UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Center For International Economics

Working Paper Series

Working Paper No. 2026-05

Forecasting economic growth with traditional methods and a simple neural network model

Shujie Li, Yuanhua Feng

März 2026

CIE Center For International Economics

Forecasting economic growth with traditional methods and a simple neural network model

Shujie Li, Yuanhua Feng
Faculty of Business Administration and Economics, Paderborn University

March 31, 2026

Abstract

Macroeconomic time series forecasting is crucial for guiding government policy decisions, business strategies, and understanding economic trends. However, predicting macroeconomic variables remains a significant challenge. The complexity of economic systems, insufficient data, high levels of volatility complicate the task of accurate forecasting. To enhance forecasting accuracy, we propose two novel models to capture both linear and nonlinear dynamics. First, we generalize the random walk model by incorporating a drift term, which is estimated using a simple neural network model. Second, a hybrid model is introduced to combine local linear regression and the neural network model. Additionally, we adopt other models from Fritz et al. (2024) for combination. These models are combined using a simple averaging method. Our results demonstrate that the newly proposed neural network-based models produce the lowest average MASE. Additionally, model combination is an effective strategy for enhancing the performance of GDP forecasting in most countries and it is less risky than relying on a single model.

Keywords: nonparametric approaches, combination of forecasting, NNAR, Random Walk

JEL Codes: C14, C51

1 Introduction

Forecasting is an essential task in the macroeconomic field. Business policies, central bank strategies, and government economic decisions all rely on accurate forecasting. However, predicting the future values of economic indicators remains a significant challenge due to insufficient data, missing values, high volatility, nonlinear relationships, and policy changes. In order to improve the forecasting accuracy, many researchers have been working on the combination of different models. According to Terui and Van Dijk (2002), the fundamental assumption of combining different forecasting models is that single models cannot identify the true data process precisely. Different models play a complementary role in modelling and forecasting the data. Hibon and Evgeniou (2005) did not confirm that a combination method is better in each case. However, they claimed that a forecast combination method is less risky than using a single model approach. Stock and Watson (2004) analyzed quarterly economic data from seven countries and found that the most effective combination methods are the simplest ones, using either equal weights or nearly equal weights. These simple combination forecasts demonstrate consistent performance across both time periods and countries. Similarly, Smith and Wallis (2009), Hendry and Clements (2004), and Fritz et al. (2024) found that simple combined models often outperformed single models.

Neural networks imitate the human brain's ability to identify patterns and make forecasts based on past experiences. Initially, neural networks were used in cognitive science and engineering. Over the past decades, they have been widely applied in finance and economics for time series forecasting (Hill et al., 1996, Azoff, 1994, Tang and Fishwick, 1993, Refenes, 1994). Zhang et al. (1998) reviewed the use of neural networks in forecasting and discussed how they can serve as a strong alternative to traditional linear models. The major advantages of neural networks are well known, including their lower sensitivity to error term assumptions and their ability to tolerate noise and chaotic components, nonlinear capabilities, and heavy tails (White, 1989). However, it has also been criticized for its "black box" nature. Faraway and Chatfield (1998) proposed not to apply neural networks directly into a "black box", but to select parameters based on traditional modelling skills. Autoregressive neural networks (NNAR) model represents a class of neural networks that use lagged observations of a time series as inputs. In this framework, the number of lags

is selected based on the Akaike Information Criterion (AIC), rather than being chosen arbitrarily. The main differences between NNAR and the autoregressive moving average (ARMA) (Box and Jenkins, 1976) are that NNAR does not impose restrictions on its parameters to ensure stationarity and can capture nonlinear relationships in the data.

Zhang (2003) proposed a hybrid model that combines ARIMA and a simple feedforward neural network, where both models are used jointly. This approach aims to capture different forms of relationships in the time series. Such a hybrid model leverages the advantages of ARIMA and neural networks in linear and nonlinear modeling. In practice, time series data often contain both linear and nonlinear components. Therefore, hybrid models have become a common approach to improve forecasting accuracy. Inspired by the hybrid model in Zhang (2003), we propose two models that incorporate NNAR with a non-parametric approach and a random walk model. The first model applies a local linear regression to estimate the trend, with NNAR used to further estimate the detrended series. The second model models the series as a random walk and the increments follow an NNAR. In this setup, the NNAR model is applied to the first-differenced stationary process. Additionally, three other types of random walk models and a local linear regression model following an ARMA process are also considered in this paper. This paper follows the idea of Fritz et al. (2024), where combining different models can improve forecast accuracy and reduce variance. The goal of this paper is to enhance GDP forecasting accuracy for the large economies in both developed and emerging countries. This is achieved through the cooperation of various models and their combinations with equal weights. To assess the robustness and accuracy of the proposed models, we select 17 major economies worldwide for evaluation.

This paper is structured as follows, section 2 introduces nonparametric regression for the trend estimation. Section 3 introduces ARMA and NNAR as the sequential models applied to the trend-adjusted process. Section 4 discusses random walk models with various types of drifts. Section 5 describes the forecasting of j-step ahead point forecasts and prediction intervals, while Section 6 discusses model evaluation. Section 7 covers the datasets and empirical results, followed by Section 8, which concludes the paper.

2 Nonparametric regression approach

One of the main advantages of nonparametric regression is that it estimates relationships between variables without a predefined functional form, allowing the data to determine the structure. The most commonly used or popular non-parametric estimators are the Nadaraya-Watson and Gasser-Müller estimators. However, the Nadaraya-Wastson estimator suffers from large bias. Fan and Gijbels (2018), Fan (1992), Fan (1993) showed that this estimator has zero minimax efficiency. On the other hand, the Gasser-Müller estimator exhibits increased variability. Additionally, both estimators perform poorly at the boundary region. Various methods, such as reflection- and boundary kernel methods, have been proposed to deal with this issue. However, they are not as efficient as automatic boundary correction of local polynomial fitting.

2.1 Local polynomial fitting

We are aiming to forecast the logarithmic transformation of the GDP of the selected countries. Such time series usually exhibit deterministic trend due to economic growth which cannot be well estimated by any specified parametric function. Following the idea of Feng et al. (2020), if the trend is assumed to be deterministic, a fully automatic nonparametric procedure is applied. Assuming that the time series data $Y_t, t = 1, \ldots, n$, was generated from the following model,

$$Y_t = m(x_t) + z_t, \tag{1}$$

where $x_t = t/n$ is the rescaled time, m is a smooth deterministic trend function. z_t refers to the zero mean stationary process with autocovariances given by $\gamma_z(l) = \text{cov}\,(z_1, z_{1+l})$. It is assumed that $\gamma_z(l)$ converge to zero sufficiently fast, such that $\sum_{k=0}^{\infty}(l+1)^4|\gamma_z(l)| < \infty$. The unknown function $m(x)$ can be estimated locally by a polynomial of degree p, when m is at least $p+1$-times differentiable at point x_0. By applying a Taylor expansion, for x is in the neighborhood of x_0, we obtain

$$\begin{aligned} m(x) &\approx m(x_0) + m'(x_0)(x - x_0) + \frac{m''(x_0)}{2!}(x - x_0)^2 + \ldots + \frac{m^p(x_0)}{p!}(x - x_0)^p \\ &= m(x_0) + \beta_1(x - x_0) + \beta_2(x - x_0)^2 + \ldots + \beta_p(x - x_0)^p. \end{aligned} \tag{2}$$

The ν-th derivative of $m(x)$ is denoted by $m^{(\nu)}(x)$. The relationship between the estimated derivative and the polynomial coefficients is given by $\hat{m}^{(\nu)}(x_0) = \nu!\hat{\beta}_\nu$, which is asymptotically equivalent to kernel regression using a k-th order kernel $K(u)$, where $k = p + 1$, with automatic boundary correction. The local polynomial estimator of $m^{(\nu)}(x_0)$ can be obtained by minimizing the locally weighted least squares problem,

$$Q = \sum_{t=1}^{n} \{Y_t - \sum_{j=0}^{p} \beta_j (x_t - x_0)^j\}^2 K(\frac{x_t - x_0}{h}), \tag{3}$$

where h denotes the bandwidth which determines the size of the local neighborhood around x_0. The k-th order kernel function is given by

$$K(u) = C_\mu (1 - u^2)^\mu \mathbf{I}_{[-1,1]}(u), \tag{4}$$

where C_μ is a normalization constant that ensures $K(u)$ integrates to 1. μ is the smooth parameter that controls the shape of the kernel. In the R package `smoots` (Feng et al., 2022), $\mu = 0$, 1, 2 and 3 correspond to uniform, Epanechnikov, bisquare and triweight kernel, respectively. This package is used for estimating the deterministic trend function, with epanechnikov kernel selected as the weight function.

Two issues have to be addressed. First, the bandwidth h plays a significant role in local polynomial regression. A large bandwidth results in a large bias, while a small bandwidth results in an overly noisy estimation. Therefore, the trade-off between bias and variance is crucial for selecting the bandwidth. There are many techniques to choose the bandwidth, for instance, smoothing cross-validation or a plug-in bandwidth selector. According to Fan and Gijbels (2018), plug-in methods are a conceptually easy technique and they perform empirically good. Feng et al. (2020) developed an iterative plug-in (IPI) algorithm which automatically determines the optimal bandwidth. Second, the choice of the order of local polynomial has to be considered. However, this problem is less important than the choice of bandwidth. This is because the modelling bias is mainly controlled by the bandwidth rather than the order of local polynomial. If a bandwidth h is given, for a higher order of p, the modelling bias is expected to be reduced. But, this would lead to a larger variance. A low order of p is recommended because the bandwidth controls the complexity of a model, i.e. $p = \nu + 1$. For more detail read Fan and Gijbels (2018) section 3.3.

2.2 IPI-algorithm for bandwidth section

Feng et al. (2020) proposed a fully automatic data-driven IPI for bandwidth selection. The accuracy of $\hat{m}^{(\nu)}(x)$ is usually valued by the mean integrated squared error (MISE), the MISE is given by

$$MISE(h) = \int_{c_b}^{d_b} E\{[\hat{m}^{(\nu)}(x) - m^{(\nu)}(x)]^2\}dx, \tag{5}$$

where $0 < c_b < d_b < 1$, when the decision of the bandwidth is considered, the boundary effects are reduced on this measure. Minimization of the MISE results in the theoretical optimal bandwidth (h_{opt}).

The IPI-algorithm is constructed for selecting the asymptotically optimal bandwidth h_A, which minimizes the asymptotic integrated squared error (AMISE). The AMISE of $\hat{m}^{(\nu)}$ is as follows,

$$AMISE(h) = h^{2(k-\nu)}\frac{I[m^{(k)}]\beta^2}{[k!]^2} + \frac{2\pi c_f(d_b - c_b)R(K)}{nh^{2\nu+1}},$$

where $I[m^{(k)}] = \int_{c_b}^{d_b}[m^{(k)}(x)]^2dx > 0$, $\beta = \int_{-1}^{1} u^k K(u)du$, $R(K) = \int_{-1}^{1} K^2(u)du$. $c_f = f(0)$ is the value of the spectral density at the frequency zero. The asymptotically h_A for estimating $m^{(v)}$ can be obtained by minimizing AMISE and it is given by

$$h_A = n^{-1/(2k+1)}\Big(\frac{2\nu+1}{2(k-\nu)}\frac{(d_b - c_b)R(K)2\pi c_f[k!]^2}{I[m^{(k)}]\beta^2}\Big)^{1/(2k+1)}, \tag{6}$$

where c_f and $I[m^{(k)}]$ are unknown. For a given bandwidth, the estimation of $I[\hat{m}^{(k)}]$ is not affected by the error correlation, which can be estimated by known approaches for independent errors. The key point is the estimation of the variance factor c_f, which can be estimated using a data-driven IPI-algorithm developed by Feng et al. (2020). The following lag-window estimator of c_f with Bartlett-Window weights is used,

$$\hat{c}_{f,M} = \frac{1}{2\pi}\sum_{l=-M}^{M} w_l\hat{\gamma}(l), \tag{7}$$

where $\hat{\gamma}(l)$ refers to the sample acf computed from $r_{t,V} = Y_t - \hat{m}_V$, which are the residuals obtained from a pilot estimate using a bandwidth h_V. M refers to the window-width of the lag-window estimator. The weights w_l are calculated by $w_l = l/(M + 0.5)$. Feng et al. (2020) proposed the IPI-algorithm for iteratively determining the window-width, which will be used for estimating the variance factor c_f. This IPI-algorithm is adjusted from that of Bühlmann (1996) as follows:

a) Set the initial window-width as $M_0 = [n/2]$, where $[\cdot]$ refers to the integer part.

b) Global steps: following Bühlmann (1996), in the j-th iteration estimate $\int (f(\lambda)^2) d\lambda$. Let $\int f^{(1)}(\lambda) d\lambda$ be the integral of the first generalized derivative of $f(\lambda)$. Let $M'_j = [M_{j-1}/n^{2/21}]$ and estimate $\int (f(\lambda)^1) d\lambda$ by this window-width. The estimates are inserted into formula (5) in Bühlmann (1996) and obtain M_j. Iterate the procedure until the optimal M converges or until maximal 20 iterations. $\hat{M}_G$ denotes the selected M.

c) Local adaptation at $\lambda = 0$: set $M' = [\hat{M}_G/n^{2/21}]$ and use this window-width to calculate $\int f^{(1)}(\lambda) d\lambda$. Inserting the estimates into equation (5) in Bühlmann (1996) to obtain the $\hat{M}$.

Feng et al. (2020) indicated that this algorithm achieves a rate of convergence $O(n^{-1/3})$. As long as the window width converges, the choice of the initial value M_0 does not affect the results. Therefore, any value between 1 and $n/2$ can be chosen. Additionally, they showed that the asymptotically optimal bandwidth for estimating c_f should be $CF \cdot h_A$ instead of h_A, even though the variance factor c_f can be well estimated from the residuals of $\hat{m}$. The correction factor CF is given by

$$CF = \left\{ 2k \left[\frac{2K(0)}{R(K)} - 1 \right] \right\}^{1/(2k+1)}, \tag{8}$$

where CF is always larger than 1.

The IPI-algorithm is proposed for selecting the bandwidth for local linear and local cubic estimators $\hat{m}$, with $k = 2$ and $k = 4$, respectively (Feng et al., 2020). The procedure is as follows,

a) Start with an initial bandwidth h_0.

b) For $j = 1, \ldots,$ estimate $m(\cdot)$ by means of the bandwidth correction factors $(CF \cdot h_{j-1})$ and calculate the corresponding residuals. $\hat{c}_f$ is obtained from those residuals by a data driven lag-window estimator.

c) Let $\alpha = 5/7$ or $5/9$ for $p = 1$, and $\alpha = 9/11$ or $9/13$ for $p = 3$. Estimated $I[m^{(k)}]$ with $h_{d,j} = h_{j-1}^{\alpha}$, and a local polynomial of order $p_d = p + 2$. We obtain

$$h_j = \left(\frac{[k!]^2}{2k\beta^2} \frac{2\pi \hat{c}_f (d_b - c_b) R(K)}{I[\hat{m}^{(k)}]} \right)^{1/(2k+1)} \cdot n^{-1/(2k+1)}. \tag{9}$$

d) Increase j by 1. Repeat step b) and c) until convergence is reached or a given number of iterations, for instance 20 iterations, is achieved. Then, let $\hat{h} = h_j$.

3 Trend stationary residuals analysis

Statistic models for time series analysis should usually be applied to stationary processes. Applying those models to a non-stationary process can lead to unreliable estimation results. After removing the estimated trend function $m(x)$ from the time series, any suitable models can be applied to analyze the zero mean stationary process $\hat{z}_t = Y_t - \hat{m}(x_t)$.

3.1 Trend stationary ARMA

Two methods are employed to further analyze the detrended series, z_t. First, z_t is assumed to follow an ARMA(p, q) process, which is given by

$$\phi(B)z_t = \psi(B)\xi_{1,t}, \tag{10}$$

where $\phi(B) = 1 - \beta_1 B - \cdots - \beta_p B^p$ and $\psi(B) = 1 + \alpha_1 B + \cdots + \alpha_q B^q$ are the AR and MA polynomials, respectively, with no common factors, and all their roots lie outside of the unit circle. Both Eq. (1) and (10) define the semiparametric ARMA model. The effect of the error caused be the nonparametric trend, $m(\cdot)$, is negligible on the subsequent parametric estimation (Feng and Zhou, 2015). $\xi_{1,t}$ are i.i.d innovations with $E(\xi_{1,t}) = 0$ and $\text{var}\,(\xi_{1,t}) = \sigma^2_{\xi_1}$. The order of the ARMA model is determined using the Bayesian Information Criterion (BIC) values. We refer to the semiparametric ARMA model as the local linear regression with an ARMA process, abbreviated as LLR.

3.2 Trend stationary NNAR

Second, z_t is further analyzed using an NNAR model, which is a FFNN with a single hidden layer, where univariate lagged values are used as input features. A basic neural network consists of an input layer and an output layer. The output is a weighted sum of the inputs, which can be interpreted as a linear model. However, the model becomes

nonlinear when a hidden layer is introduced. The activation function in the hidden layer enables the network to capture complex, nonlinear relationships between the inputs and the output, making it much more powerful for tasks like time series forecasting. The mathematical representation of the NNAR is given by

$$z_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j g(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} z_{t-i}) + \xi_{2,t}, \tag{11}$$

where $\alpha_j, j = 1, ..., q$ and $\beta_{ij}, i = 1, ..., p; j = 1, ..., q$ are the parameters, referred to as weights in the NNAR. α_0 and β_{0j}, $j = 1, \ldots, q$ are biases. p denotes lagged inputs or the number of input nodes and q denotes the number of hidden nodes. In the hidden layer, the outputs from the previous layer are modified by a sigmoid function $g(\cdot)$, and the transformed values are then passed to the next layer. The sigmoid function is defined as

$$g(x) = \frac{1}{1 + e^{-x}}. \tag{12}$$

$\xi_{2,t}$ is assumed to be i.i.d with $N(0, \sigma^2_{\xi_2})$ innovations. The NNAR model is equivalent to a nonlinear autoregressive model. For non-seasonal time series, p is equal to the optimal number of lags for an AR(p) model according to Akaike Information Criterion (AIC). The rule of thumb for selecting the value for q is $(p + 1)/2$ according to Hyndman and Athanasopoulos (2018). Eq. (1) and (11) define the local linear NNAR (LLNNAR) model.

4 Random walk variants

All selected GDP data in this paper exhibit a clear upward trend, therefore, a random walk without drift will not be considered. For random walk models, the time series achieve stationarity through differencing. In the literature such a trend is considered as a stochastic trend. Following Fritz et al. (2024), four variants of random walk models with drifts are introduced. These variants include random walk with a constant drift (RWC), random walk with a linear drift (RSL), random walk with a nonparametric smooth drift function (RLL) and random walk with an NNAR drift (RNNAR).

The RWC model is defined as,

$$Y_t = Y_{t-1} + c + \xi_{3,t}, \qquad t = 1, ..., n, \tag{13}$$

where c is a constant drift parameter and ΔY_t are the first differences of the log-GDP. c can be estimated by the sample mean of ΔY_t, i.e., $c = \frac{1}{n-1}\sum_{t=1}^{n-1}\Delta Y_t$. The RWC model assumes that Log-GDP follows a stochastic trend with a constant expected growth rate over time. The error term $\xi_{3,t}$ is assumed to be i.i.d., $\xi_{3,t} \sim N(0, \sigma^2_{\xi_3})$.

Unlike the RWC model, which assumes a constant expected growth rate over time, the linear drift model assumes that the expected growth rate changes linearly over time. In this case, the first differences ΔY_t, can be modeled using a simple linear model. The RSL is given by,

$$Y_t = Y_{t-1} + \beta_0 + \beta_1 t + \xi_{4,t}, \qquad t = 1, ..., n, \tag{14}$$

where β_0 is the initial average growth rate and β_1 measures how much that growth rate increase or decrease each period. The error term $\xi_{4,t}$ is assumed to be i.i.d. with $\xi_{4,t} \sim N(0, \sigma^2_{\xi_4})$.

The linear drift in the RSL model can be further generalized to a local linear drift, which is modeled using a nonparametric smooth function. The growth rate ΔY_t are then estimated nonparametrically. This model is referred to as the RLL. It is defined as follows,

$$Y_t = Y_{t-1} + \delta(x_t) + \xi_{5,t}, \qquad t = 1, ..., n, \tag{15}$$

where $\delta(x_t)$ is the time varying drift function estimated by local linear regression. $\xi_{5,t}$ are i.i.d with $N(0, \sigma^2_{\xi_5})$ innovations.

The last variant of random walk models uses the NNAR model to estimate the growth rate and is referred to as the RNNAR model. The aim of this model is to detect nonlinear relationships in the series that cannot be captured by linear models. This model is given by,

$$Y_t = Y_{t-1} + f(x_t) + \xi_{6,t}, \qquad t = 1, ..., n, \tag{16}$$

where $f(x_t)$ is the nonlinear drift function estimated using the NNAR model. $\xi_{6,t}$ are again assumed to be an i.i.d $N(0, \sigma^2_{\xi_6})$ process.

5 Forecasting approaches

5.1 j-step ahead point forecasts for trend-stationary models

Point forecasts for the LLR and LLNNAR models contain two components, including a forecast of the trend component and a forecast of the ARMA or NNAR component. The trend function $m(x_t)$ in Eq. (1) is estimated using the IPI-algorithm. A simple linear extrapolation is used to conduct the j-step ahead trend forecasting, given by

$$\hat{m}(x_{n+j}) = \hat{m}(x_n) + j\Delta m, \quad j = 1, 2, ..., k, \tag{17}$$

where $\Delta m = \hat{m}(x_n) - \hat{m}(x_{n-1})$ represents the estimated local slope at the end of the series. The LLR point forecasts combine trend forecasts and ARMA forecasts of the trend-stationary process. The j-step ahead forecasting is as follows,

$$\hat{Y}_{i,n+j} = \hat{m}(x_{n+j}) + \hat{z}_{i,n+j}, \quad i = 1, 2; \quad j = 1, 2, ..., k. \tag{18}$$

The point forecasts of the LLNNAR follow the same method as in Eq. (18), with the key difference being that $\hat{z}_{n+j}$ is obtained from the NNAR model predictions.

5.2 j-step ahead point forecasts for random walk models

The RWC, RSL, RLL and RNNAR models are different variants of the random walk model, in which j-step ahead forecasts of log-GDP are obtained from the most recent observation Y_n by adding the cumulative predicted growth over j steps. Different models vary in how they estimate the cumulative growth rate or the drift. The forecast equations are defined as follows,

$$\hat{Y}_{3,n+j} = Y_n + j\hat{c}, \qquad j = 1, ..., k, \tag{19}$$

$$\hat{Y}_{4,n+j} = Y_n + j\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{j} (n+i), \qquad j = 1, ..., k, \tag{20}$$

$$\hat{Y}_{5,n+j} = Y_n + \sum_{i=1}^{j} \hat{\delta}(x_{n+i}), \qquad j = 1, ..., k, \tag{21}$$

$$\hat{Y}_{6,n+j} = Y_n + \sum_{i=1}^{j} \hat{f}(x_{n+i}), \qquad j = 1, ..., k. \tag{22}$$

The RWC model utilizes the last observed value and adds a constant drift $\hat{c}$ cumulatively at each horizon. The RSL forecasts incorporate cumulative predicted linear changes in the growth rate. For the RLL and RNNAR models, the time-varying and non-linear changes in the growth rate are estimated via local linear regression and the NNAR model, respectively. In all cases, the final point forecasts are obtained by summing the last observed log-GDP and the aggregated estimated growth increments.

5.3 Prediction interval estimation

A prediction interval quantifies the uncertainty in the forecast. Point forecasts alone without accounting for this uncertainty are less valuable. To construct the prediction intervals, a few assumptions are required. For all models, the errors are assumed to be i.i.d. and approximately normally distributed with zero mean and constant variance. This assumption is not always valid for neural networks. In practice, distribution-free methods, such as the bootstrap interval, are commonly used to construct intervals for neural networks. Nevertheless, constructed intervals under the normality assumption for feedforward neural networks have been studied in Chryssolouris et al. (1996). Papadopoulos et al. (2001) and Dybowski and Roberts (2001) compared different methods for estimating intervals, including maximum likelihood, approximate bayesian and the bootstrap technique. For simplicity, in this study, all intervals are estimated under the normality assumption. Additionally, it is assumed that the observed time series can be reasonably approximated by all considered models. For instance, if Y_t follows the RWC model, then the other random walk variants, such as, RSL, RLL and RNNAR are also valid. It can also be shown that the LLR and LLNNAR are good approximations. Moreover, the cross-correlation between the errors of the six models is assumed to be stationary. A theoretical study on these assumptions is beyond the scope of this study.

The error z_t in the LLR model is assumed to follow an ARMA(p, q) process and it can be represented as a MA(∞) process by $z_t = \sum_{i=0}^{\infty} \alpha_i \xi_{1,t-i}$. The point forecasts of the LLR model is then $Y_{1,n+j} = \hat{Y}_{1,n+j} + \sum_{i=0}^{j-1} \alpha_i \xi_{1,n+j-i}$ with the approximate variance $\text{Var}\,(Y_{1,n+j}) = \sigma_1^2 \sum_{i=0}^{j-1} \alpha_i^2$. The error in the smooth trend function $\hat{m}(x_{n+j})$, which is of the order $O_p(n^{-2/5})$, is omitted for simplicity (Fritz et al., 2024). Individual forecasts under the other five models have an approximate variance $\text{Var}\,(Y_{i,n+j}) = j\sigma_i^2, i = 2, 3, 4, 5, 6$

and $j = 1, ..., k$ respectively. For the combined models, let $S_{comb,j}$ represent the vector of standard errors of the individual forecasts, and let R_{comb} denote the matrix of cross-correlations between the errors of different models, which remains constant over time. The estimated variance of the forecast for a specific observation is expressed as $\text{Var}\left(Y_{comb,n+j}\right) = S^T_{comb,j} R_{comb} S_{comb,j}$. If a combined model includes independent errors or even negative correlated errors from individual models, the variance of the combined model can be significantly reduced through the combination. Consequently, the bias can also be reduced through an appropriate combination. This is also known as the ensemble effect of forecast combinations. Our results provide strong evidence of this effect and further details will be discussed in the next chapter.

6 Model evaluation

The six proposed models are combined using an equal-weighted average, resulting in a total of 63 distinct models. Common scale-dependent measures, such as the mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and median absolute error (MdAE), are effective for comparing methods on the same data series. However, since their evaluation is solely based on the forecasting error, they are not appropriate for comparing different models across various datasets. One remedy is using a scaled error measurement, which is a standard method to evaluate the forecasting accuracy of different scaled datasets. Hyndman and Koehler (2006) proposed the mean absolute scaled error (MASE) as a measure of forecast accuracy that can be applied across different time series. Moreover, the MASE is less sensitive to outliers and is straightforward to interpret. For the non-seasonal time series, the MASE is defined as,

$$\begin{aligned} \text{MASE} &= \frac{1}{k} \sum_{t=n+1}^{n+k} |q_t| \quad \text{with} \\ q_t &= \frac{e_t}{\frac{1}{n-1} \sum_{i=2}^{n} |Y_i - Y_{i-1}|} \end{aligned} \tag{23}$$

where e_t is the forecasting error for the given period. The term $\sum_{i=2}^{n} |Y_i - Y_{i-1}|$ represents the sum of absolute errors from the in-sample one-step naive forecasts. The denominator of q_t therefore corresponds to the mean absolute error of the in-sample one-step naive

forecast. The forecast error e_t is scaled by this mean absolute error. This method assumes the best forecast for the next time step is simply the last observed value, i.e., $\hat{Y}_t = Y_{t-1}$. When $q_t < 1$, it indicates a better forecast than one-step ahead naïve forecast. If $q_t > 1$, it indicates that the forecast is worse than the one-step ahead naïve forecast. Another similar measure for non-seasonal data is the root mean squared scaled error (RMSSE), which is given by,

$$\begin{aligned} \text{RMSSE} &= \sqrt{\frac{1}{k}\sum_{t=n+1}^{n+k} q_t^2} \quad \text{with} \\ q_t^2 &= \frac{e_t^2}{\frac{1}{n-1}\sum_{i=2}^{n}(Y_i - Y_{i-1})^2}. \end{aligned} \tag{24}$$

According to Fritz et al. (2024), the results of the MASE and RMSSE are very similar, hence, only MASE is used as the accuracy measure of forecast accuracy. The MASE of all 63 models across 17 countries is summarized in Table A1.

7 Data and Empirical Results

7.1 Data

The dataset used in this paper consists of the annually logarithm of real GDP at constant 2017 national prices (in millions of 2017 US dollars) for the following countries: the United Kingdom (UK), India (IN), Mexico (MX), China (CN), the United States (US), Vietnam (VN), France (FR), Japan (JP), Turkey (TR), Germany (DE), South Korea (KR), Israel (IL), the Netherlands (NL), Canada (CA), Australia (AU), Switzerland (CH), and Indonesia (ID). The data spans the period from 1950 to 2019, except for Vietnam (1970 to 2019), China (1952 to 2019), Indonesia (1960 to 2019), and South Korea (1953 to 2019). The data is sourced from the Penn World Table Version 10.0. j-step ahead predictions are conducted with $j = 5$ and the remaining observations are used as the in-sample set. The analysis is conducted by means of the R package `smoots` (Feng et al., 2022), which is applied to estimate the local linear trend, while the NNAR model is estimated using `forecast` package (Hyndman and Khandakar, 2008).

7.2 Empirical analysis

In total, 63 models are applied to the log-GDP of the selected 17 countries, including 6 single models and their combinations. All models are evaluated based on their MASE. Additionally, the standard error and mean MASE values across all countries for each model are considered to assess consistency and robustness. The results of single models are reported in Table 1, while the results of all model combinations are presented in Table A1 in the appendix. Models are coded as: 1 = RWC, 2 = RSL, 3 = LLR, 4 = RLL, 5 = RNNAR, 6 = LLNNAR. Model notations use M1 for RWC, M12 for the combination of RWC and RSL, and similarly for other combinations. The graphical representation of the forecast points and prediction intervals for each single model across all countries are also provided in the appendix (see, figures A1 to A17). The black line represents the real log-GDP, the red dashed lines indicate the upper and lower 95% prediction intervals, and the blue line shows the predictions.

Table 1: Summary of MASE across countries for single models.

Models	UK	IN	MX	CN	US	VN	FR	JP	TR	DE	KR	IL	NL	CA	AU	CH	ID	SD	Mean
M1: RWC	0.58	1.37	1.02	**0.81**	0.60	0.35	1.54	2.14	0.27	1.23	1.69	1.34	0.83	1.51	0.91	0.68	**0.16**	0.55	1.00
M2: RSL	**0.18**	0.33	0.84	2.22	0.44	0.42	0.86	1.29	0.50	1.85	1.21	0.60	1.20	0.26	0.31	1.07	0.26	0.59	0.82
M3: LLR	0.99	0.36	0.47	2.43	0.92	0.27	0.38	0.46	0.42	0.57	0.53	0.42	0.76	**0.20**	0.50	0.38	0.37	0.51	0.61
M4: RLL	0.62	0.29	0.37	2.14	0.96	0.60	1.06	0.35	1.00	0.51	**0.35**	0.38	2.31	0.32	**0.25**	**0.27**	0.78	0.62	0.74
M5: RNNAR	1.00	1.33	**0.31**	1.05	**0.08**	0.56	**0.25**	**0.11**	**0.26**	**0.14**	1.07	0.82	**0.08**	1.14	0.59	**0.27**	0.39	**0.43**	**0.56**
M6: LLNNAR	0.77	**0.24**	0.51	2.51	0.53	**0.16**	0.39	0.48	0.33	0.55	0.55	**0.32**	1.06	0.28	0.41	0.29	0.49	0.54	0.58

Table 1 shows that, among the six proposed single models, the RWC model performs the worst based on the mean MASE across all countries, with a value of approximately 1. This indicates that the RWC performs no better than the one-step ahead naïve forecast. In contrast, the RNNAR is the best performer, with a mean MASE of 0.56. Additionally, the RNNAR has the lowest standard error among all single models, indicating greater consistency and robustness. The second-best performing model is the LLNNAR among single models, with a mean MASE of 0.58. These results highlight the effectiveness of the NNAR-based models in improving GDP forecasting accuracy.

The mean MASE of RSL and RLL across all countries is comparable, with values of 0.82 for RSL and 0.74 for RLL. Both models outperform the average one-step ahead näive forecast, the RLL slightly outperforms the RSL in forecasting the growth rate. A com-

parable estimation pattern for these two models can be observed in the figures in the appendix. The RSL and RLL models failed to capture the GDP growth in several countries, including the USA, Mexico, France, Japan, Germany, the Netherlands and India, where they tend to underestimate the log-GDP. In contrast, the RWC model systematically overestimates GDP forecasts in most countries, including the UK, USA, Mexico, France, Japan, Germany, the Netherlands, China, Korea, Israel, Canada, Australia and Switzerland. This overestimation likely occurs because economic growth in the test set slowed down, however, the model assumes that the average growth rate remains the same over the forecast period. As the estimation results differ considerably, combining the RSL or RLL with the RWC has great potential to improve the forecasting accuracy as positive and negative forecasting errors can offset each other.

As expected, the LLR and LLNNAR models also produce similar results, and a similar pattern can be observed in almost all the countries.This similarity arises from the fact that both models rely on the same smooth trend function for estimation. The mean MASE across all countries for the LLR model is 0.61, slightly higher than that of the LLNNAR model, indicating that the NNAR performs slightly better on the detrended series than the ARMA model.

Table A1 confirms that combining models using a simple average is an efficient approach to improve forecast accuracy. The overall mean MASE across all countries and its standard error decrease significantly when different models are combined. There are a total of 15 different two-model combinations. Some of these combinations perform well in specific cases, such as in the UK, Vietnam, Turkey, South Korea, Australia, and Canada. However, two-model combinations often lack generalizability due to limited model diversity, as the mean value of MASE across all countries drops further when more models are combined. When two similar models are combined, forecasting accuracy may not improve. In fact, combining similar models can reinforce the same errors rather than cancel them out. Furthermore, including more models does not necessarily lead to better forecasting performance. Incorporating too many models may reduce the efficiency of the combination, as redundant or highly correlated models add complexity without improving accuracy. As a result, the benefits of combination start to diminish. The combination of five or six single models fails to achieve the best performance for any of the selected countries.

Based on the average MASE values across all counties, the most accurate method, in

general, is M126. It combines the RWC, RSL, and LLNNAR models, achieving an average mean MASE of 0.44. Other combinations that produce slightly larger mean MASE value are M1235 (RWC, RSL, LLR and RNNAR), M1256 (RWC, RSL, RNNAR and LLNNAR), M12356 (RWC, RSL, LLR, RNNAR and LLNNAR) and M1245 (RWC, RSL, RLL and RNNAR). The results for each country of the combined model M126 are illustrated in Figures 1, 2 and 3.

All proposed models and their combinations fail to predict GDP growth in China (Figure 1 and A5). The best-performing model is the RWC model for China, with a MASE of 0.8, which is slightly better than the one-step ahead naïve forecast. China experienced a turning point in 2014, when the growth rate slowed considerably (see, Figure A5). This change cannot be predicted by the proposed models. All single models make the same error by overestimating and the combination of models cannot compensate for these errors. One potential remedy is to reduce the forecast horizon from 5 to 3 steps. For this shorter forecasting period, the RNNAR model outperforms all other methods and their combinations, with a MASE value of 0.20 (a forecast results for $j = 3$ for all countries were computed and are available upon request).

The combined models also fail to improve forecasts for Korea, India, the USA, and the Netherlands. For Korea, the RLL delivers the most accurate results. The RNNAR is the top performing model for the US and the Netherlands, while the LLNNAR model delivers the best forecasting results for Indian. Additionally, each single model is selected at least once as either the best- or worst-performing model for specific countries. Although the RNNAR model is the best individual performer overall, it is not the best one in every case. For countries like UK, Indian, China, Koran, Canada, its MASE is around 1 or larger. Conversely, among the six models, the RWC model is overall the worst-performing model, but it is the best model in China. These phenomena support the idea that there is no single model, which performs well in every case. Even though model combination failed in a few countries, it is still an efficient way to improve forecasting accuracy and reduce risk of relying on a single model for most cases.

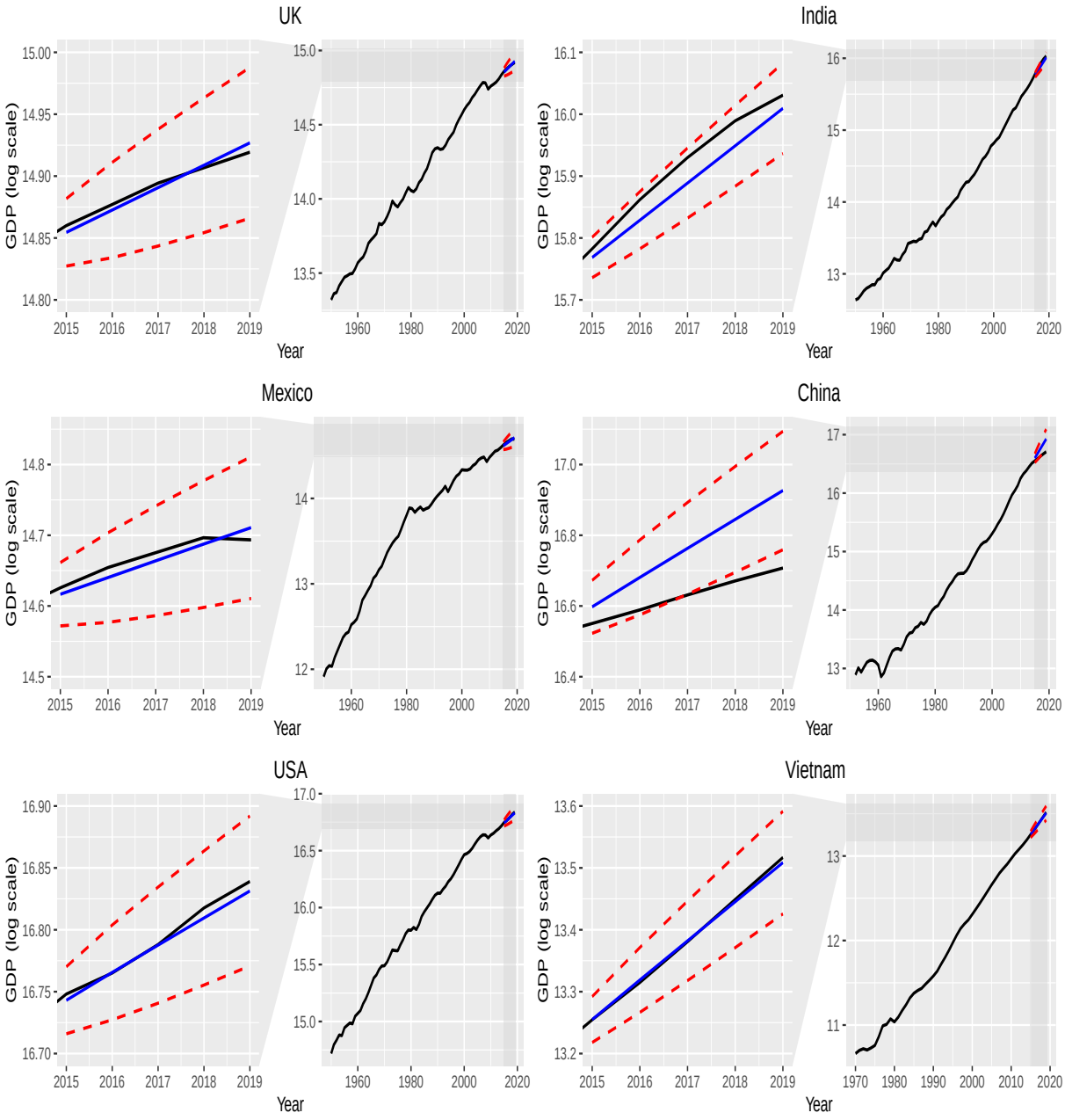

Figure 1: Point forecasts and 95% prediction intervals from the combined model M126 for the UK, India, Mexico, China, the USA, and Vietnam.

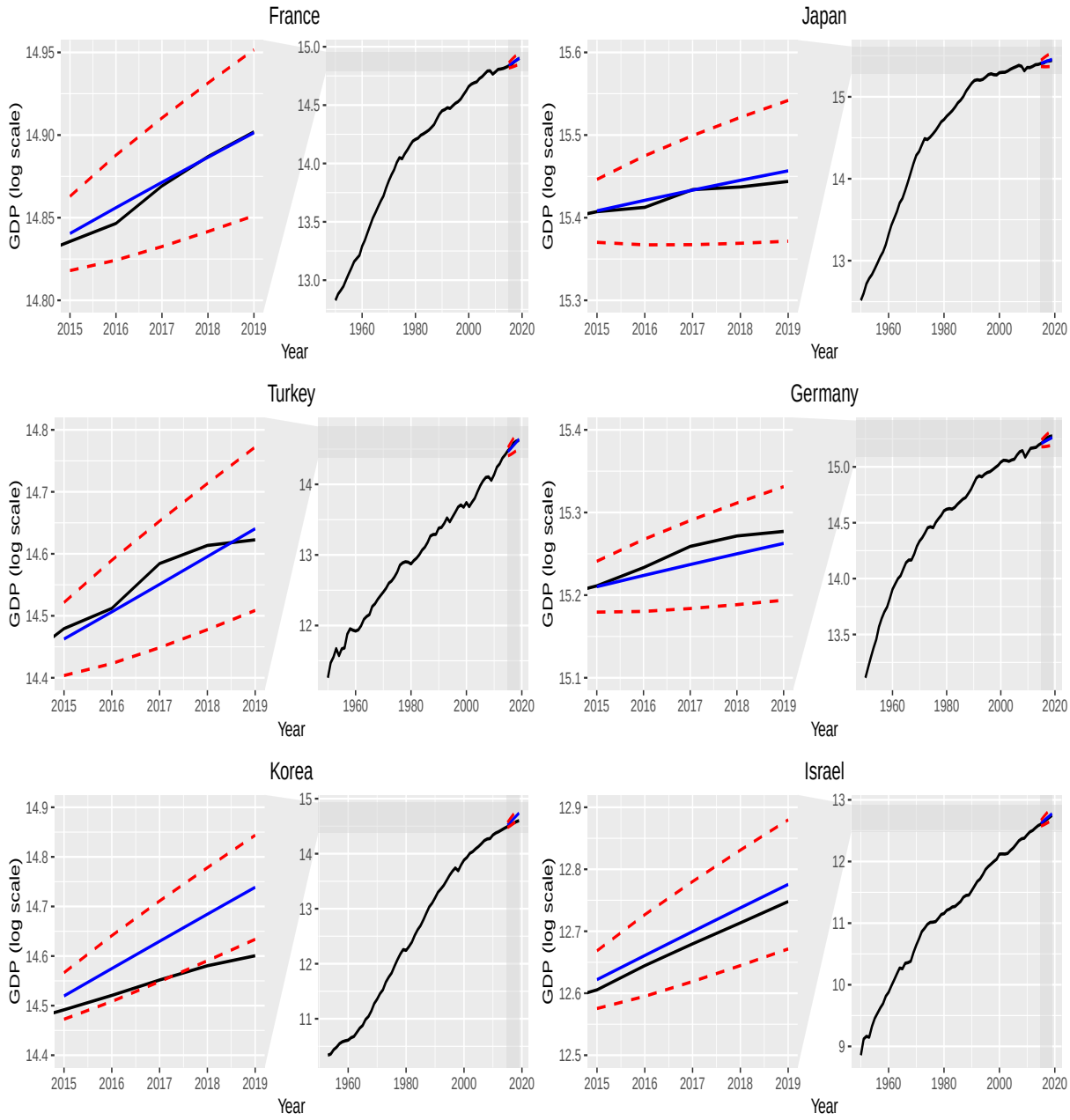

Figure 2: Point forecasts and 95% prediction intervals from the combined model M126 for France, Japan, Turkey, Germany, Korea and Israel.

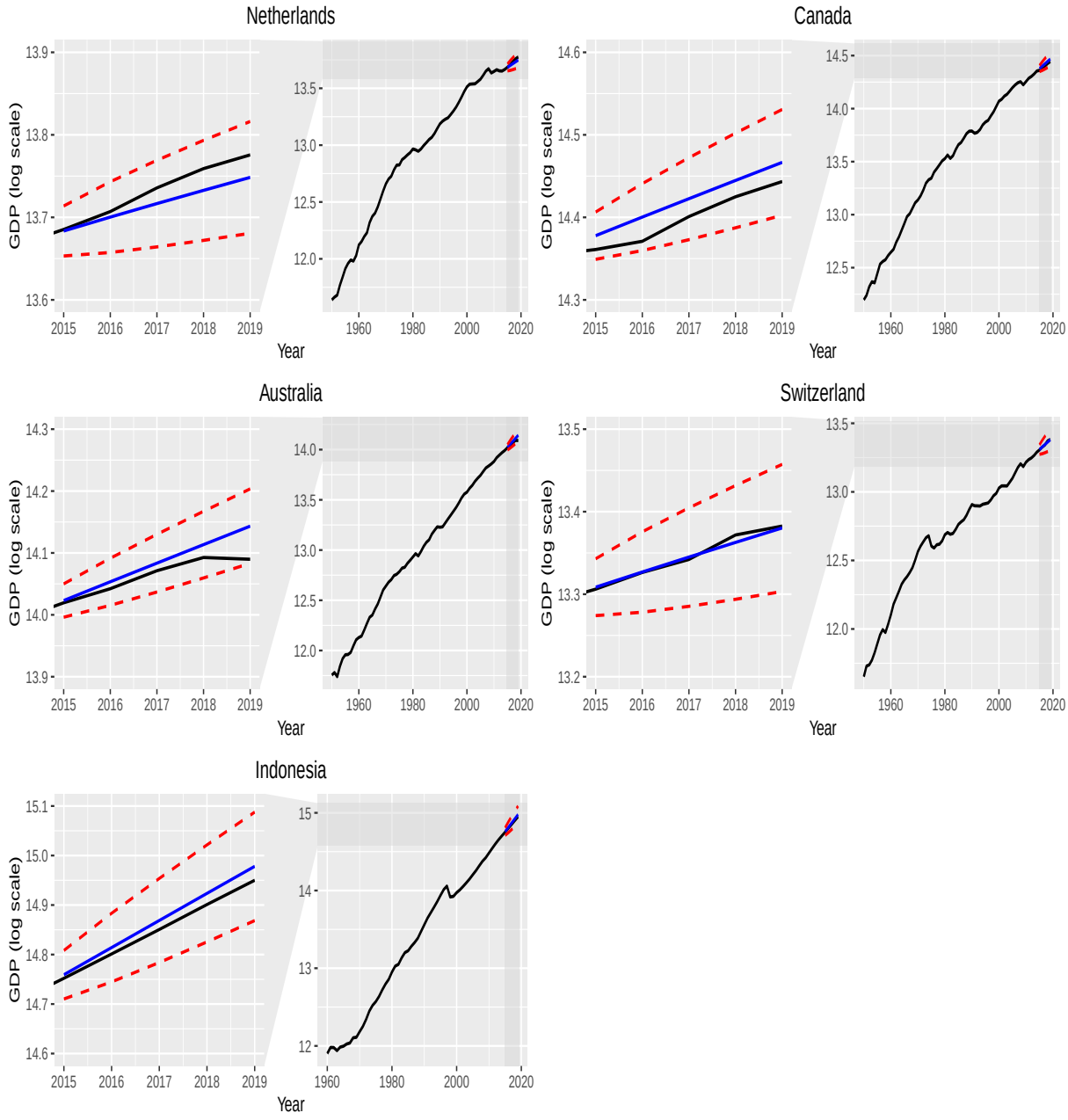

Figure 3: Point forecasts and 95% prediction intervals from the combined model M126 for the Netherlands, Canada, Australia, Switzerland and Indonesia.

8 Conclusion

This paper follows the idea of Fritz et al. (2024) and adopts the four models proposed therein. To further enhance forecasting accuracy, we introduce two additional models– RNNAR and LLNNAR–that integrate NNAR to capture nonlinear relationships that linear models cannot detect. In the RNNAR model, the drift is estimated using NNAR, while LLNNAR applies a local linear regression to the series with NNAR modeling of the residuals. The models are combined using an equal weight averaging method, resulting in a total of 63 different forecasting models presented in this paper.

The RNNAR is identified as the best-performing model among the six single models across all countries, with the lowest average MASE, followed by the LLNNAR. This demonstrates that forecasting performance benefits significantly from incorporating the NNAR model. Additionally, single models generally perform worse than their combinations. In most cases, the combined methods significantly improve forecasting accuracy and reduce standard errors. On average, the combination of RWC, RWL and LLNNAR outperforms all other models. However, it is worth noting that combining more models does not necessarily yield better results. The combination of five and six single models does not provide the best performance for any of the selected countries. Our results show that there is no superior model (single or combination) that outperforms all other models across all time series. However, combining models with equal weights proves to be an efficient way to improve prediction accuracy. Even in countries where combined models do not enhance the final forecasts, it is still less risky than relying on a single model, as there is always the risk of selecting a poorly performing model.

In this study, we focused on a very simple FFNN with a single input node and one hidden layer. For future research, more complex neural network architectures could be explored to potentially improve predictive performance.

References

Azoff, E. M. (1994). *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc.

Box, G. E. P. and G. M. Jenkins (1976). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Bühlmann, P. (1996). "Locally adaptive lag-window spectral estimation". In: *Journal of Time Series Analysis* 17 (3), pp. 247–270.

Chryssolouris, G., M. Lee, and A. Ramsey (1996). "Confidence interval prediction for neural network models". In: *IEEE Transactions on neural networks* 7 (1), pp. 229–232.

Dybowski, R. and S. J. Roberts (2001). "Confidence intervals and prediction intervals for feed-forward neural networks". In: Cambridge University Press.

Fan, J. (1992). "Design-adaptive nonparametric regression". In: *Journal of the American statistical Association* 87 (420), pp. 998–1004.

– (1993). "Local linear regression smoothers and their minimax efficiencies". In: *The annals of Statistics*, pp. 196–216.

Fan, J. and I. Gijbels (2018). *Local polynomial modelling and its applications*. Routledge.

Faraway, J. and C. Chatfield (1998). "Time series forecasting with neural networks: a comparative study using the air line data". In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 47 (2), pp. 231–250.

Feng, Y., T. Gries, and M. Fritz (2020). "Data-driven local polynomial for the trend and its derivatives in economic time series". In: *Journal of Nonparametric Statistics* 32 (2), pp. 510–533.

Feng, Y. and C. Zhou (2015). "Forecasting financial market activity using a semiparametric fractionally integrated Log-ACD". In: *International Journal of Forecasting* 31 (2), pp. 349–363.

Feng, Y. et al. (2022). "The smoots Package in R for Semiparametric Modeling of Trend Stationary Time Series." In: *R Journal* 14 (1).

Fritz, M. et al. (2024). "Forecasting economic growth by combining local linear and standard approaches". In: *Journal of Applied Statistics*, pp. 1–19.

Hendry, D. F. and M. P. Clements (2004). "Pooling of forecasts". In: *The Econometrics Journal* 7 (1), pp. 1–31.

Hibon, M. and T. Evgeniou (2005). "To combine or not to combine: selecting among forecasts and their combinations". In: *International journal of forecasting* 21 (1), pp. 15–24.

Hill, T., M. O'Connor, and W. Remus (1996). "Neural network models for time series forecasts". In: *Management science* 42 (7), pp. 1082–1092.

Hyndman, R. J. and G. Athanasopoulos (2018). *Forecasting: principles and practice*. OTexts.

Hyndman, R. J. and Y. Khandakar (2008). "Automatic time series forecasting: the forecast package for R". In: *Journal of Statistical Software* 27 (3), pp. 1–22.

Hyndman, R. J. and A. B. Koehler (2006). "Another look at measures of forecast accuracy". In: *International journal of forecasting* 22 (4), pp. 679–688.

Papadopoulos, G., P. J. Edwards, and A. F. Murray (2001). "Confidence estimation methods for neural networks: A practical comparison". In: *IEEE transactions on neural networks* 12 (6), pp. 1278–1287.

Refenes, A.-P. (1994). *Neural networks in the capital markets*. John Wiley & Sons, Inc.

Smith, J. and K. F. Wallis (2009). "A simple explanation of the forecast combination puzzle". In: *Oxford bulletin of economics and statistics* 71 (3), pp. 331–355.

Stock, J. H. and M. W. Watson (2004). "Combination forecasts of output growth in a seven-country data set". In: *Journal of forecasting* 23 (6), pp. 405–430.

Tang, Z. and P. A. Fishwick (1993). "Feedforward neural nets as models for time series forecasting". In: *ORSA journal on computing* 5 (4), pp. 374–385.

Terui, N. and H. K. Van Dijk (2002). "Combined forecasts from linear and nonlinear time series models". In: *International Journal of Forecasting* 18 (3), pp. 421–438.

White, H. (1989). "Learning in artificial neural networks: A statistical perspective". In: *Neural computation* 1 (4), pp. 425–464.

Zhang, G. P. (2003). "Time series forecasting using a hybrid ARIMA and neural network model". In: *Neurocomputing* 50, pp. 159–175.

Zhang, G., B. E. Patuwo, and M. Y. Hu (1998). "Forecasting with artificial neural networks: The state of the art". In: *International journal of forecasting* 14 (1), pp. 35–62.

Appendix

Table A1: Summary of MASE across countries for all models.

Models are coded as: 1 = RWC, 2 = RSL, 3 = LLR, 4 = RLL, 5 = RNNAR, 6 = LLNNAR.

Notation: M1 refers to RWC, M12 indicates the combination of RWC and RSL, etc.

Models	UK	IN	MX	CN	US	VN	FR	JP	TR	DE	KR	IL	NL	CA	AU	CH	ID	SD	Mean
M1	0.58	1.37	1.02	**0.81**	0.60	0.35	1.54	2.14	0.27	1.23	1.69	1.34	0.83	1.51	0.91	0.68	0.16	0.55	1.00
M2	0.18	0.33	0.84	2.22	0.44	0.42	0.86	1.29	0.50	1.85	1.21	0.60	1.20	0.26	0.31	1.07	0.26	0.59	0.82
M3	0.99	0.36	0.47	2.43	0.92	0.27	0.38	0.46	0.42	0.57	0.53	0.42	0.76	0.20	0.50	0.38	0.37	0.51	0.61
M4	0.62	0.29	0.37	2.14	0.96	0.60	1.06	0.35	1.00	0.51	0.35	0.38	2.31	0.32	0.25	0.27	0.78	0.62	0.74
M5	1.00	1.33	0.31	1.05	**0.08**	0.56	0.25	0.11	0.26	0.14	1.07	0.82	**0.08**	1.14	0.59	0.27	0.39	0.43	0.56
M6	0.77	**0.24**	0.51	2.51	0.53	0.16	0.39	0.48	0.33	0.55	0.55	0.32	1.06	0.28	0.41	0.29	0.49	0.54	0.58
M12	0.37	0.80	0.20	1.52	0.13	**0.04**	0.34	0.43	0.29	0.31	1.45	0.37	0.20	0.77	0.61	0.21	0.21	0.43	0.48
M13	0.21	0.86	0.32	1.62	0.16	0.31	0.60	0.84	0.31	0.33	1.11	0.88	0.13	0.78	0.70	0.53	0.26	0.40	0.59
M14	**0.15**	0.73	0.34	1.47	0.18	0.47	0.24	0.90	0.56	0.36	0.67	0.86	0.74	0.92	0.37	0.47	0.47	**0.33**	0.58
M15	0.79	1.35	0.65	0.93	0.32	0.46	0.89	1.09	**0.25**	0.62	1.38	1.08	0.43	1.32	0.75	0.21	0.12	0.41	0.74
M16	0.16	0.77	0.31	1.66	0.08	0.24	0.57	0.83	0.26	0.34	1.12	0.83	0.15	0.89	0.66	0.48	0.32	0.41	0.57
M23	0.45	0.35	0.66	2.32	0.68	0.08	0.60	0.87	0.41	1.21	0.87	0.15	0.97	0.23	0.40	0.35	0.32	0.53	0.64
M24	0.26	0.31	0.61	2.18	0.70	0.09	0.96	0.82	0.35	1.18	0.43	0.16	1.76	0.23	0.18	0.40	0.52	0.58	0.66
M25	0.55	0.78	0.30	1.64	0.23	0.08	0.31	0.63	0.33	0.92	1.14	0.11	0.59	0.58	0.45	0.66	0.07	0.41	0.55
M26	0.34	0.29	0.68	2.37	0.49	0.15	0.63	0.89	0.36	1.20	0.88	0.19	1.13	0.23	0.36	0.39	0.38	0.55	0.64

Table continued...

Models are coded as: 1 = RWC, 2 = RSL, 3 = LLR, 4 = RLL, 5 = RNNAR, 6 = LLNNAR.

Notation: M1 refers to RWC, M12 indicates the combination of RWC and RSL, etc.

Models	UK	IN	MX	CN	US	VN	FR	JP	TR	DE	KR	IL	NL	CA	AU	CH	ID	SD	Mean
M34	0.81	0.32	0.42	2.28	0.94	0.43	0.70	0.40	0.48	0.54	**0.09**	0.40	1.53	**0.19**	0.18	0.33	0.57	0.55	0.63
M35	0.17	0.84	0.26	1.74	0.47	0.42	0.16	0.21	0.34	0.30	0.80	0.62	0.37	0.59	0.54	0.14	0.05	0.40	0.47
M36	0.88	0.30	0.49	2.47	0.72	0.20	0.37	0.47	0.38	0.56	0.54	0.37	0.90	0.20	0.45	0.34	0.43	0.52	0.59
M45	0.23	0.71	0.23	1.59	0.49	0.58	0.41	0.16	0.47	0.26	0.36	0.60	1.14	0.73	0.22	0.10	0.20	0.39	0.50
M46	0.69	0.27	0.44	2.32	0.75	0.36	0.72	0.42	0.47	0.53	0.10	0.35	1.69	0.30	**0.17**	0.28	0.63	0.57	0.62
M56	0.20	0.75	0.27	1.78	0.27	0.34	0.15	0.23	0.30	0.28	0.81	0.57	0.51	0.71	0.50	0.11	0.05	0.41	0.46
M123	0.16	0.65	0.24	1.82	0.25	0.07	0.12	0.13	0.34	0.40	1.14	0.39	0.38	0.53	0.57	0.13	0.26	0.44	0.45
M124	0.16	0.57	0.22	1.72	0.27	0.18	0.23	0.17	0.32	0.38	0.85	0.37	0.90	0.62	0.35	**0.10**	0.40	0.40	0.46
M125	0.56	0.98	0.23	1.36	0.11	0.17	0.31	0.29	0.28	0.22	1.32	0.52	0.13	0.89	0.60	0.23	**0.03**	0.42	0.48
M126	0.17	0.59	0.25	1.85	0.13	0.06	0.11	0.13	0.30	0.39	1.15	0.35	0.48	0.60	0.55	0.11	0.30	0.45	**0.44**
M134	0.34	0.61	0.21	1.79	0.43	0.41	0.14	0.45	0.40	0.15	0.62	0.71	0.74	0.63	0.41	0.44	0.44	0.37	0.52
M135	0.23	1.02	0.31	1.43	0.11	0.39	0.48	0.57	0.29	0.24	1.09	0.86	0.09	0.90	0.67	0.27	0.05	0.40	0.53
M136	0.39	0.63	0.22	1.92	0.28	0.25	0.27	0.40	0.32	0.15	0.92	0.69	0.34	0.61	0.61	0.45	0.34	0.41	0.52
M145	0.35	0.93	0.33	1.33	0.13	0.50	0.24	0.61	0.39	0.24	0.80	0.84	0.48	0.99	0.44	0.23	0.18	0.34	0.53
M146	0.27	0.55	0.21	1.82	0.30	0.36	0.13	0.44	0.39	**0.14**	0.63	0.68	0.85	0.70	0.39	0.41	0.48	0.39	0.51
M156	0.31	0.96	0.31	1.46	0.08	0.35	0.47	0.56	0.26	0.23	1.10	0.83	0.11	0.98	0.64	0.24	0.09	0.41	0.53
M234	0.51	0.33	0.56	2.26	0.77	0.15	0.75	0.70	0.31	0.98	0.46	0.07	1.42	0.21	0.21	0.17	0.47	0.55	0.61
M235	0.17	0.64	0.35	1.90	0.46	0.14	0.33	0.57	0.36	0.80	0.94	0.21	0.64	0.40	0.47	0.32	0.08	0.43	0.52

Table continued...

Models are coded as: 1 = RWC, 2 = RSL, 3 = LLR, 4 = RLL, 5 = RNNAR, 6 = LLNNAR.

Notation: M1 refers to RWC, M12 indicates the combination of RWC and RSL, etc.

Models	UK	IN	MX	CN	US	VN	FR	JP	TR	DE	KR	IL	NL	CA	AU	CH	ID	SD	Mean
M236	0.56	0.31	0.61	2.39	0.63	0.05	0.53	0.74	0.38	0.99	0.77	0.07	1.00	0.22	0.41	0.17	0.37	0.54	0.60
M245	0.21	0.55	0.32	1.80	0.47	0.25	0.56	0.54	0.29	0.78	0.64	0.20	1.16	0.50	0.25	0.35	0.22	0.41	0.53
M246	0.43	0.29	0.58	2.29	0.65	0.11	0.77	0.71	0.30	0.97	0.47	**0.06**	1.52	0.22	0.20	0.18	0.51	0.57	0.60
M256	0.20	0.58	0.36	1.93	0.33	0.10	0.33	0.58	0.33	0.80	0.94	0.18	0.74	0.48	0.44	0.35	0.12	0.43	0.52
M345	0.21	0.59	0.28	1.87	0.63	0.48	0.38	0.26	0.37	0.36	0.42	0.54	1.01	0.50	0.30	0.17	0.25	0.41	0.51
M346	0.79	0.30	0.45	2.36	0.80	0.33	0.60	0.43	0.38	0.55	0.24	0.37	1.37	0.22	0.25	0.31	0.55	0.53	0.61
M356	0.25	0.62	0.31	2.00	0.49	0.32	0.22	0.30	0.34	0.37	0.72	0.52	0.59	0.49	0.50	0.17	0.16	0.42	0.49
M456	0.16	0.53	0.29	1.90	0.50	0.43	0.40	0.27	0.36	0.35	0.42	0.51	1.11	0.58	0.28	0.15	0.29	0.42	0.50
M1234	0.23	0.51	0.26	1.90	0.43	0.20	0.26	0.07	0.30	0.42	0.77	0.38	0.86	0.48	0.39	0.13	0.39	0.42	0.47
M1235	0.22	0.82	0.21	1.63	0.19	0.19	0.15	0.12	0.32	0.31	1.12	0.49	0.28	0.68	0.58	0.14	0.10	0.42	0.44
M1236	0.27	0.53	0.29	1.99	0.32	0.09	0.15	0.07	0.33	0.43	1.00	0.37	0.54	0.46	0.53	0.14	0.32	0.45	0.46
M1245	0.30	0.76	**0.20**	1.56	0.21	0.27	0.15	0.14	0.29	0.29	0.90	0.48	0.66	0.75	0.41	0.13	0.20	0.37	0.45
M1246	0.18	0.47	0.27	1.92	0.33	0.16	0.26	**0.07**	0.29	0.42	0.77	0.36	0.94	0.53	0.37	0.12	0.42	0.44	0.46
M1256	0.27	0.78	0.22	1.65	0.11	0.16	0.13	0.12	0.29	0.29	1.13	0.47	0.35	0.74	0.56	0.13	0.13	0.42	0.44
M1345	0.16	0.79	0.22	1.61	0.32	0.45	0.11	0.34	0.34	0.15	0.73	0.74	0.55	0.75	0.46	0.27	0.23	0.37	0.48
M1346	0.45	0.50	0.25	1.97	0.45	0.34	0.19	0.21	0.35	0.17	0.60	0.61	0.82	0.54	0.41	0.40	0.45	0.41	0.51
M1356	0.16	0.81	0.23	1.70	0.22	0.33	0.26	0.31	0.30	0.14	0.96	0.73	0.25	0.74	0.60	0.27	0.16	0.41	0.48
M1456	0.17	0.74	0.22	1.63	0.23	0.41	**0.10**	0.34	0.34	0.14	0.74	0.71	0.63	0.81	0.44	0.24	0.26	0.38	0.48

Table continued...

Models are coded as: 1 = RWC, 2 = RSL, 3 = LLR, 4 = RLL, 5 = RNNAR, 6 = LLNNAR.

Notation: M1 refers to RWC, M12 indicates the combination of RWC and RSL, etc.

Models	UK	IN	MX	CN	US	VN	FR	JP	TR	DE	KR	IL	NL	CA	AU	CH	ID	SD	Mean
M2345	0.17	0.50	0.35	1.96	0.58	0.25	0.50	0.52	0.27	0.73	0.62	0.25	1.06	0.38	0.31	0.19	0.26	0.43	0.52
M2346	0.57	0.31	0.55	2.32	0.71	0.14	0.66	0.65	0.28	0.87	0.49	0.13	1.33	0.21	0.26	0.13	0.48	0.54	0.59
M2356	0.19	0.52	0.39	2.05	0.48	0.14	0.34	0.55	0.35	0.74	0.84	0.24	0.74	0.37	0.45	0.19	0.18	0.45	0.52
M2456	0.15	0.46	0.36	1.98	0.49	0.22	0.52	0.52	0.27	0.73	0.62	0.23	1.14	0.44	0.29	0.21	0.29	0.45	0.52
M3456	0.35	0.49	0.31	2.03	0.61	0.39	0.38	0.32	0.33	0.41	0.45	0.49	1.02	0.45	0.33	0.18	0.31	0.43	0.52
M12345	0.16	0.68	0.24	1.73	0.35	0.27	0.20	0.08	0.27	0.34	0.83	0.47	0.68	0.61	0.43	0.11	0.24	0.39	0.45
M12346	0.34	0.45	0.29	2.02	0.45	0.19	0.28	0.10	0.28	0.45	0.73	0.37	0.90	0.44	0.39	0.15	0.41	0.44	0.48
M12356	0.17	0.69	0.25	1.80	0.26	0.18	0.12	0.07	0.32	0.35	1.01	0.46	0.43	0.60	0.54	0.12	0.18	0.43	0.44
M12456	0.17	0.64	0.24	1.75	0.27	0.24	0.20	0.07	0.27	0.33	0.83	0.45	0.74	0.65	0.41	0.11	0.26	0.40	0.45
M13456	0.18	0.66	0.22	1.79	0.37	0.38	0.14	0.18	0.31	0.16	0.70	0.66	0.65	0.66	0.45	0.27	0.28	0.39	0.47
M23456	0.26	0.44	0.38	2.07	0.57	0.23	0.48	0.51	0.27	0.70	0.60	0.27	1.06	0.36	0.33	0.14	0.30	0.45	0.53
M123456	0.16	0.59	0.27	1.86	0.38	0.25	0.23	0.08	0.26	0.37	0.78	0.45	0.74	0.55	0.43	0.12	0.28	0.41	0.46

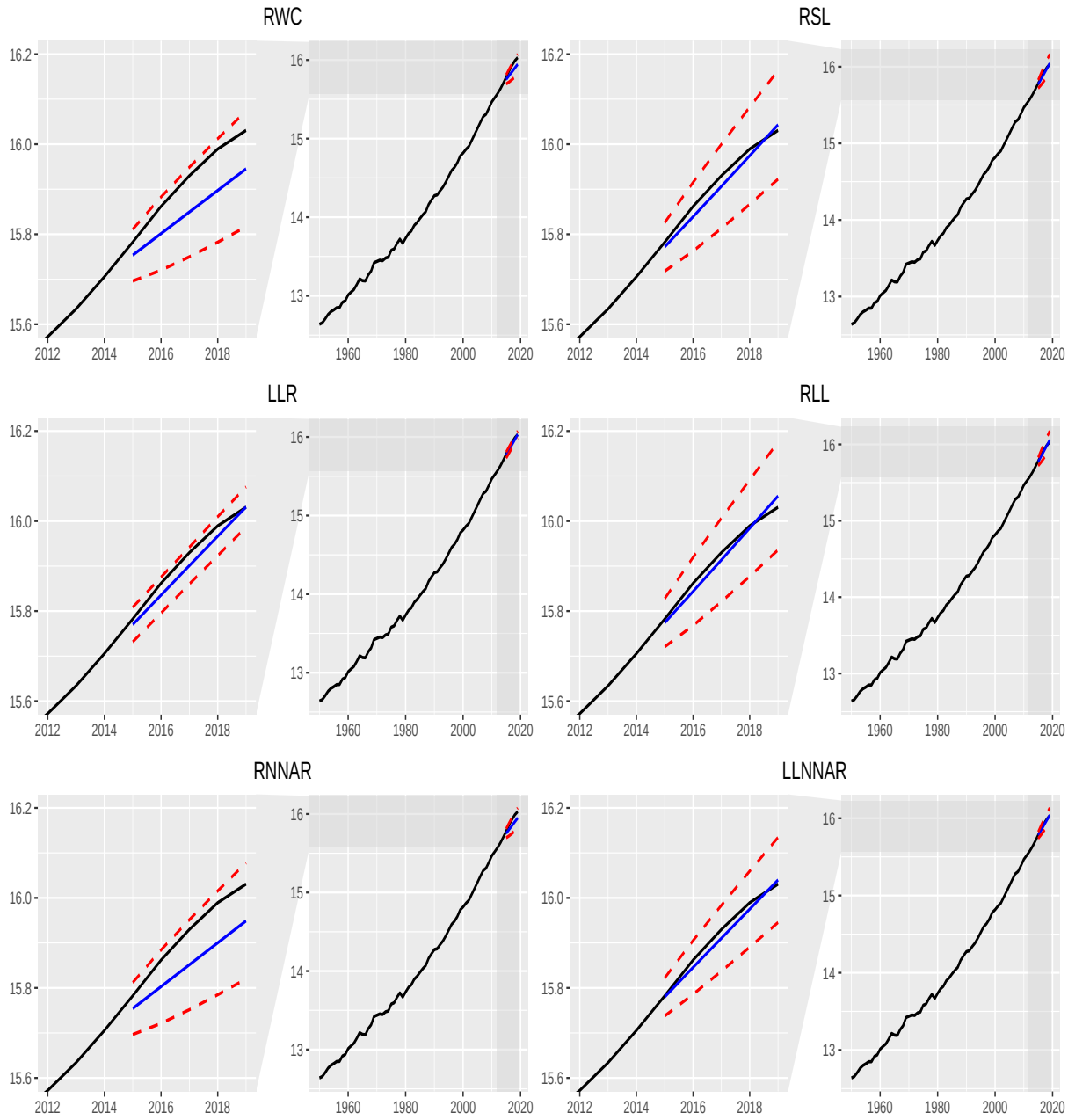

Figure A1: Point forecasts and 95% prediction intervals from all single models for Indian.

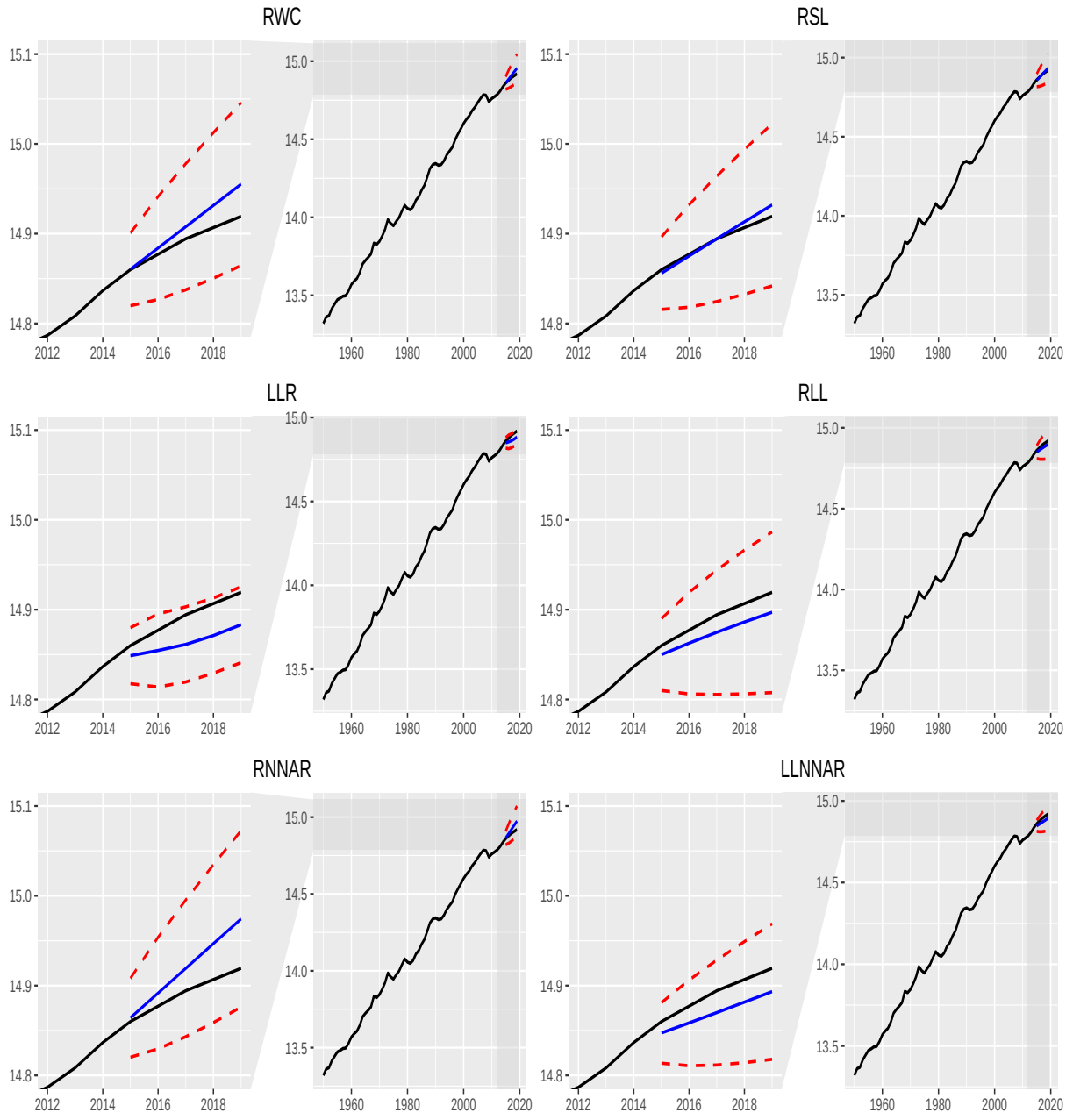

Figure A2: Point forecasts and 95% prediction intervals from all single models for the UK.

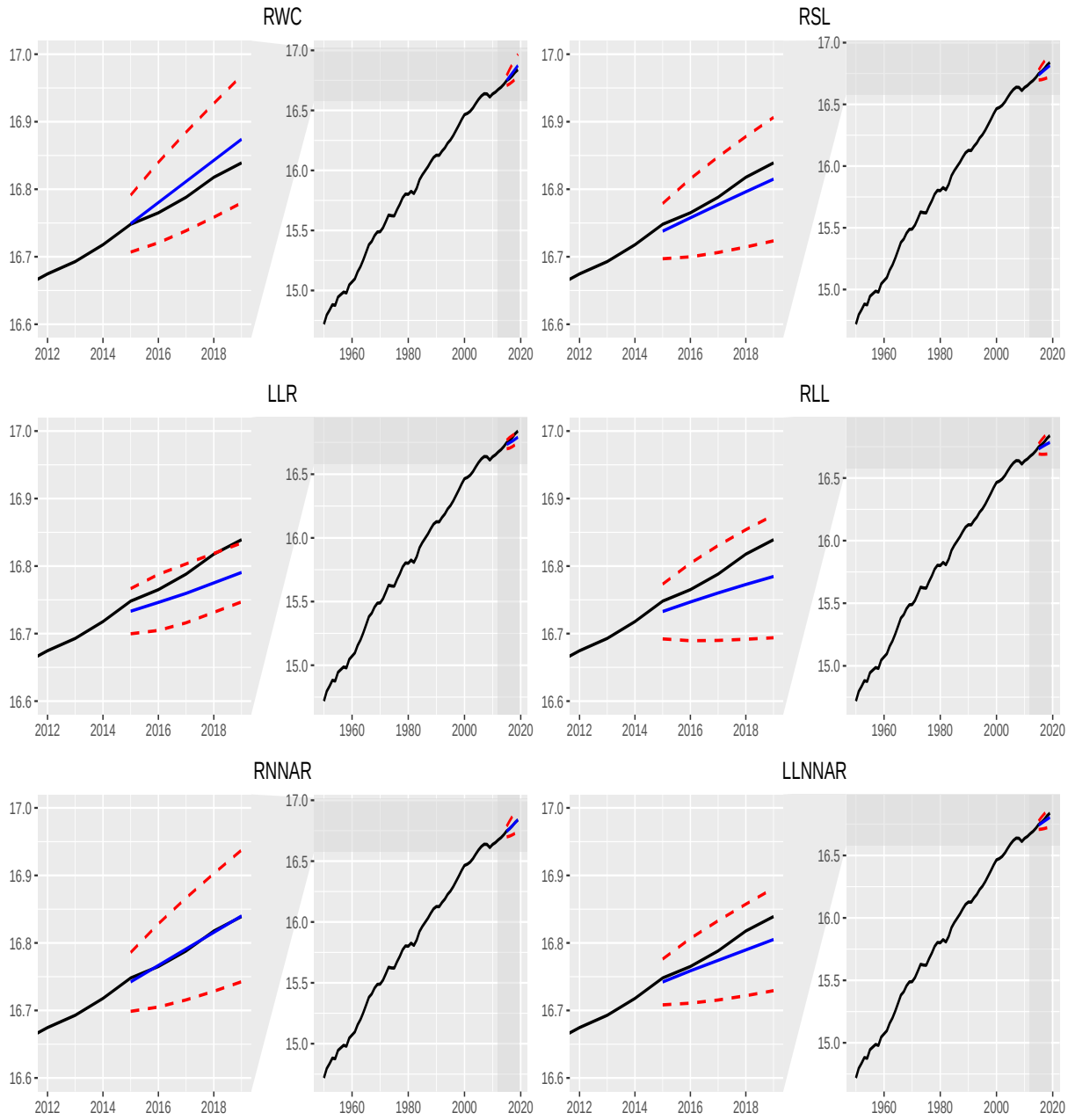

Figure A3: Point forecasts and 95% prediction intervals from all single models for the USA.

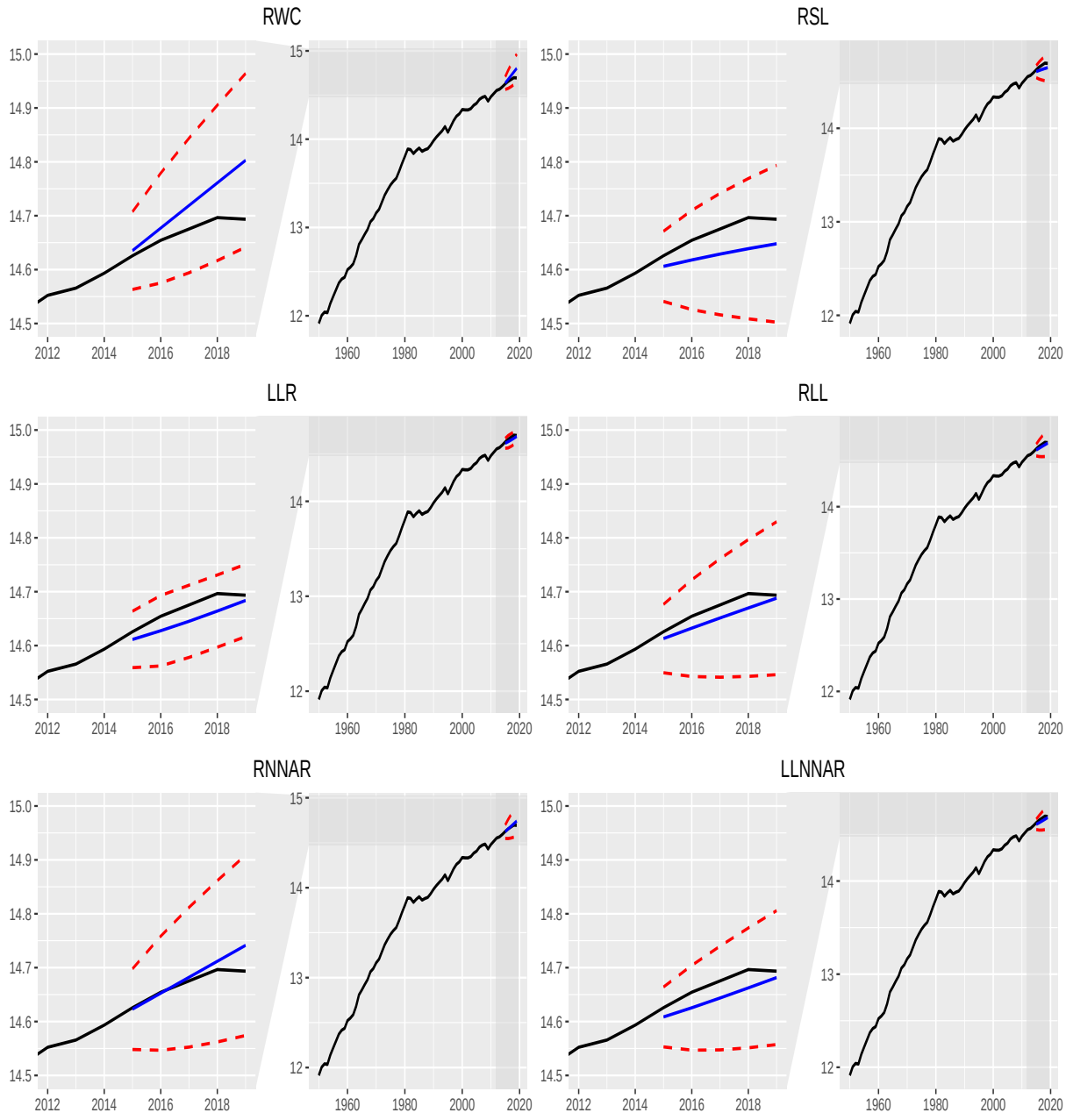

Figure A4: Point forecasts and 95% prediction intervals from all single models for Mexico.

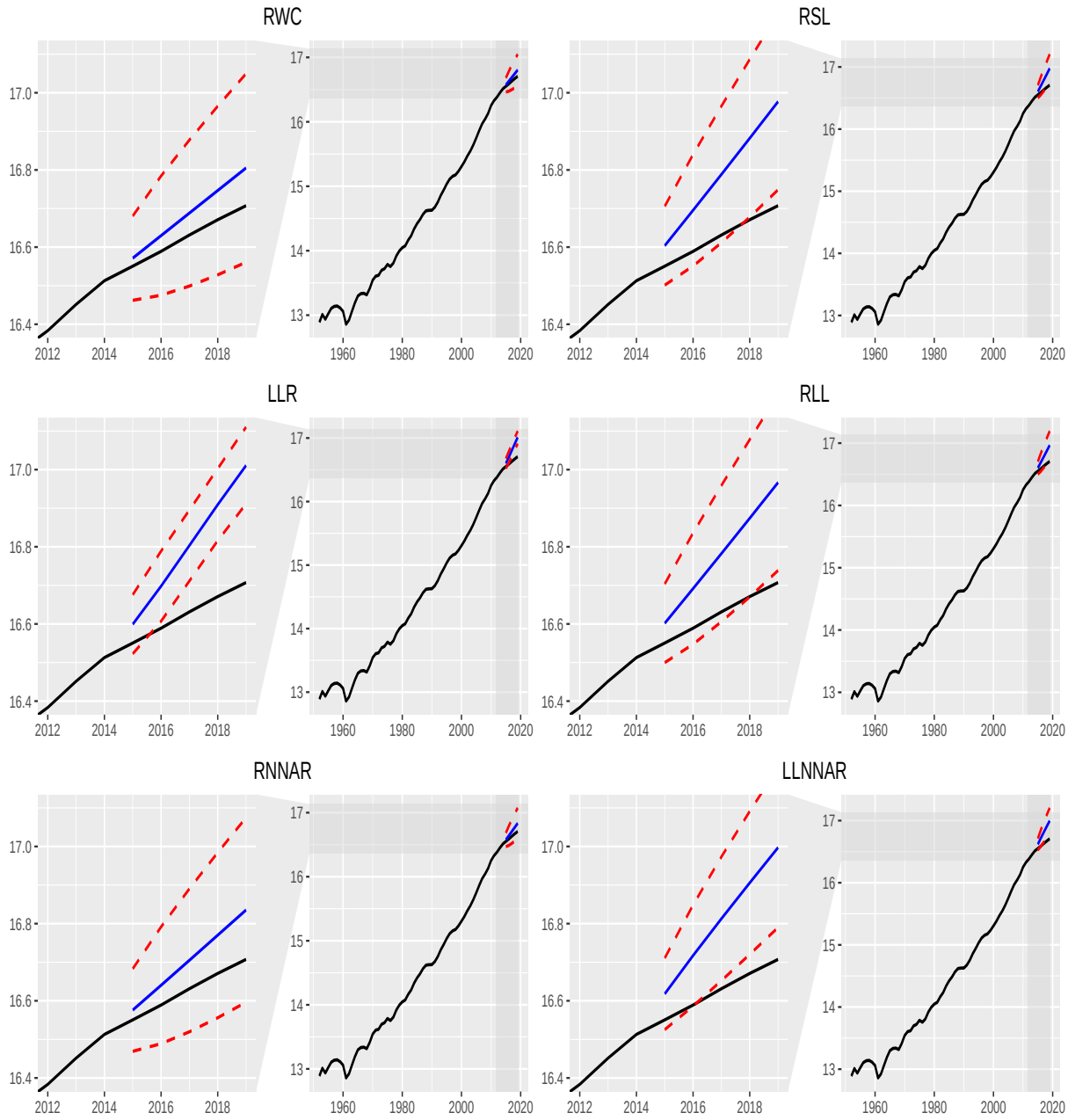

Figure A5: Point forecasts and 95% prediction intervals from all single models for China.

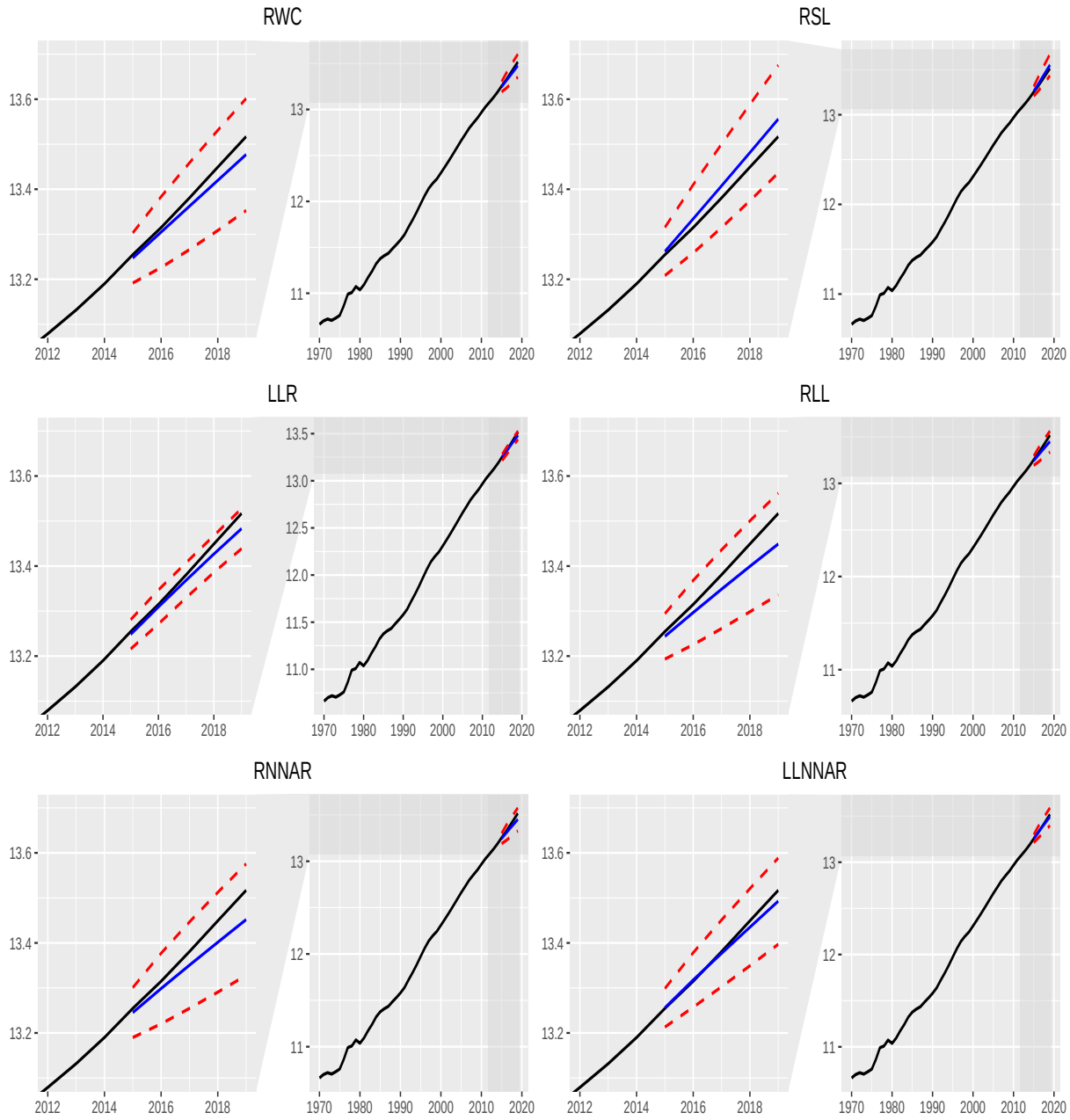

Figure A6: Point forecasts and 95% prediction intervals from all single models for Vietnam.

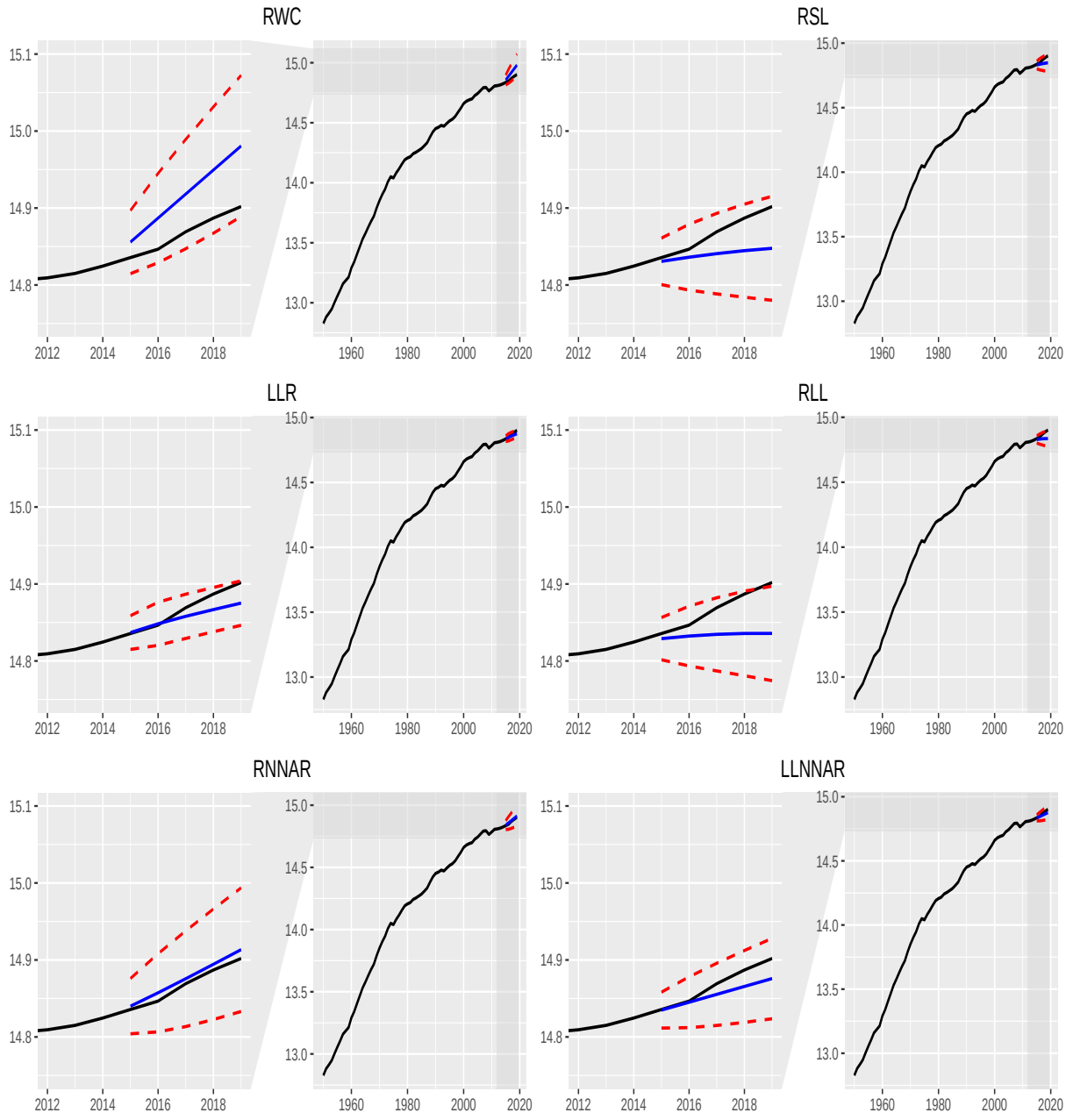

Figure A7: Point forecasts and 95% prediction intervals from all single models for France.

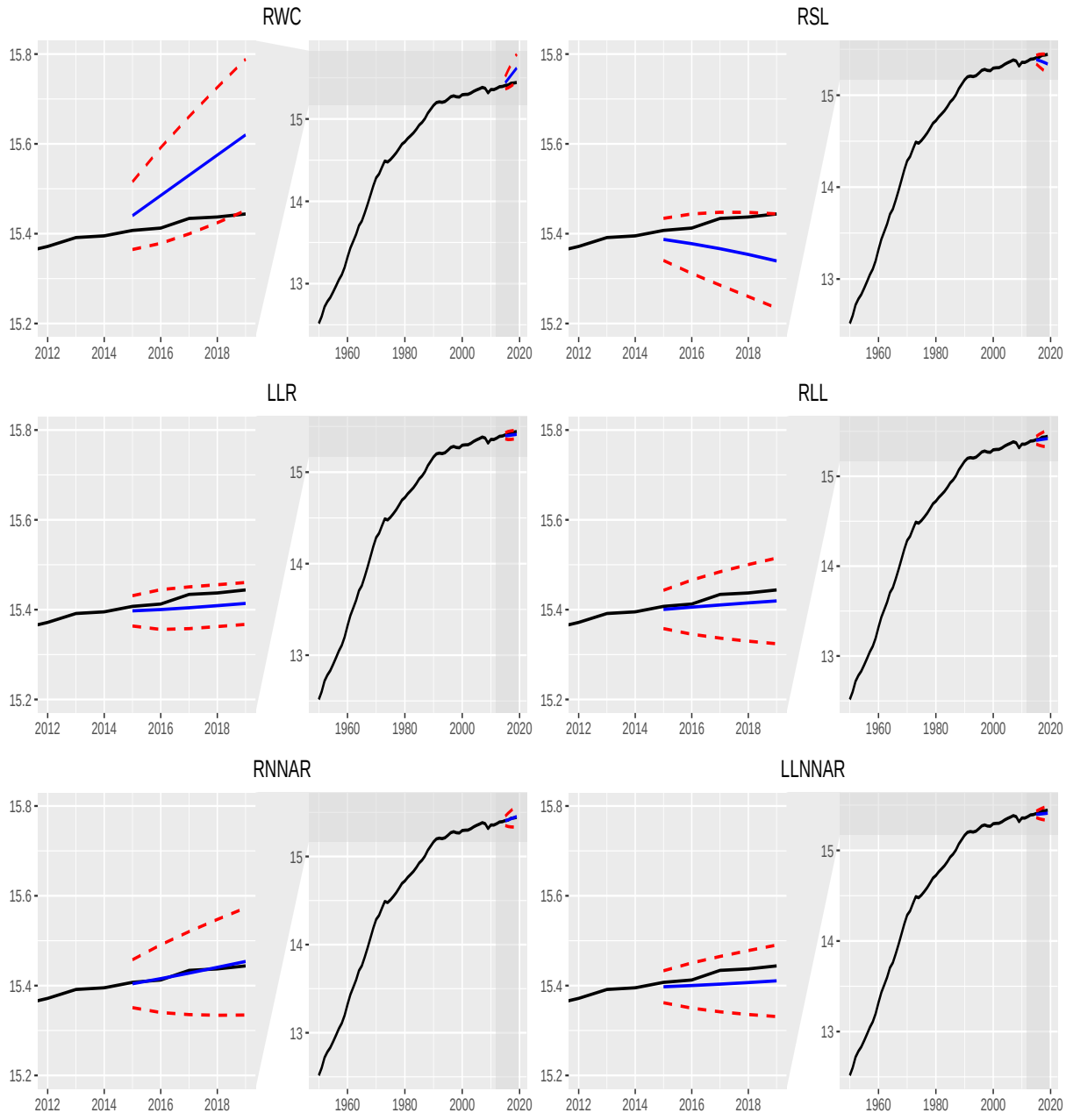

Figure A8: Point forecasts and 95% prediction intervals from all single models for Japan.

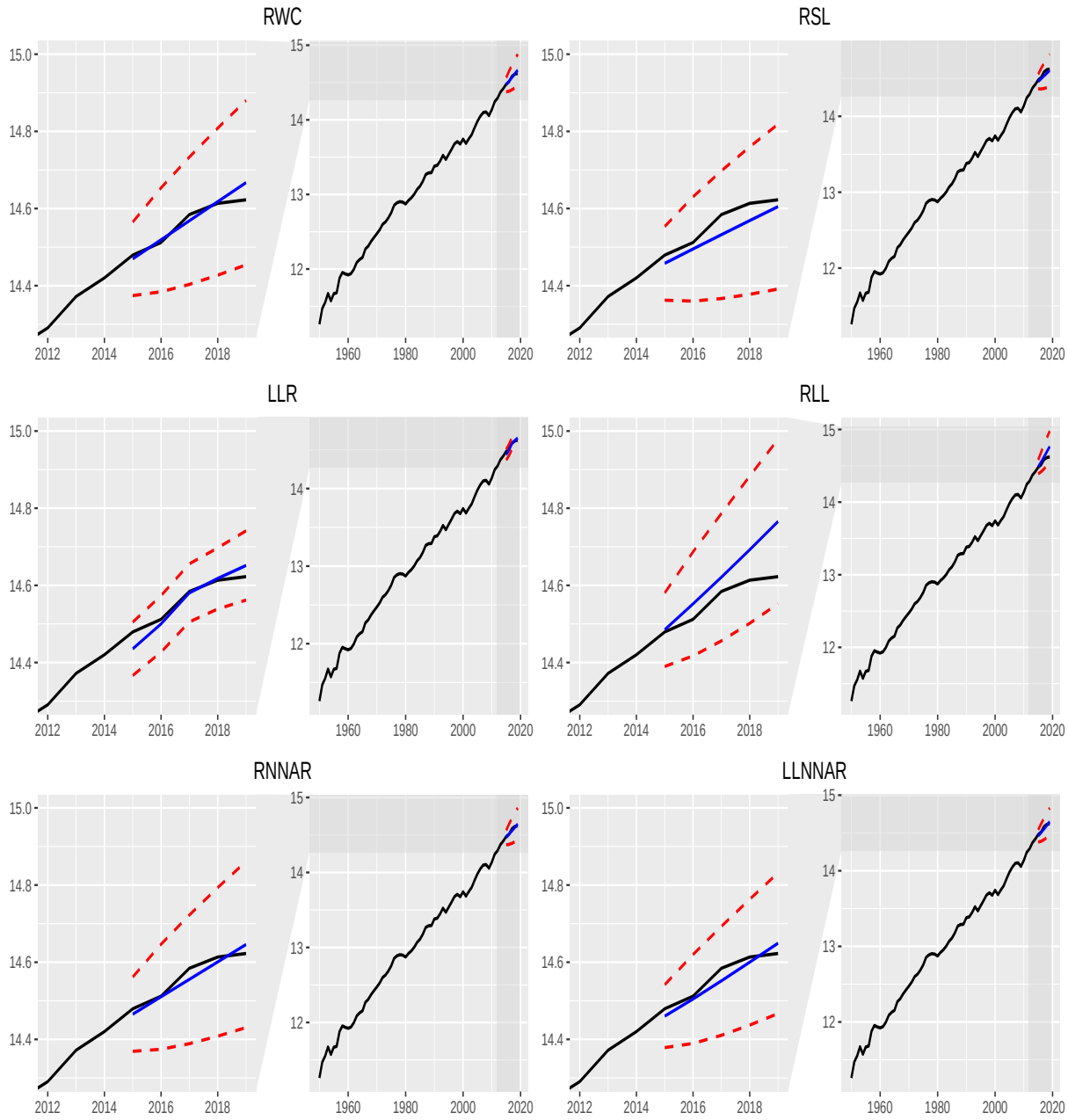

Figure A9: Point forecasts and 95% prediction intervals from all single models for Turkey.

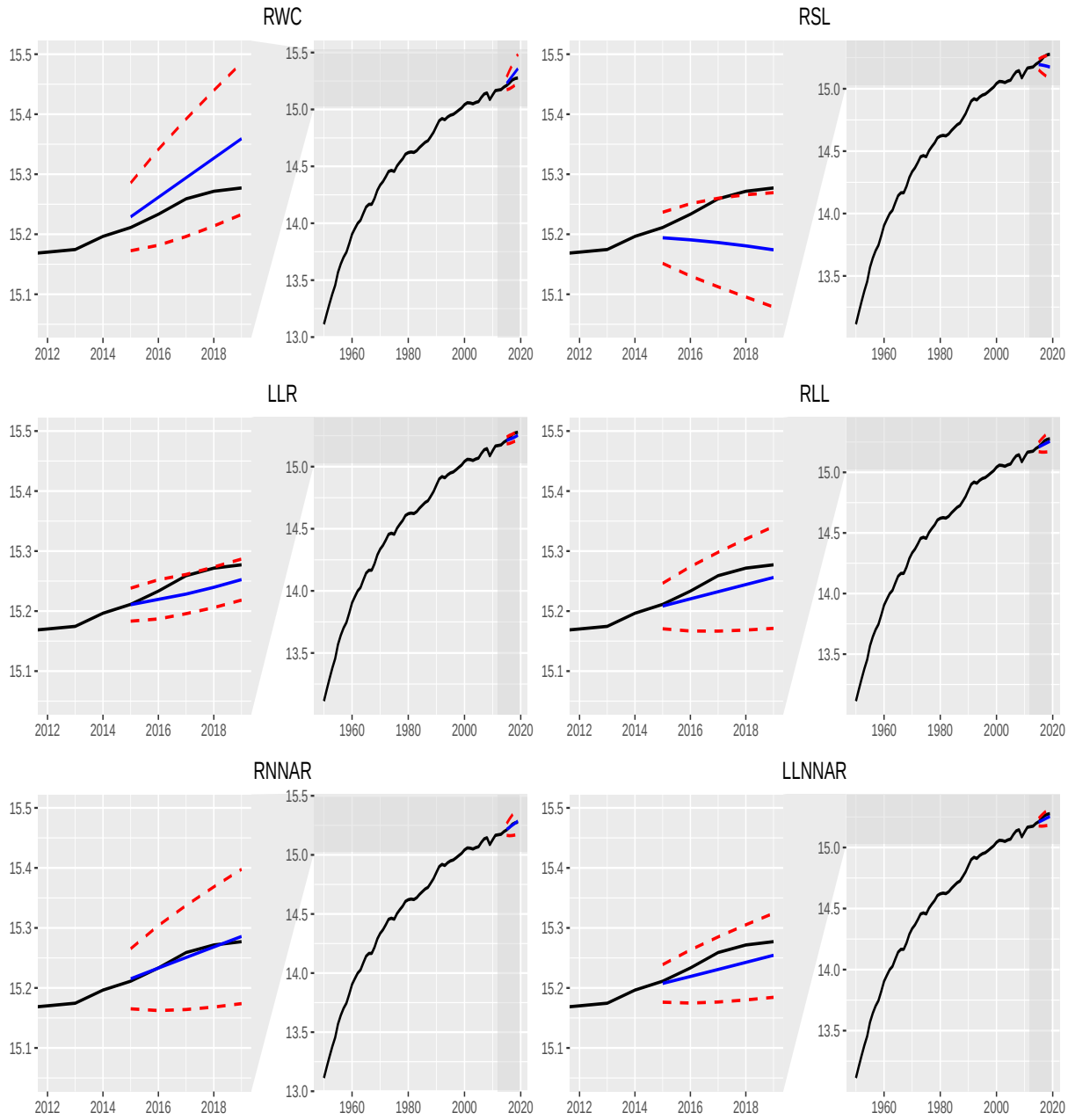

Figure A10: Point forecasts and 95% prediction intervals from all single models for Germany.

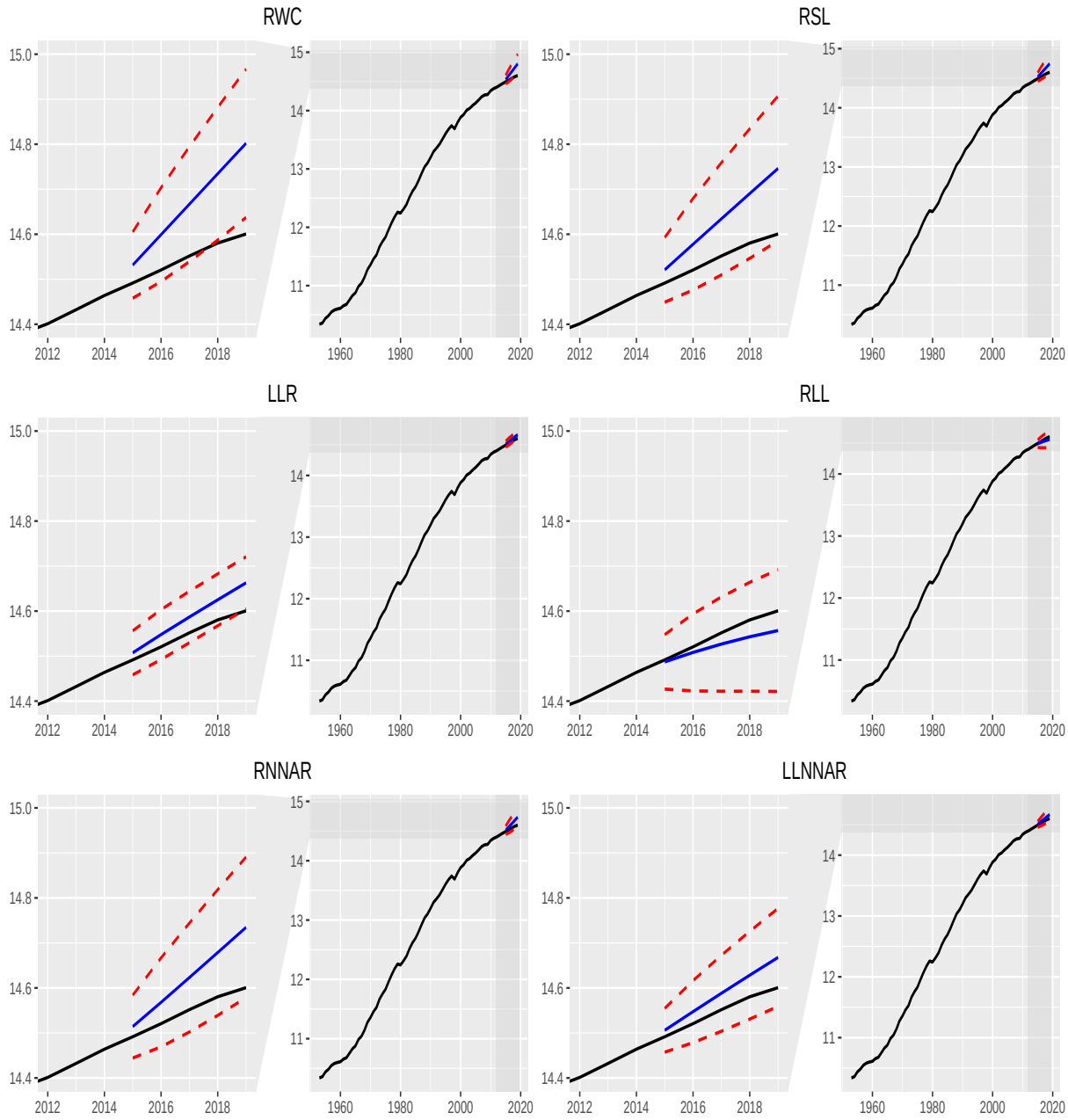

Figure A11: Point forecasts and 95% prediction intervals from all single models for Korea.

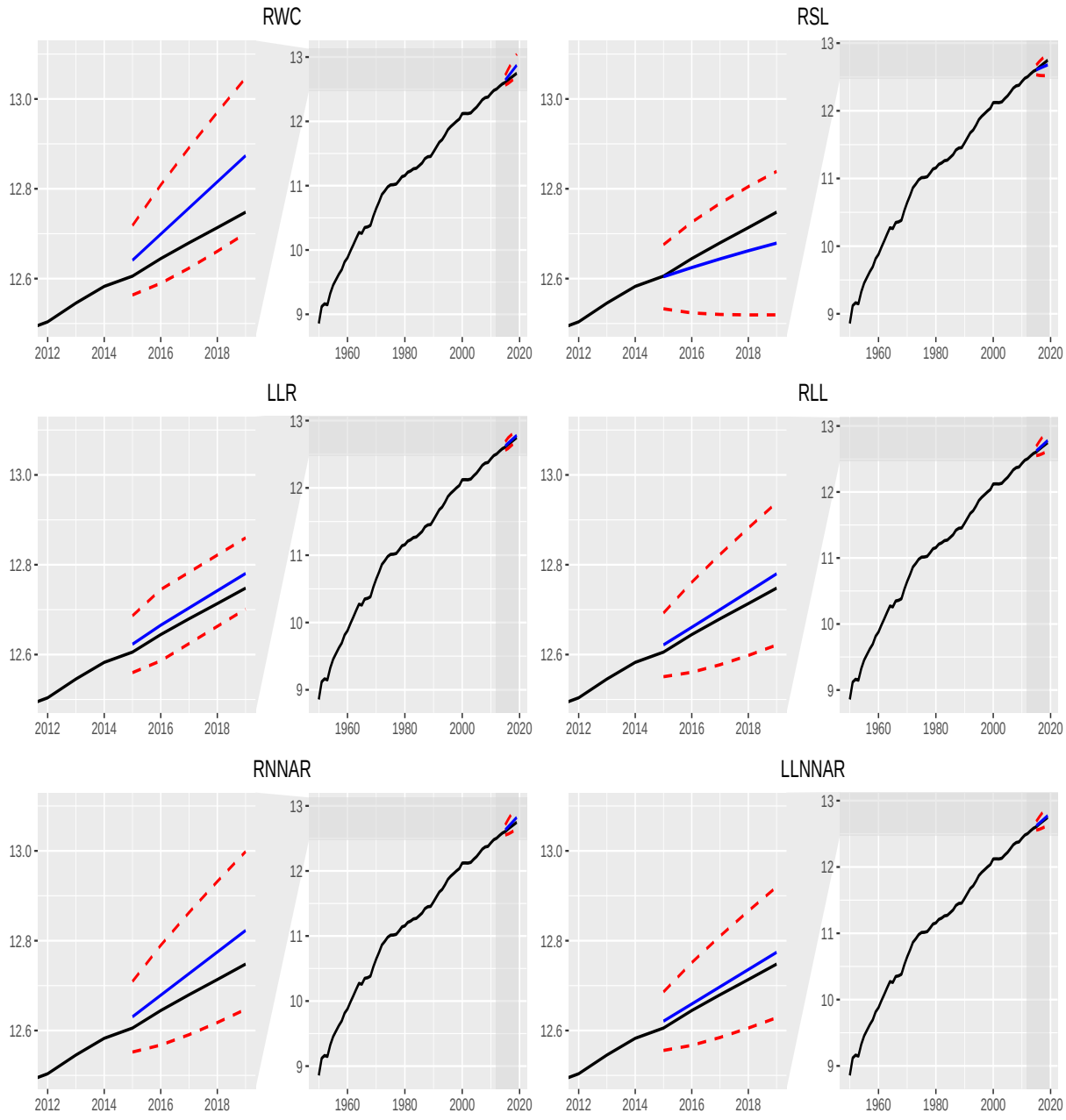

Figure A12: Point forecasts and 95% prediction intervals from all single models for Israel.

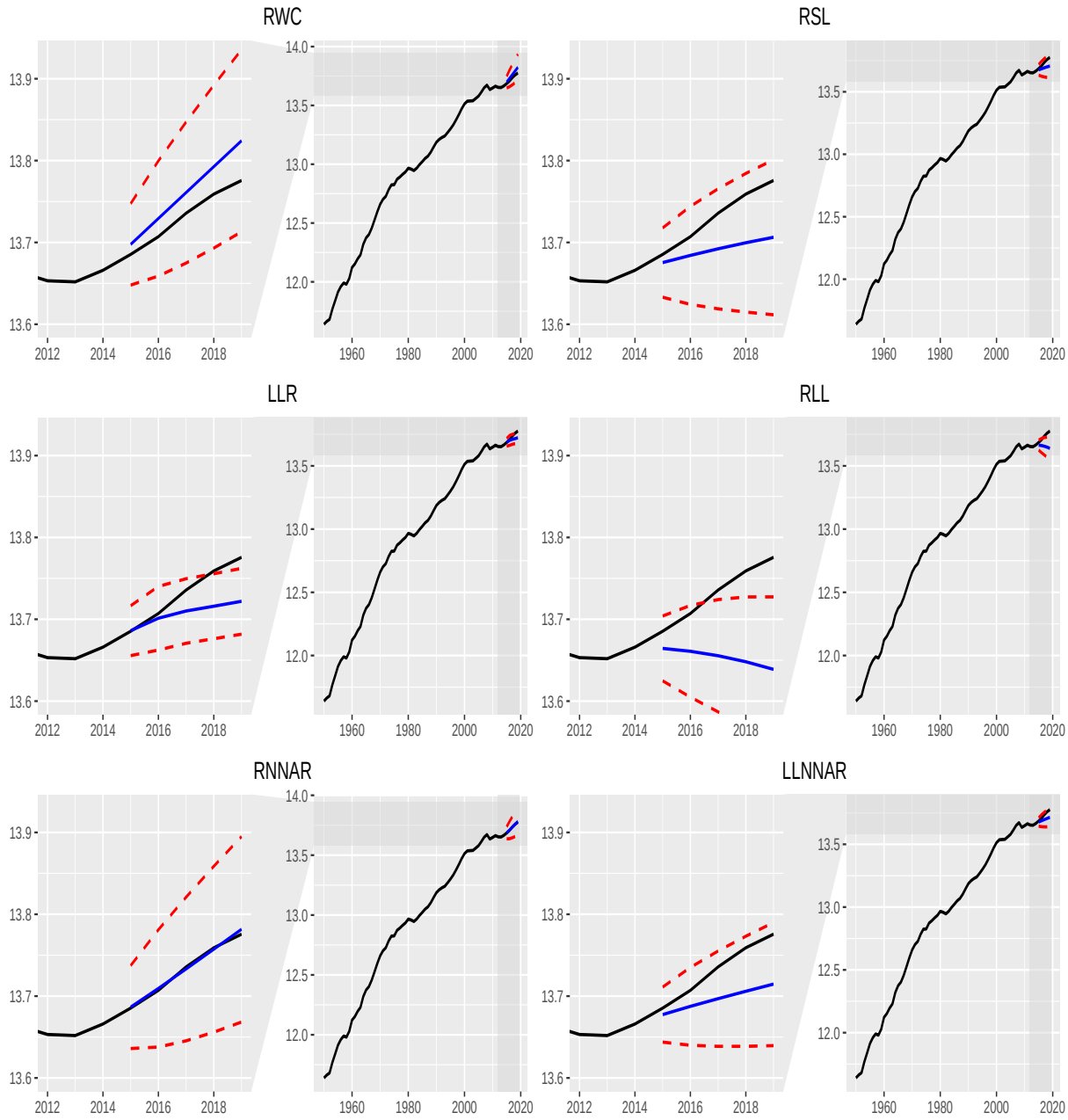

Figure A13: Point forecasts and 95% prediction intervals from all single models for the Netherlands.

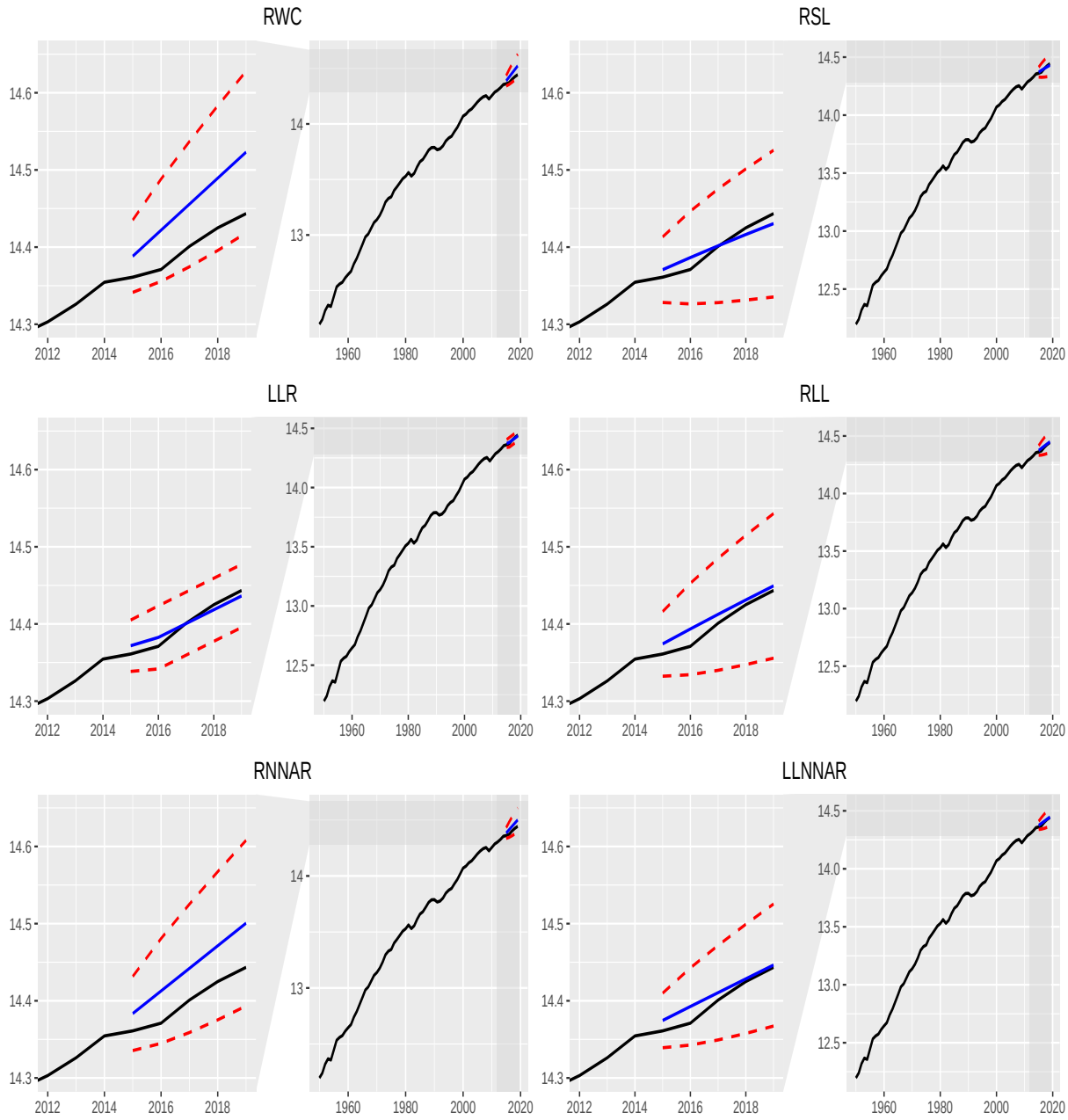

Figure A14: Point forecasts and 95% prediction intervals from all single models for Canada.

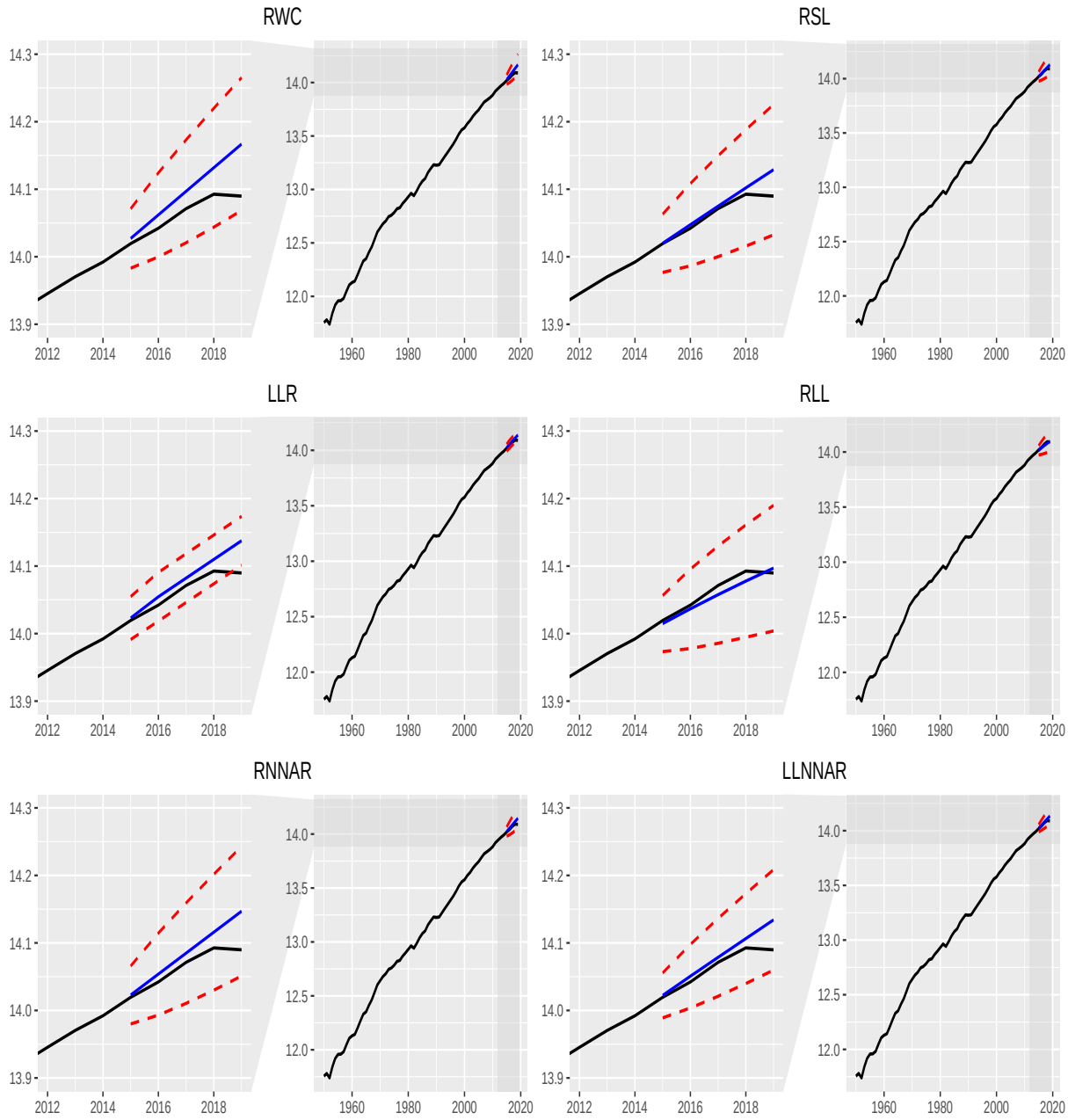

Figure A15: Point forecasts and 95% prediction intervals from all single models for Australia.

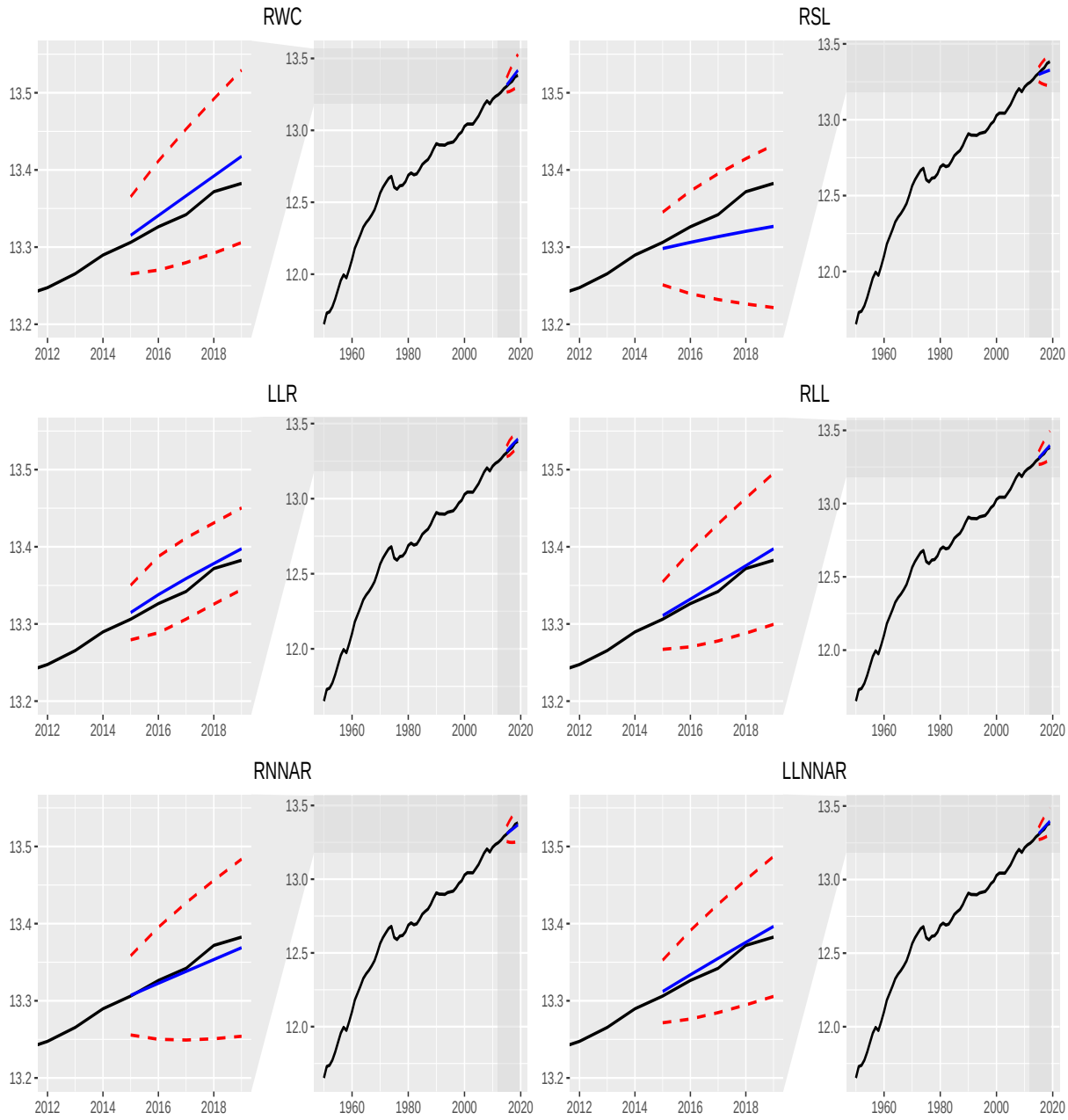

Figure A16: Point forecasts and 95% prediction intervals from all single models for Switzerland.

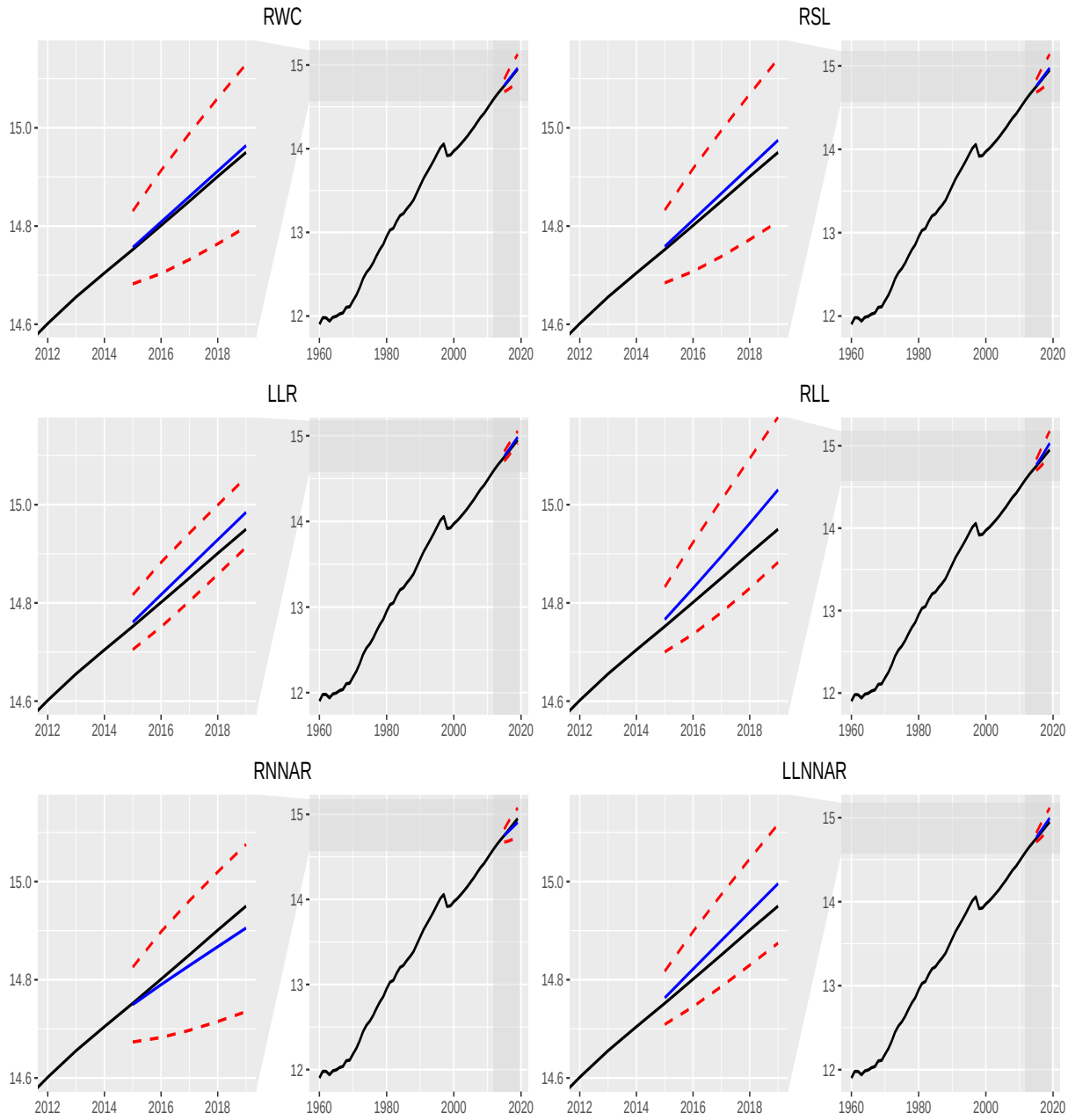

Figure A17: Point forecasts and 95% prediction intervals from all single models for Indonesia.

Recent discussion papers

2026-05	Shujie Li Yuanhuan Feng	Forecasting economic growth with traditional methods and a simple neural network model
2026-04	Dominik Schulz Yuanhua Feng Christian Peitz Oliver Kojo Ayensu	Estimating, Forecasting and Backtesting a Family of Exponential and Other GARCH Models Using the fEGarch Package
2026-03	Dominik Schulz Thi Thu Huong Do Yuanhua Feng	A semiparametric spatial FARIMA applied in the presence of spatial seasonality
2026-02	Dominik Schulz	The R Package deseats for Data-Driven Trend and Seasonality Estimation in Time Series
2026-01	Dominik Schulz Yuanhua Feng Thomas Gries Marlon Fritz Sebastian Letmathe	Diagnosing the trend and bootstrapping the forecasting intervals using a semiparametric ARMA
2025-06	Sarah Kühn Nadja Stroh-Maraun	Community Costs in Neighborhood Help Problems
2025-05	Sarah Kühn	Child Care Allocation Mechanisms: Navigating Incomplete Preference Elicitation
2025-04	Sarah Kühn Papatya Duman Britta Hoyer Thomas Streck Nadja Stroh-Maraun	Non-induced Preferences in Matching Experiments
2025-03	Marlon Fritz Thomas Gries Lukas Wiechers	Growth with Mismatch: Theory and Evidence from TFP Estimates
2025-02	Lukas Wiechers	A Real-Time Analysis of Fundamentals and Bubbles in the S&P 500
2025-01	Claus-Jochen Haake Martin R. Schneider	An Axiomatization of the Banzhaf Index to Measure Influence in Qualitative Comparative Analysis
2024-05	Iuliia Myroshnychenko Thomas Gries	Application of a dynamic welfare-centered energy transition model
2024-04	Claus-Jochen Haake Thomas Streck	A Measure for Contestedness of a Two-Person Bargaining Problem
2024-03	Youssouf Sangare Edem Adotey	Institutions and challenges to development: An Analysis of Economic Growth, Governance, and Development in Sub-Saharan Africa
2024-02	Papatya Duman Claus-Jochen Haake	Size Reduction Reform in German Parliament: a game theoretic analysis of power indices in the Bundestag
2024-01	Iuliia Myroshnychenko Thomas Gries	Developing a dynamic welfare-centered energy transition model: methodological challenges and perspectives
2023-03	Yuanhua Feng Thomas Gries Sebastian Letmathe	FIEGARCH, modulus asymmetric FILog-GARCH and trend-stationary dual long memory time series
2023-02	Claus-Jochen Haake Thomas Streck	Distortion through modeling asymmetric bargaining power
2023-01	Carina Burs	A Model of Cycles and Bubbles under Heterogeneous Beliefs in Financial Markets