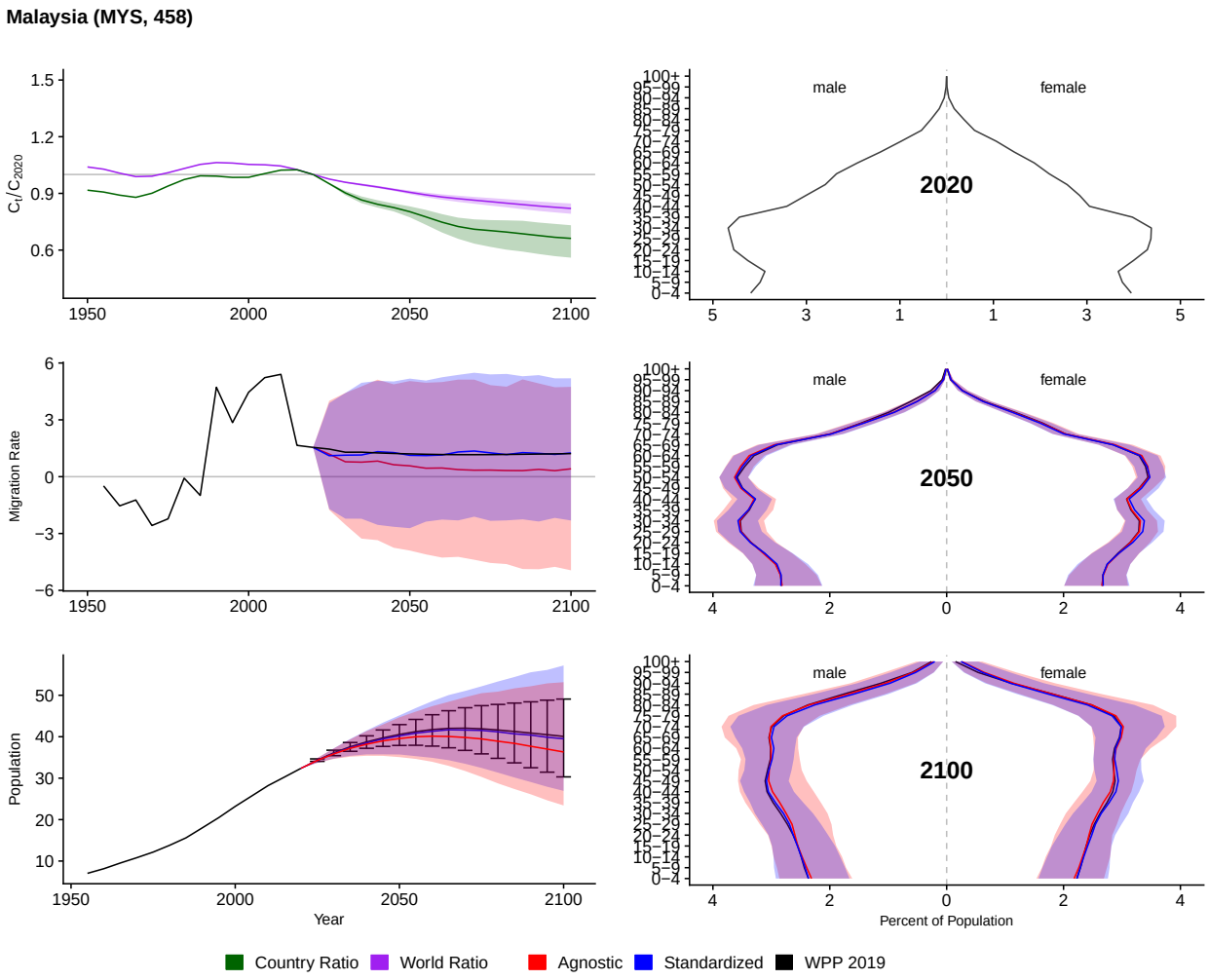


Figure A.182: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

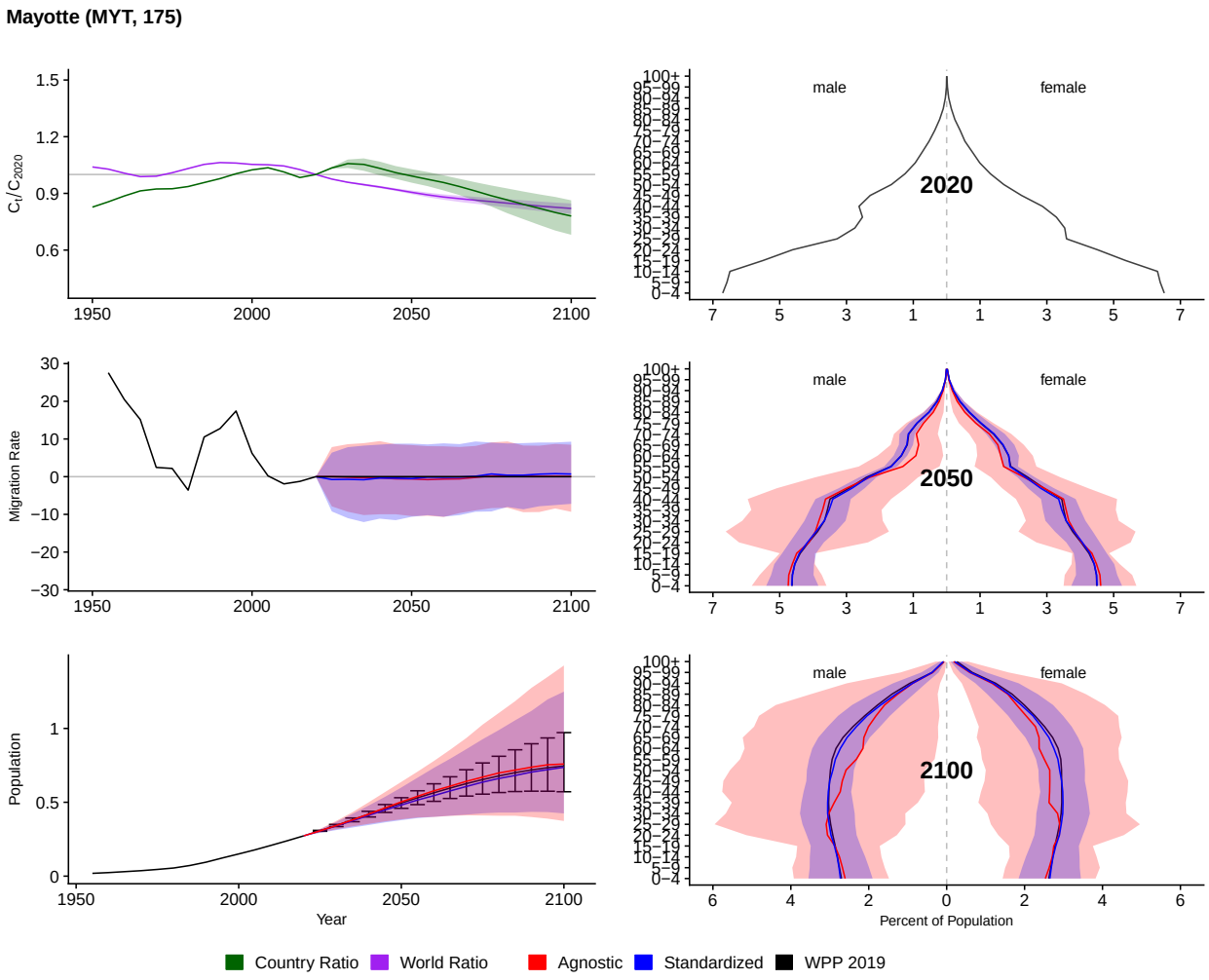


Figure A.183: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

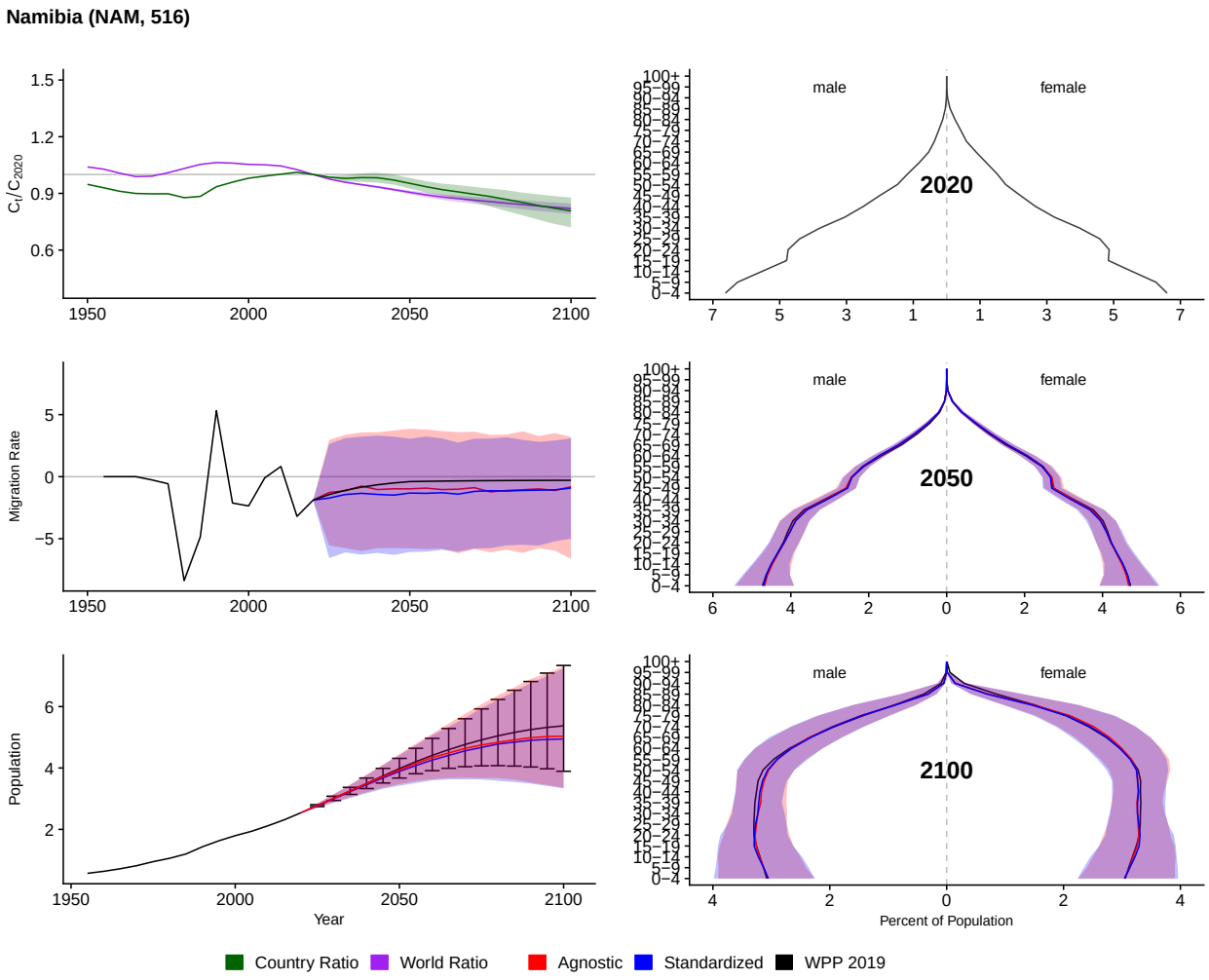


Figure A.184: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

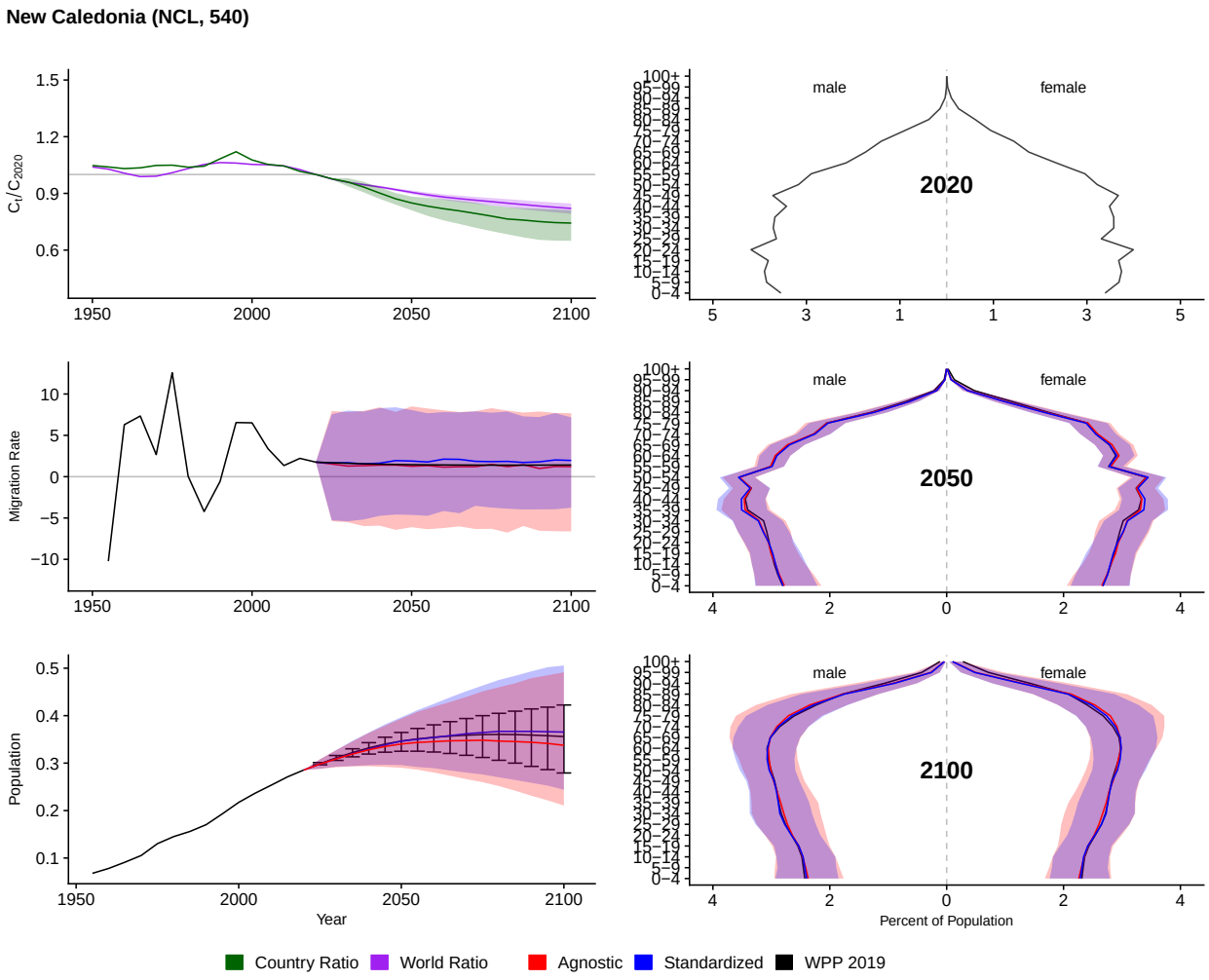


Figure A.185: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

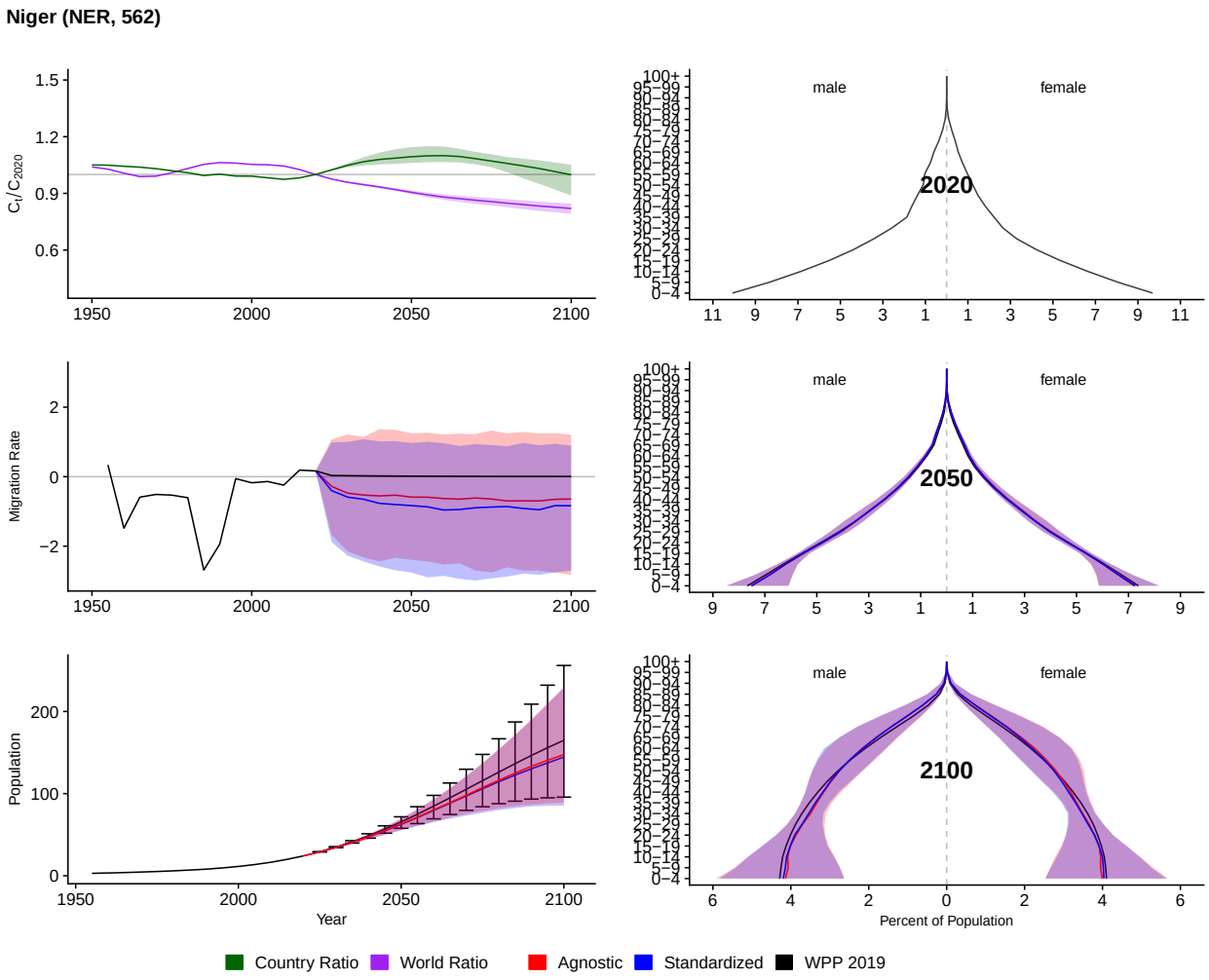


Figure A.186: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

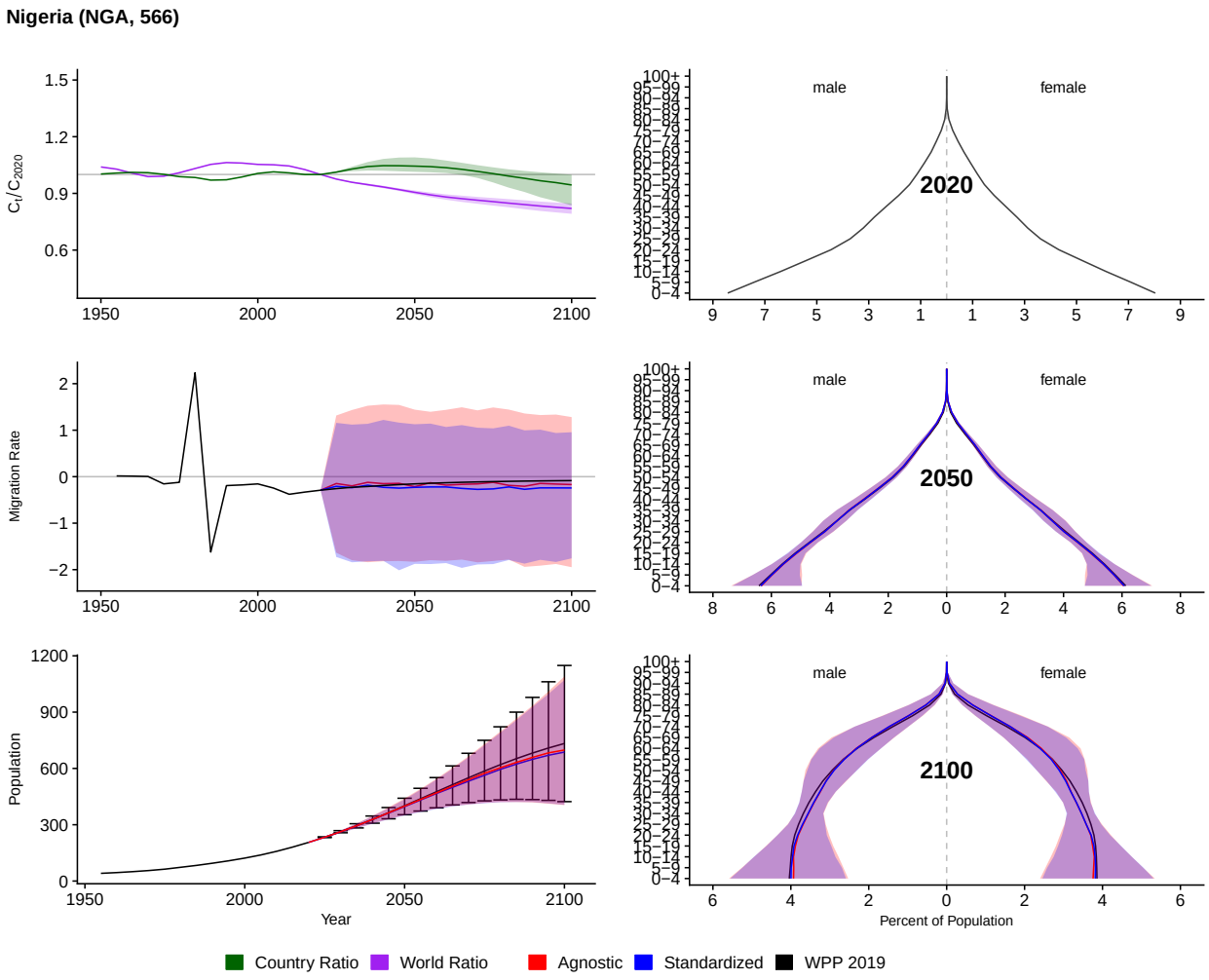


Figure A.187: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

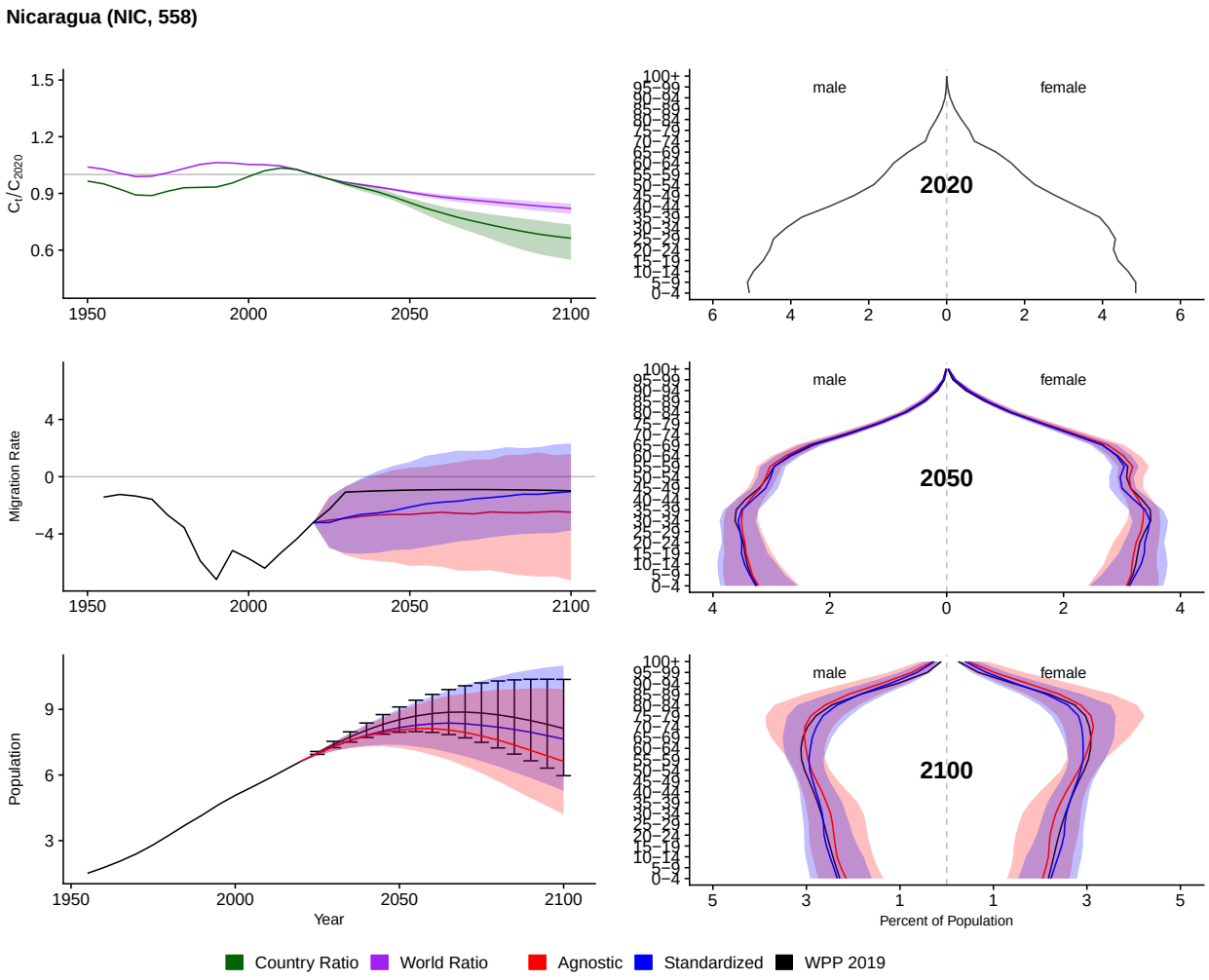


Figure A.188: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

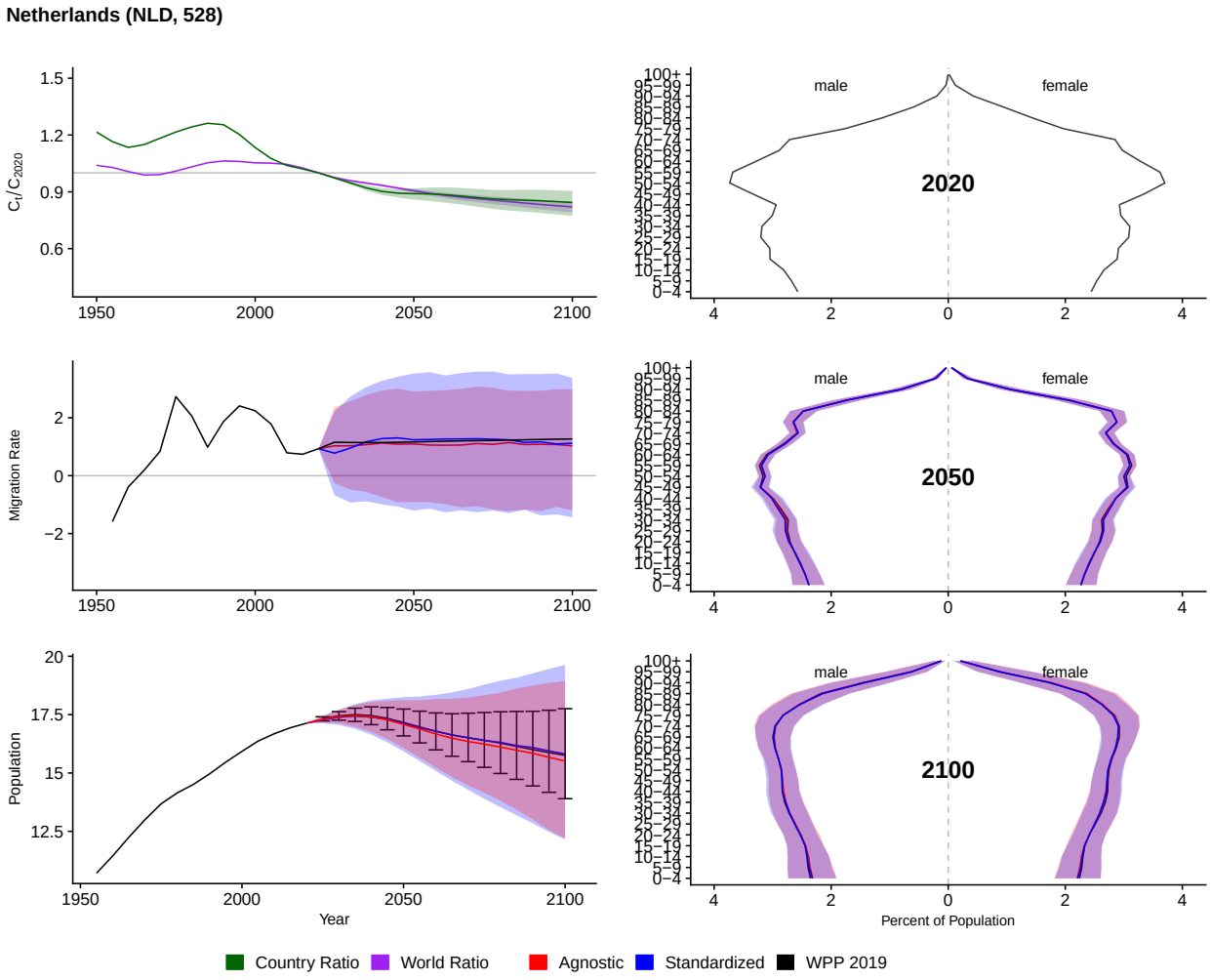


Figure A.189: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

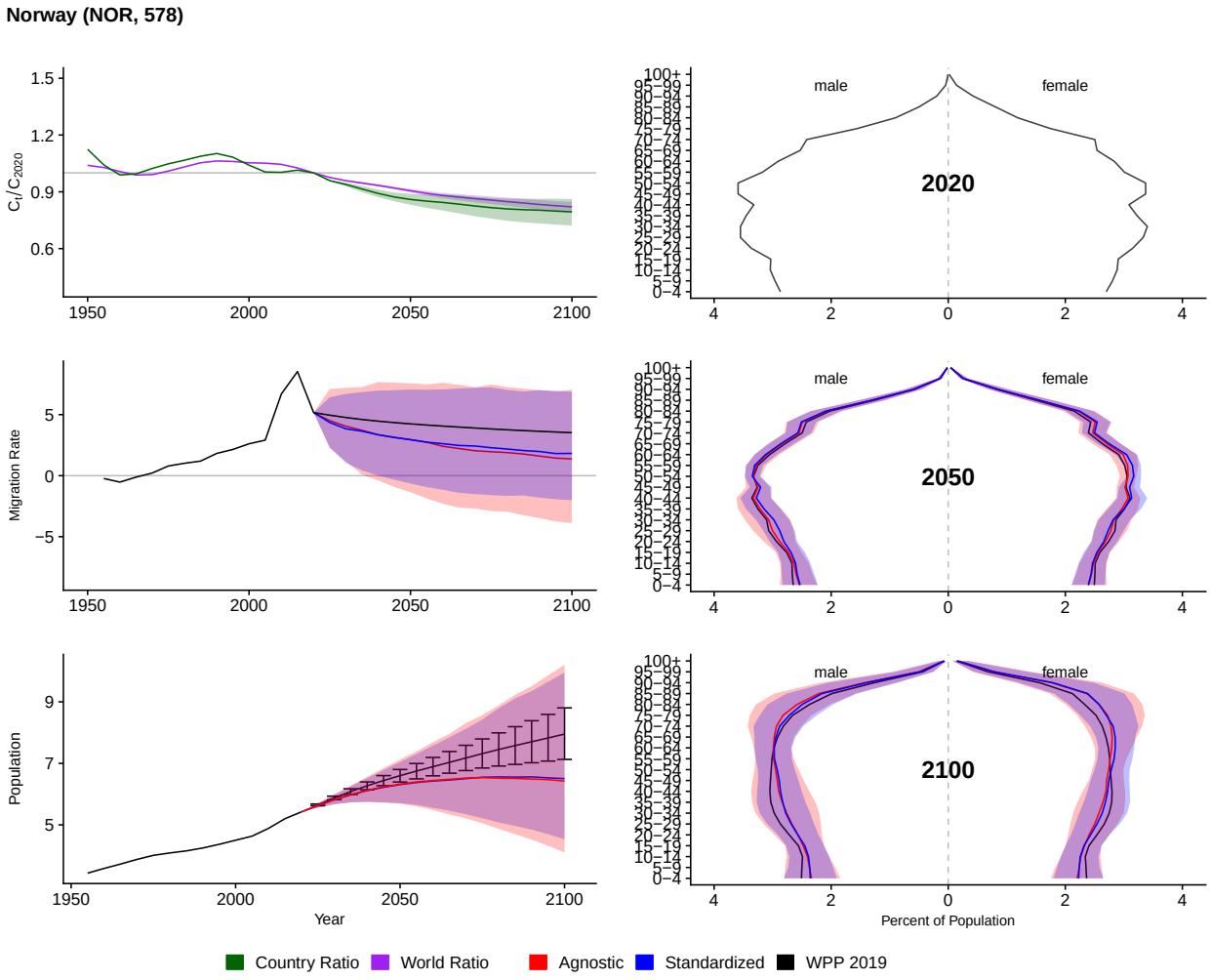


Figure A.190: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

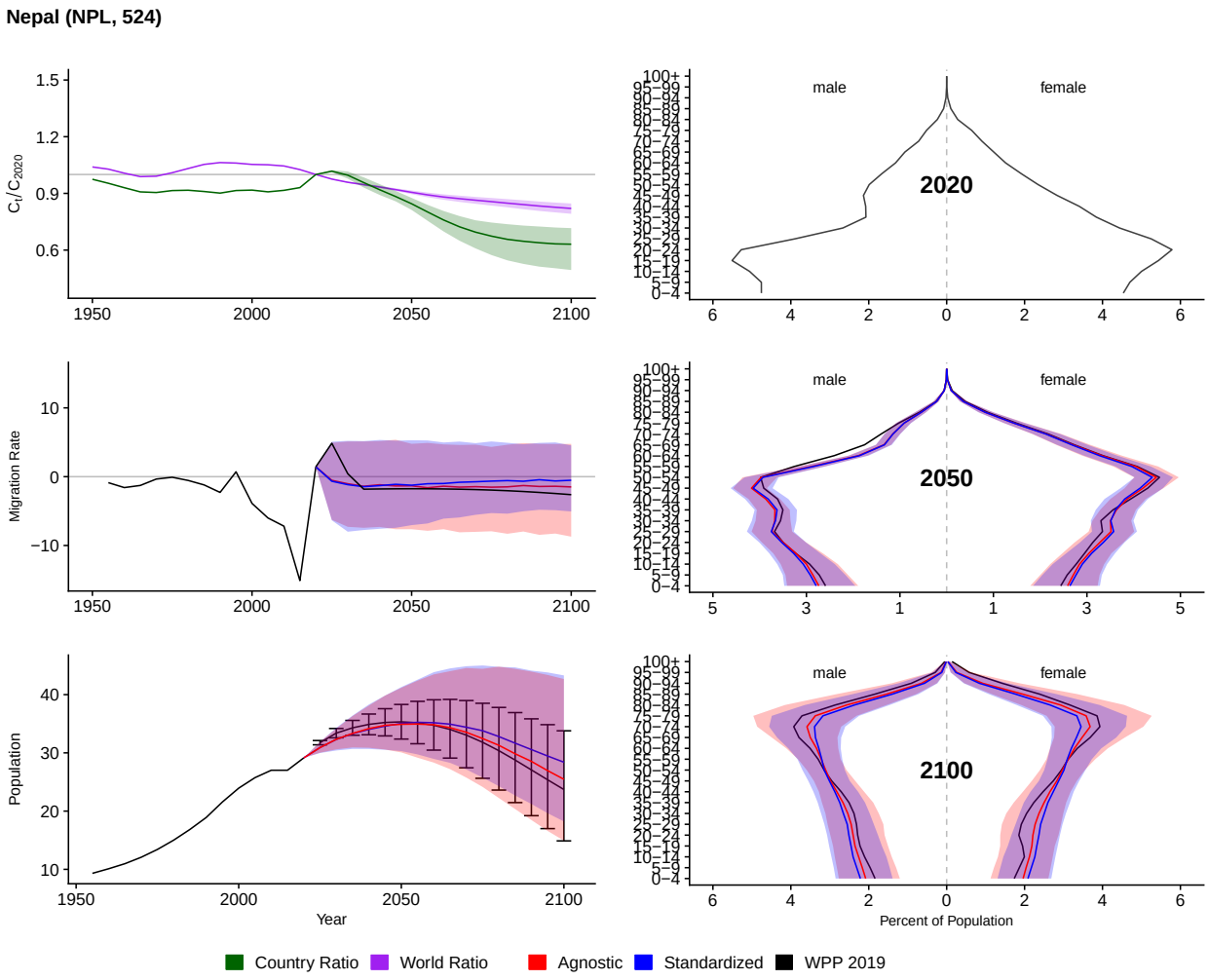


Figure A.191: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

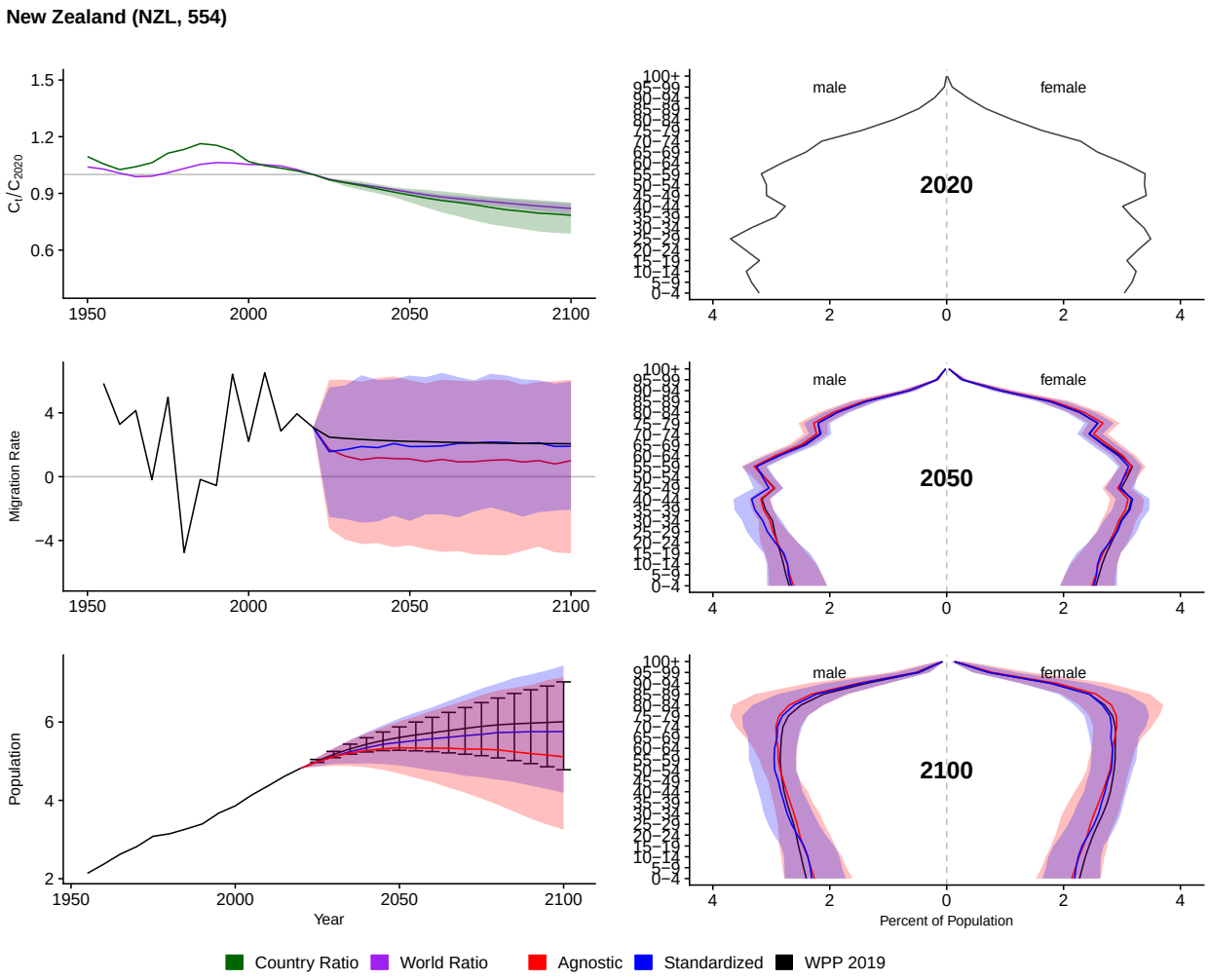


Figure A.192: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

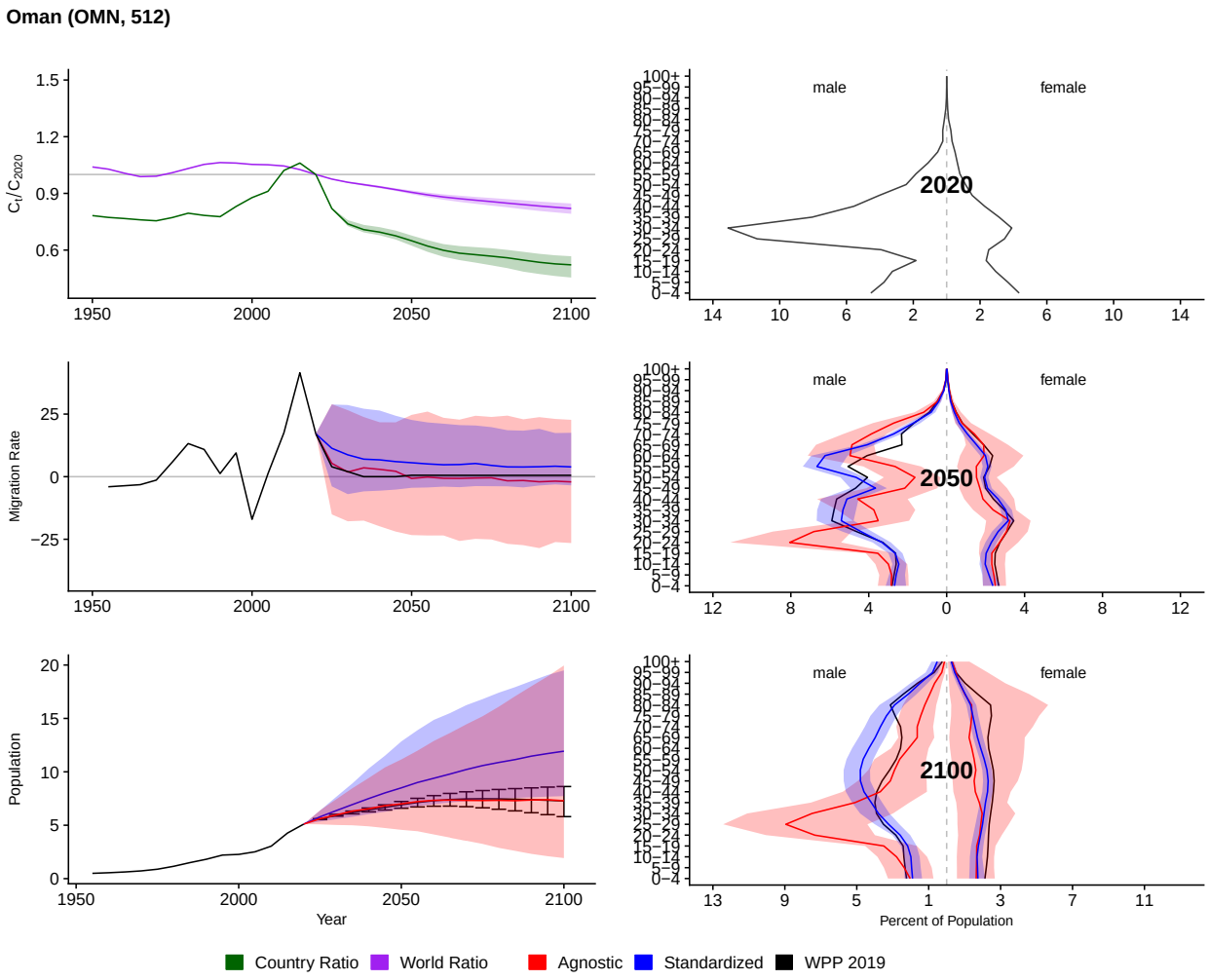


Figure A.193: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

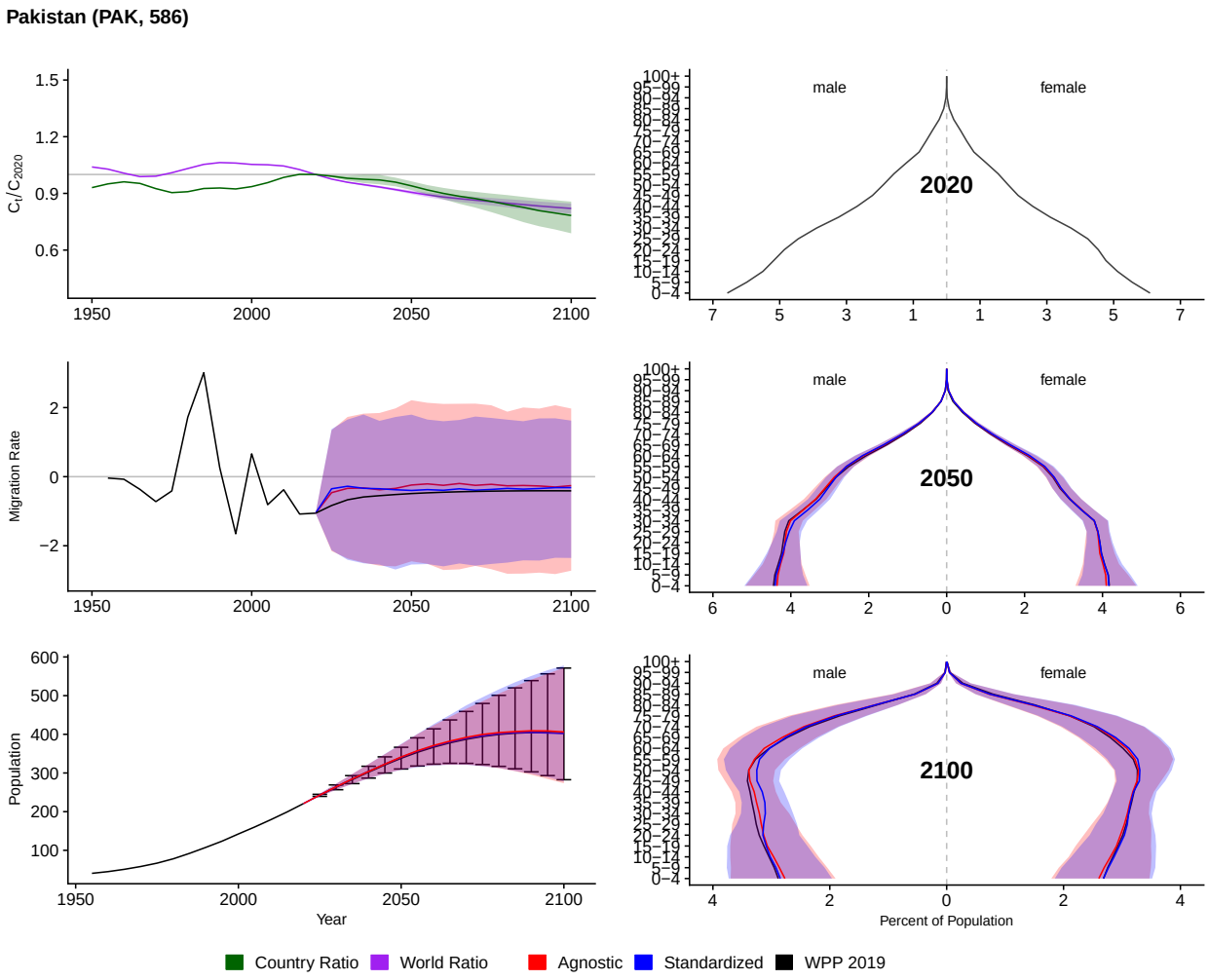


Figure A.194: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

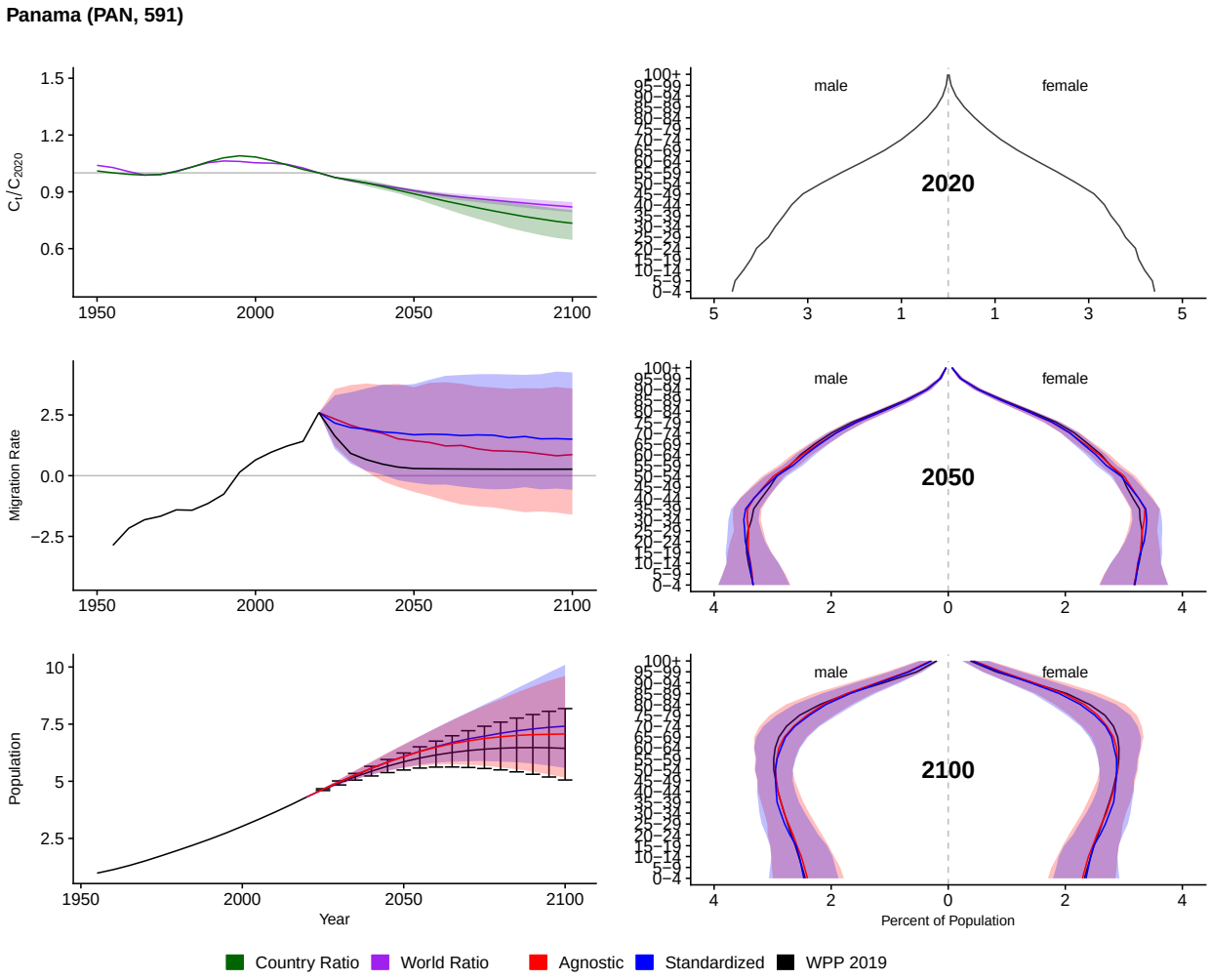


Figure A.195: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

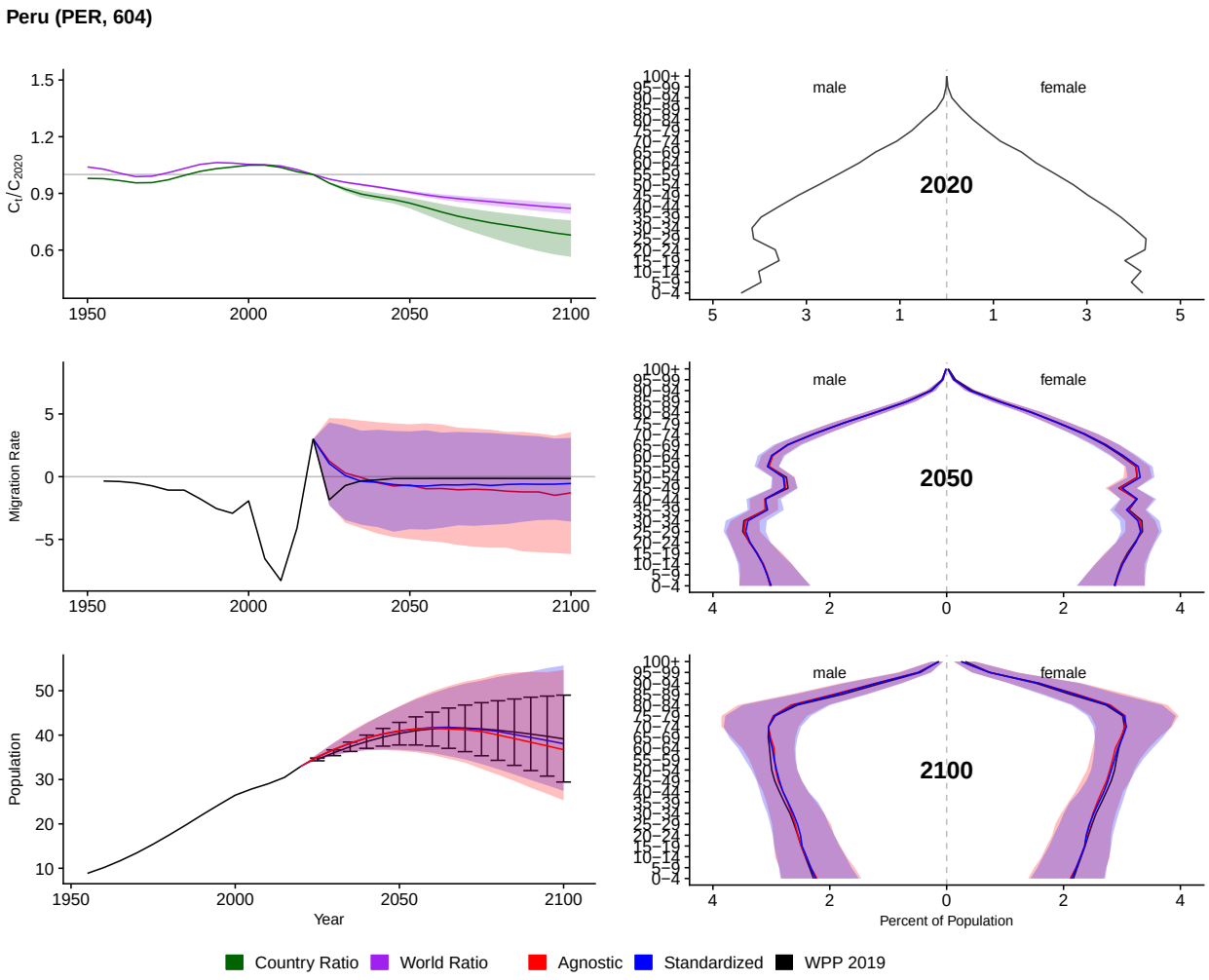


Figure A.196: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

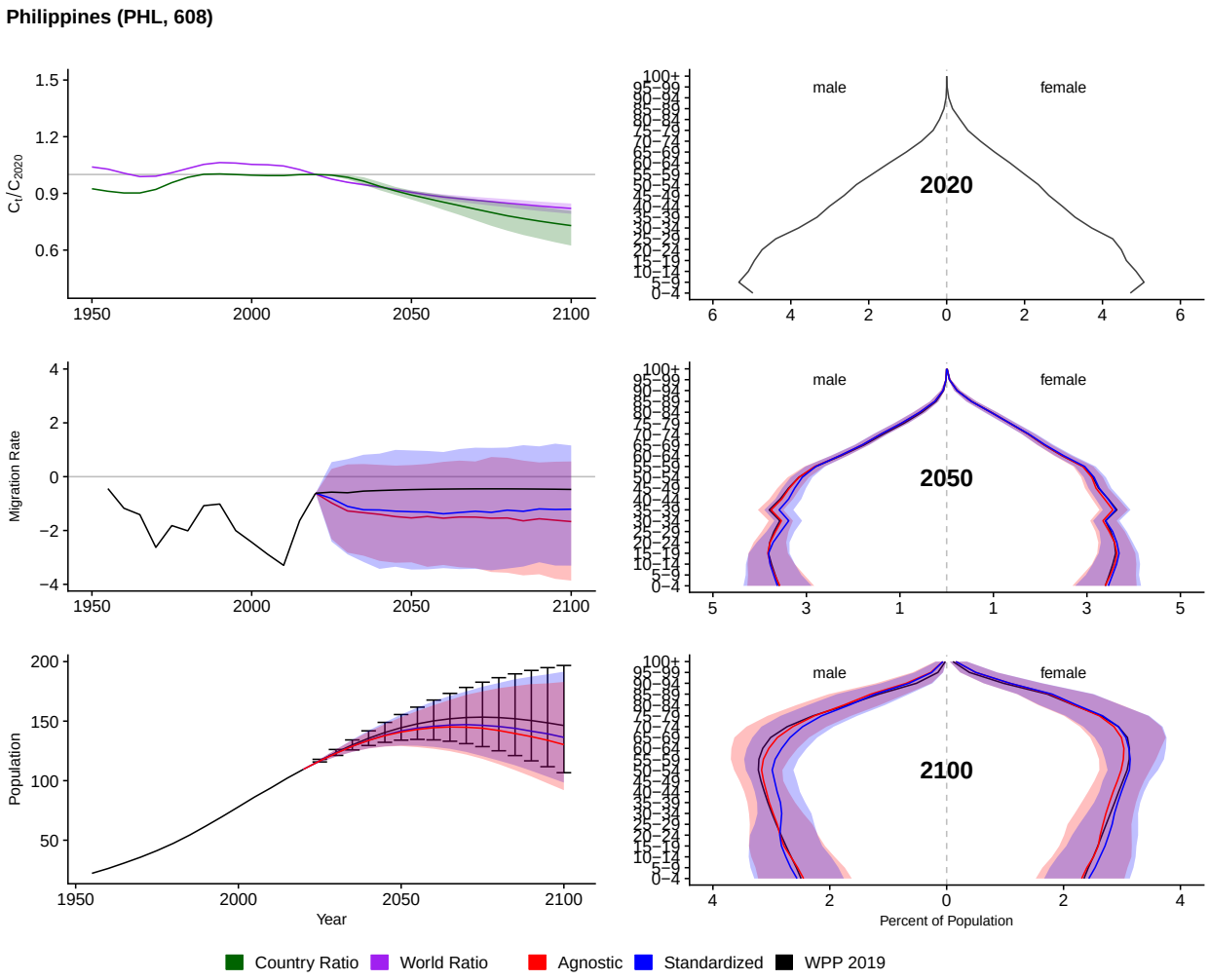


Figure A.197: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

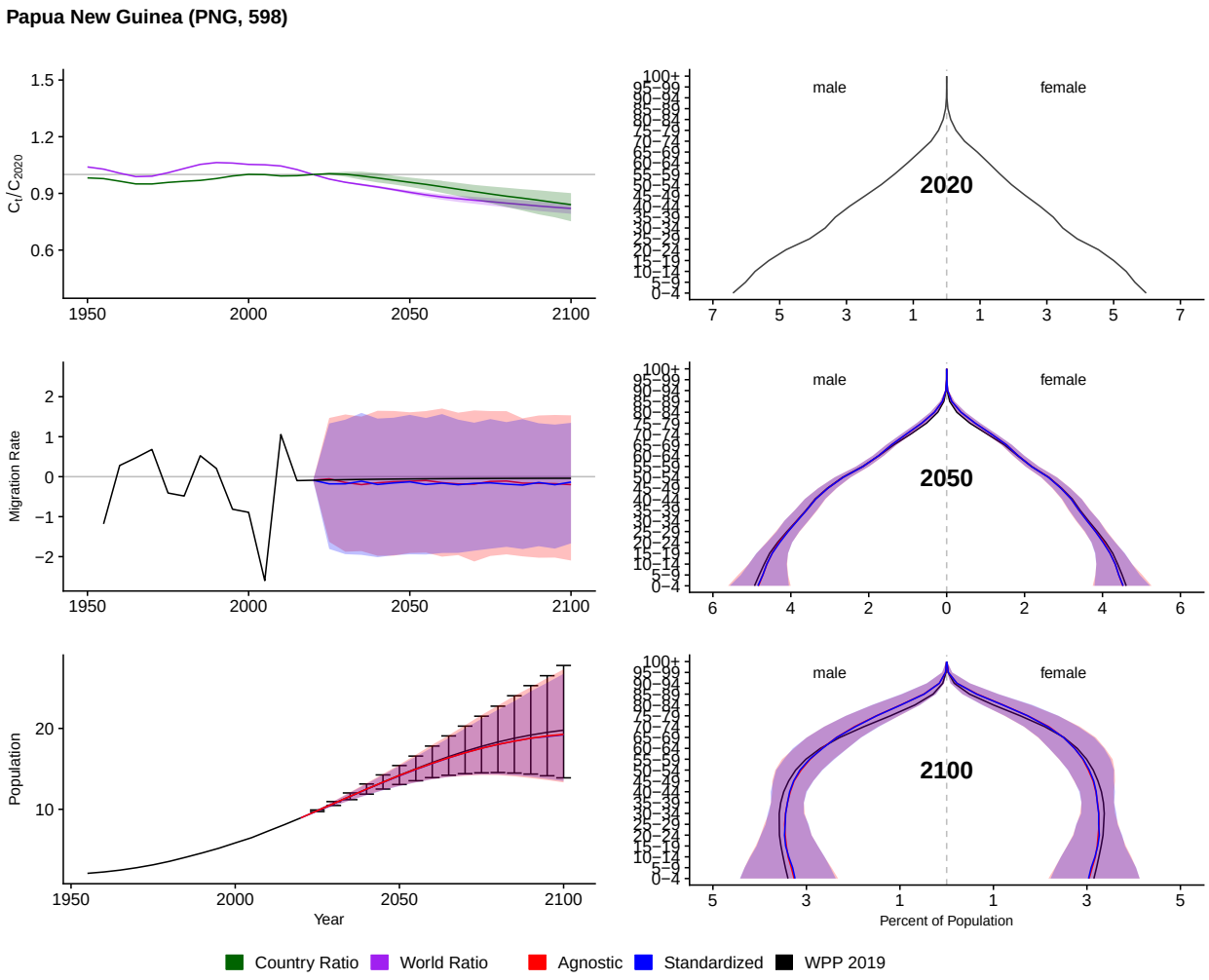


Figure A.198: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

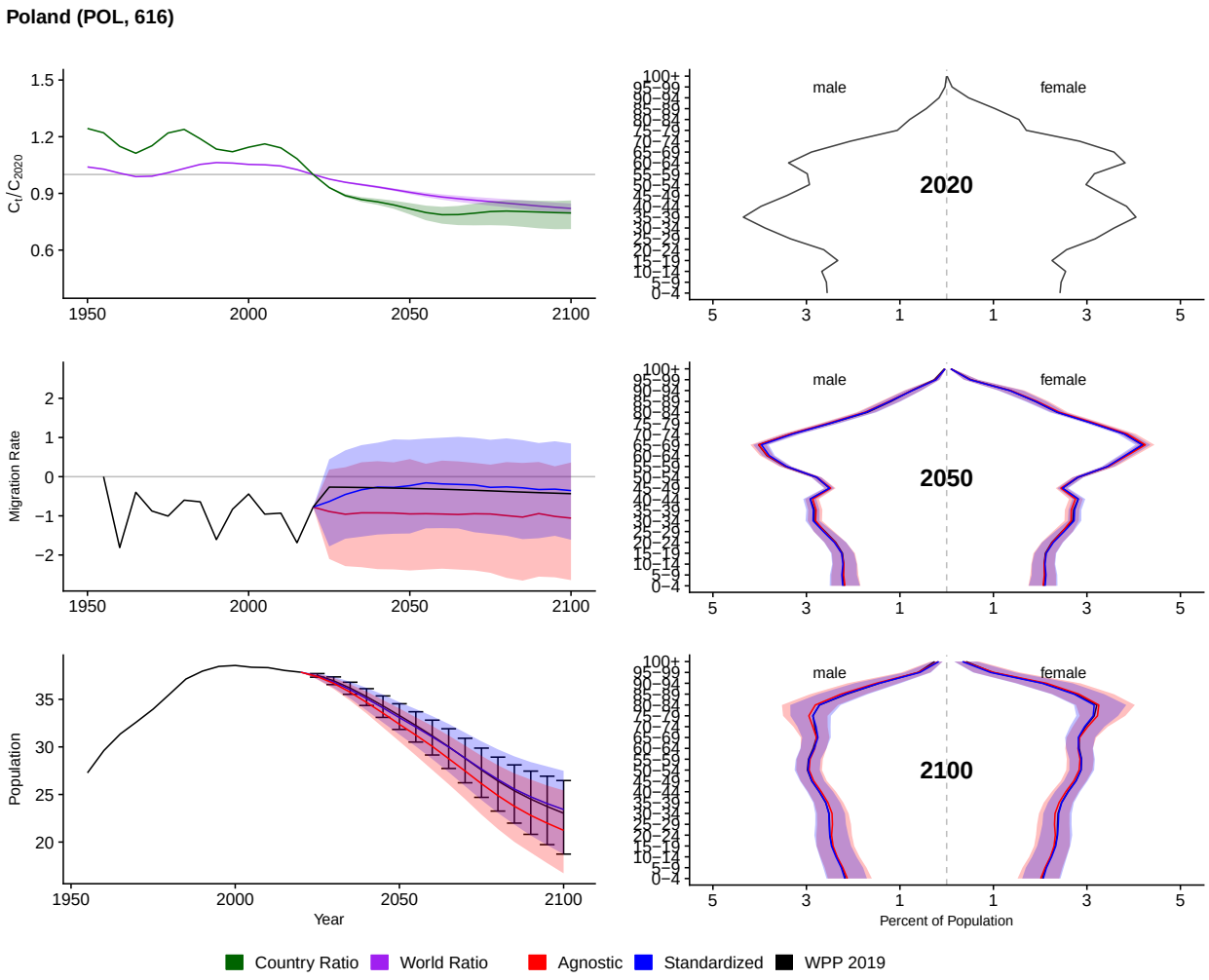


Figure A.199: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

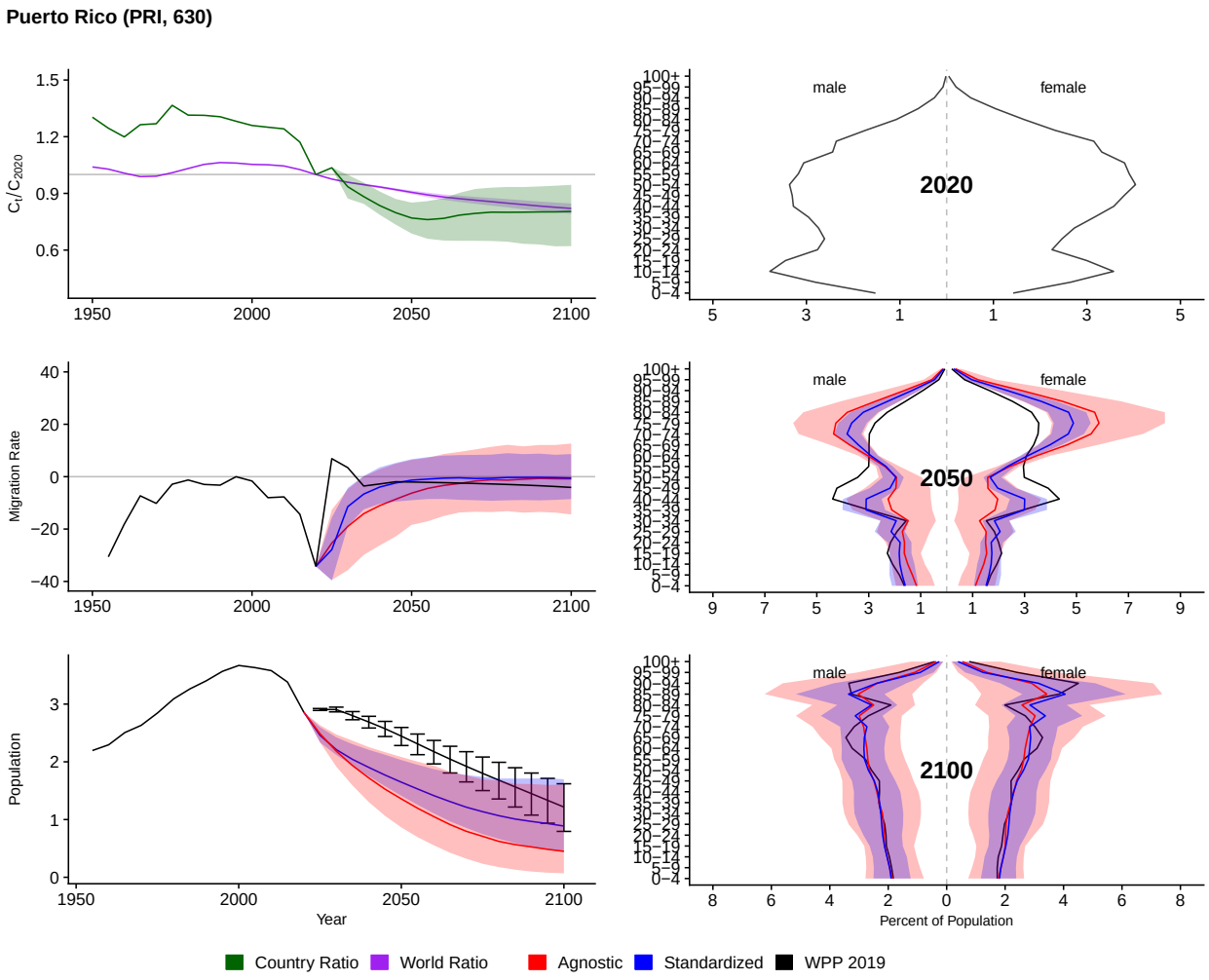


Figure A.200: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

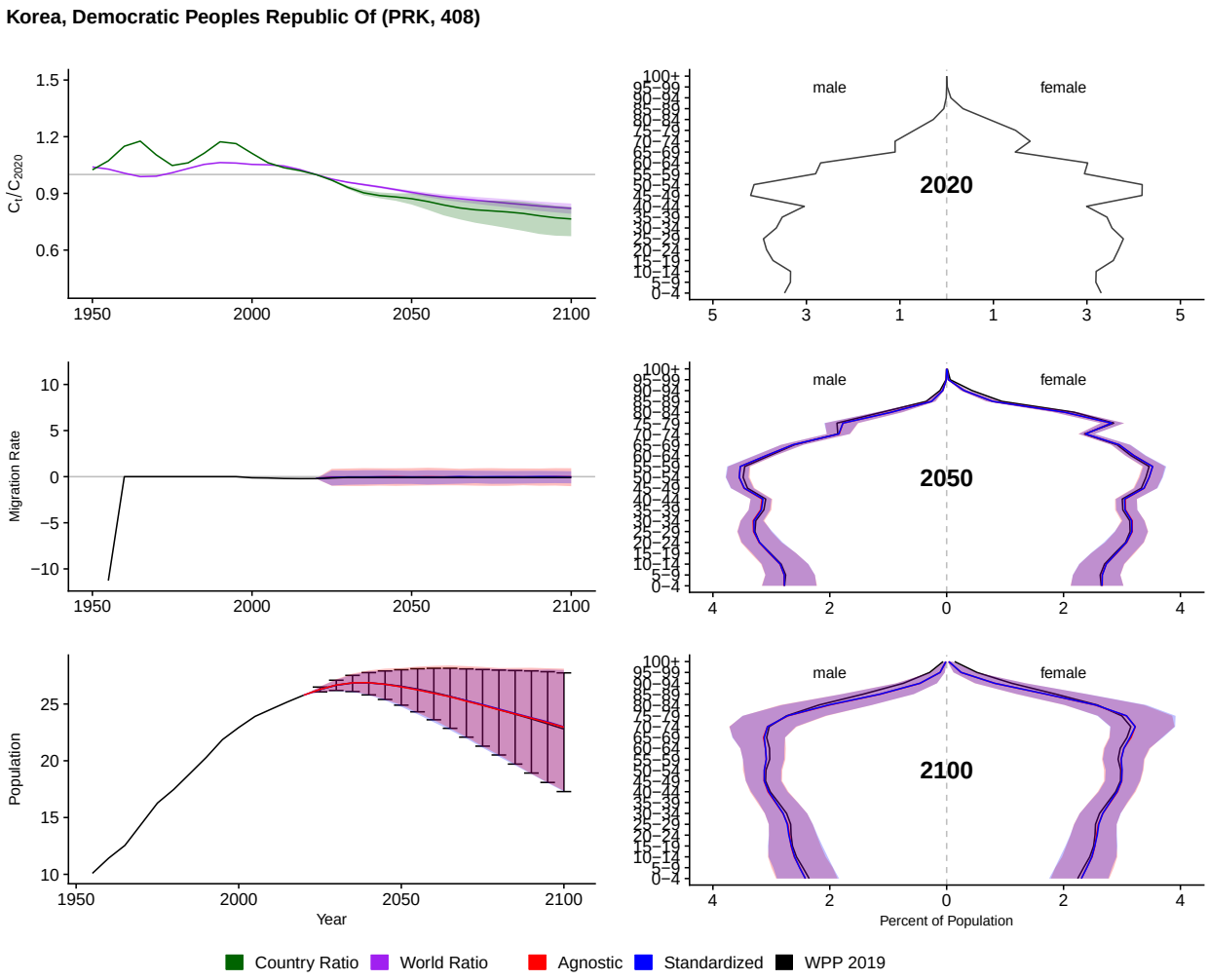


Figure A.201: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

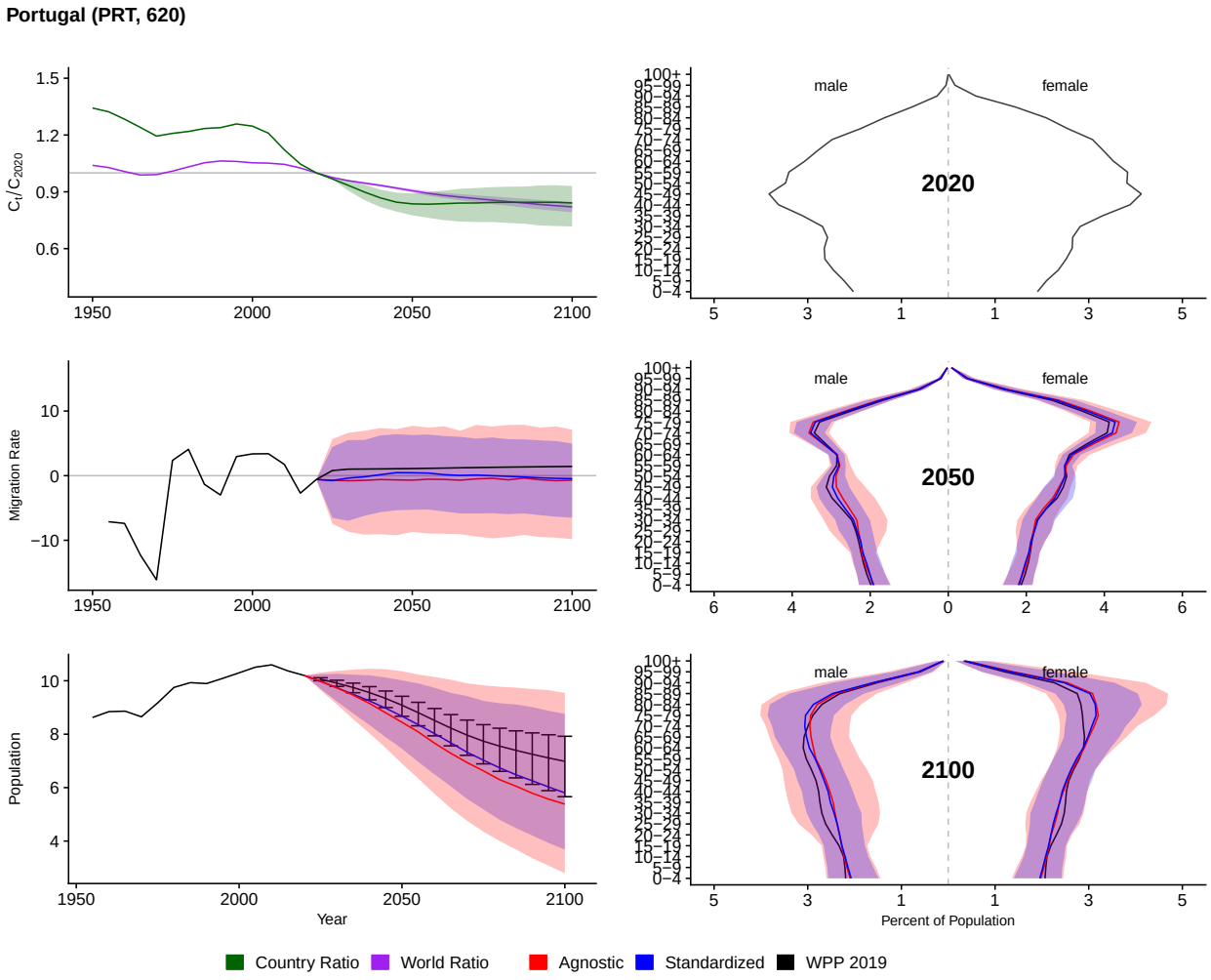


Figure A.202: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

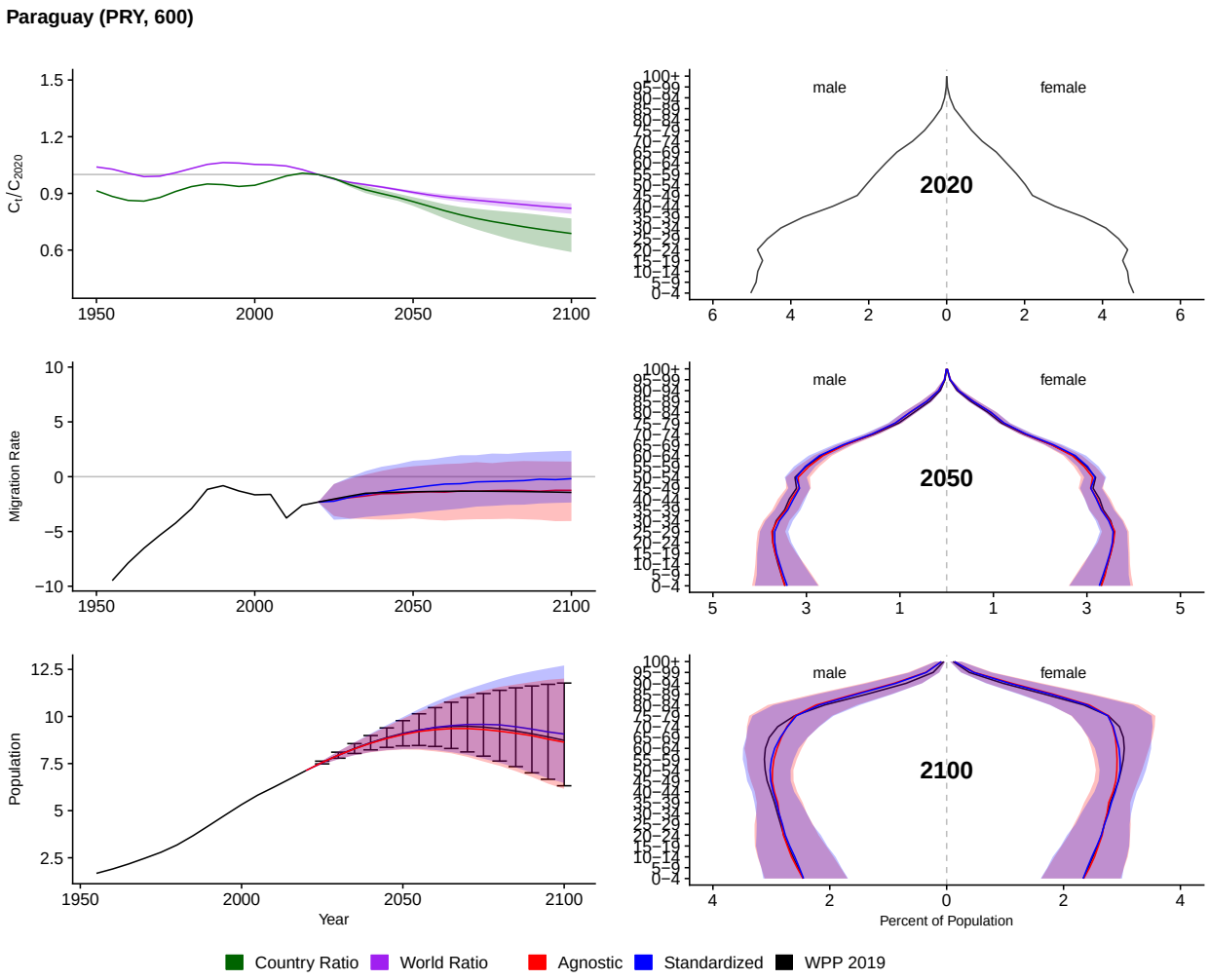


Figure A.203: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

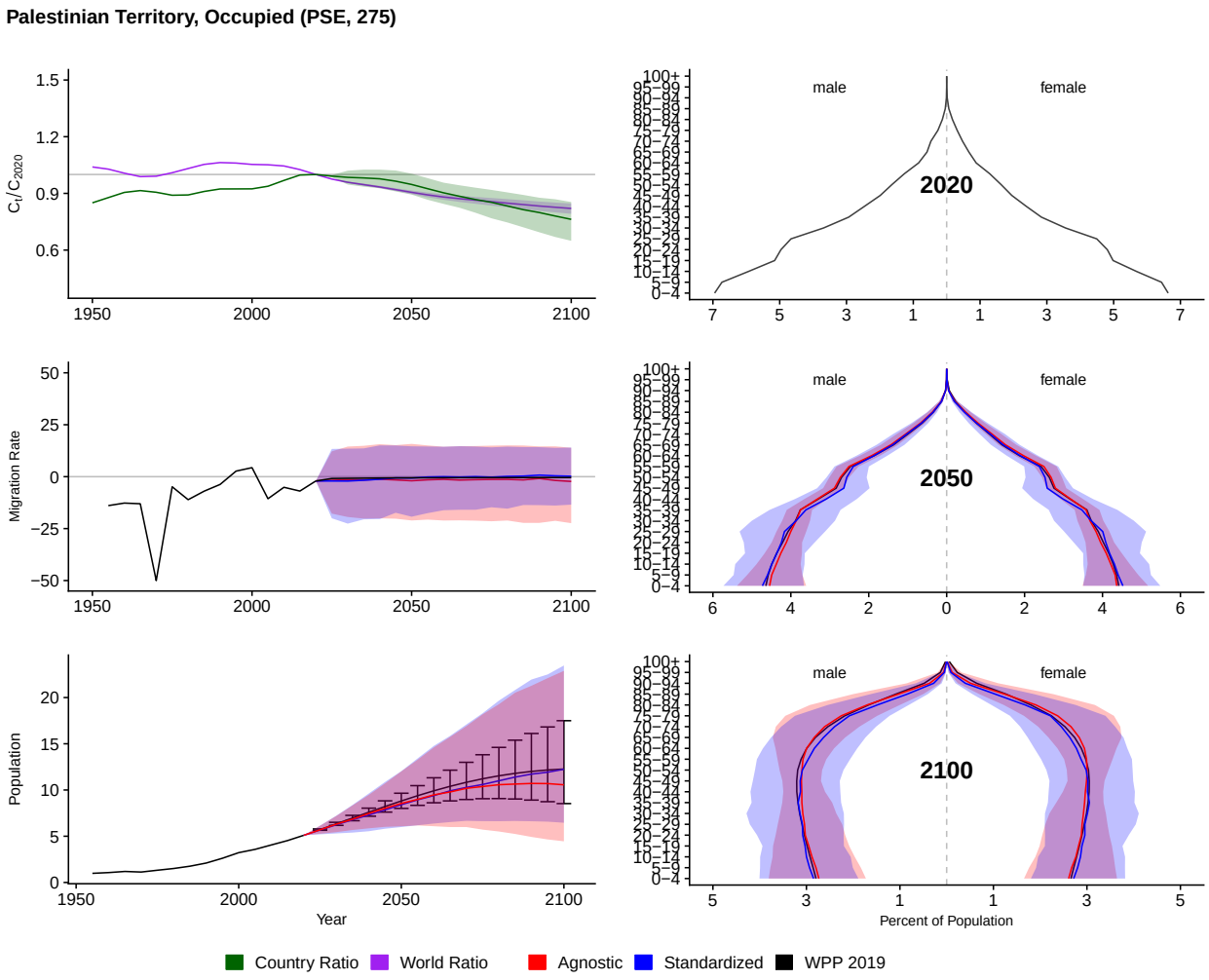


Figure A.204: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

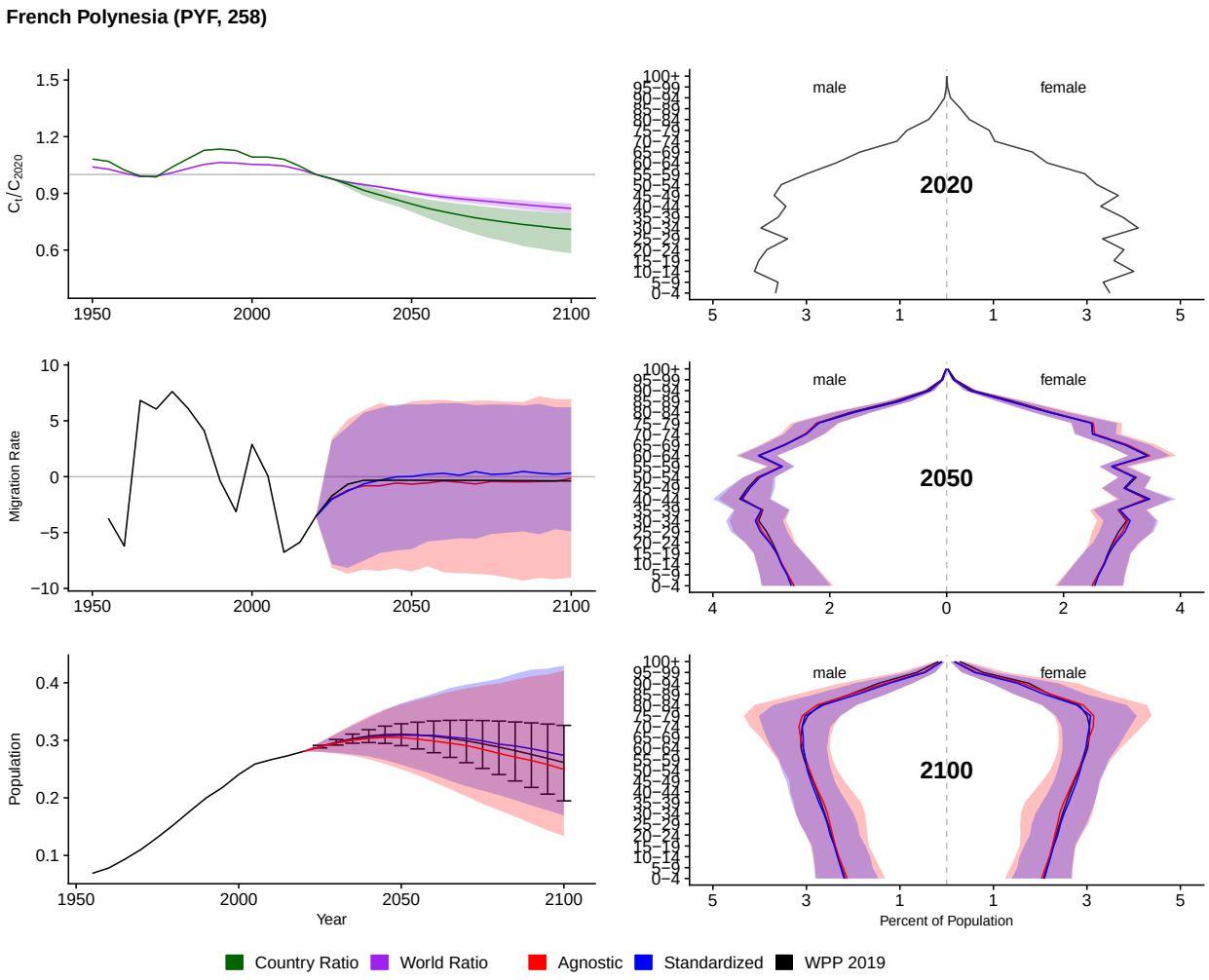


Figure A.205: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

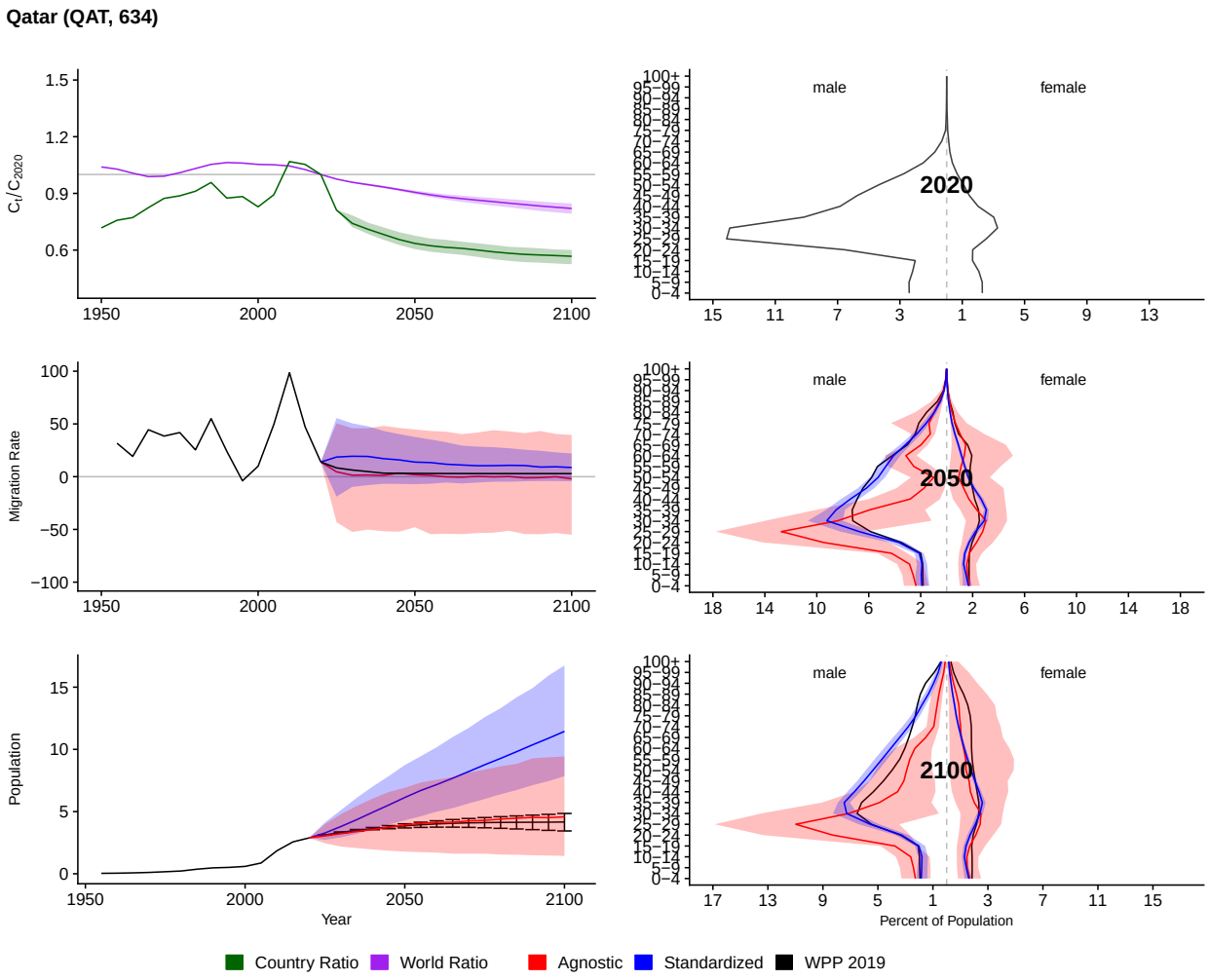


Figure A.206: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

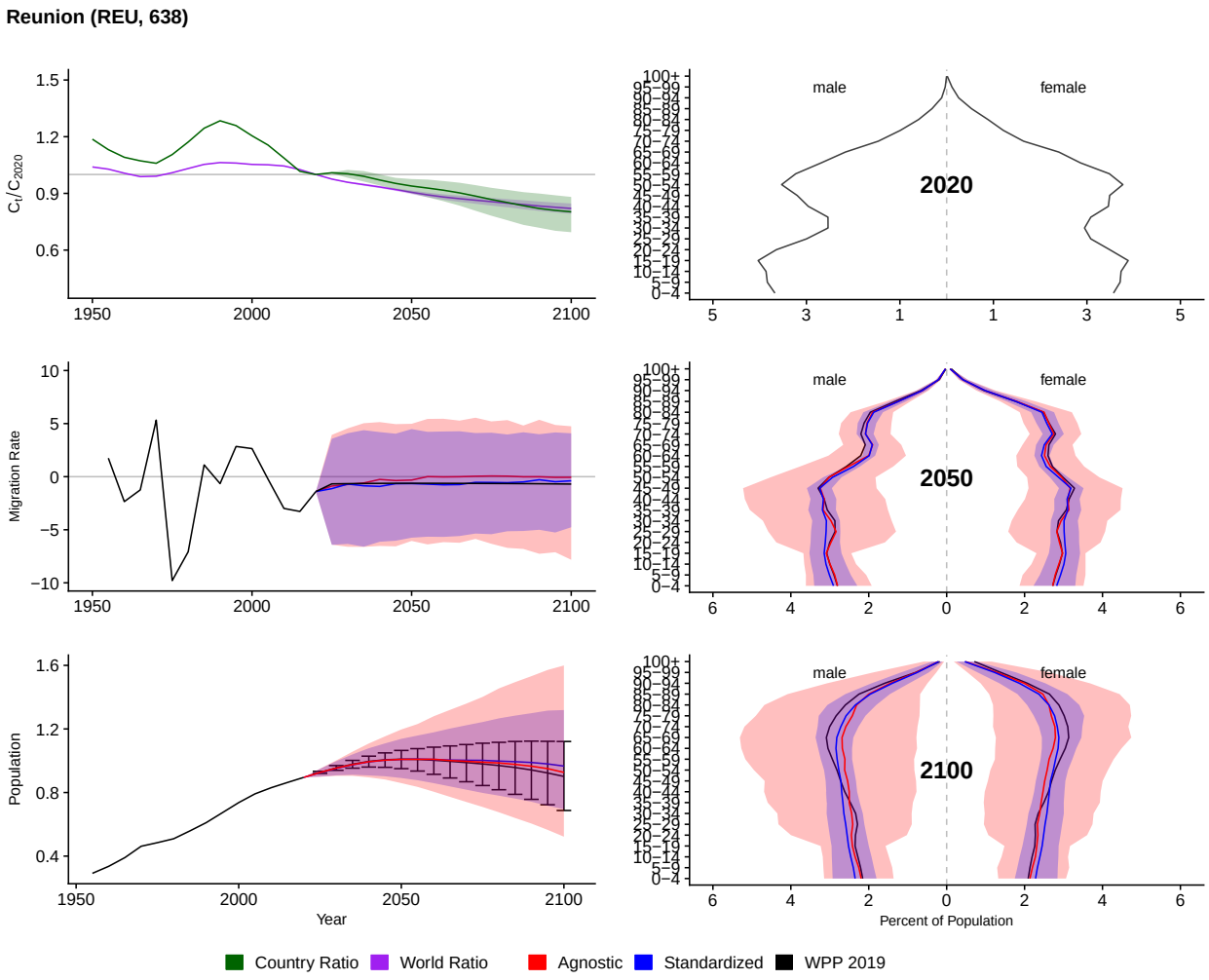


Figure A.207: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

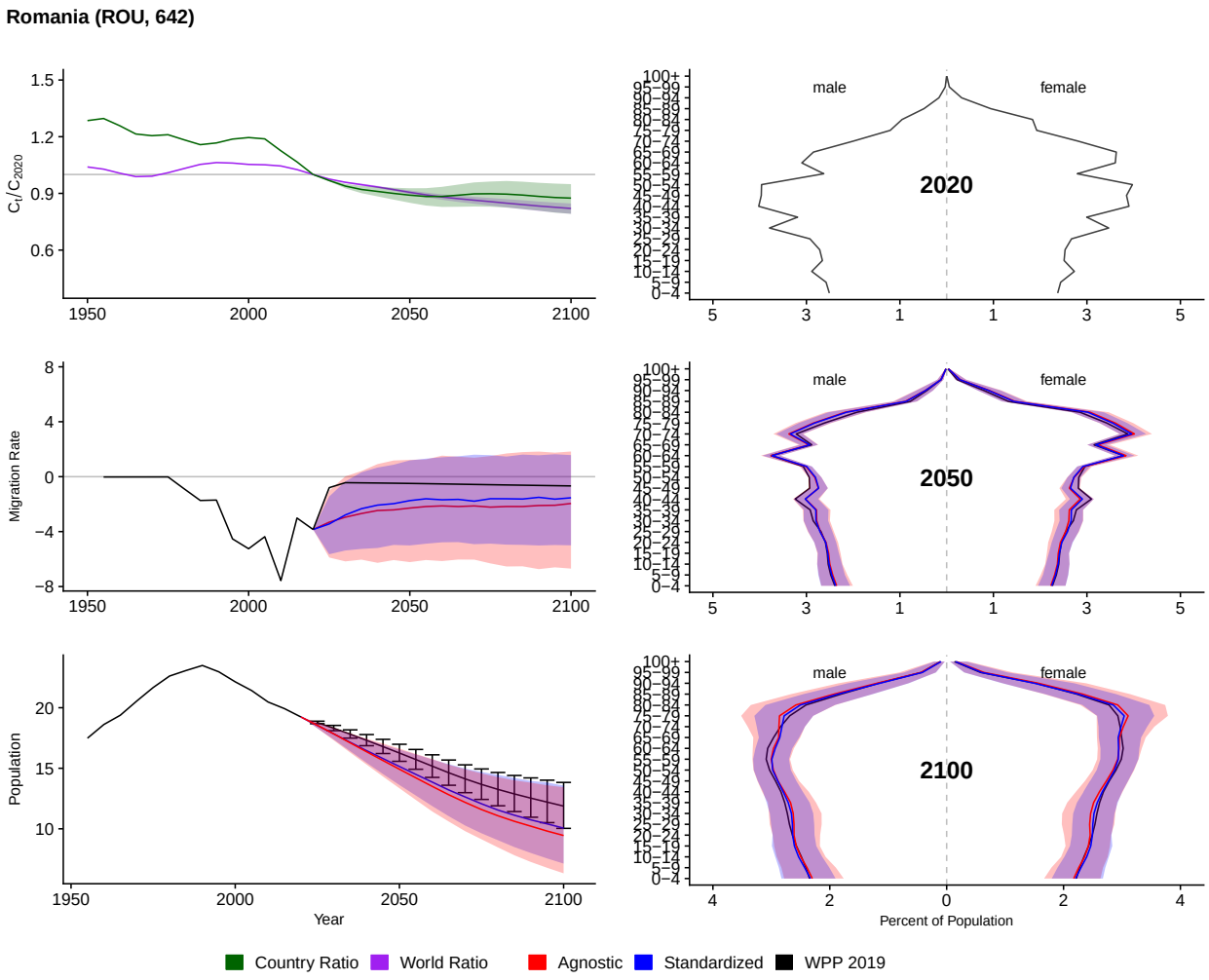


Figure A.208: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

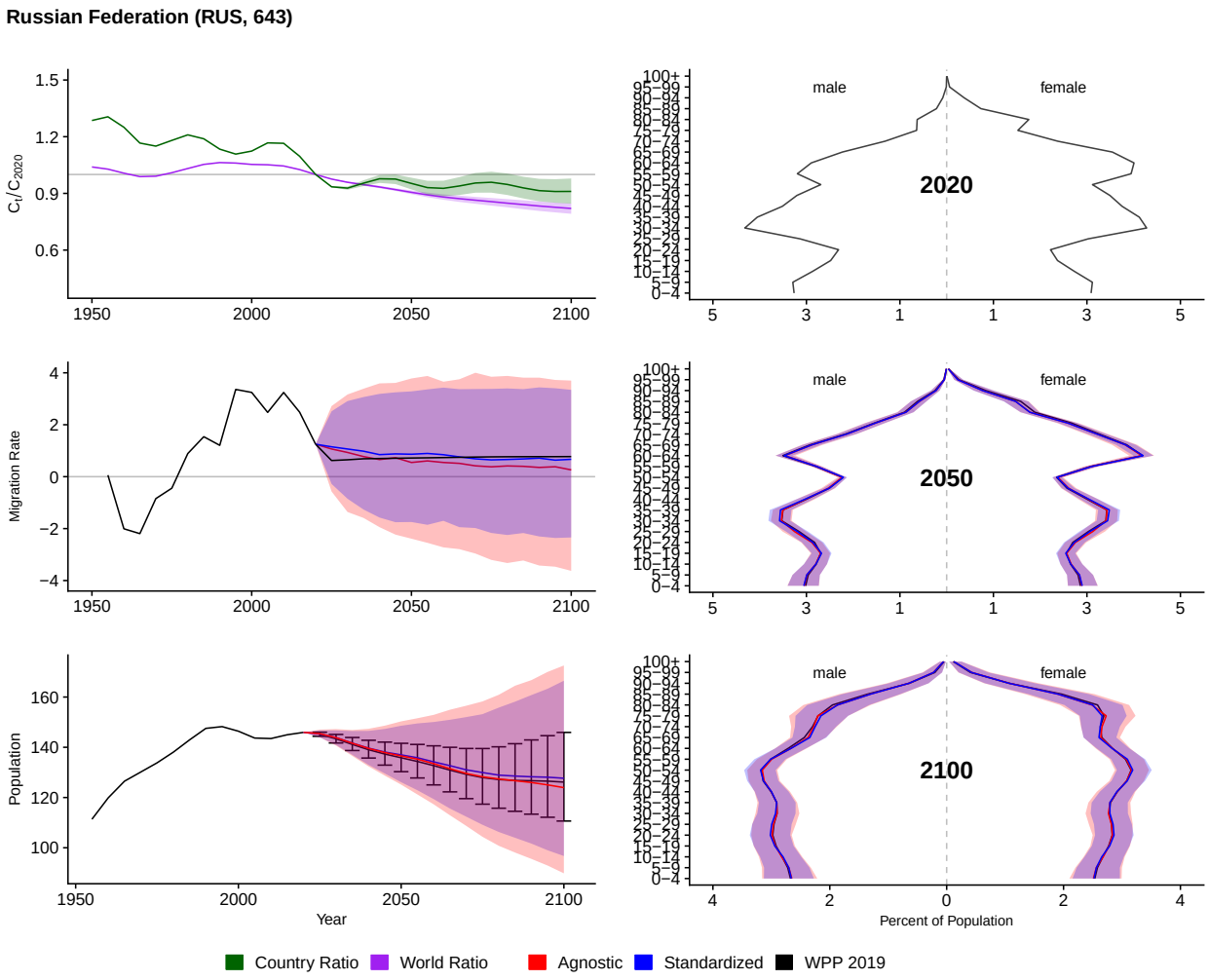


Figure A.209: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

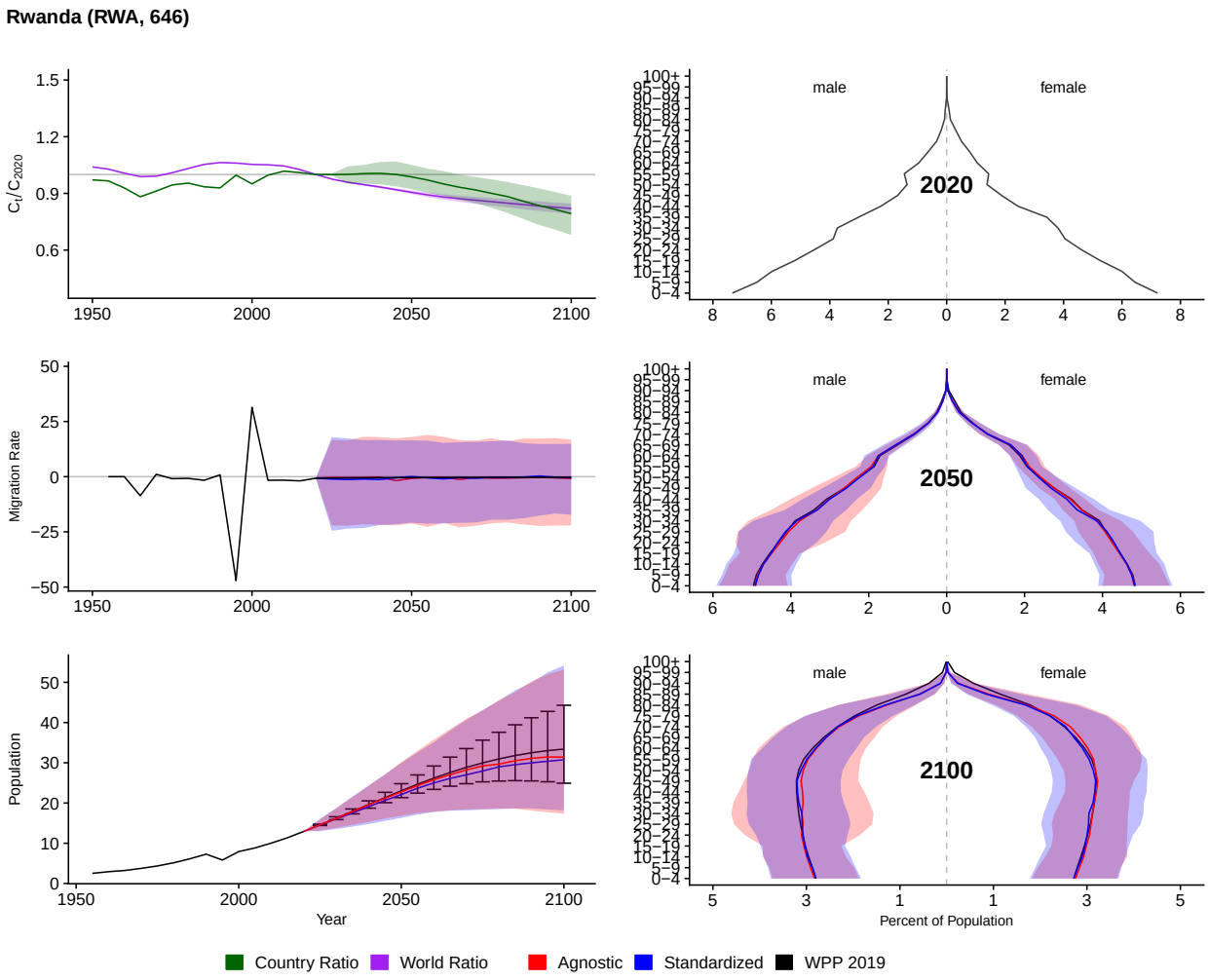


Figure A.210: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

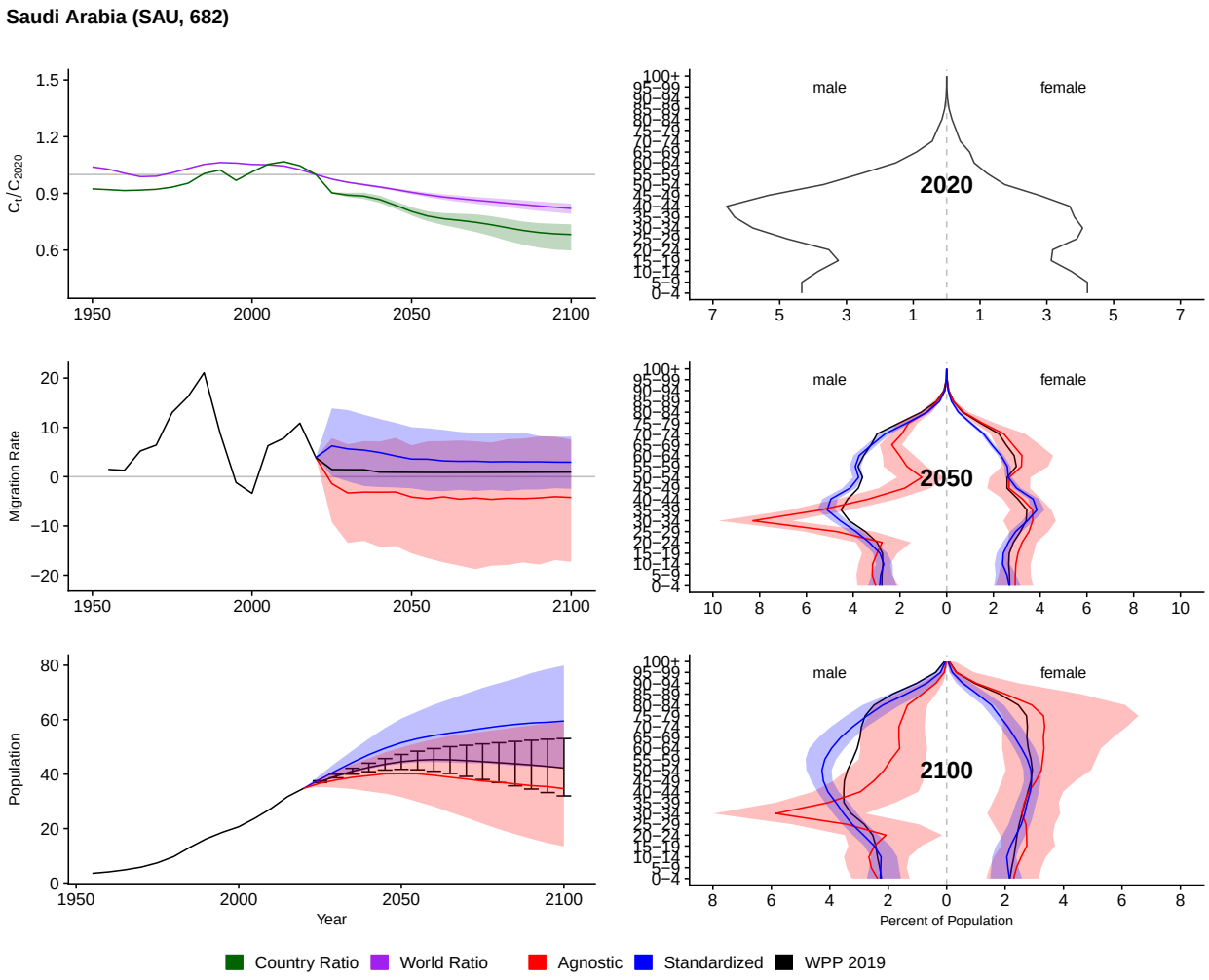


Figure A.211: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

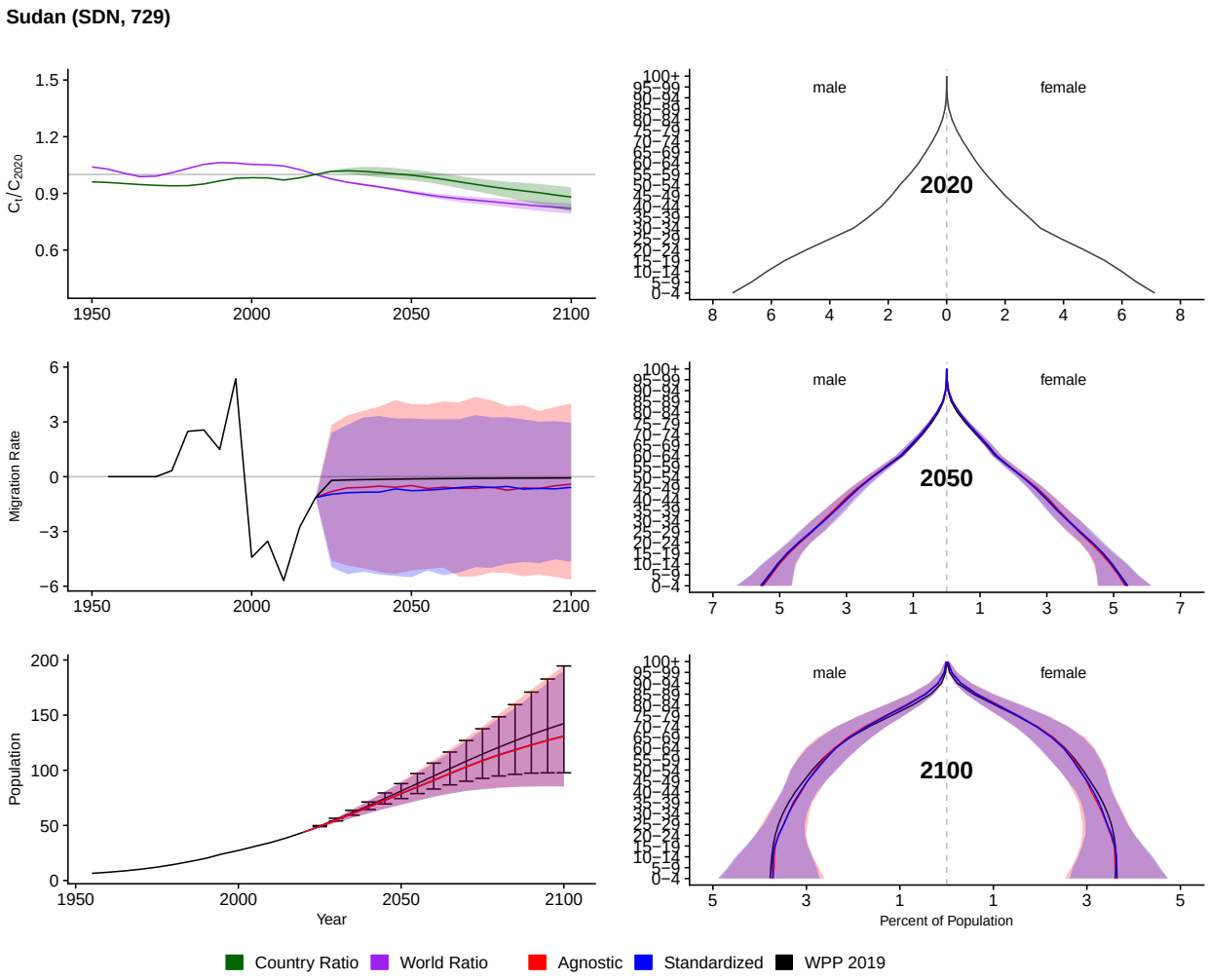


Figure A.212: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

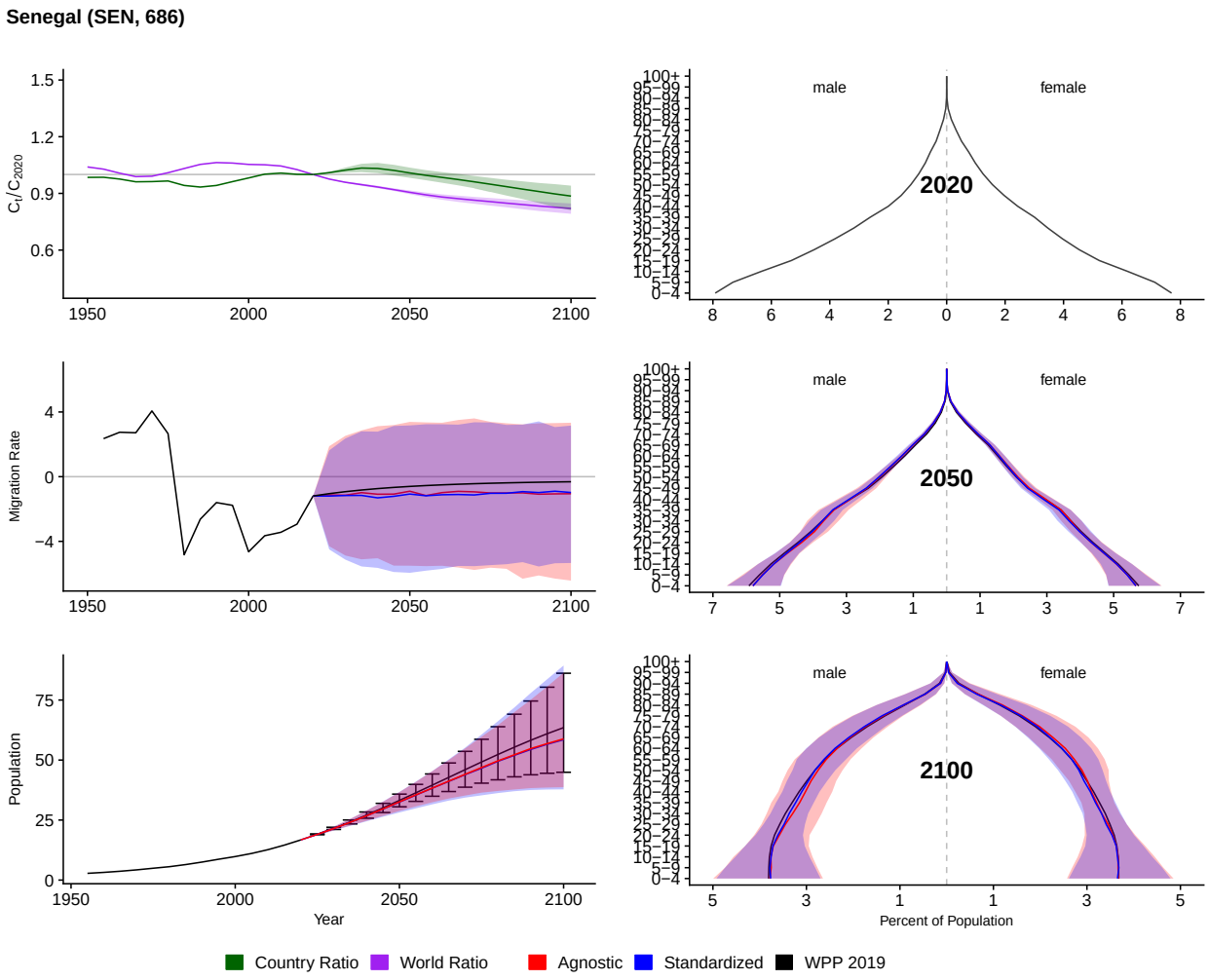


Figure A.213: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

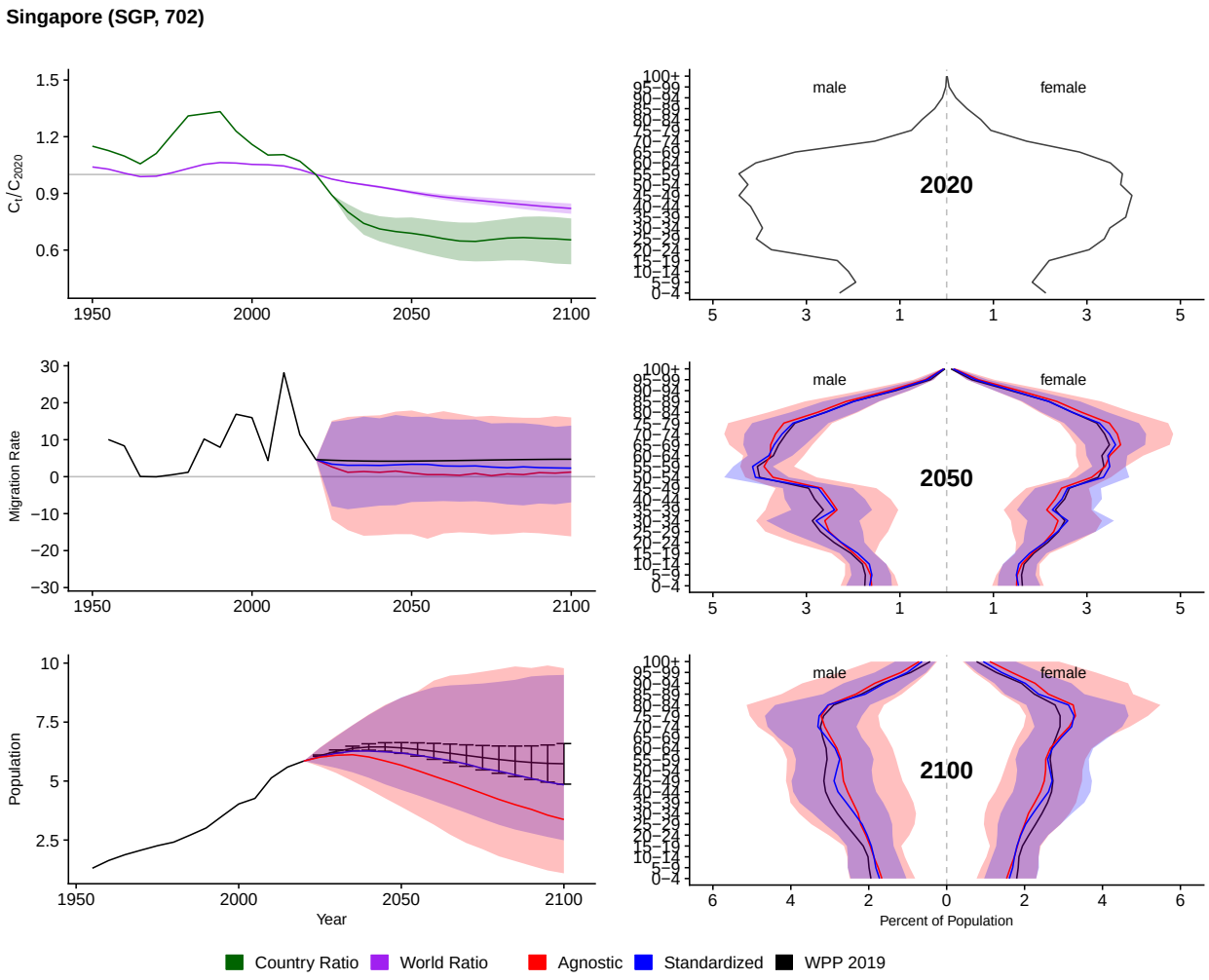


Figure A.214: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

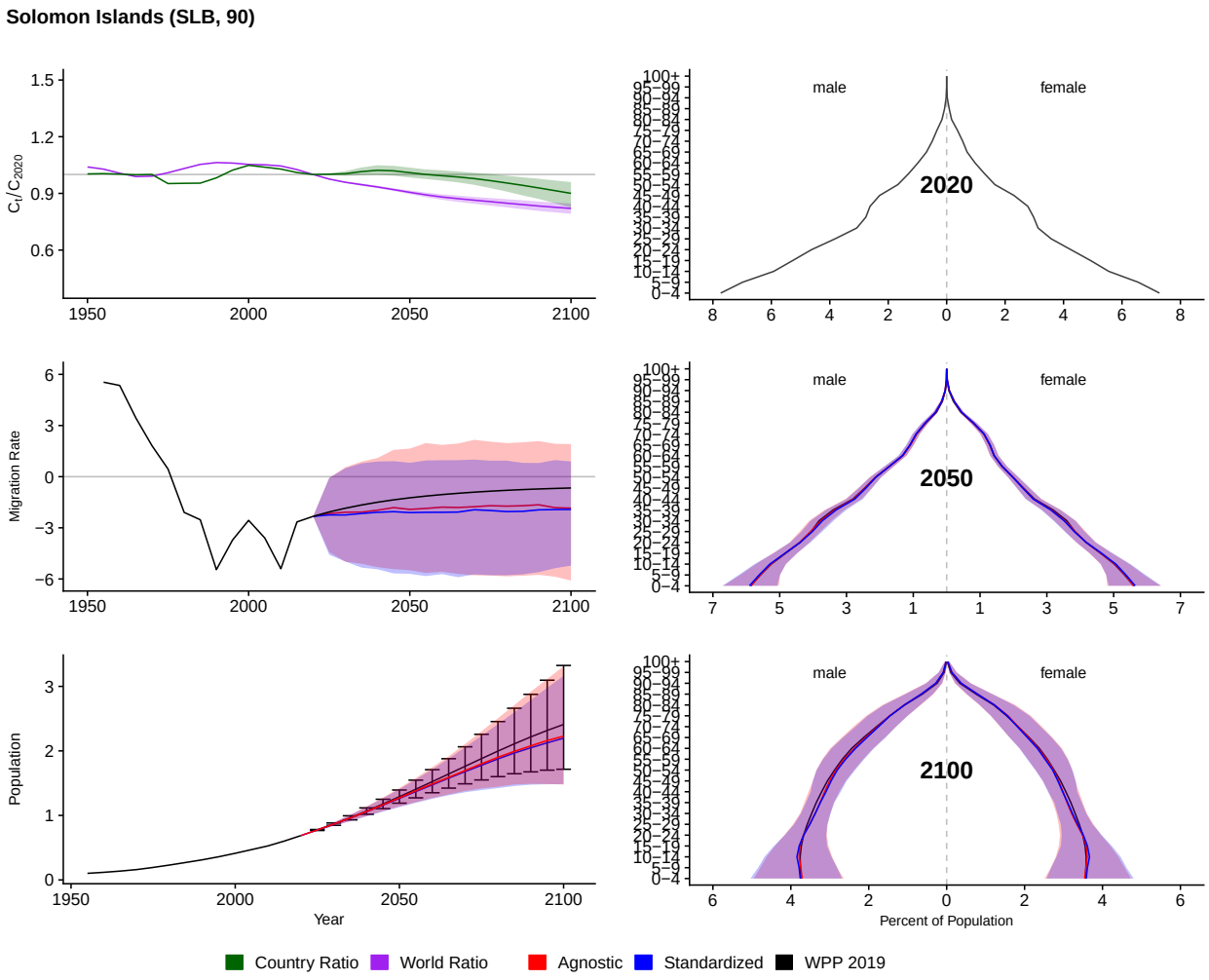


Figure A.215: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

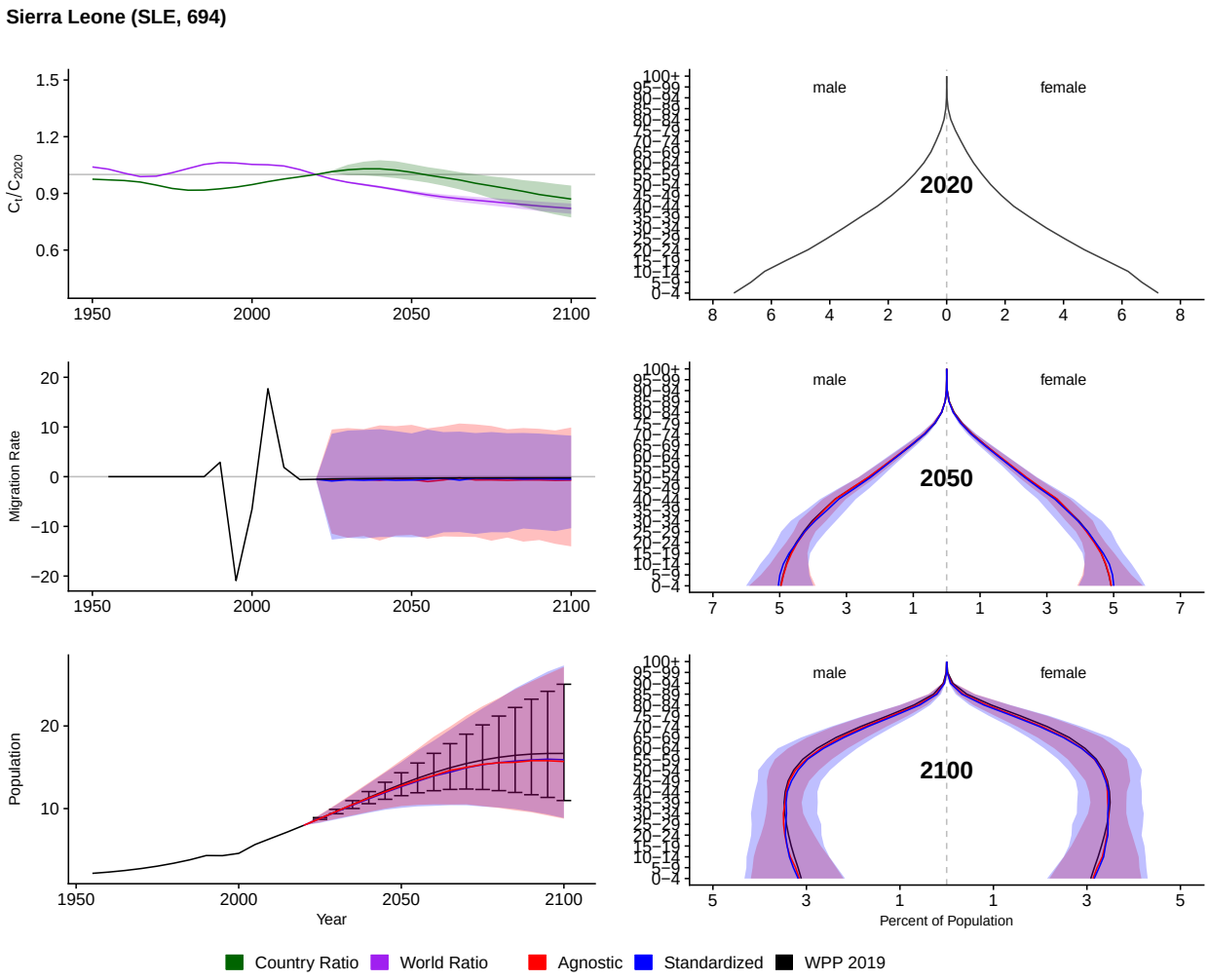


Figure A.216: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

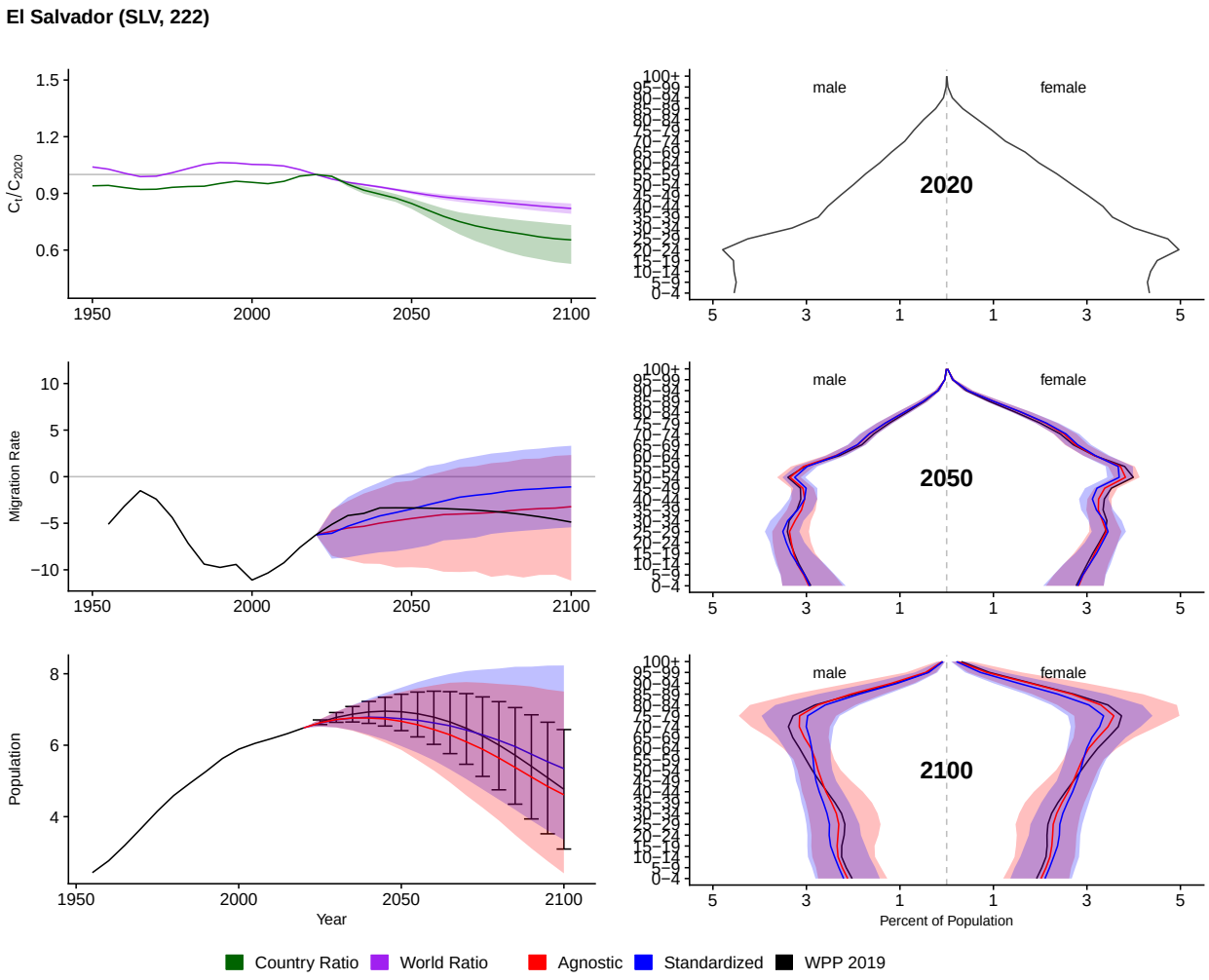


Figure A.217: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

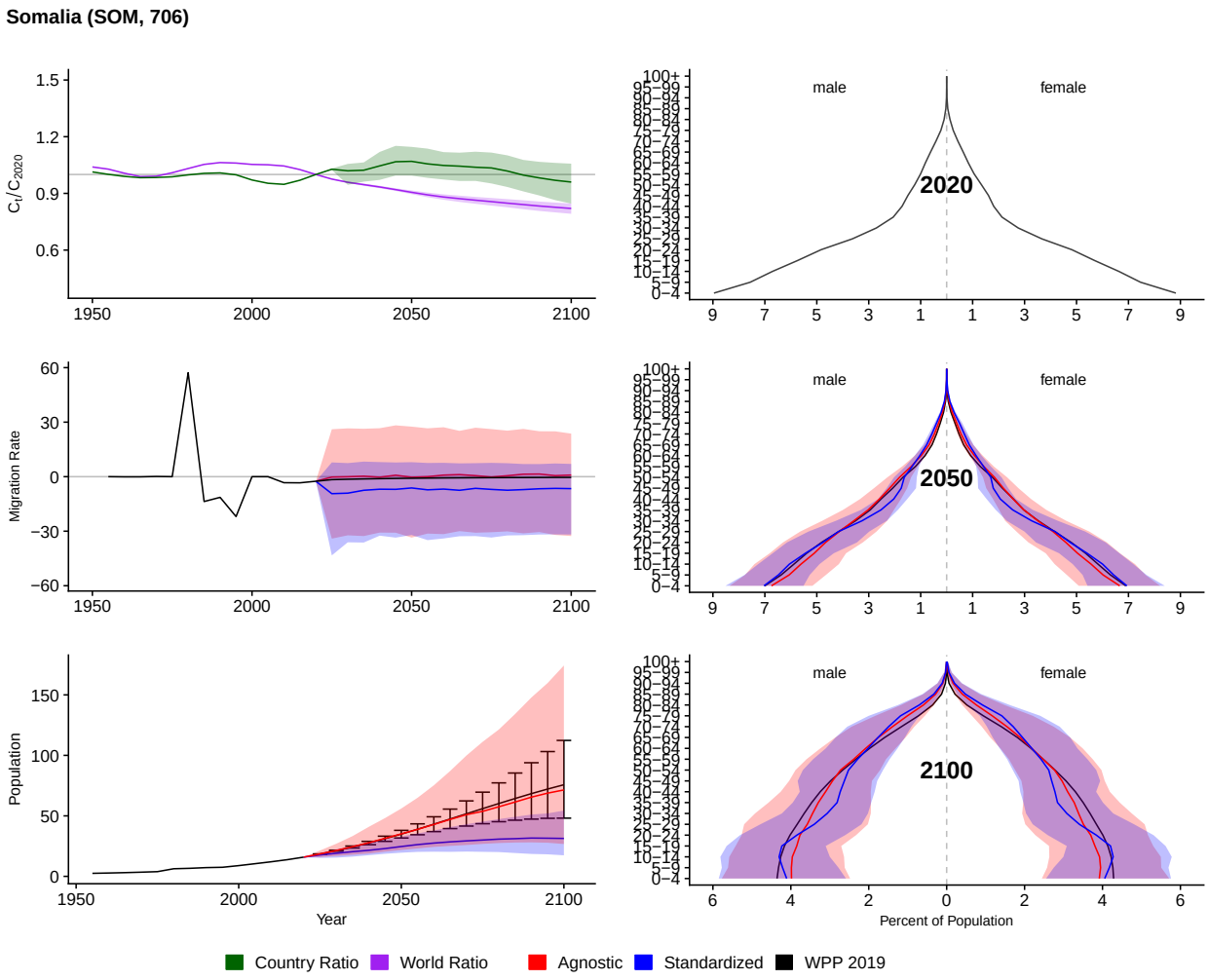


Figure A.218: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

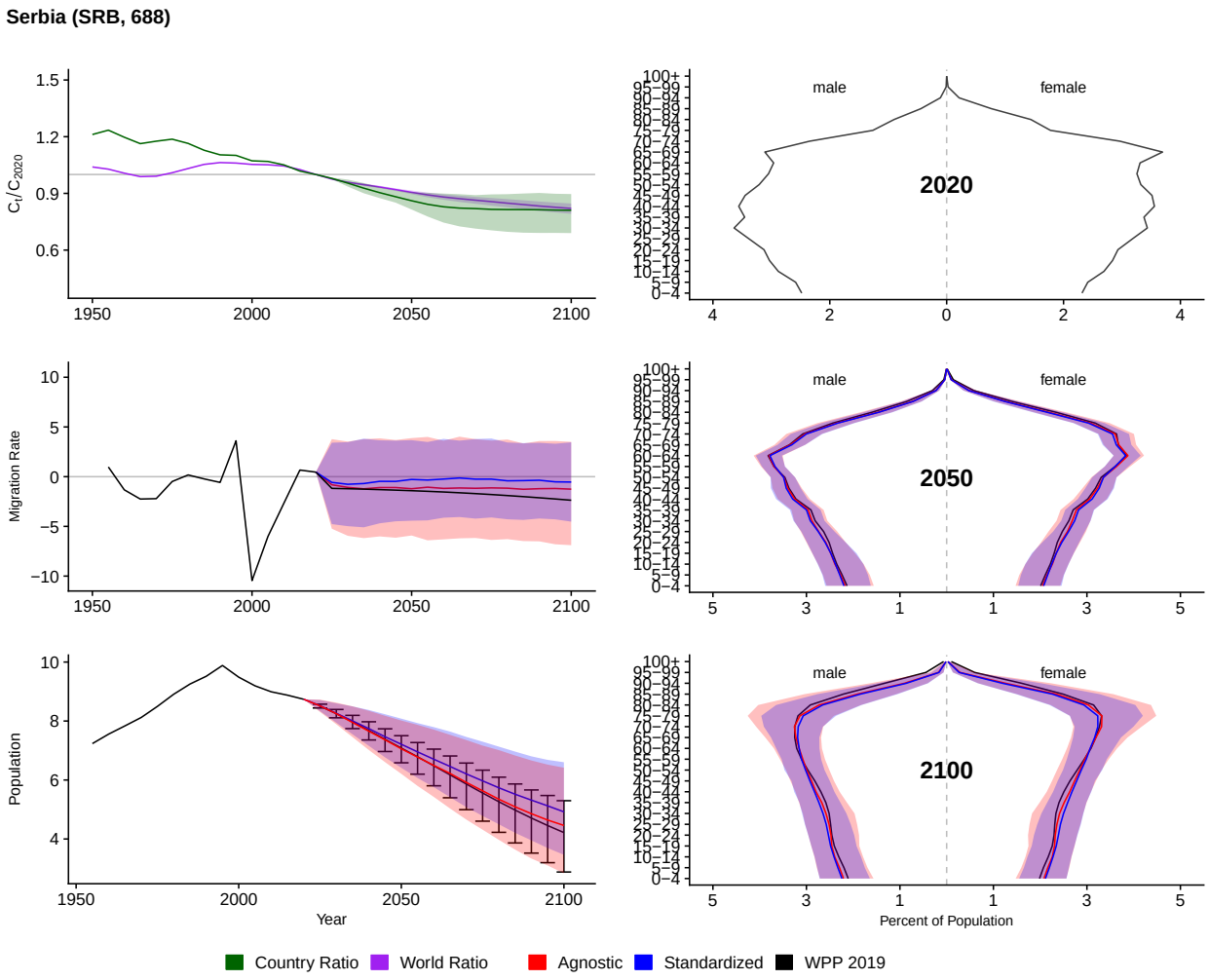


Figure A.219: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

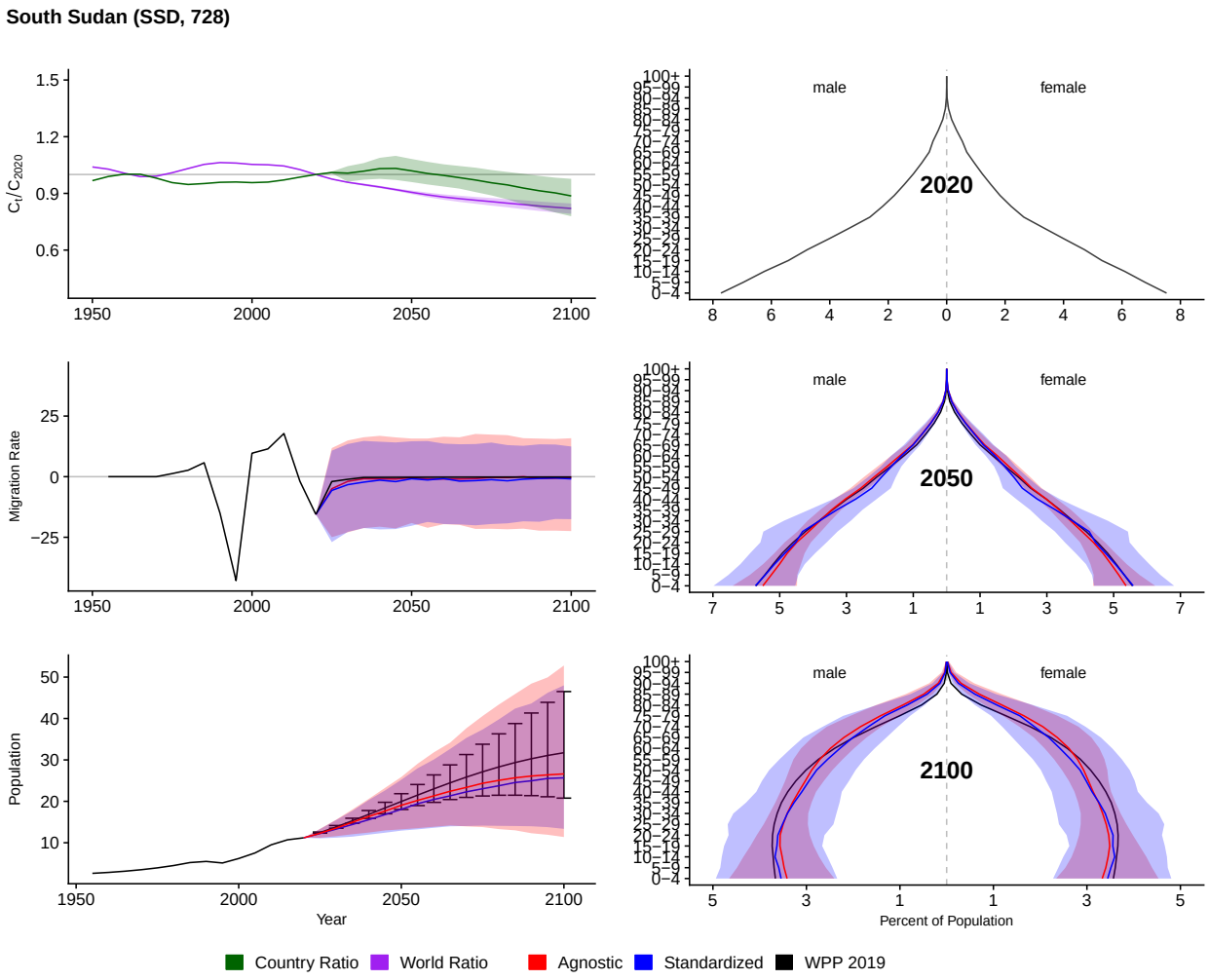


Figure A.220: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

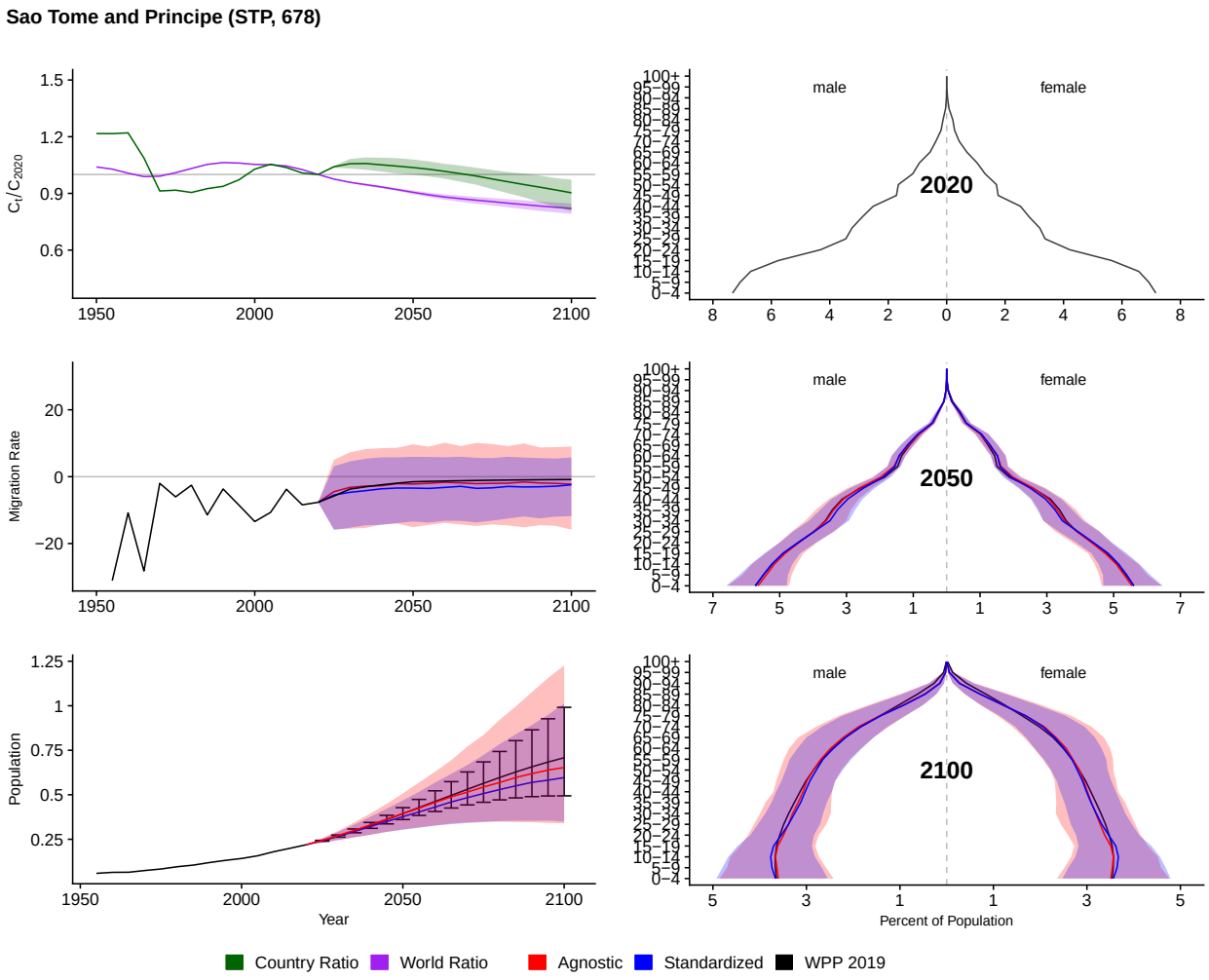


Figure A.221: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

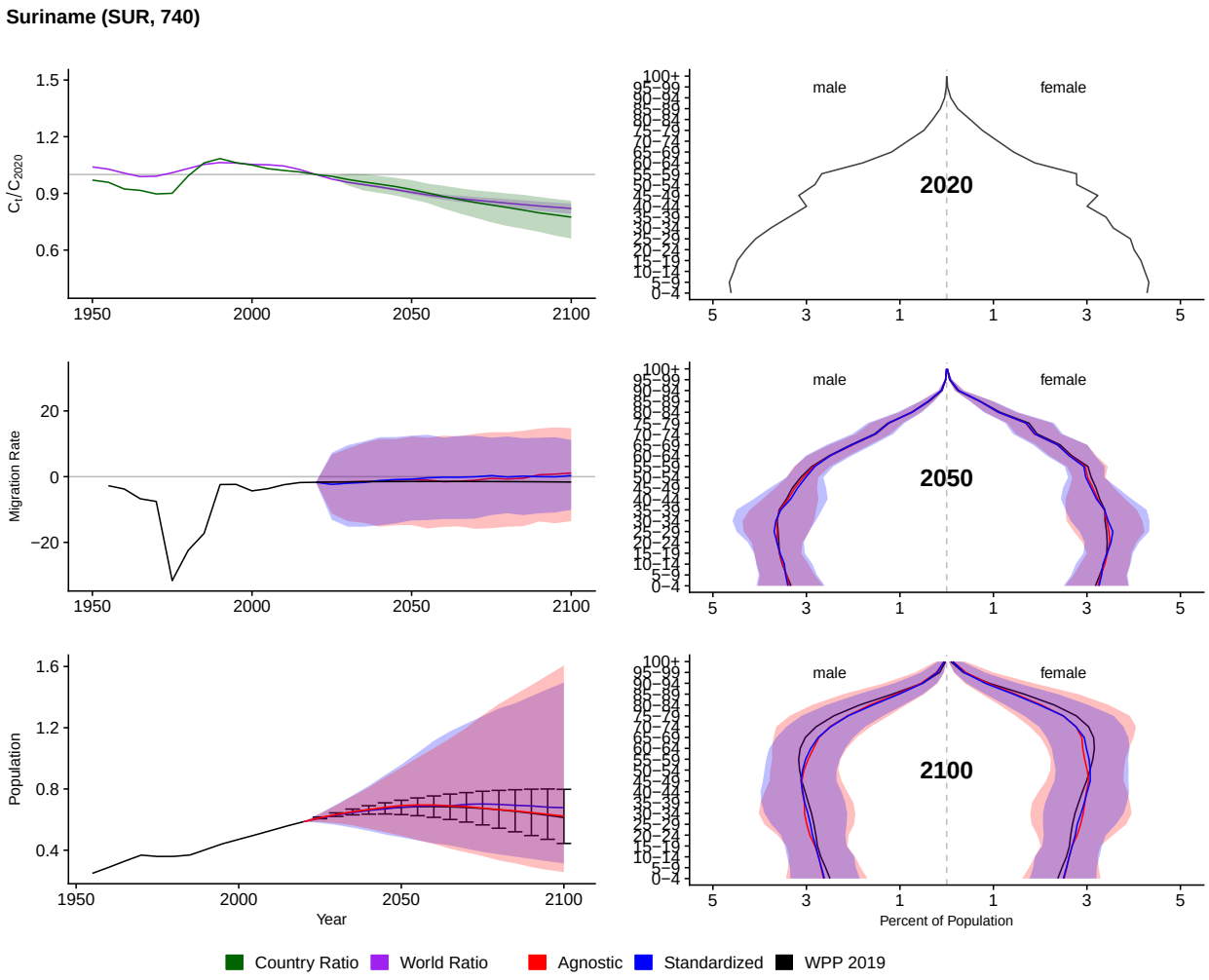


Figure A.222: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

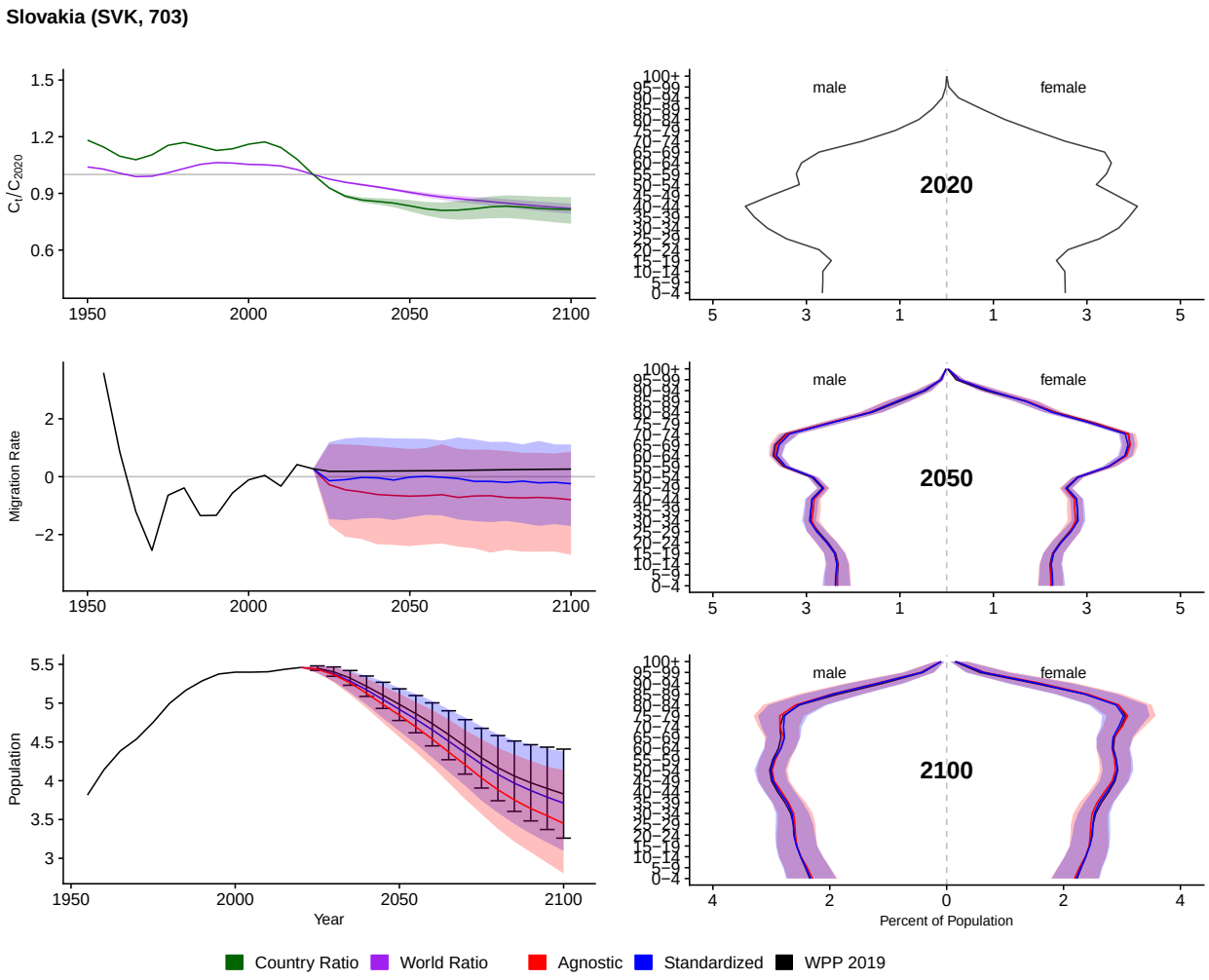


Figure A.223: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

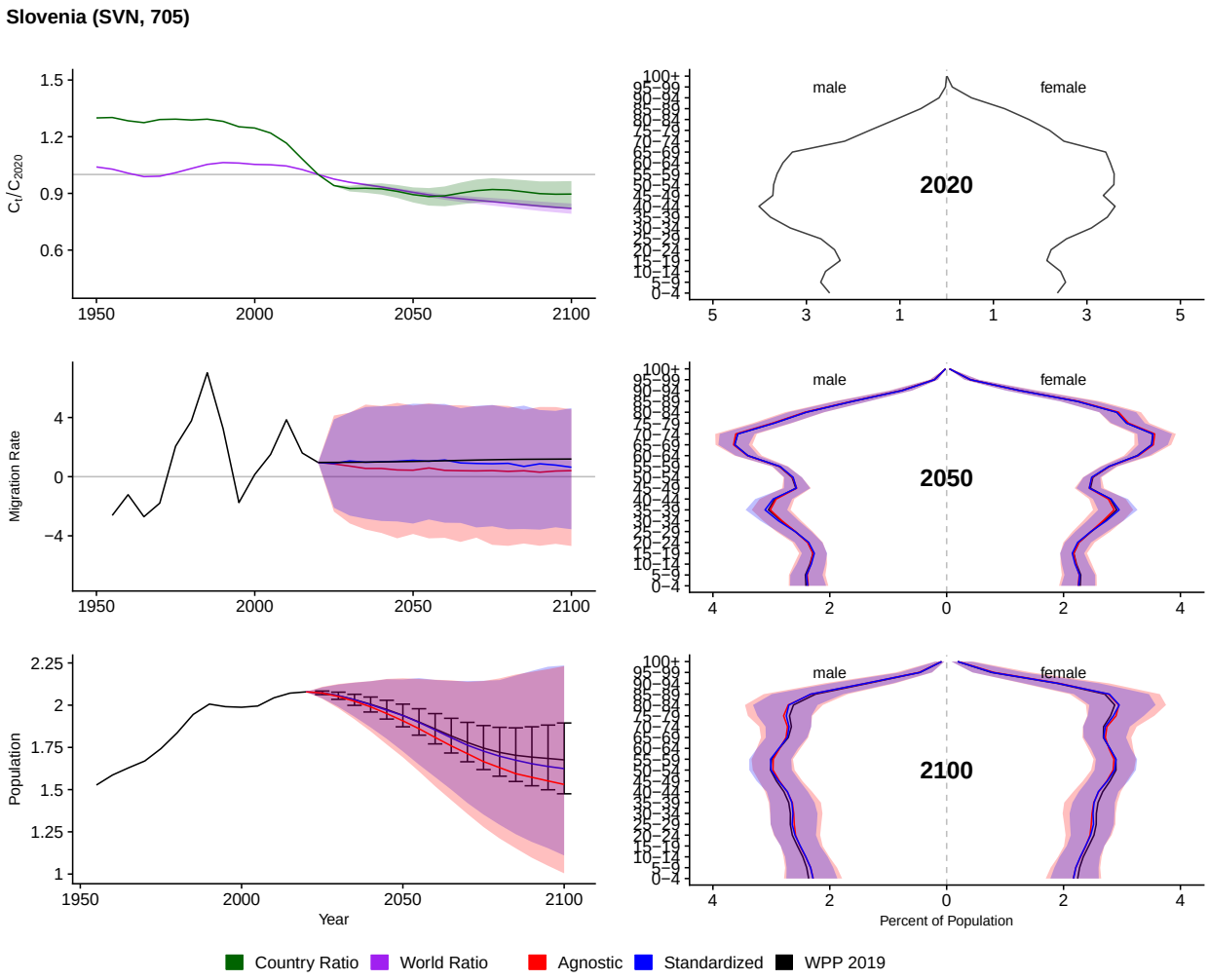


Figure A.224: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

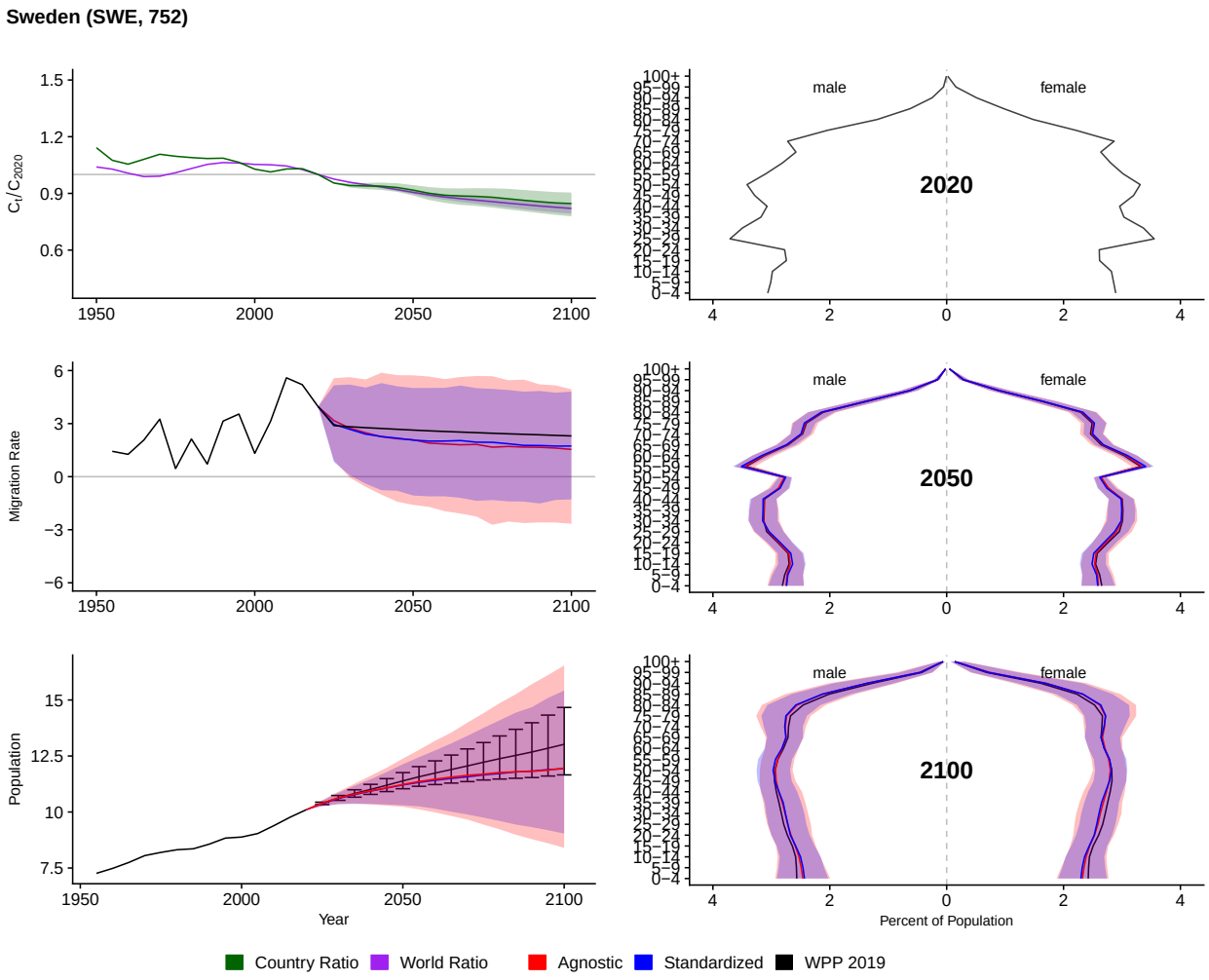


Figure A.225: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

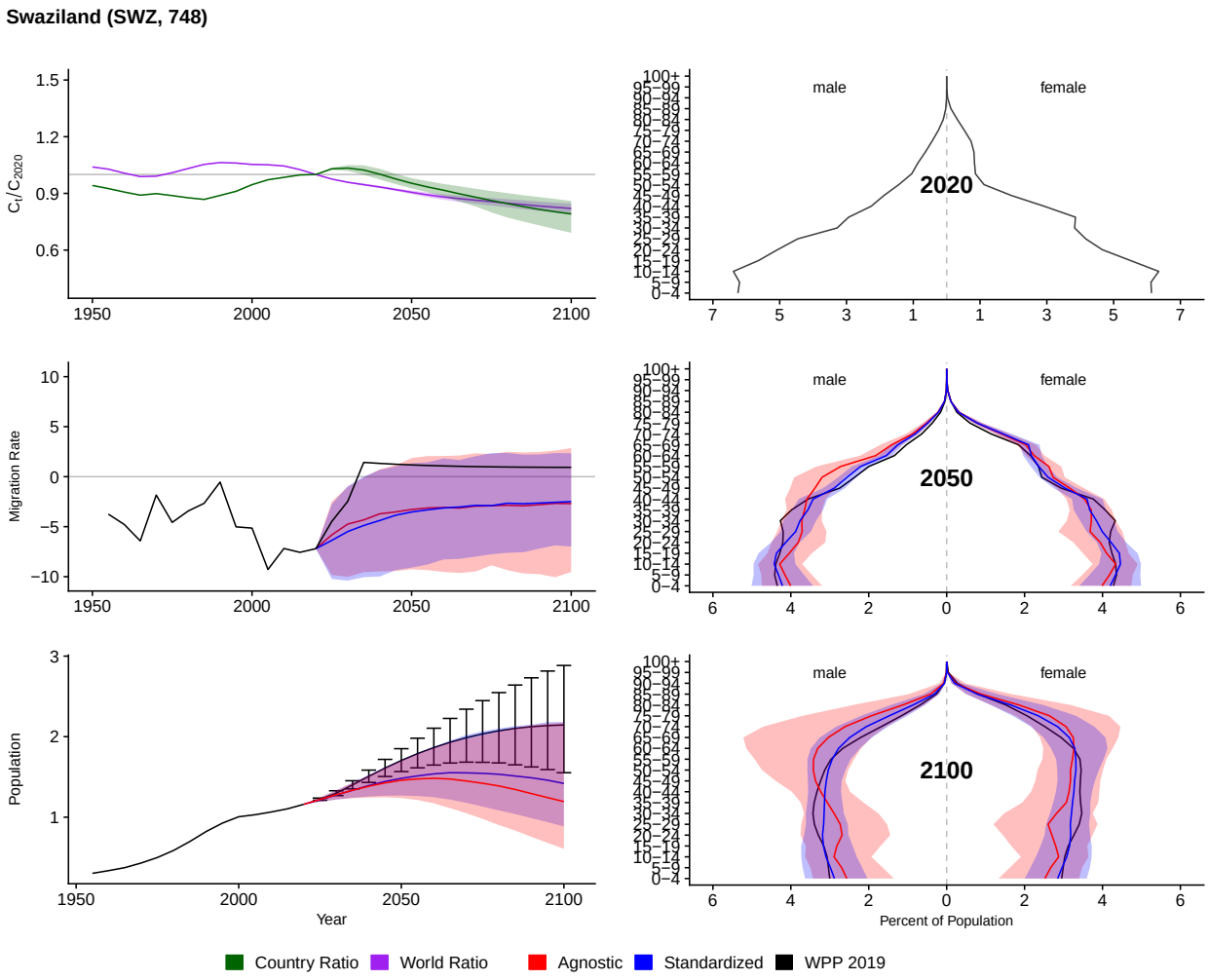


Figure A.226: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

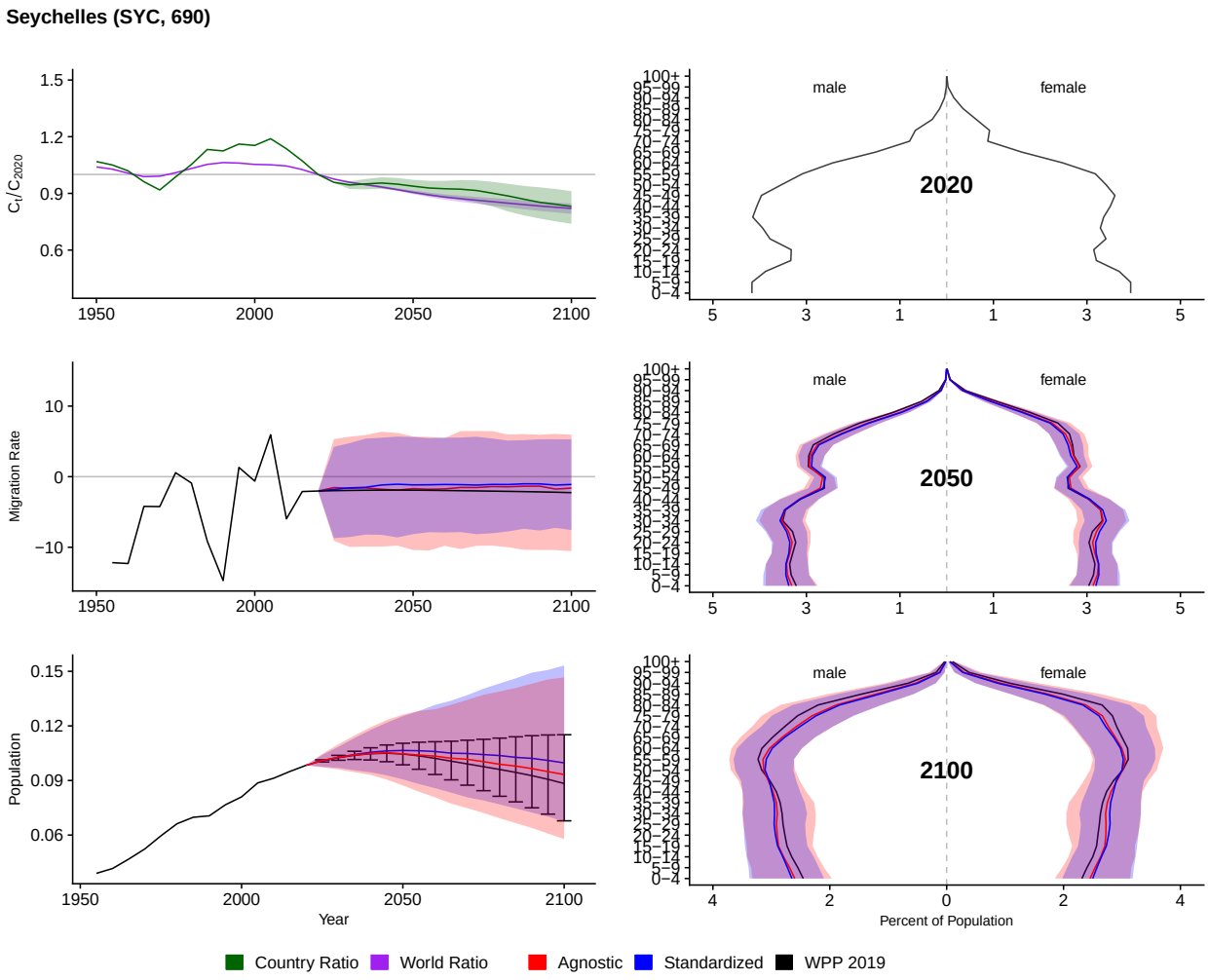


Figure A.227: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

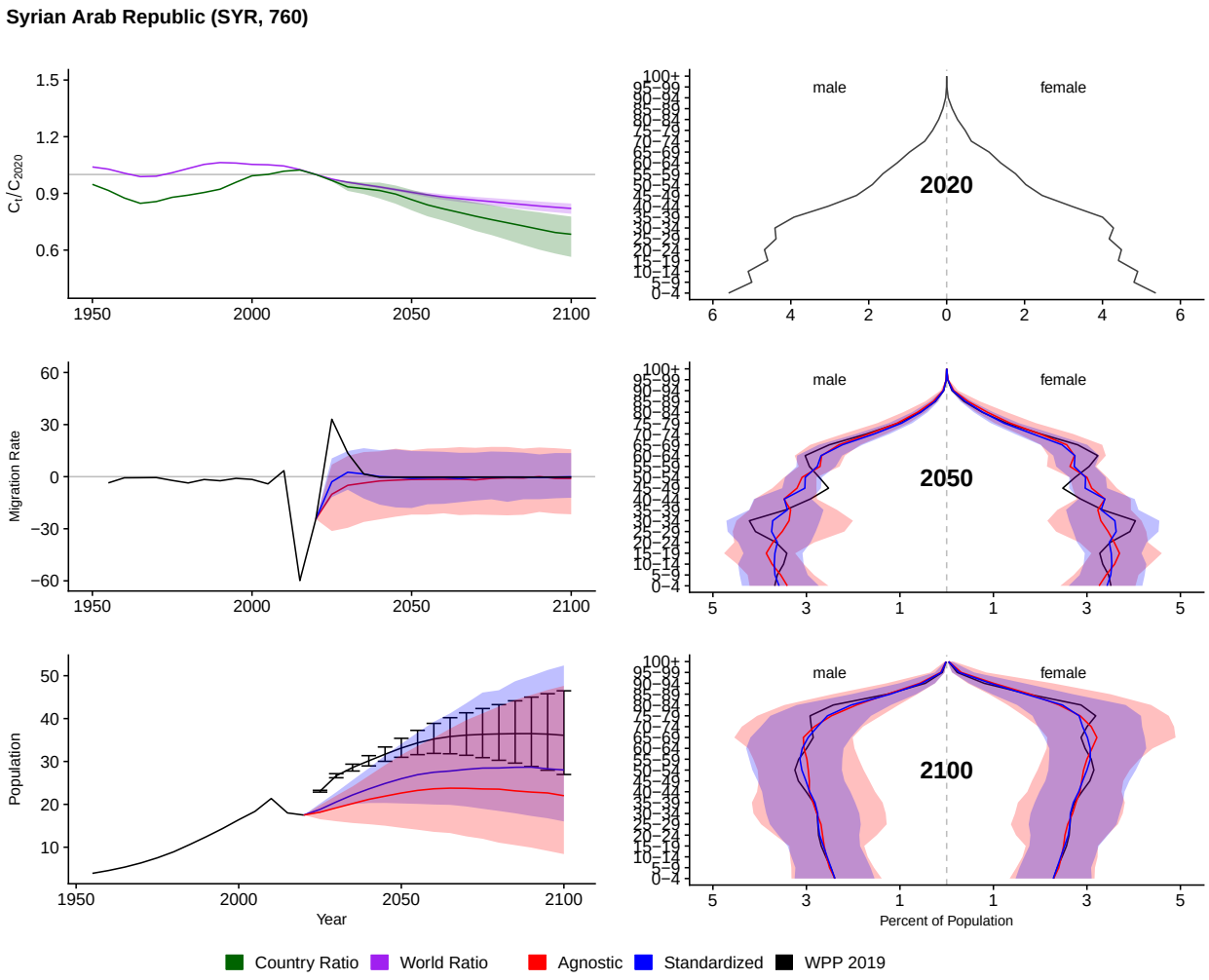


Figure A.228: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

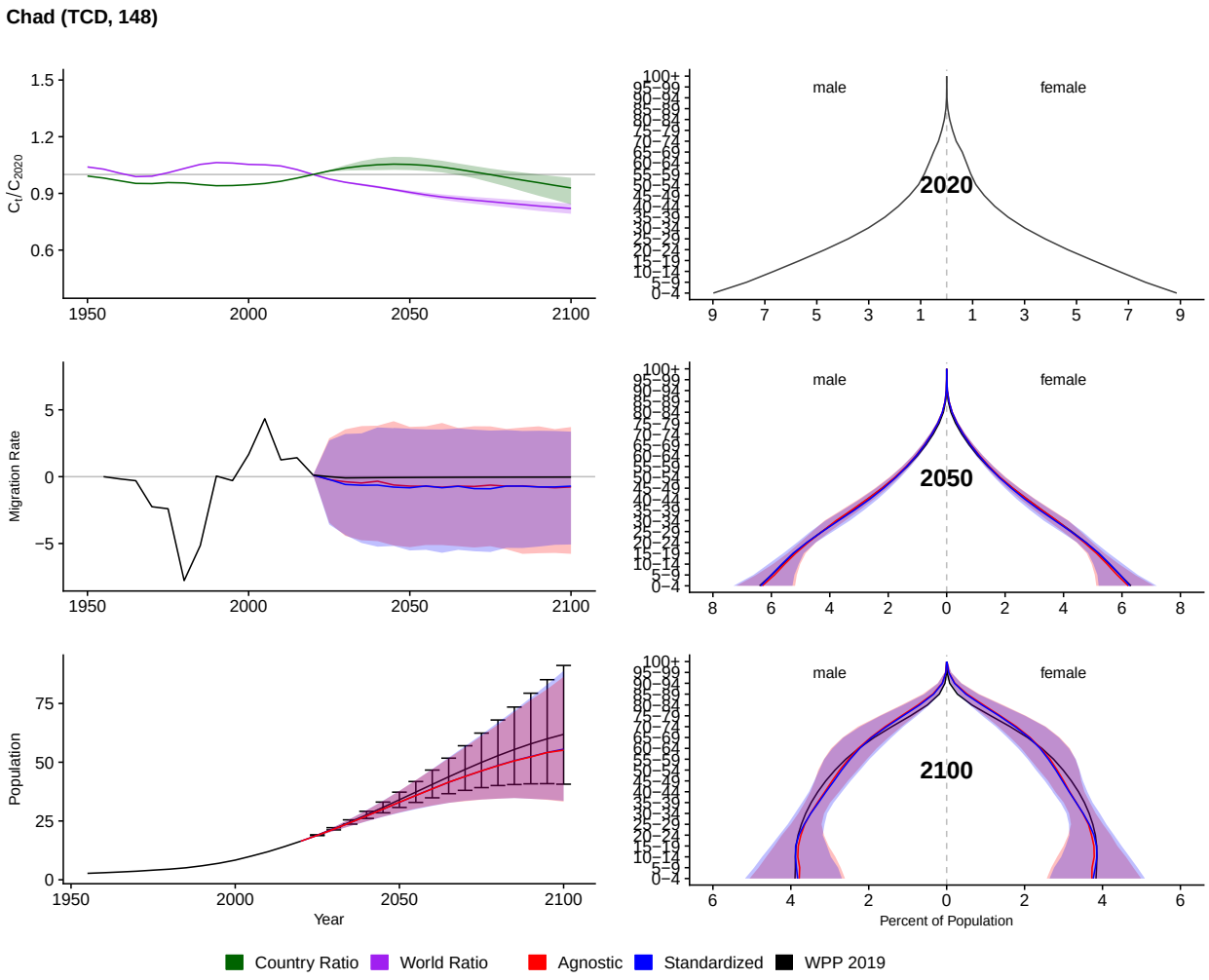


Figure A.229: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

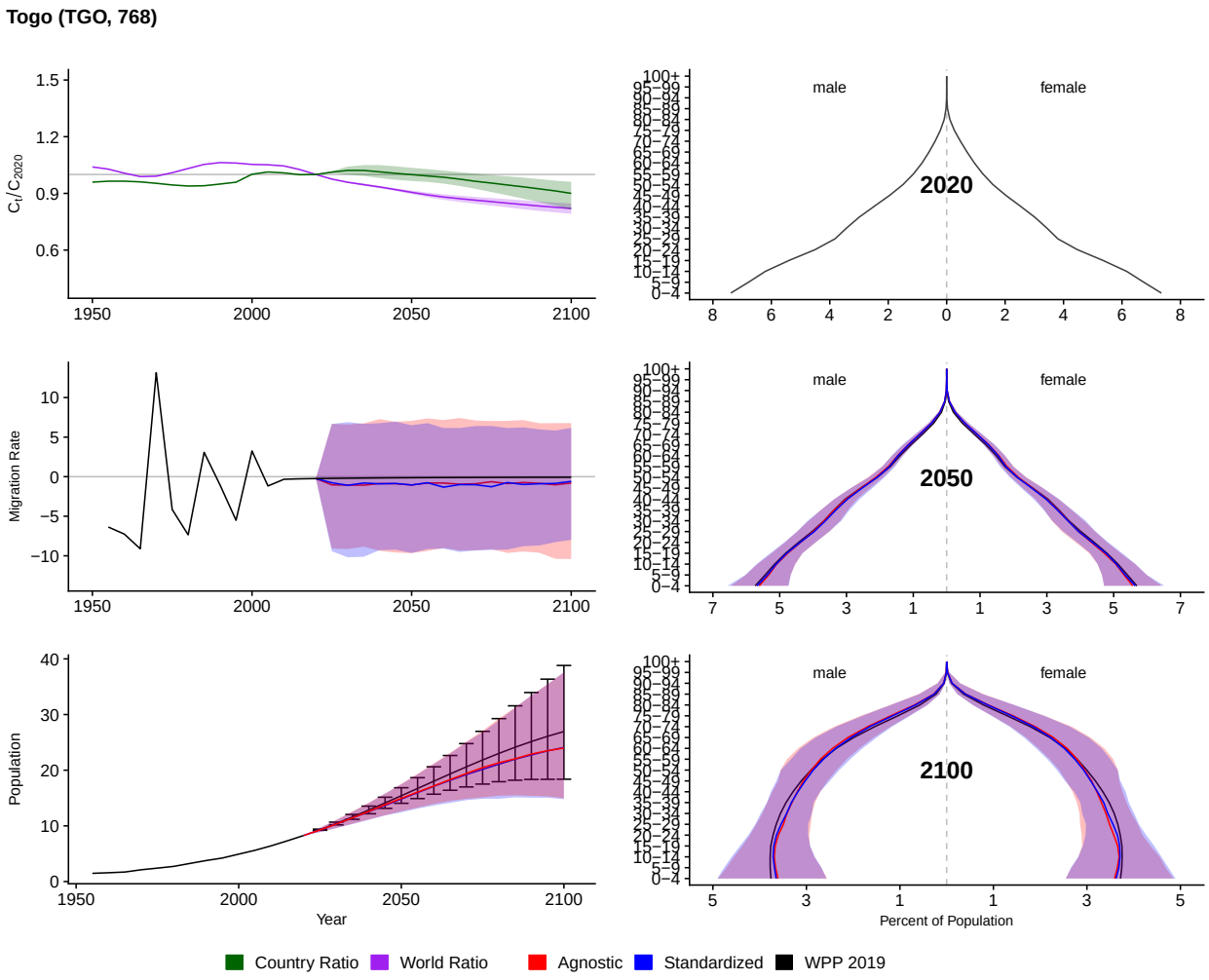


Figure A.230: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

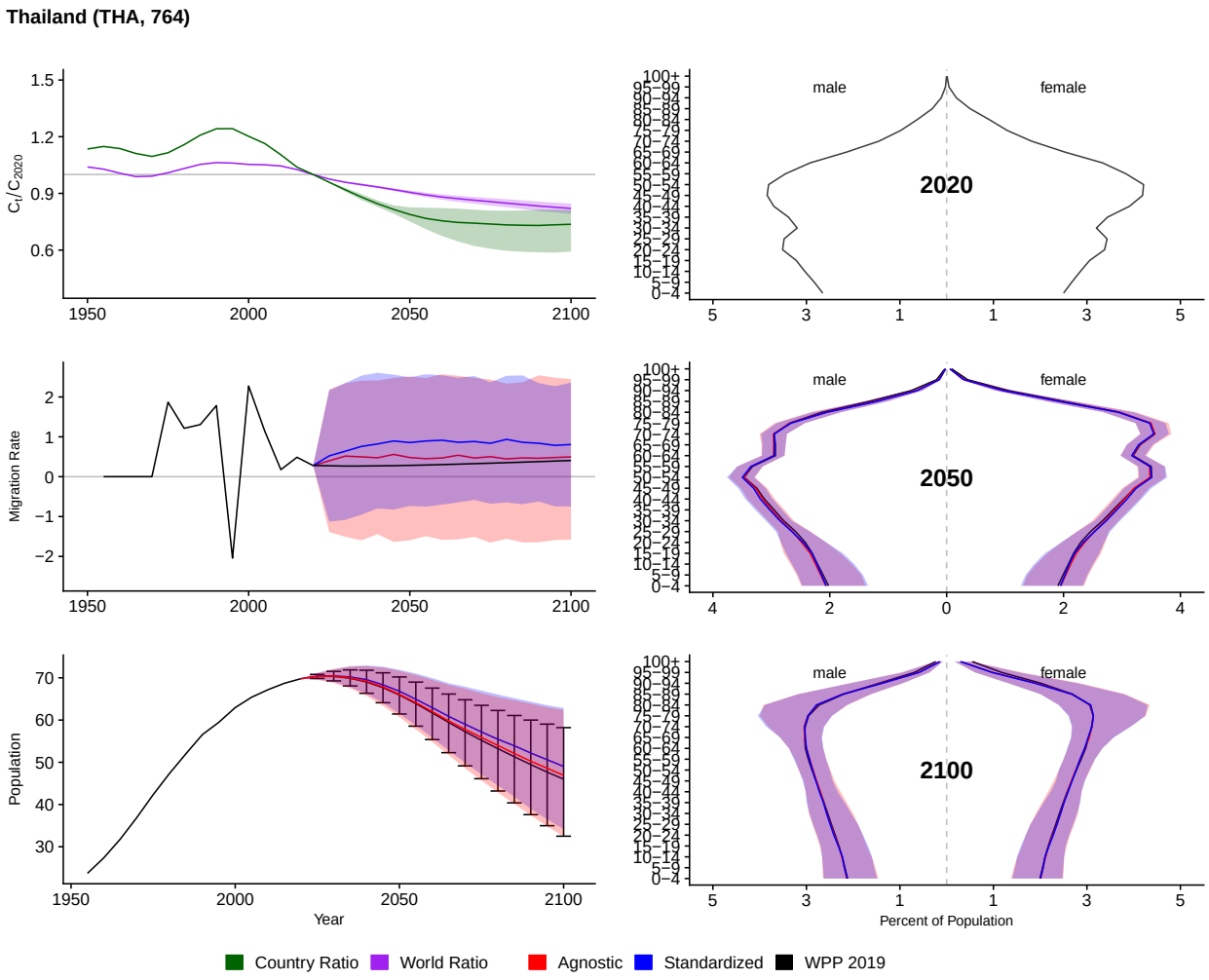


Figure A.231: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

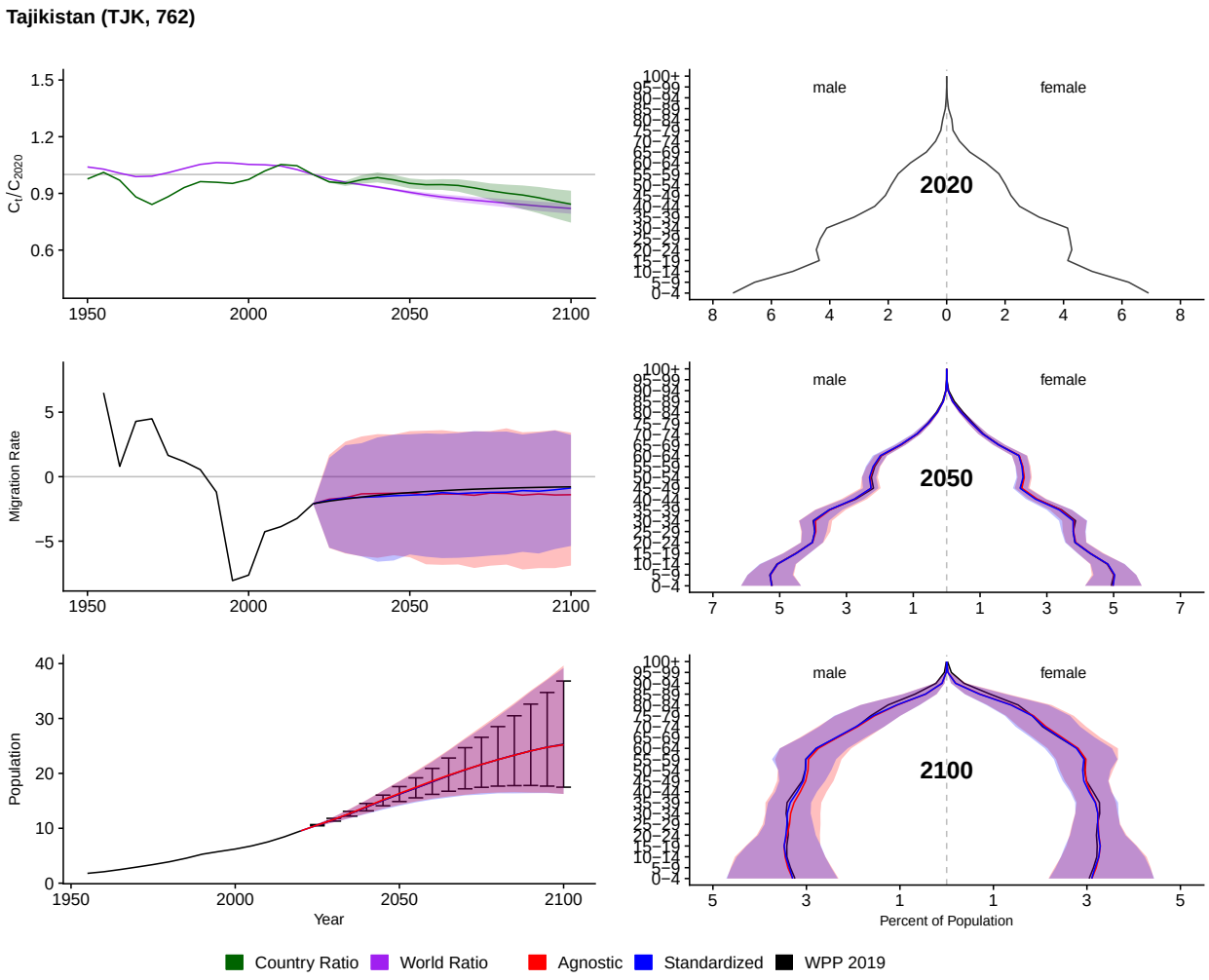


Figure A.232: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

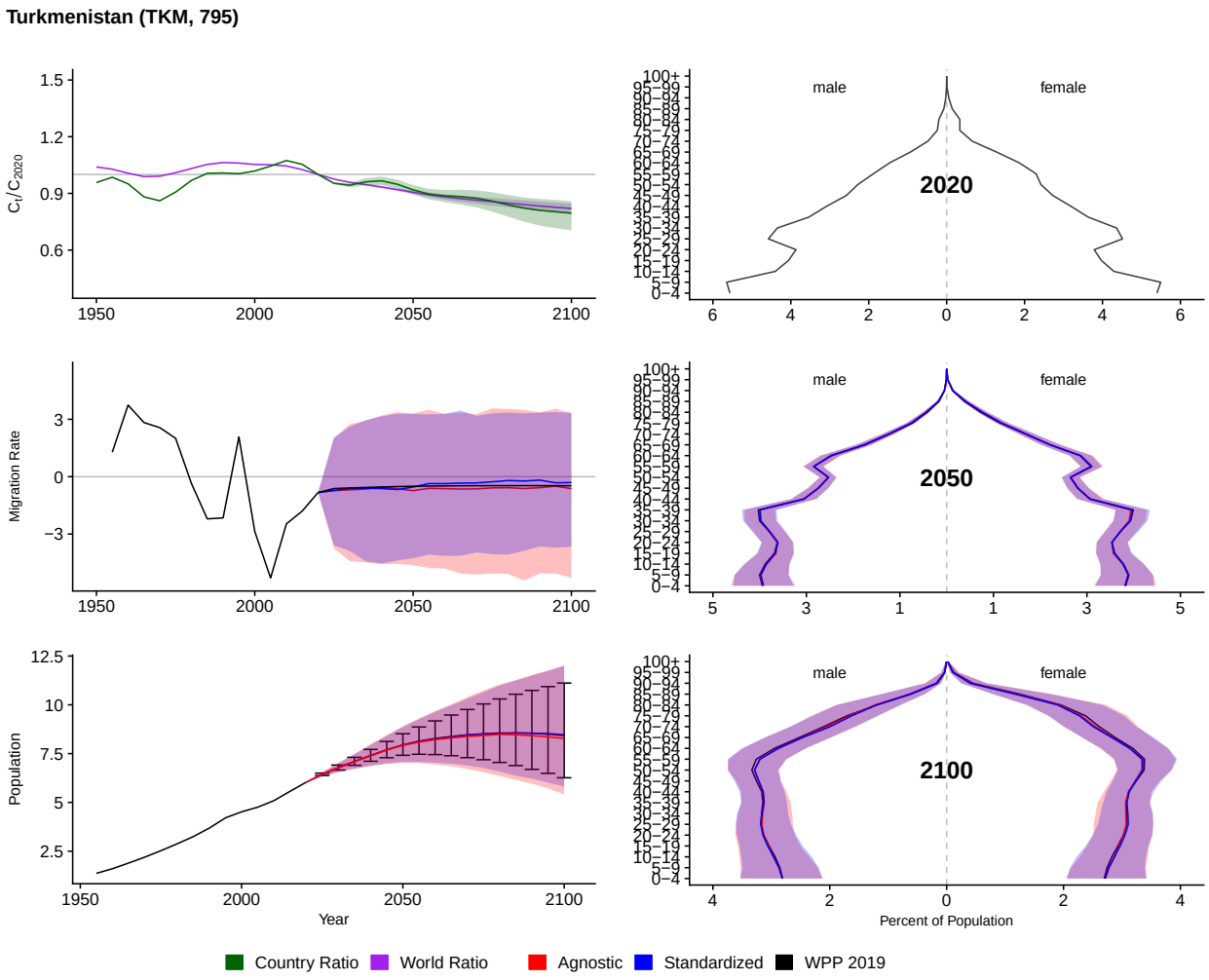


Figure A.233: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

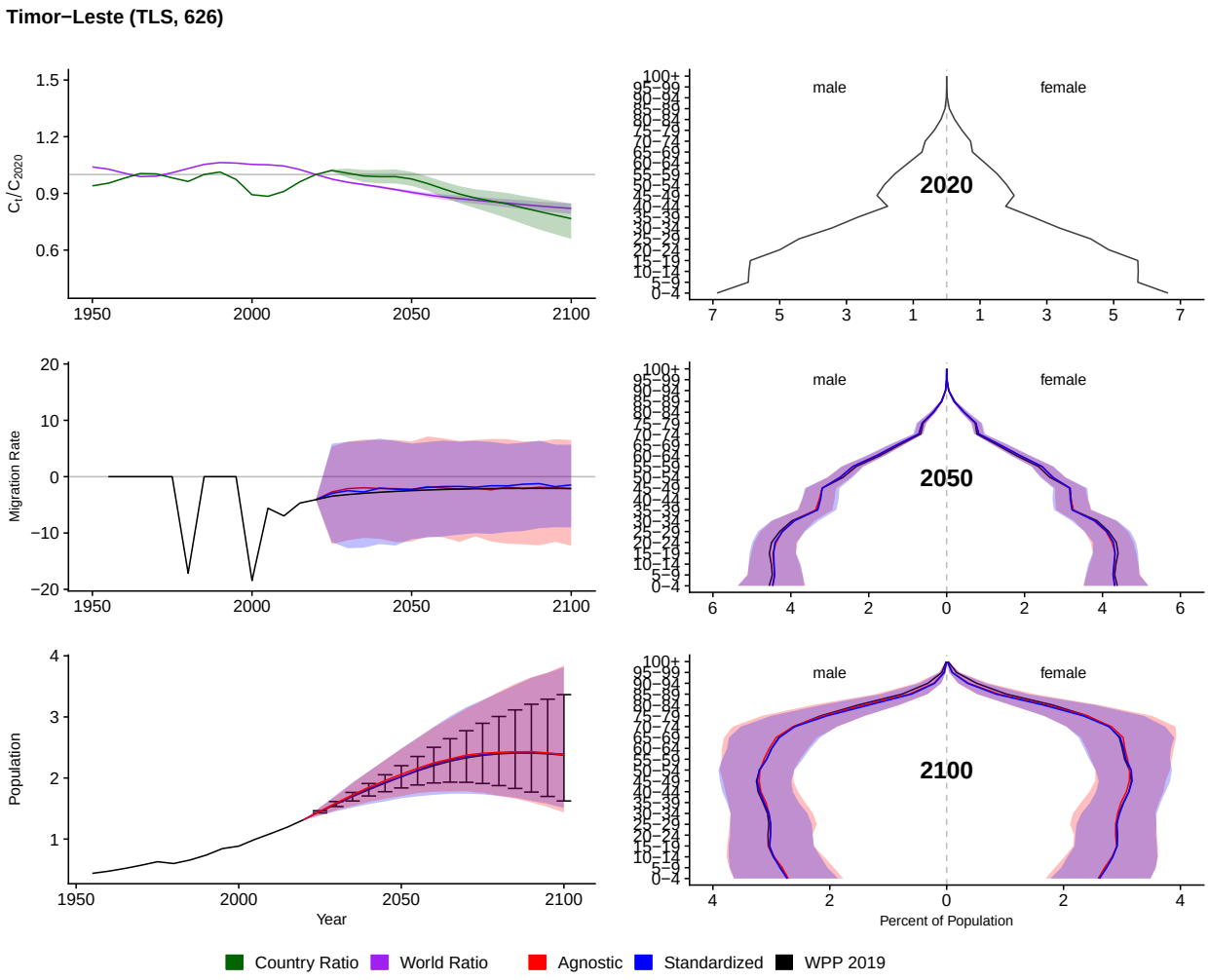


Figure A.234: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

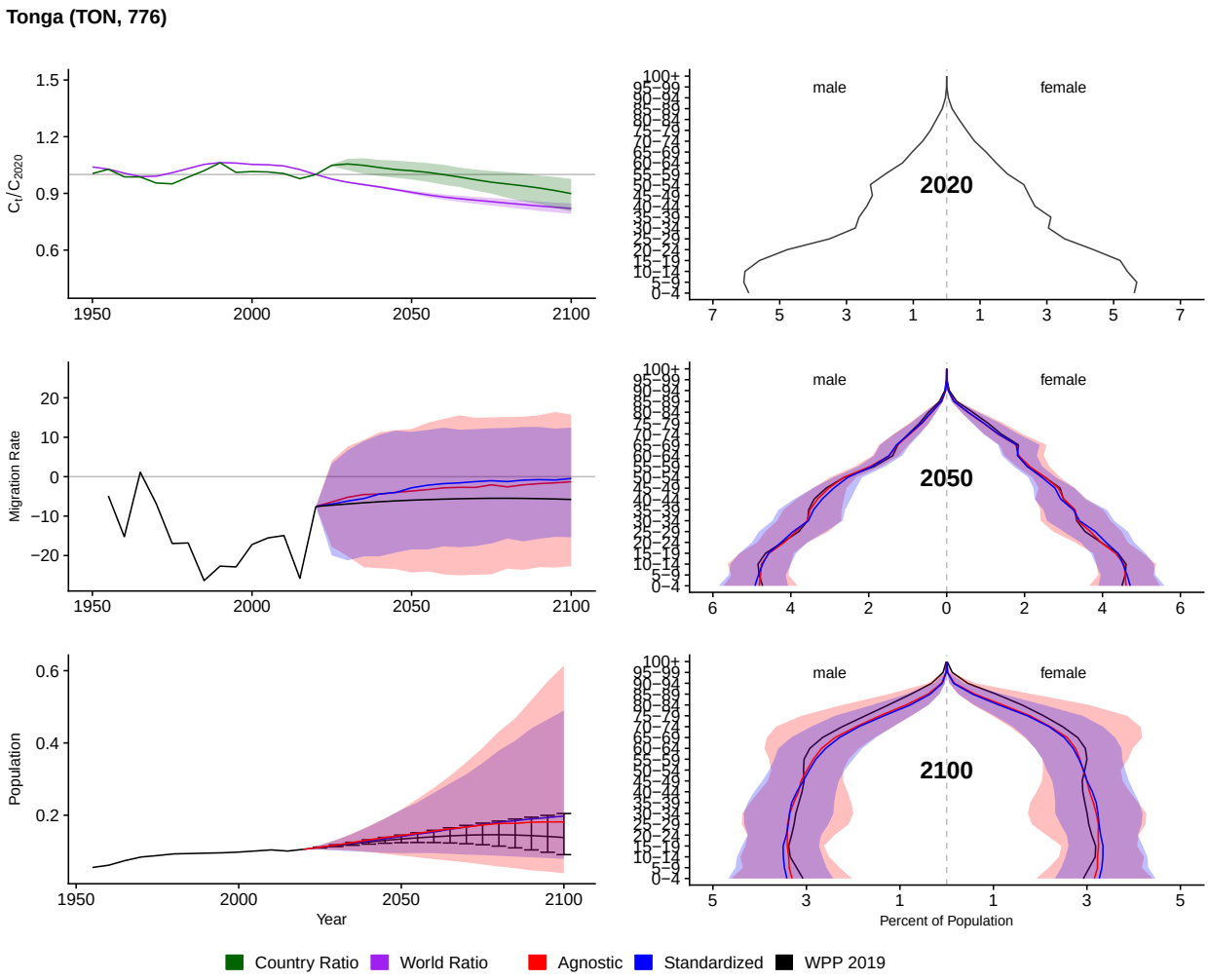


Figure A.235: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

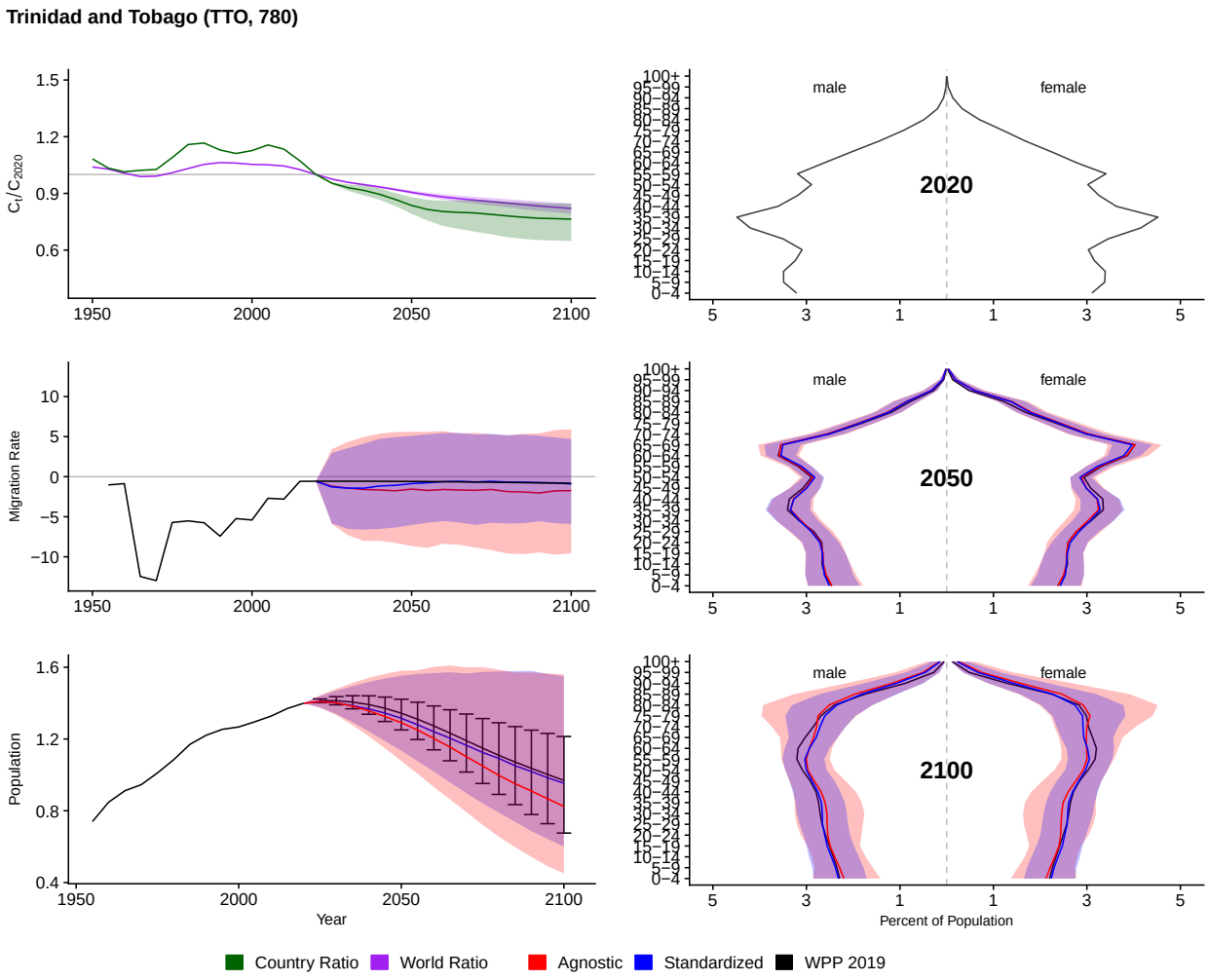


Figure A.236: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

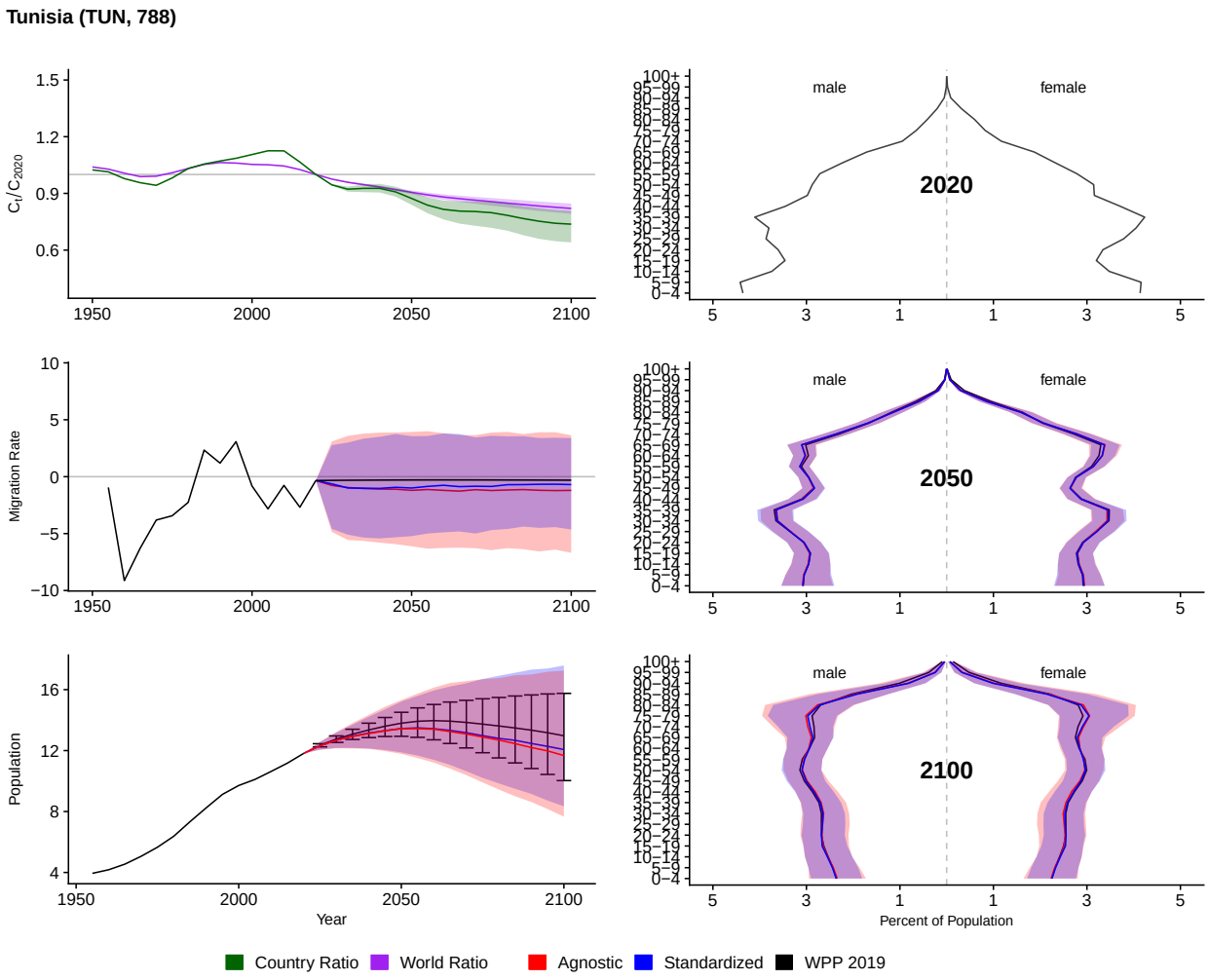


Figure A.237: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

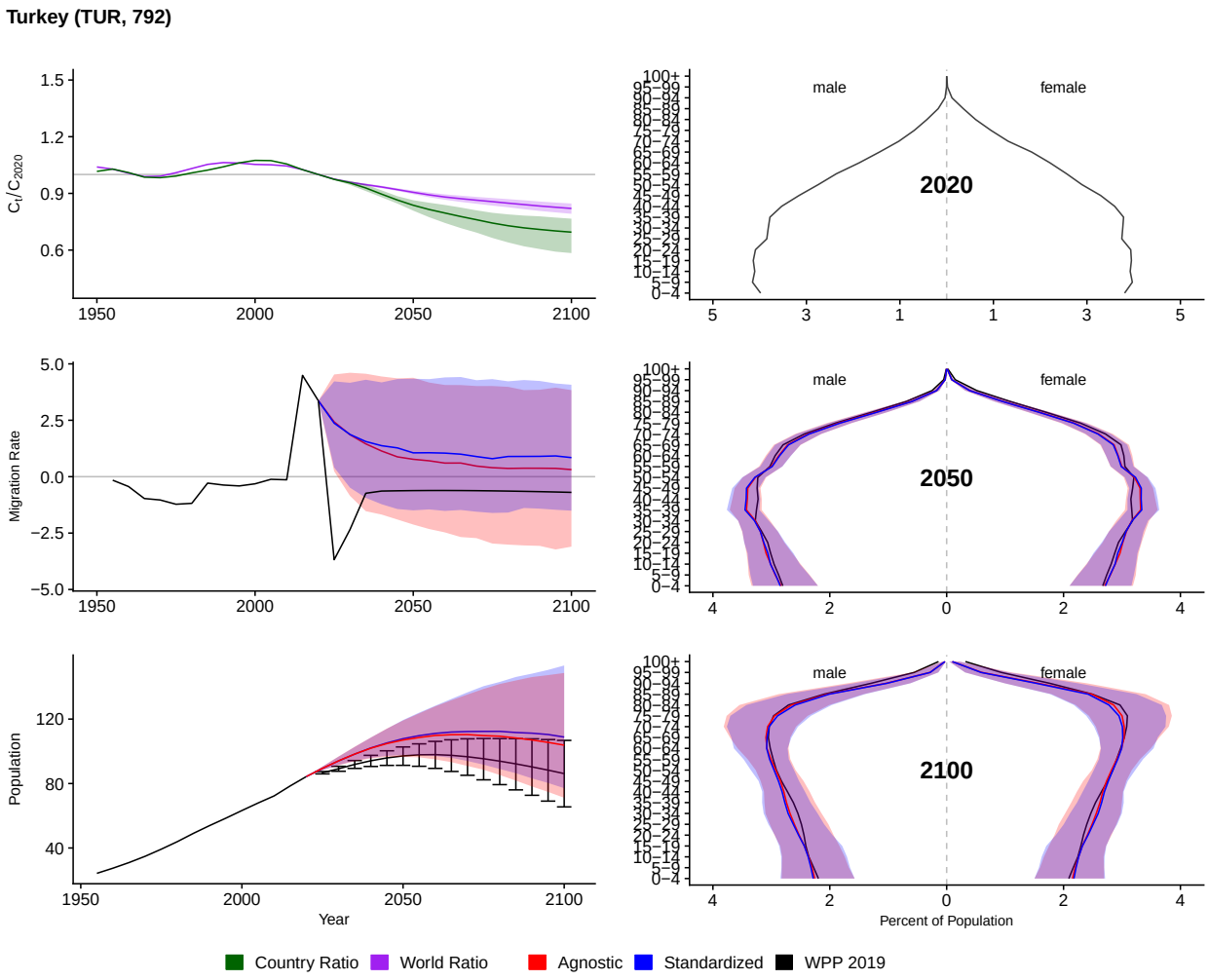


Figure A.238: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

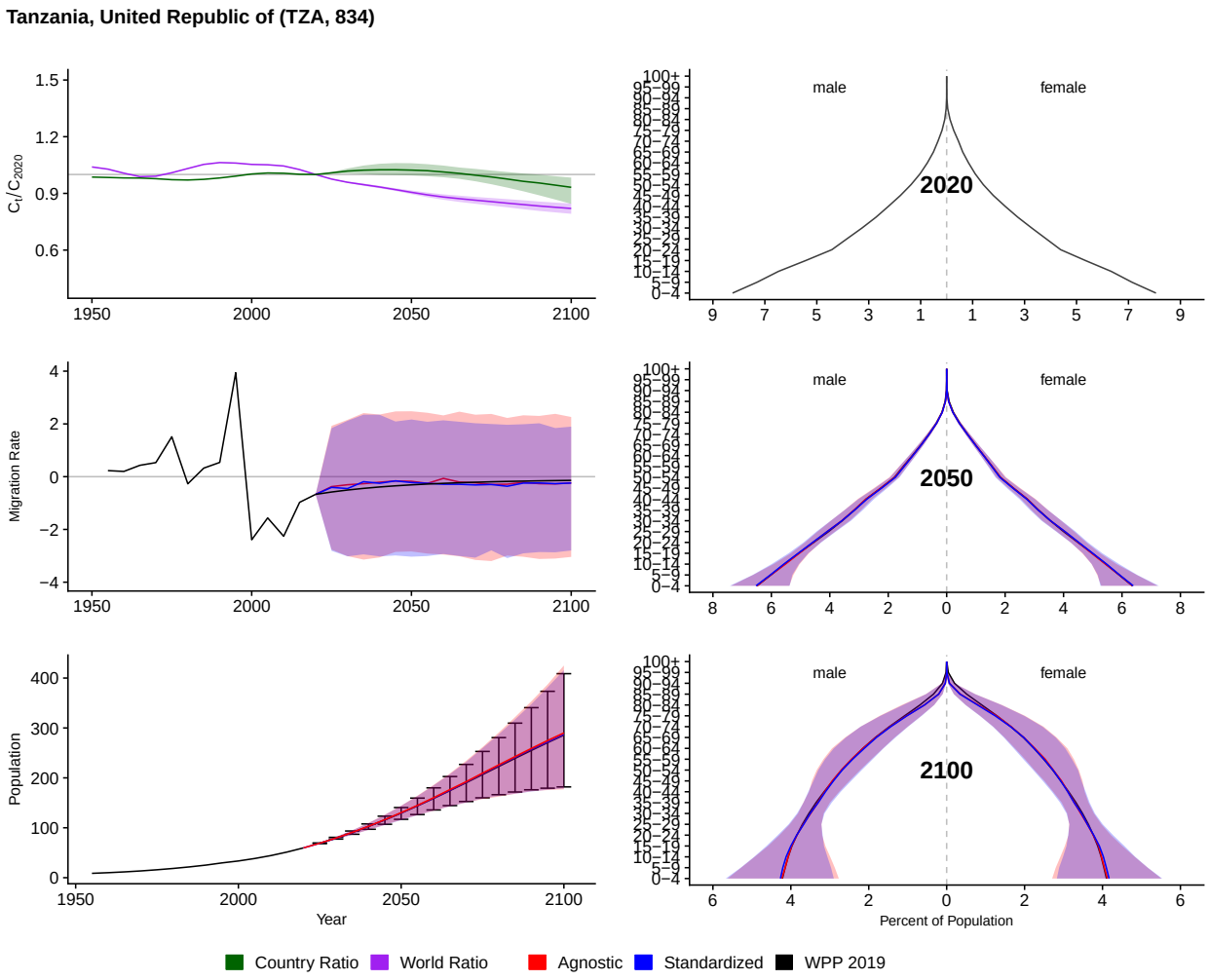


Figure A.239: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

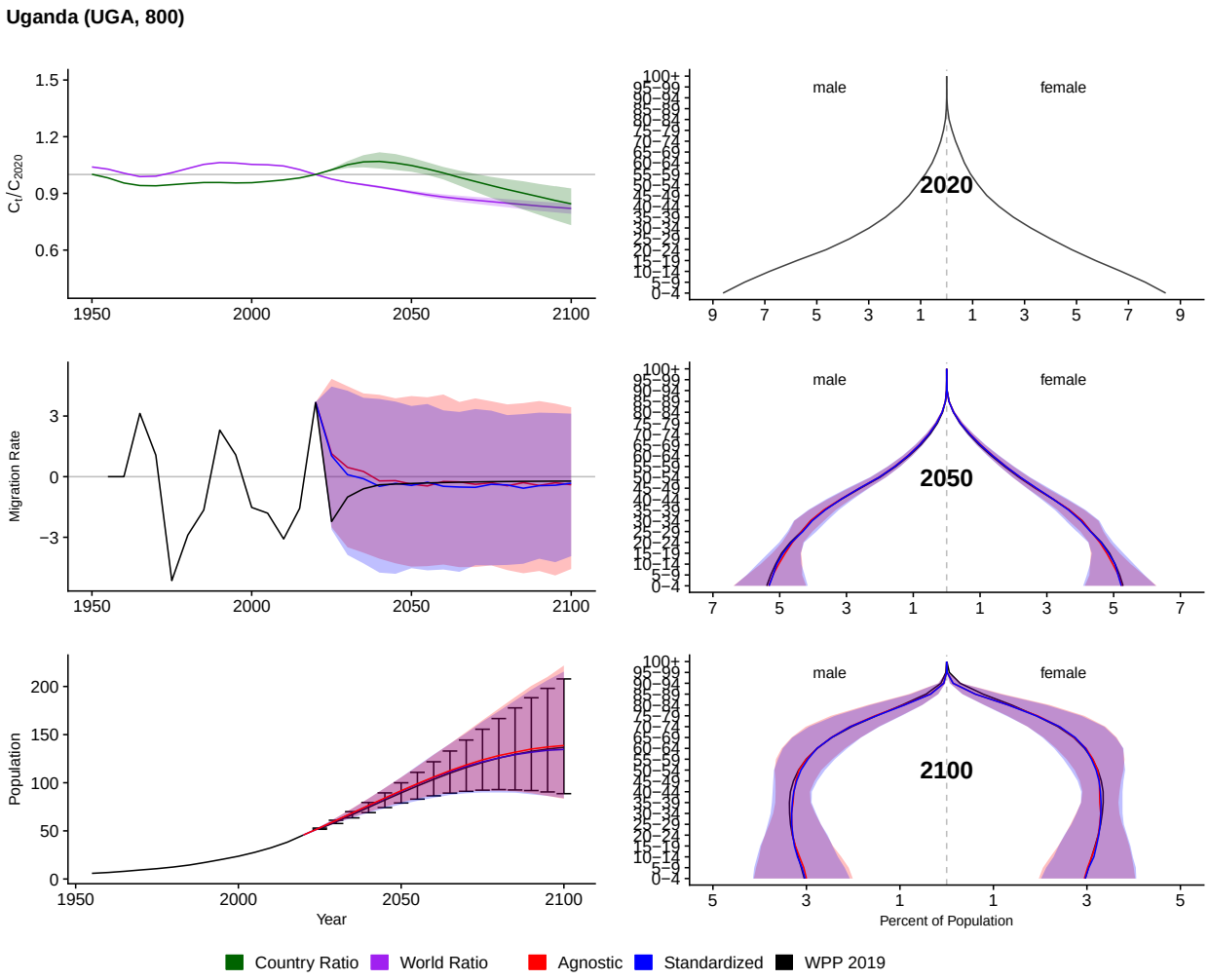


Figure A.240: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

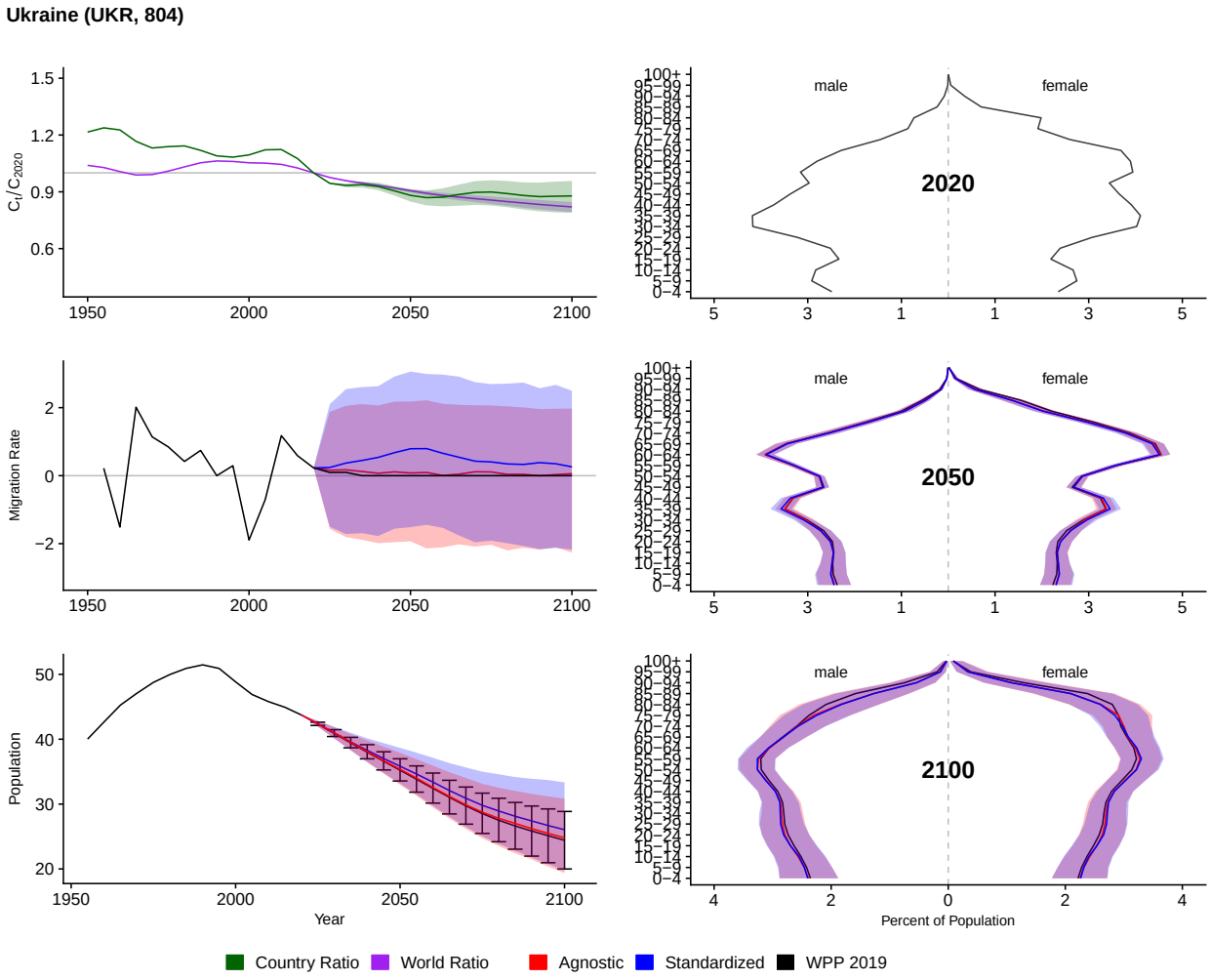


Figure A.241: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

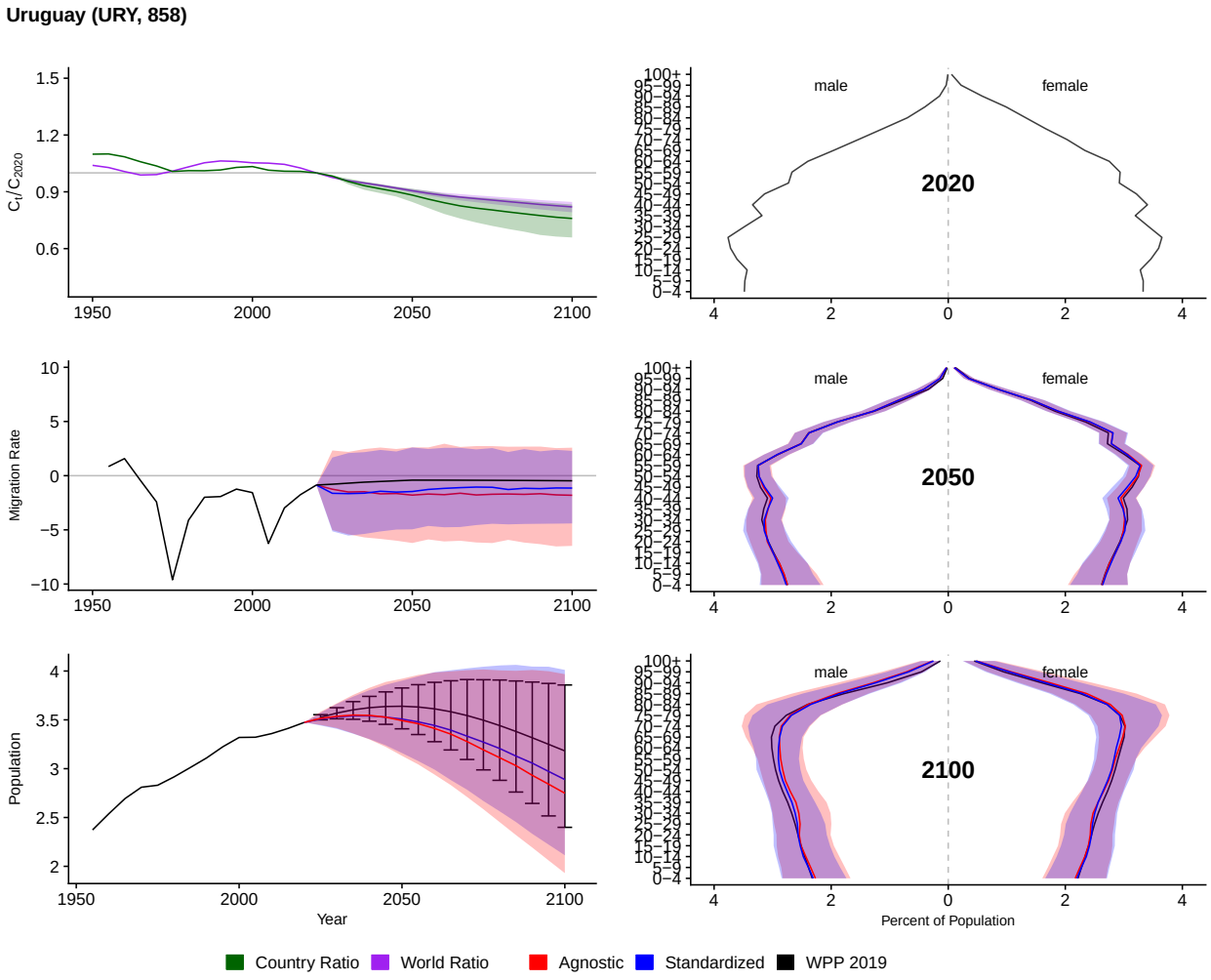


Figure A.242: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

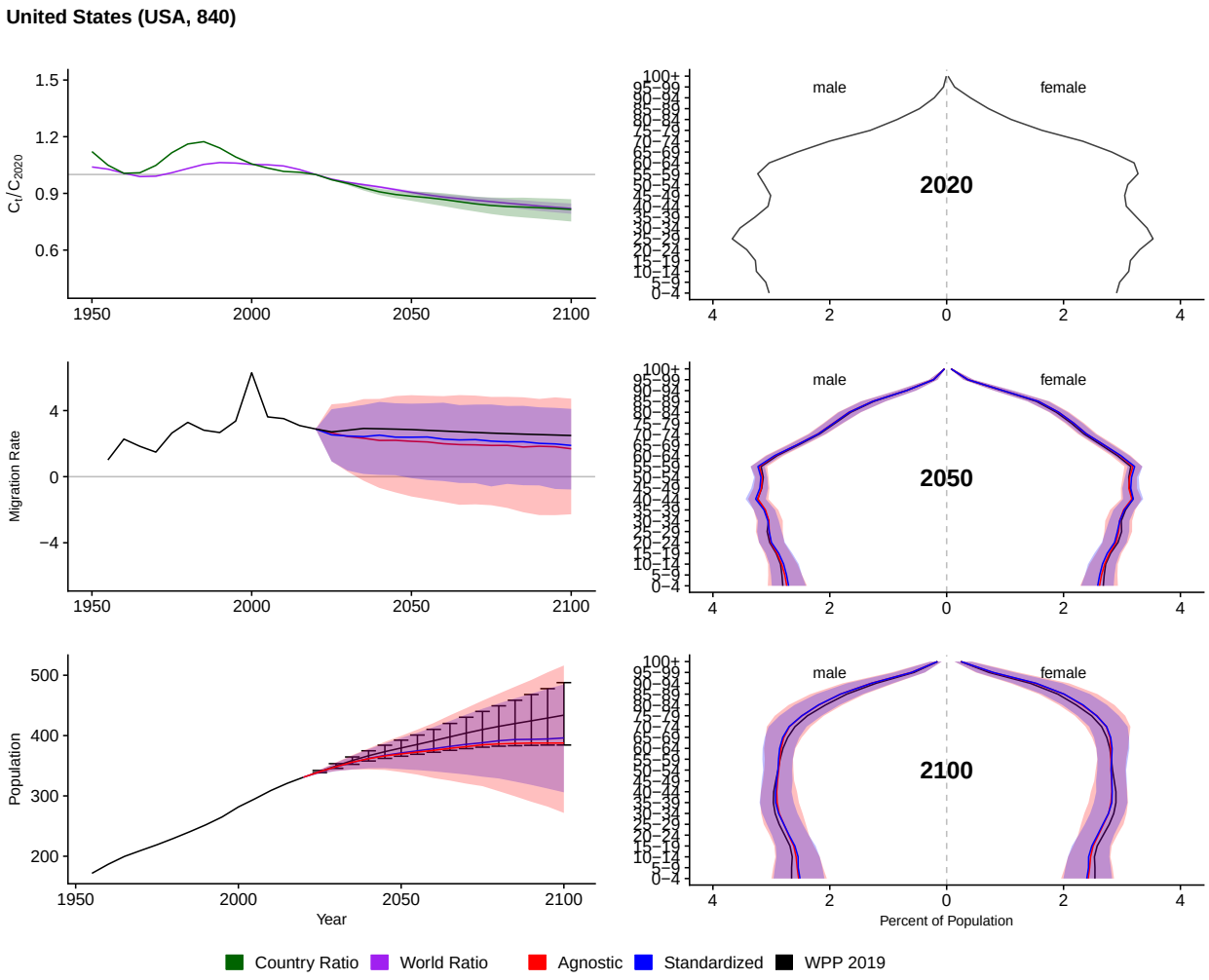


Figure A.243: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

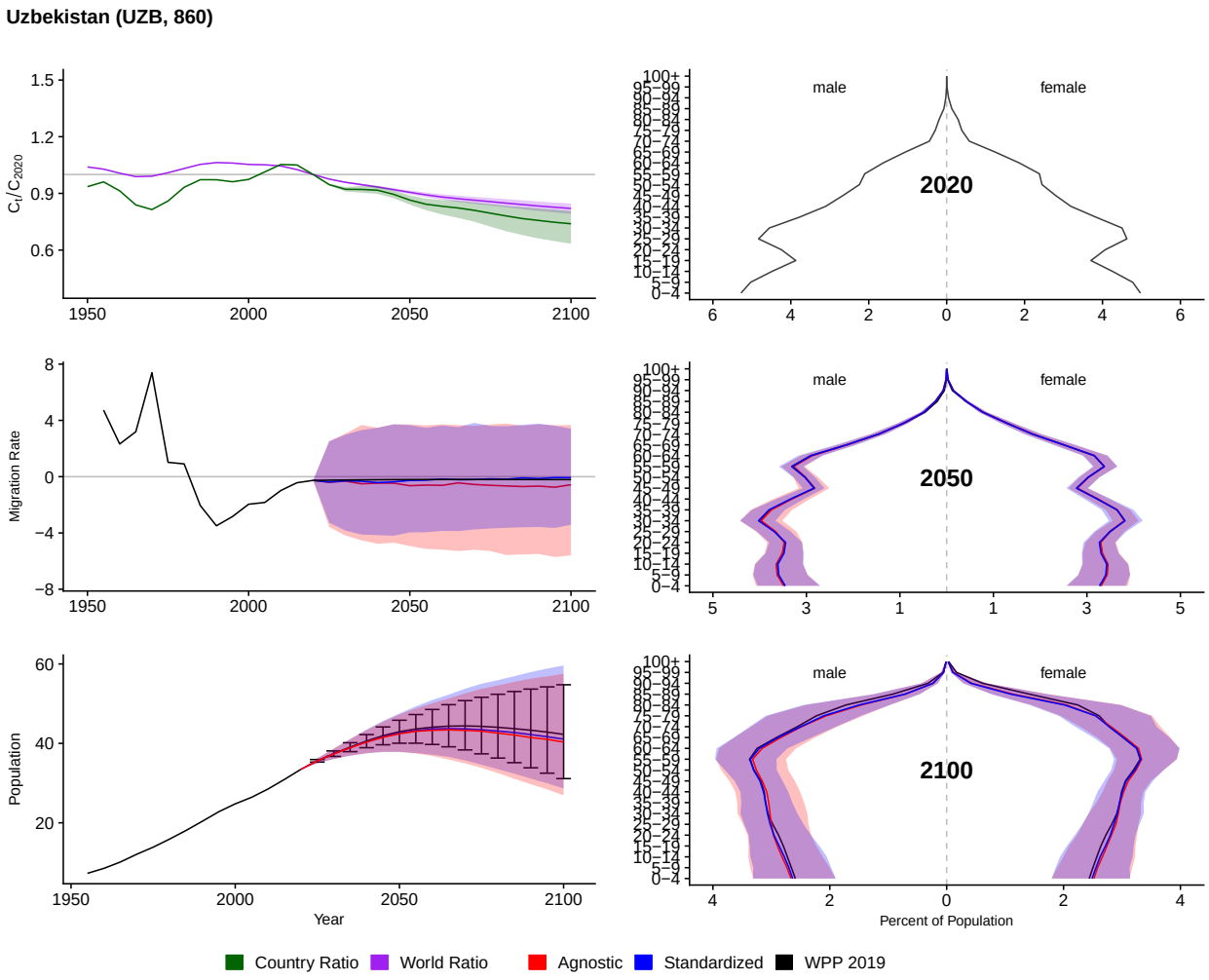


Figure A.244: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

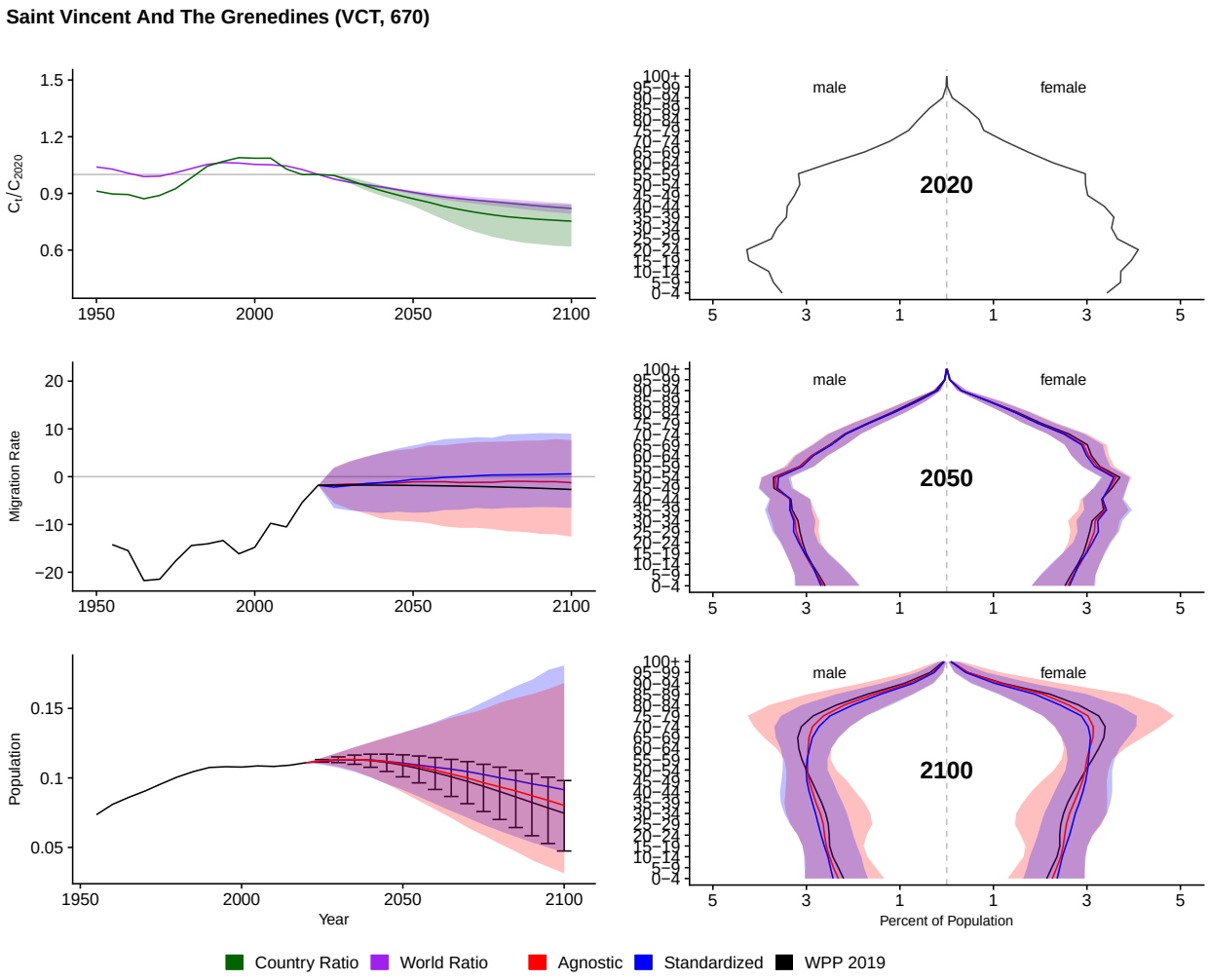


Figure A.245: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

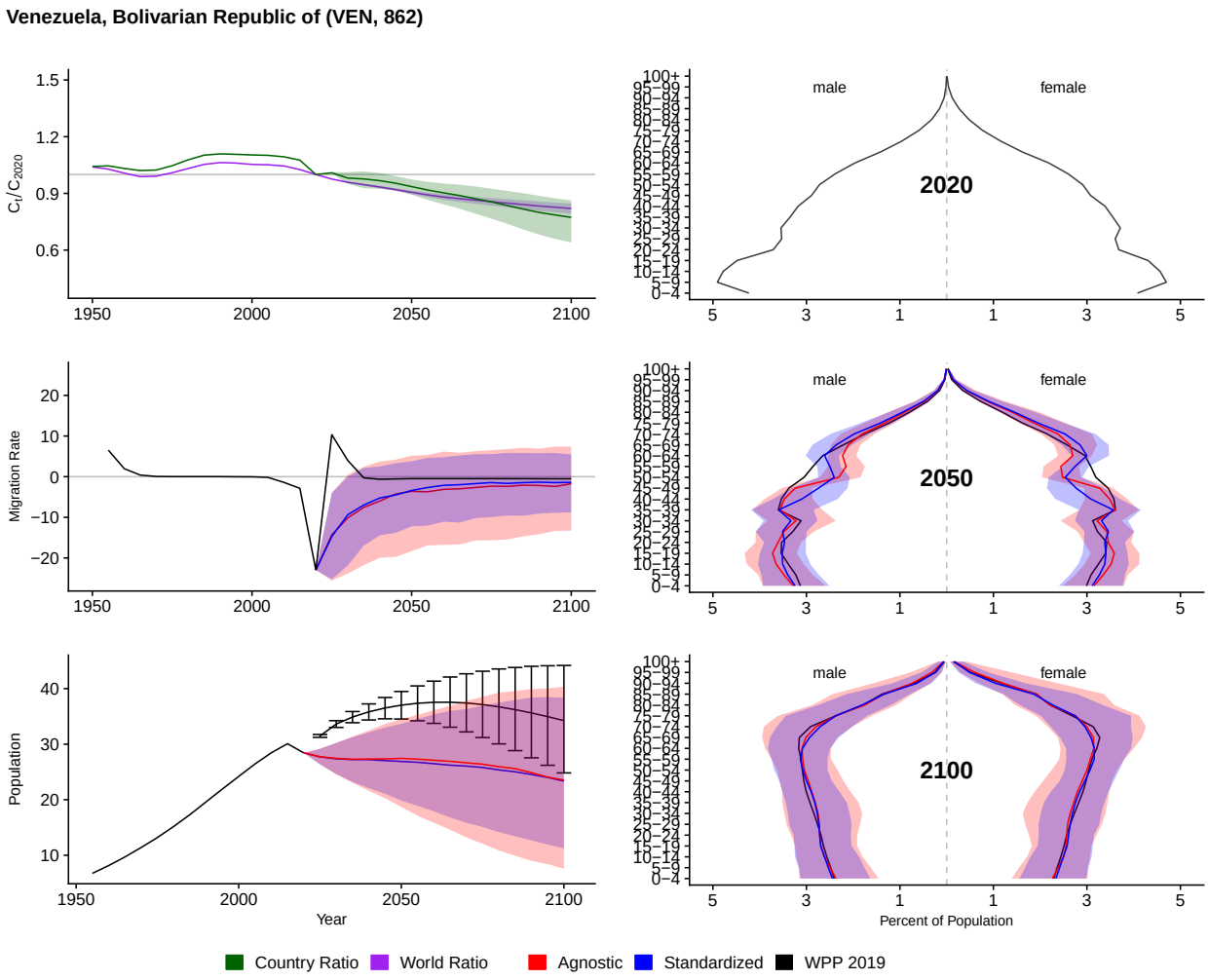


Figure A.246: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

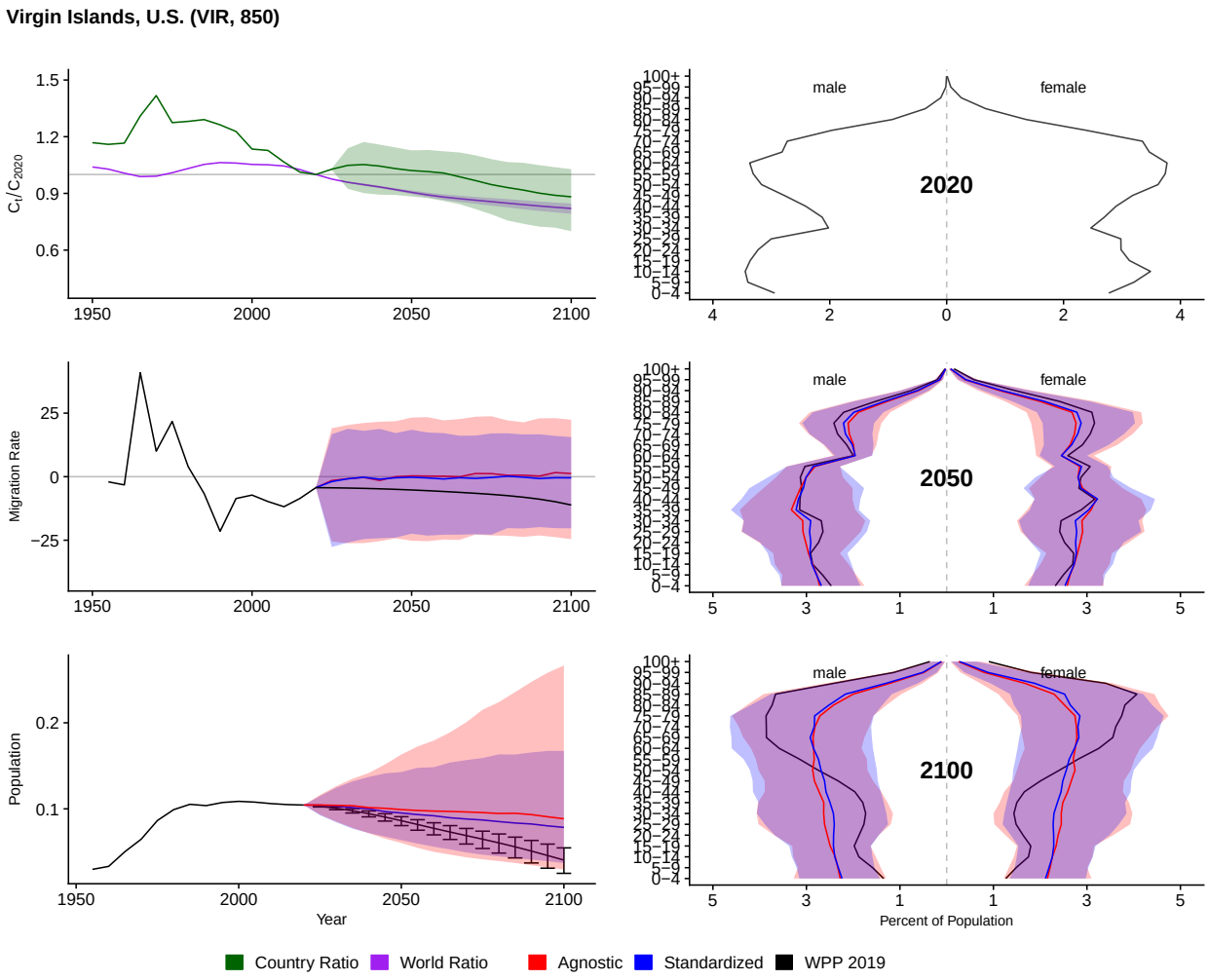


Figure A.247: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

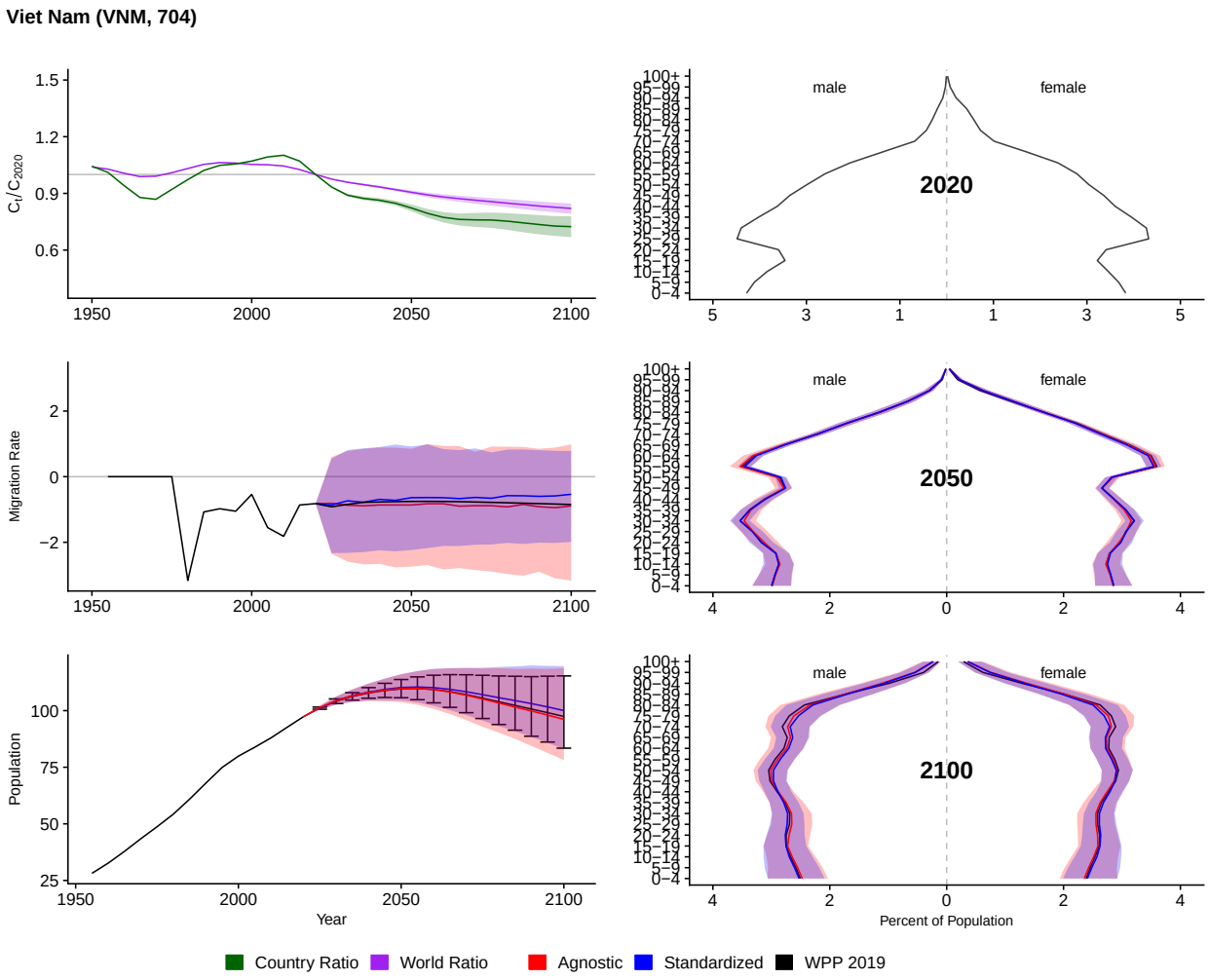


Figure A.248: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

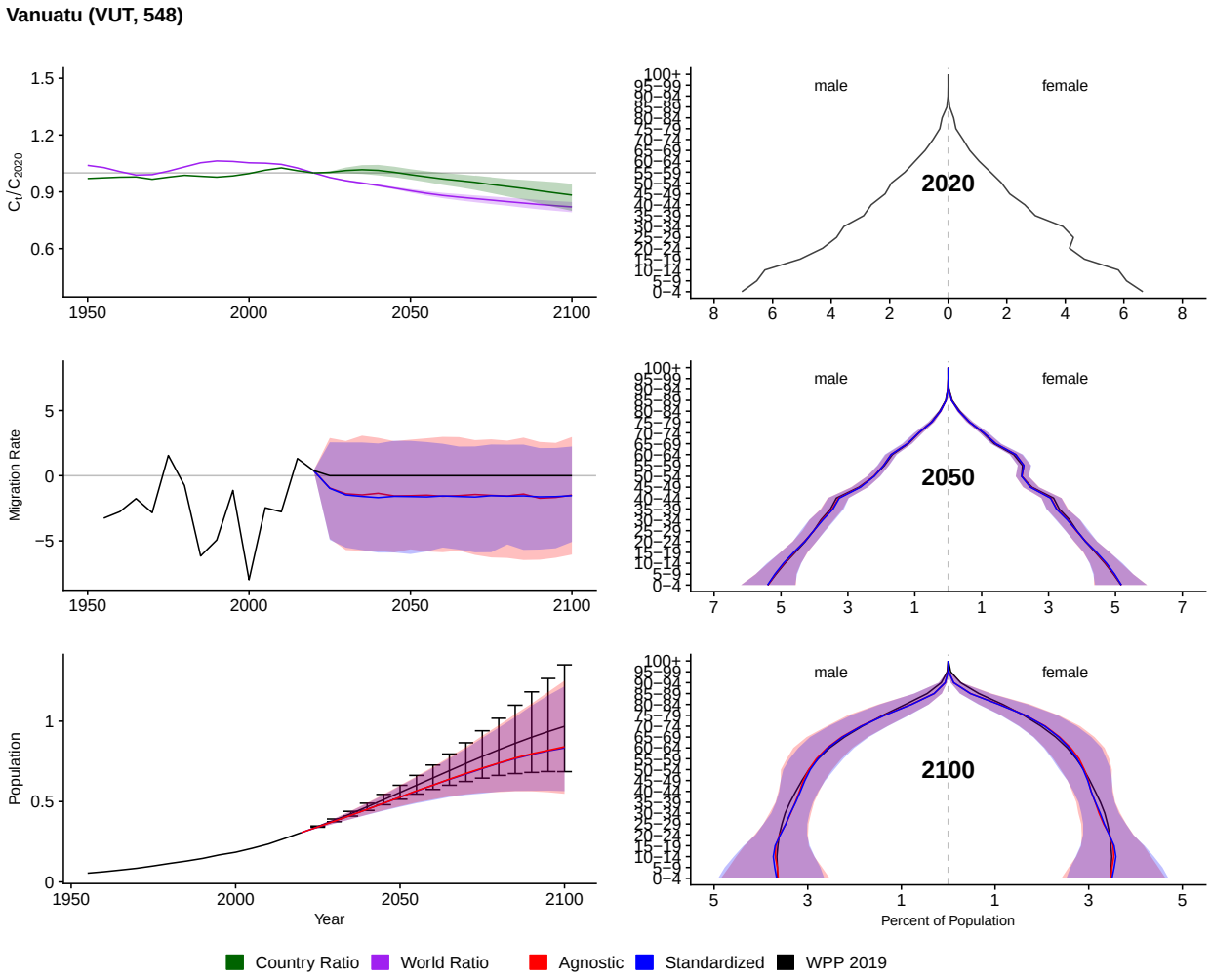


Figure A.249: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

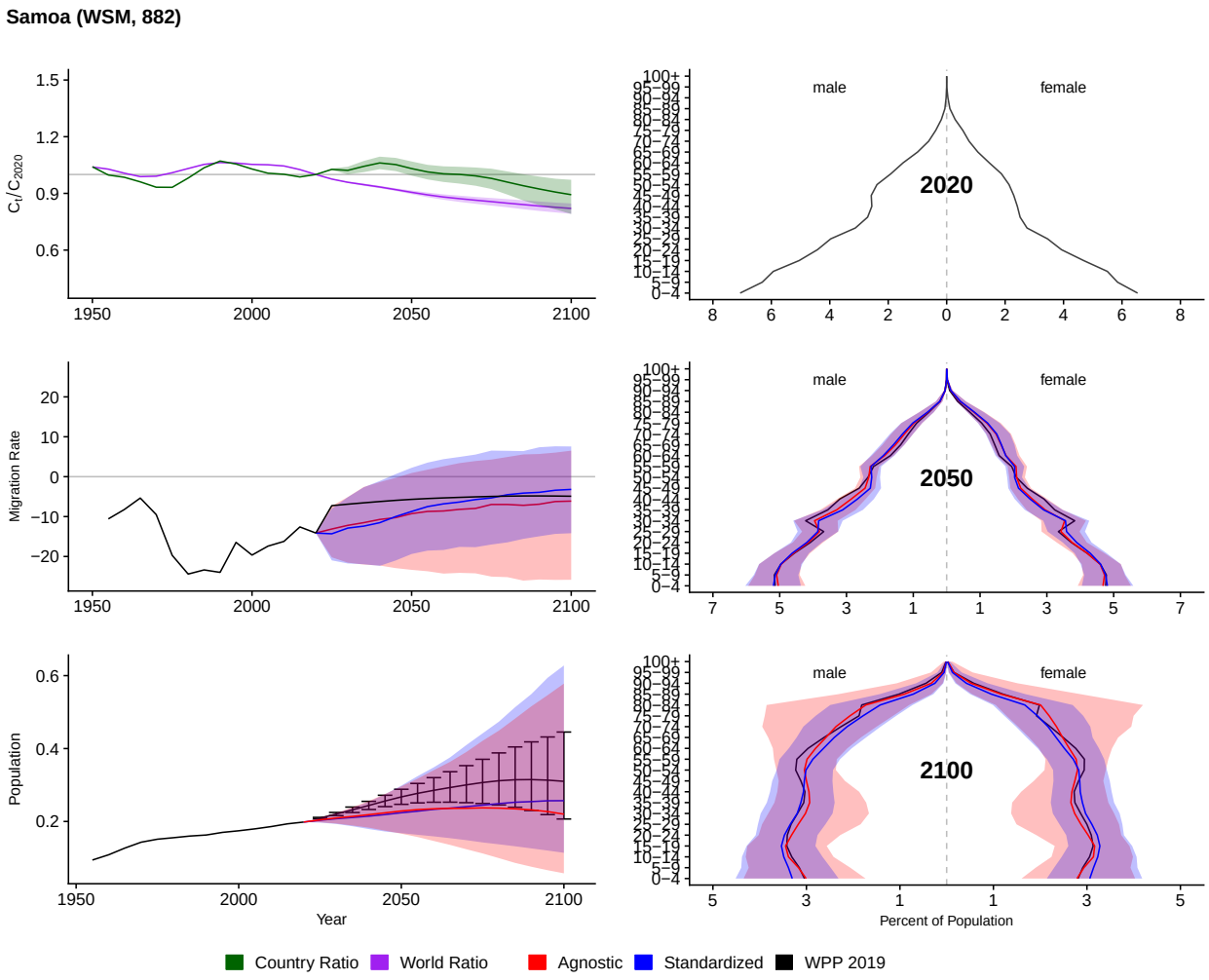


Figure A.250: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

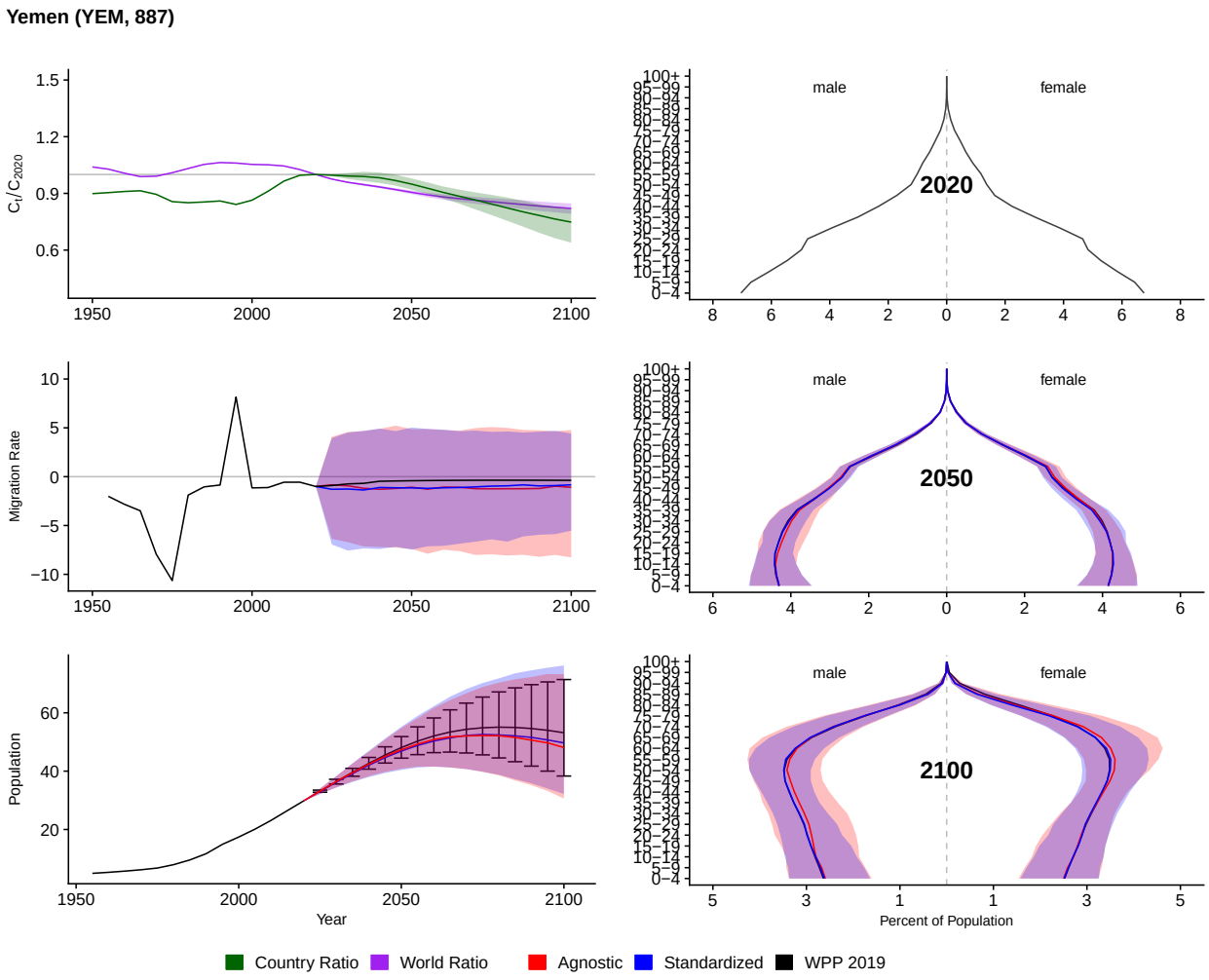


Figure A.251: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

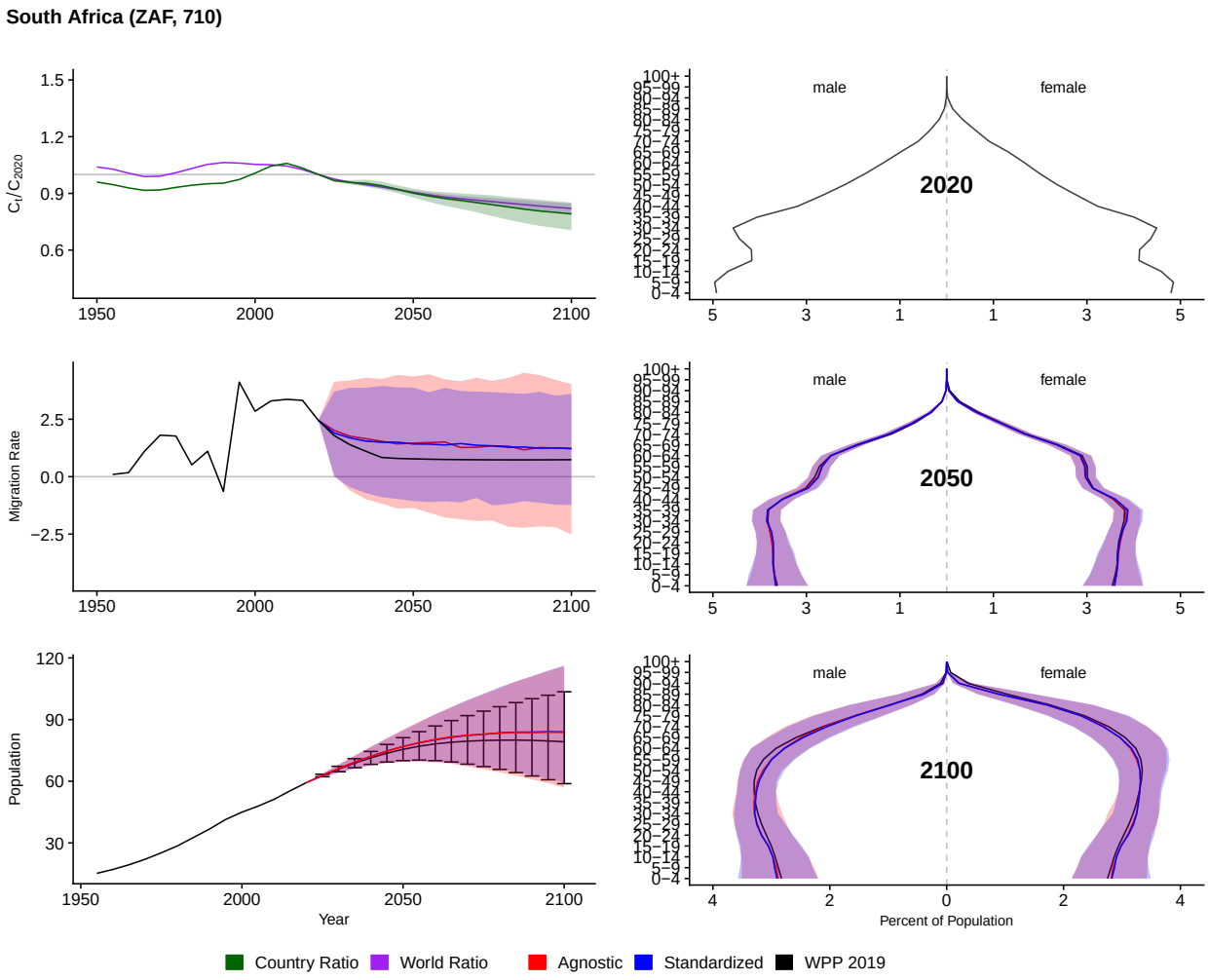


Figure A.252: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

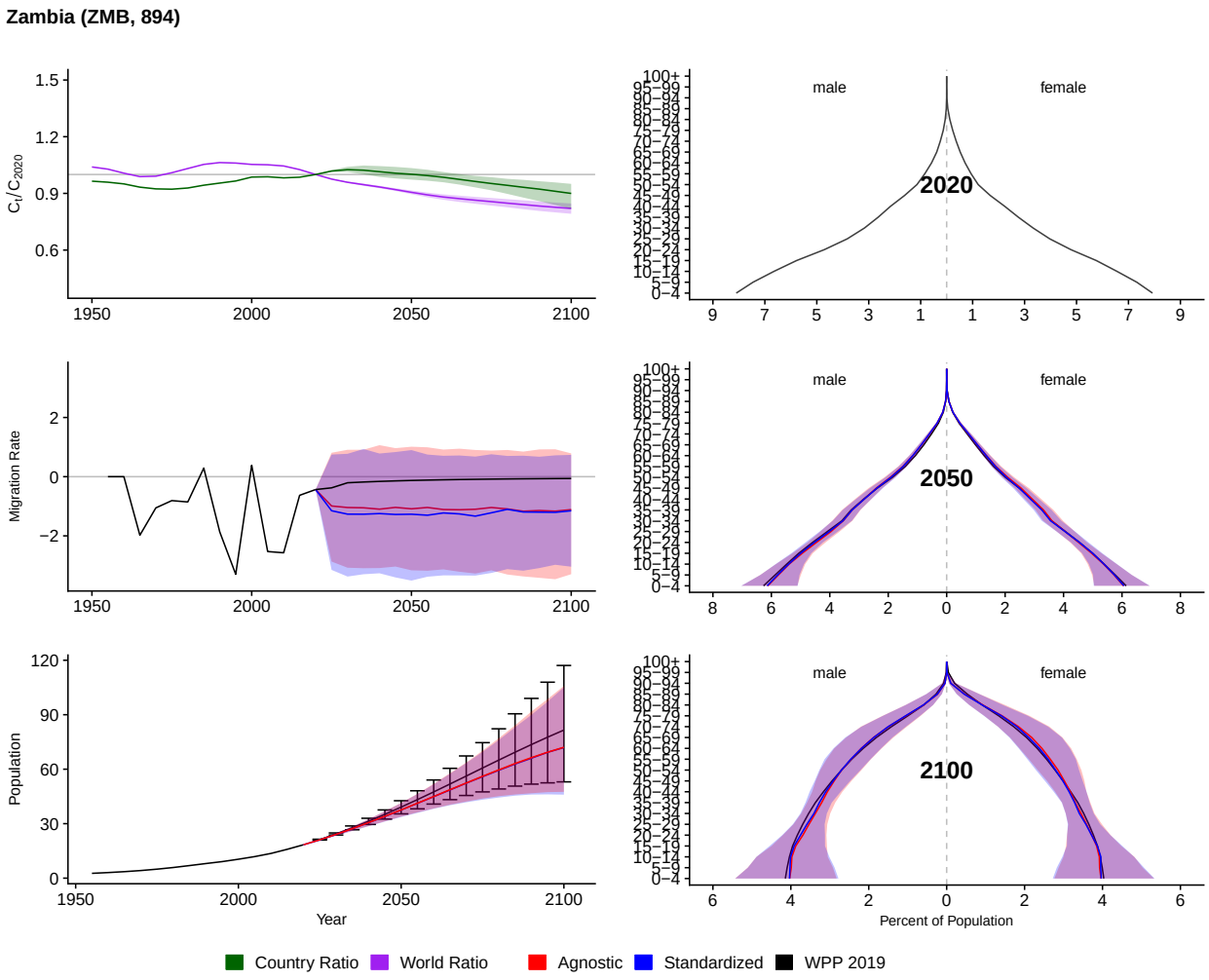


Figure A.253: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.

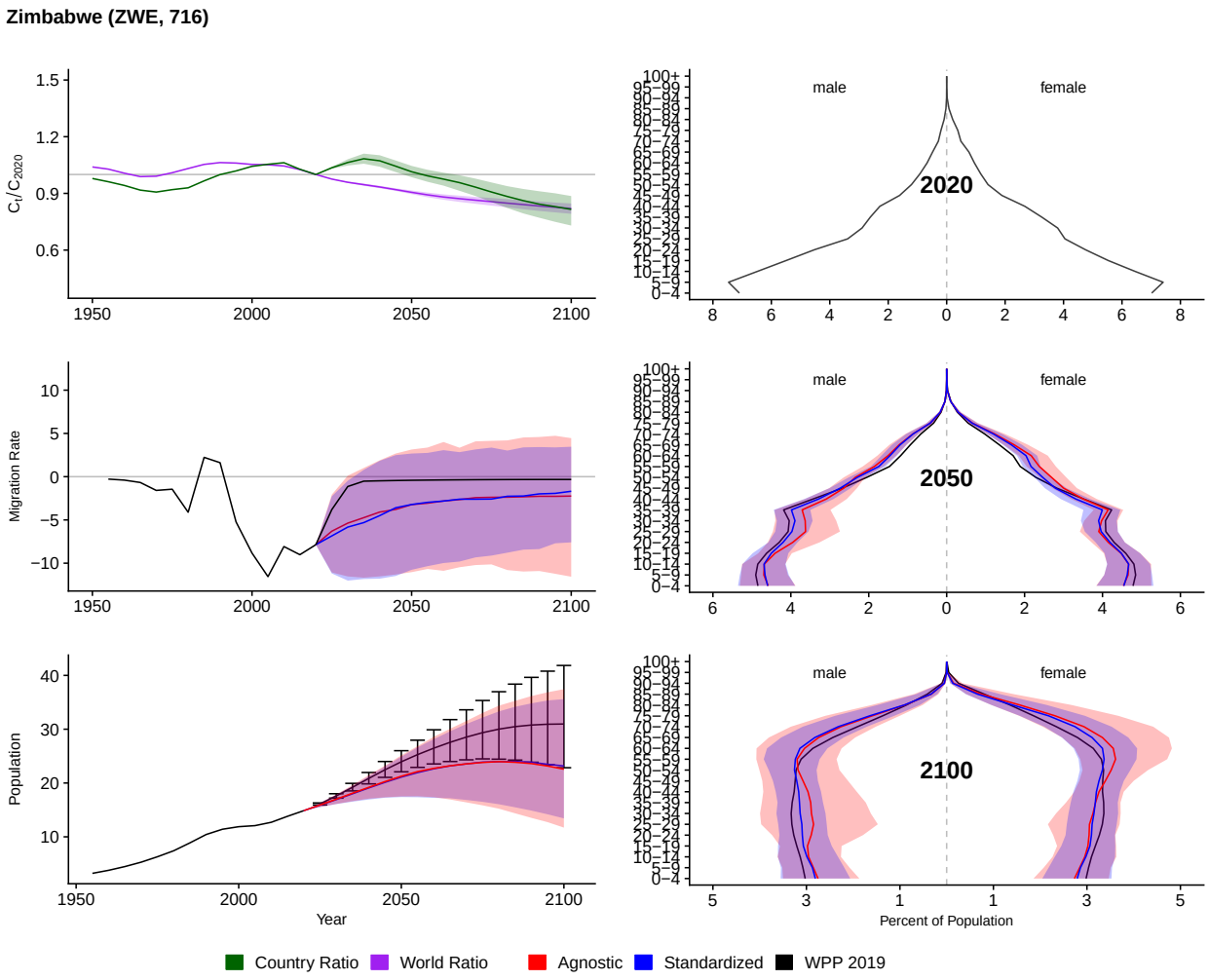


Figure A.254: **Left Column:** Probabilistic forecasts of 2020 base-year Migration Age Structure Index (MASI) for each country (■) and the globe (■), age-standardized and age-agnostic net migration rate (net annual migrants per thousand), and population (millions of people) through 2100. **Right Column:** Observed and forecast population age pyramids for 2020, 2050, and 2100 using age-standardized or age-agnostic migration method. Forecasts use probabilistic age-standardized net migration (■), probabilistic age-agnostic net migration (■), fertility, and mortality. Solid lines in each plot indicate the observed and median forecasts. World Population Prospects (WPP 2019) net migration and population forecasts (■). Shaded regions show the 80% prediction interval. Forecasts start in the 2020-2025 period.