INTERNATIONAL MONETARY FUND

When Forecasts Meet Reality: Assessing Climate Damage Functions

Johannes Emmerling, Paul Waidelich, Matthieu Bellon, and Emanuele Massetti

WP/26/85

2026
MAY

WORKING PAPER

 WP/26/85

IMF Working Paper
Fiscal Affairs Department

When Forecasts Meet Reality:
Assessing Climate Damage Functions
Prepared by: Johannes Emmerling[1,2], Paul Waidelich[3], Matthieu Bellon[4], and Emanuele Massetti[4,*]

Authorized for distribution by Dora Benedek
May 2026

ABSTRACT: This paper assesses estimates of the economic impacts of climate change by leveraging the IMF's World Economic Outlook (WEO) forecasts (1990-2023) as climate-free counterfactuals. Placebo tests confirm WEO forecasts do not capture climate effects. By adding climate damage estimates to forecasts and comparing with actual GDP growth, we find climate damage functions explain only a small share of forecast errors—reducing mean absolute errors by up to 0.4 percentage points (about 6% of the forecast error). The most severe damage functions predict contractions in some countries that are inconsistent with observed growth, suggesting overstated near-term climate impacts.

JEL Classification Numbers:	O44, Q54, Q56, E66
Keywords:	climate impacts; validation; forecasts; damage function

Author's E-Mail Address:	Johannes Emmerling (johannes.emmerling@cmcc.it), Paul Waidelich (paul.waidelich@gess.ethz.ch), Matthieu Bellon (mbellon@imf.org), Emanuele Massetti (emassetti@imf.org)

1 CMCC Foundation - Euro-Mediterranean Center on Climate Change, Lecce, Italy.
2 RFF-CMCC European Institute on Economics and the Environment, Via Bergognone 34, 20144 Milan, Italy
3 Energy Technology and Policy Group, ETH Zurich, Clausiusstr. 37, 8092 Zurich, Switzerland
4 International Monetary Fund, Washington D.C.

* We thank participants at EAERE 2025, Bergen, and ECEMP 2025, Brussels, the Malmsten Workshop in Sustainability, Gothenburg, Francesco Granella, Louis Daumas, Marta Mastropietro, and Romain Fillon for very valuable comments. J. Emmerling acknowledges funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101081604 (PRISMA) and No. 101081369 (SPARCCLE); P. Weidelich acknowledges funding from the Swiss State Secretariat for Education, Research and Innovation (SERI, No. 22.00541). The views expressed herein are those of the authors and should not be attributed to the IMF, its executive boards, or its management team.

WORKING PAPERS

When Forecasts Meet Reality: Assessing Climate Damage Functions

Prepared by Johannes Emmerling[1,2], Paul Waidelich[3], Matthieu Bellon[4], and Emanuele Massetti[4,*]

1 CMCC Foundation - Euro-Mediterranean Center on Climate Change, Lecce, Italy.
2 RFF-CMCC European Institute on Economics and the Environment, Via Bergognone 34, 20144 Milan, Italy
3 Energy Technology and Policy Group, ETH Zurich, Clausiusstr. 37, 8092 Zurich, Switzerland
4 International Monetary Fund, Washington D.C.
* We thank participants at EAERE 2025, Bergen, and ECEMP 2025, Brussels, the Malmsten Workshop in Sustainability, Gothenburg, Francesco Granella, Louis Daumas, Marta Mastropietro, and Romain Fillon for very valuable comments. J. Emmerling acknowledges funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101081604 (PRISMA) and No. 101081369 (SPARCCLE); P. Weidelich acknowledges funding from the Swiss State Secretariat for Education, Research and Innovation (SERI, No. 22.00541). The views expressed herein are those of the authors and should not be attributed to the IMF, its executive boards, or its management team

Contents

1 Introduction

Quantifying the economic impacts of climate change has become critical for climate policy issues, including estimating the Social Cost of Carbon (Rennert et al., 2022), quantifying the potential impact on the financial system (Ranger et al., 2022), or the economic implications of overshooting the 1.5 degree target (Drouet et al., 2021). A growing body of literature has complemented the bottom-up approach linking sectoral studies and macroeconomic output used by Integrated Assessment Models (IAMs) (Barrage and Nordhaus, 2024) with empirically estimated links between economic growth and weather shocks. Most studies focus on the effects of random mean annual temperature variations (Dell et al., 2012; Burke et al., 2015; Kalkuhl and Wenz, 2020; Newell et al., 2021; Nath et al., 2024; Desbordes and Eberhardt, 2024; Bilal and Känzig, 2026; Kahn et al., 2021), with an emerging literature including the impact of droughts and weather extremes (Kotz et al., 2021; Akyapı et al., 2025; Kotz et al., 2022; Callahan and Mankin, 2022). However, despite similar empirical strategies and similar data, empirical estimates diverge dramatically (IPCC, 2022; Tol, 2024; Howard and Sterner, 2025)—with implications spanning orders of magnitude for key economic outcomes, such as the social cost of carbon (SCC) (Tol, 2023). Some recent climate-econometric damage functions have produced SCC estimates of over 1,000 US-$ per ton of CO_2 (Ricke et al., 2018; Bilal and Känzig, 2026), which would render immediate, substantial climate action unilaterally rational for most major economies (Bilal and Känzig, 2025), and is one order of magnitude higher than alternative SCC approaches (Rennert et al., 2022) or impact estimates from bottom-up studies (van der Wijst et al., 2023; Takakura et al., 2019). The divergence of the economic impacts of climate change has also been acknowledged and initial efforts to combine and reconcile the figures have been undertaken (Morris et al., 2025). While meta-studies exist (Howard and Sterner, 2025; Tol, 2024; Barrage and Nordhaus, 2024), there is still considerable uncertainty about the impacts and the appropriate damage functions to be used for impact assessment in the short- and long-term (Lemoine et al., 2025).

This uncertainty creates both a scientific challenge and a dilemma for policymakers: how should researchers and policymakers choose between competing damage functions when each claims empirical or model-based support? Validating forward-looking impact projections for the distant and fundamentally uncertain future, such as end-of-century estimates under a "business as usual" scenario, is difficult, as even small climate-induced dents in annual economic growth could plausibly compound to severe impacts over time (Piontek et al., 2021). At the same time, a backward-looking validation of damage functions based on historical data is complicated by several factors. First, the true counterfactual growth path in the absence of any climate impacts is and will be unknown. Second, regression models that link economic growth to climate outcomes explain only a small fraction of the total growth variation. As a result, methods optimizing out-of-sample performance, such as cross-validation, struggle to identify significant performance differences between models suggested by different studies (Newell et al., 2021). Third, since no state-of-the-art causal inference methods can be applied to climate and broad GDP measures at the macro level, there is a key identification concern that temperature-growth relationships estimated with country fixed effects may partially reflect omitted country-specific trends that correlate with temperature, rather than causal climate impacts (Kahn et al., 2021).

This paper contributes to addressing these three challenges in the literature, by utilizing the International Monetary Fund (IMF)'s economic growth forecasts from the World Economic Outlook (WEO) as a baseline counterfactual for estimating historical climate change impacts and comparing a large set of recently proposed damage functions. The WEO forecasts GDP growth and other macroeconomic variables for each country over a five-year window and are updated every six months. These forecasts are developed based on country-level economic modeling of different complexities (Genberg et al., 2014a). The key principle is that these professional forecasts incorporate extensive information on economic fundamentals as well as structural and institutional dynamics, but they account for neither future weather shocks nor their expected impacts. They presumably capture the influence of existing climate change trends to the extent that WEO projections are anchored in historical growth rates. However, they do not account for changes in these existing trends.

Combining this dataset with a consistent implementation of available damage functions, we analyze the implied results to address three research questions: first, we use the forecast to test whether the forecasted GDP per capita growth is indeed not significantly related to climate variables, as implied by our assumption that it reflects counterfactual growth without climate impacts. Second, we calculate historical country-level climate impacts as per various prominent damage functions for multiple five-year evaluation windows between 2011-2019 using historical climate data.[1] We then add them to corresponding WEO growth forecasts. These "reconstructed predictions" of GDP per capita growth including climate impacts allow us to assess to what extent damage functions reduce the difference between WEO forecasts and actually observed growth, indicating how much of the growth difference was potentially caused by climate change. Third, we provide meta-estimates of the already "baked-in" climate impacts into current (2023) GDP in Europe by combining different, consistently implemented damage functions according to their relative performance in reducing forecast

[1]We focus on the 2011–2019 period as the most recent window that excludes widespread crisis years, such as the COVID-19 pandemic. We exclude the 2008–2010 financial crisis period and 2020–2023 pandemic period to avoid confounding from extreme non-climate shocks.

errors.

We find several key results. First, our placebo test supports that temperature shifts affect actual GDP per capita growth rather than IMF forecasts, consistent with the notion that climate impacts are contained in the forecast error but not in the forecast itself. This result validates the use of the WEO forecasts as a climate-naïve counterfactual. Second, we demonstrate that different climate damage functions can explain a low but positive share of IMF forecast errors. The best-performing damage functions reduce the WEO's mean absolute error over a five-year horizon by 0.1–0.4 percentage points of GDP per capita growth rates—representing up to 6% of the total forecast error or at most 3% of average annual growth. Most of the forecasting error is due to the WEO's systematic positive bias and the good performance of some damage functions is essentially due to the fact that they predict systematic damages on average across countries. However, a few damage functions are also able to reduce the dispersion of forecast errors around the average error. Third, our approach enables a country-specific reality check of some of the recent very strong (positive and negative) expected economic climate damage estimates, which reveals important limitations. The most severe damage functions would have predicted GDP per capita contractions in several countries over the last decade that are inconsistent with observed growth patterns, suggesting these may overstate the magnitude or the year-to-year fluctuations of near-term climate impacts. This finding does not necessarily invalidate long-term projections, but it raises questions about short-term validity in some specifications and countries. Finally, we estimate the cumulative impact of observed warming on European GDP from 2010 to 2023 using a weighted average of impacts predicted by the best-performing damage functions. We find that European GDP is approximately 1.0% lower in 2023 compared to a counterfactual with no additional warming after 2010.

2 Data

We build a large dataset of country-level GDP forecasts and actual growth rates from the IMF World Economic Outlook (WEO) covering the period 1990–2023. The WEO is published twice per year (April and October), but we use only the Fall edition for each year. The forecasts cover the ongoing year plus five subsequent years for more than 180 countries and include growth rates of total real GDP in local currency units, public debt, current account balance, and inflation. The WEO forecasting process combines "top-down" and "bottom-up" approaches (Genberg et al., 2014a): an IMF global econometric model first forecasts initial global conditions transmitted to country desk economists, who then incorporate these conditions with inputs from country authorities and other forecasters. The methodology varies substantially across countries depending on data availability and forecasting horizon. Specifically, the more data a country has available, the more likely it is that sophisticated econometric techniques are used, ranging from structural multi-equation models to reduced-form equations and historical values (Genberg et al., 2014a,b). Consequently, forecasts for low-income countries and at longer horizons tend to rely on less sophisticated approaches. Decentralized reviews and centralized checks before final publication aim to ensure consistency across forecasts.

Crucially for our methodology, there is no evidence of systematic use of country-level climate projections, especially before 2016. Subsequent IMF guidance (IMF, 2016) explains that three- to five-year ('long-term') country projections should include the expected impact of natural disasters for the region times the historical probability of these disasters occurring in that country each year. It also recommends that short-term projections should not account for possible natural disaster effects. In contrast, there is no such recommendation for less salient weather shocks or for using climate projections. Moreover, there is no mention of climate in the 2014 IMF forecast evaluation report (Genberg et al., 2014b), and a careful review of all the citations in the 2022 stock-taking IMF Staff Climate Note (Aligishiev et al., 2022) does not show any reference to the use of climate variable projections before 2018.

The IMF WEO forecast errors and their patterns have been well studied (Genberg et al., 2014b; Celasun et al., 2021; Gatti et al., 2024), typically finding that their accuracy is broadly similar to that of the private sector. Nevertheless, short-term and medium-term WEO forecasts typically over-predict GDP growth and underestimate inflation. Forecast errors in WEO forecasts are also not evenly distributed. They are dependent on the global and country cycles: growth over-predictions are more likely to occur during recessions, while there is weaker evidence of systematic errors outside of these episodes. Forecast errors are highest for emerging-market and low-income economies. Further, predictive accuracy drops as the forecasting horizon gets longer, but not linearly: errors increase more steeply over the three-month to two-year horizons. Given the IMF's political and financial influence, these forecasts have been also studied from a political economic perspective, revealing small effects related to ongoing policy discussions (Dreher et al., 2008). Research on other potential sources of biases (e.g., Eicher and Rollinson (2023)) indicates only minor effects. Because climate projections are not systematically incorporated, the forecasts provide a robust climate-naïve counterfactual for short- and medium-term analysis across a broad set of countries over recent decades.

Starting with GDP growth rate forecasts, we derive per-capita growth rates, which represent the conventional dependent variable in climate-econometric modeling. We use observed population growth rates for the study period from the WEO Spring 2024 edition to convert

GDP growth forecasts into GDP per capita growth forecasts. Since population growth is only marginally affected by climate during the sample period covered, using GDP forecasts seems reasonable. We also use the historical data from the latest WEO (Oct 2024) for actual observed GDP and population levels.

We combine this dataset on GDP growth and forecasts with climatic variables used in the studies included in our sample. For mean temperature and annual precipitation, the two main variables used in the literature, we use population-weighted annual averages based on the ERA5 reanalysis dataset from a peer-reviewed repository of climate indicators with global coverage (Gortan et al., 2024). This approach helps avoid inconsistencies arising from the fact that many damage function studies use different climate data sources and aggregation weighting schemes. For the damage function by Bilal and Känzig (2026) that is driven by global mean surface temperature (GMST) change instead of country-level climate variables, we take GMST anomaly data from the widely used Berkeley Earth dataset (Rohde and Hausfather, 2020). For Akyapı et al. (2025), whose LASSO-informed damage function is driven by measures of heat, drought, and moderate temperatures, we use their original ERA5-based data set.

The summary statistics of our sample are displayed in Table 1. Country-level population-weighted temperatures range between -2 (in Mongolia) and 30 °C (in the Sahel), while annual precipitation ranges between less than 10 mm (in Qatar) to over 5,000 mm (in Costa Rica). Global warming, measured as the change of GMST relative to the pre-industrial 1850–1900 baseline, averages approximately 0.98 °C during our sample period and reaches 1.5 °C in 2023. Real GDP growth averages 3.1%, with large variation across countries and over time. GDP contractions greater than -50% occurred in conflict periods (Libya 2011, Iraq 1991, Armenia 1992) and during the COVID-19 pandemic (Macao 2020). Large growth spikes of up to +104.5 % occur in subsequent recovery years (e.g., Libya in 2012, Kuwait in 1992 following the Gulf War, and Macao in 2023). Compared to observed growth rates, the one-year-ahead IMF forecast is, on average, about 0.9 percentage points higher and exhibits far less extreme lower tails (the worst forecast being only -10 % during our sample period), reflecting the difficulty of predicting extreme crises such as the pandemic or wars. This difference vanishes for same-year forecasts, which incorporate information available by the fall.

In Appendix C, we list and discuss countries that are included in the WEO but are excluded from our analysis due to missing climate, damage function, or GDP data.

Table 1: Summary statistics

Statistic	N	Mean	St. Dev.	Min	Max
Annual mean temperature (°C)	6208	19.805	7.114	−1.841	29.843
Baseline temperature 1980-2010 (°C)	6208	19.450	7.206	−0.866	28.427
Temperature anomaly (detrended, °C)	5475	0.021	0.120	−0.189	0.287
Annual precipitation (mm)	6208	1232.910	842.079	7.540	5144.120
GMST change over 1850-1900 (°C)	6252	0.999	0.233	0.551	1.535
GMST shift vs. 1986-2005 baseline (°C)	6252	0.229	0.233	−0.218	0.766
Drought index (PDSI < -4, z-score)	5222	0.005	1.044	−5.668	5.427
Temperature stress 9-12 months (z-score)	5222	−0.004	1.019	−8.370	10.048
Extreme heat days (>35°C, z-score)	5222	0.010	1.036	−12.594	9.429
Temperature variability (°C^2)	6252	0.175	0.532	0.000	4.121
Wet days (days)	6252	17.199	50.795	0.000	330.876
Extreme precipitation (mm)	6252	2.276	10.648	0.000	221.000
Population growth	6124	0.015	0.018	−0.186	0.349
Real GDP per capita growth	6124	0.017	0.056	−0.613	0.928
IMF per-capita growth forecast (1 yr ahead)	5886	0.026	0.036	−0.246	0.851
IMF total growth forecast (1 yr ahead)	5886	0.041	0.034	−0.101	0.854
Annual growth rate MAE (1 yr ahead, pp)	5886	0.028	0.043	0.000	0.925

3 Empirical strategy

Let g_{it} represent the actual growth rate of real GDP per capita for a given year t and country $i = 1, .., N$. Let $\hat{g}_{it}^{\tau}$ denote the IMF forecast made in year τ, calculated using the real GDP growth forecast adjusted for actual population growth. As the IMF forecasts GDP growth five years ahead, for each year t up to 2020 we have five different Fall WEO forecasts from previous years ($\tau = t-1, ..., \tau = t-5$, and one same-year forecast ($\tau = t$). The counterfactual path of GDP per capita predicted by the IMF in year τ can be reconstructed by applying forecasts of annual GDP growth to the GDP per capita observed in the forecast year, $y_{i\tau}$:

$$\hat{y}_{it}^{\tau} = y_{i\tau} \prod_{k=\tau+1}^{t} (1 + \hat{g}_{ik}^{\tau}). \tag{1}$$

$\hat{y}_{it}^{\tau}$ is our counterfactual scenario. It assumes "average" rather than "observed" weather. The ex-post "weather-augmented" forecast is derived by adding the impact of "observed" weather shocks, estimated using climate damage functions $s = 1..S$ from the literature:

$$\hat{y}_{it}^{\tau s} = \hat{y}_{it}^{\tau}(1 + \Delta_{it}^{\tau s}). \tag{2}$$

where $\Delta_{it}^{\tau s}$ is the annual country-specific GDP per capita impact of all weather shocks from year τ to year t, estimated using damage function s.

Typically, damage functions are derived empirically by regressing GDP per capita growth on the set of climate indicators C_{it} after applying the functional transformation f, including other controls and fixed effects:

$$g_{it} = f(\mathbf{C_{it}}) + \gamma_i + \alpha_t + \varepsilon_{it} \tag{3}$$

However, some of the macroeconomic damage functions in the literature that we use in this study are derived by aggregating sectoral damages, the cost of sea level rise, and risks from global climate tipping points, which are not covered by econometric studies (notably, COACCH and DICE 2023).

The annual country-specific GDP per capita impact of all weather shocks from year τ to year t, denoted by $\Delta_{it}^{\tau s}$, is then calculated as:

$$\Delta_{it}^{s} = f(C_{it}) - f(C_{i\tau}) \tag{4}$$

where the functional forms f and its coefficients are exactly as those used by the authors of each study. $C_{i\tau}$ is the "reference" climate in the year of the forecast and follows the definition adopted by each study. Realizations of climatic variables observed in years prior to the forecast year τ are not considered, because they are assumed to be already incorporated into baseline growth expectations and, therefore, into the forecasts. Mathematical expressions for the calculation of Δ_{it}^{s} for each damage function are provided in Appendix B.

By comparing the forecasted GDP per capita level without climate impacts ($\hat{y}_{it}^{\tau}$) and with climate impacts ($\hat{y}_{it}^{\tau s}$) to the actual GDP per capita (y_{it}), we can assess if incorporating climate impacts improves (or deteriorates) the forecasting error. Figure 1 visualizes this approach for an example country (South Africa) and an example damage function (Burke et al., 2015).

Specifically, we define the forecast error for a forecast made in year τ for country i in year t as

$$e_{it}^{\tau} \equiv \frac{\hat{y}_{it}^{\tau} - y_{it}}{y_{i\tau}}. \tag{5}$$

Similarly, we write the error of the climate impact-adjusted forecast as

$$e_{it}^{\tau s} \equiv \frac{\hat{y}_{it}^{\tau s} - y_{it}}{y_{i\tau}}. \tag{6}$$

We normalize forecast errors by GDP per capita in the base year τ, to interpret all our forecast error results as a percentage of the base-year GDP per capita. For instance, $e_{it}^{\tau} = 0.05$ means that the WEO fall edition from year τ overestimated actual GDP per capita in year $t > \tau$ by 5% relative to the base year's GDP per capita. If, for damage function s,

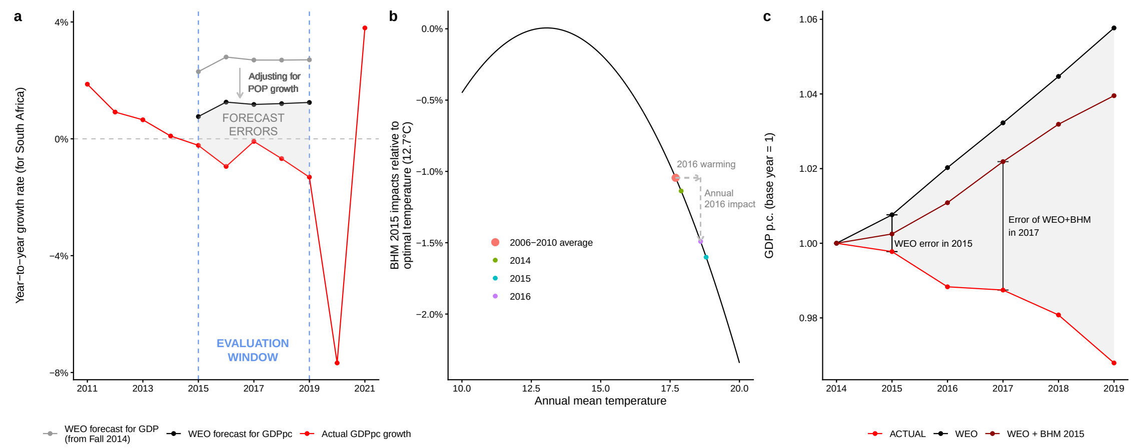


Figure 1: Illustration of the forecast, climate impacts, and construction of the three key time series, here as an example depicted for the "BHM 2015" damage function and South Africa.

$e_{it}^{\tau s} = 0.03$, then incorporating climate impacts reduced the error by three percentage points compared to the baseline forecast. Note that this definition of the forecast errors differs from the simple period-by-period growth rate differences, as it accounts for the cumulative chaining of forecasts over time.[2] The statistical properties of these two indicators and their relationship are key results of this exercise. If adding climate damages systematically reduces the gap between forecast and actual GDP per capita, the estimated impacts have out-of-sample predictive power, suggesting that they partially explain past IMF forecast errors.

3.1 Climate damage function implementations

To estimate the growth effects of climate impacts, we apply empirical and model-based damage functions from recent studies. Specifically, we implement the specifications of nine widely used studies, summarized in Table 2. Three damage functions, one based on the DICE 2023 model, one based on bottom-up impact estimates implemented in a CGE model (COACCH), and one based on a downscaled version of a global meta study by Howard and Sterner (2025) express GDP impacts directly in levels, and are modeled as quadratic functions of GMST. The remaining specifications are estimated using GDP per capita growth rates and incorporate various climatic drivers, including local annual temperature, precipitation, and indices of climate extremes.

[2] This approach allows us to examine whether the reduction in forecast errors from adding damage functions increases over time, as expected if omitted climate impacts are persistent (Nath et al., 2024). We also implement two alternative definitions of the forecast error. First, we compute the error only for the last period (e.g., when $\tau = t - 5$), which evaluates the forecast error in the last year and thus captures the full cumulative effect. Second, we compute the simple forecast error in terms of growth rates, defined as $\tilde{e}_{it}^{\tau s} = \frac{\hat{y}_{it}^{\tau s} - \hat{y}_{it-1}^{\tau s}}{\hat{y}_{it-1}^{\tau s}} - \frac{y_{it} - y_{it-1}}{y_{it-1}}$, which represents the annual forecast error including climate impacts of damage specification s.

We use the original coefficients and specifications and implement all functions on our compiled dataset applied to the quasi counterfactual GDP per capita trajectory $\hat{y}_{it}^{s}$ as "Reference" or no-climate baseline.[3]

Table 2: Climate damage function specifications

Specification	Impact on	Climate drivers	Reference
ABM 2025	growth	Heat (Tx35), drought (PDSI), mean temperature	Akyapı et al. (2025)
BK 2026	growth	GMST	Bilal and Känzig (2026)
BHM 2015	growth	Mean temperature	Burke et al. (2015)
COACCH	level	GMST	van der Wijst et al. (2023)
DICE 2023	level	GMST	Barrage and Nordhaus (2024)
HS 2025	level	GMST	Howard and Sterner (2025); Tol (2024)[4]
KW 2020	growth	Mean temperature	Kalkuhl and Wenz (2020)
KLW 2022	growth	Mean temperature (others omitted)	Kotz et al. (2022)
KMNPRY 2021	growth	Mean temperature	Kahn et al. (2021)

3.2 Forecast error and climate impact evaluation measures

We now evaluate how well climate impacts from each damage function specification s explain forecast errors using several metrics computed across countries i and years t.

Sign matching and directional accuracy. We assess whether climate impacts move forecasts in the "correct" direction — that is, toward the actual GDP per capita — by computing the share of country-year observations where the sign of the climate impact is opposite to the sign of the forecast error. In such cases, adding climate impacts could potentially reduce the WEO forecast error:

$$\text{SignMatch}^{\tau s} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=\tau+1}^{\tau+T} \mathbb{1}\{\text{sgn}(-\Delta_{it}^{\tau s}) = \text{sgn}(e_{it}^{\tau})\} \tag{7}$$

This metric captures only the direction of climate impacts, not their magnitude. A value above 0.5 indicates climate impacts explain the direction of forecast errors better than chance, assuming errors are centered around zero.

Forecast accuracy metrics. To what extent can the overall forecast error be reduced by incorporating climate impacts? To answer this question, we compute three standard

[3] Bilal and Känzig (2026) is included here, as the paper has been accepted at the Quarterly Journal of Economics as of March 2026. We use the regional coefficients from their replication package.

metrics for each specification s: mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE).

$$\text{MAE}^{\tau s} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=\tau+1}^{\tau+T} |e_{it}^{\tau s}| \tag{8}$$

$$\text{MSE}^{\tau s} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=\tau+1}^{\tau+T} (e_{it}^{\tau s})^2 \tag{9}$$

$$\text{RMSE}^{\tau s} = \sqrt{\text{MSE}^{\tau s}} \tag{10}$$

Similarly, we define the performance metrics for the standard forecast without impacts (based on e_{it}^{τ}) as MAE^{τ}, MSE^{τ} and RMSE^{τ}. The improvement in forecast accuracy from adding climate impacts under specification s is calculated as:

$$\begin{aligned} \Delta\text{MAE}^{\tau s} &= \text{MAE}^{\tau} - \text{MAE}^{\tau s} \\ \Delta\text{RMSE}^{\tau s} &= \text{RMSE}^{\tau} - \text{RMSE}^{\tau s} \end{aligned} \tag{11}$$

where positive (negative) values indicate improved (deteriorated) forecast accuracy when climate impacts are included.

Forecast window selection. Limiting our sample to the 2010s to avoid distortions from the financial crisis and the COVID-19 pandemic leaves us with five base years, in which the IMF issued forecasts up to five years out, from $\tau = 2010$ to $\tau = 2014$. To use as much data as possible, we repeat our exercise for each evaluation window separately and then present average performance scores across all five evaluation windows. For instance, the average MAE improvement across the five evaluation windows for a damage specification s is calculated as

$$\Delta\text{MAE}^{s} = \frac{1}{5} \sum_{\tau=2010}^{2014} \Delta\text{MAE}^{\tau s} \tag{12}$$

Overall, we rely primarily on MAE to evaluate damage functions because, for some countries, severe economic contractions that were not anticipated by the WEO generate extreme outliers in forecast errors. When using MSE, these outliers dominate overall performance, which is undesirable given that the largest forecast errors in the sample (e.g., Libya in 2011) are not attributable to climate impacts.

Forecast error decomposition. To understand the sources of forecast improvements, we decompose MAE and MSE into two components: (i) the average error across countries and (ii) the error in country-specific deviations from this average.

We compute the global MAE (and similarly MSE) as:

$$\mathrm{MAE}^{\tau s}_{GLOBAL} = \frac{1}{T}\sum_t \left| \frac{\overline{\hat{y}^s_t} - \overline{y_t}}{\overline{y_\tau}} \right| \tag{13}$$

where $\overline{\hat{y}^s_t} = \frac{1}{N}\sum_i \hat{y}^s_{it}$ and $\overline{y_t} = \frac{1}{N}\sum_i y_{it}$ denote the global averages of forecasted and actual GDP per capita, respectively.

The country-level variation, abstracting from global average GDP per capita growth, is measured by:

$$\mathrm{MAE}^{\tau s}_{COUNTRY} = \frac{1}{TN}\sum_t \sum_i \left| \frac{\widetilde{\hat{y}^s_{it}} - \widetilde{y_{it}}}{\widetilde{y_{i\tau}}} \right|. \tag{14}$$

where $\widetilde{\hat{y}^s_{it}} = \hat{y}^s_{it} - \overline{\hat{y}^s_t}$ and $\widetilde{y_{it}} = y_{it} - \overline{y_t}$ represent country-level deviations from the global average.

This approach allows us to decompose forecast errors into to components: (i) the error in global average per capita GDP growth and (ii) variation in country-level forecasts with climate impacts, thereby isolating regional heterogeneity in forecasts and climate effects. By construction, these components sum to the total error for MSE but not for MAE. Nevertheless, this decomposition provides insight into whether forecast errors and climate impacts stem primarily from global growth trends or from regional heterogeneity and, as shown below, yields qualitatively similar results for both MAE and MSE. As before, we compute improvements in forecast errors relative to the raw WEO forecast without climate impacts:

$$\begin{aligned} \Delta\mathrm{MAE}^{\tau s}_{GLOBAL} &= \mathrm{MAE}^{\tau}_{GLOBAL} - \mathrm{MAE}^{\tau s}_{GLOBAL} \\ \Delta\mathrm{MAE}^{\tau s}_{COUNTRY} &= \mathrm{MAE}^{\tau}_{COUNTRY} - \mathrm{MAE}^{\tau s}_{COUNTRY} \end{aligned} \tag{15}$$

We use this global average and country heterogeneity decomposition to show the source of the error and the potential improvement of the different climate impact specifications.

4 A placebo test for economic climate impact regressions

Before evaluating damage functions, we first assess if regressing WEO's GDP per capita forecasts, instead of actual GDP per capita growth, on climate variables yields similar findings regarding climate impacts or null effects. This placebo test addresses a key identification concern in climate econometric studies: if the estimated temperature-growth relationship were driven by omitted country-specific trends (e.g., institutional changes, sectoral transformations) that correlate with temperature, then IMF forecasts—which incorporate these economic fundamentals—should partially capture these trends and hence these misidentified climate impacts. By contrast, if estimated climate effects vanish when using forecasts as the dependent variable, this supports two conclusions: first, the IMF forecasts do not systematically incorporate climate projections, validating their use as climate-naïve counterfactual; second, the identifying variation in climate-growth regressions stems indeed from deviations from expected growth paths, not from predictable country-specific trajectories that happen to correlate with temperature.

To see this point, consider the true data-generating process for GDP per capita growth:

$$g_{it} = f(C_{it}) + g(X_{it}) + \gamma_i + \alpha_t + \varepsilon_{it} \tag{16}$$

where C_{it} is the vector of climate variables, X_{it} represents other economic determinants, γ_i are country fixed effects, and α_t are time fixed effects. A typical approach to estimate this relationship in the climate econometrics literature would be given by equation 3 above, effectively omitting X_{it}. If C_{it} and X_{it} are correlated (i.e., $\mathrm{Cov}(C_{it}, X_{it}) \neq 0$)—for example, if countries with higher temperatures systematically differ in their institutional development—then omitting X_{it} leads to biased estimates of the effect of climate.

As argued above, the IMF forecast $\hat{g}_{it}^{\tau}$ incorporates country-specific and global trends, but does not so far systematically account for climate projections. Under ideal circumstances of predicting all non-climate factors correctly, we can therefore represent the forecast as:

$$\hat{g}_{it}^{\tau} = g(X_{it}) + \gamma_i + \alpha_t + \eta_{it} \tag{17}$$

The key insight is that the forecast error $g_{it} - \hat{g}_{it}^{\tau}$ isolates the climate shock:

$$g_{it} - \hat{g}_{it}^{\tau} = f(C_{it}) + (\varepsilon_{it} - \eta_{it}) \tag{18}$$

This difference no longer contains $g(X_{it})$. The IMF's forecasting model effectively serves as a sophisticated control function that captures complex relationships between economic determinants without requiring explicit specification in the climate impact regression. We can now perform a placebo test by regressing the IMF forecast itself on climate variables:

$$\hat{g}_{it}^{\tau} = f(C_{it}) + \gamma_i + \alpha_t + \nu_{it} \tag{19}$$

If the forecasts are truly climate-naïve, we expect no significant relationship between C_{it} and $\hat{g}^{\tau}_{it}$ after controlling for fixed effects, except for short-term forecasts where climate impacts may already be manifesting in macroeconomic variables available to forecasters. Finding a significant relationship would indicate that either the forecasts incorporate climate information, or a systematic correlation between climate and economic variables not fully captured by fixed effects in the model, which could indicate an identification problem in the climate econometrics model.

Table 3 presents results from estimating the Burke et al. (2015) specification using actual growth and IMF forecasts at different horizons. While the temperature-growth relationship is statistically significant in actual data (Column 1), it is not present when forecasted growth is used (Columns 2-6).

Dependent Variables:	Actual GDP growth	year-0	year-1	year-2	year-3	year-4	year-5
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
Tmean	0.0133***	0.0084**	-0.0033	0.0003	-0.0020	-0.0030**	-0.0007
	(0.0041)	(0.0032)	(0.0025)	(0.0017)	(0.0023)	(0.0014)	(0.0016)
$Tmean^2$	-0.0005***	-0.0002**	0.0001	3.23×10^{-5}	2.56×10^{-5}	7.87×10^{-5}	9.46×10^{-6}
	(0.0001)	(0.0001)	(8.22×10^{-5})	(6.16×10^{-5})	(7.19×10^{-5})	(5.05×10^{-5})	(6.1×10^{-5})
Pt	2.15×10^{-6}	-5.3×10^{-6}	-3.45×10^{-6}	-3.92×10^{-6}	-3.47×10^{-6}	4.67×10^{-6}	4.8×10^{-6}
	(6.98×10^{-6})	(5.56×10^{-6})	(3.88×10^{-6})	(3.32×10^{-6})	(3.44×10^{-6})	(5.55×10^{-6})	(6.42×10^{-6})
Pt^2	-6.47×10^{-10}	8.28×10^{-10}	1.88×10^{-10}	-1.29×10^{-10}	3.17×10^{-10}	-5.38×10^{-10}	-6.18×10^{-10}
	(1.46×10^{-9})	(1.05×10^{-9})	(7.24×10^{-10})	(7.91×10^{-10})	(8.08×10^{-10})	(8.72×10^{-10})	(8.11×10^{-10})
Fixed-effects							
Country	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Varying Slopes							
Year (squared) (Country)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics							
Observations	6,351	6,024	5,846	5,665	5,478	5,288	5,095
R^2	0.29408	0.31039	0.34661	0.40060	0.47310	0.43114	0.42003
Within R^2	0.00300	0.00116	0.00088	0.00211	0.00094	0.00101	0.00057

Clustered (Country) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3: Regression Results: Real GDP Growth and Forecasts, Burke et al. (2015) specification.

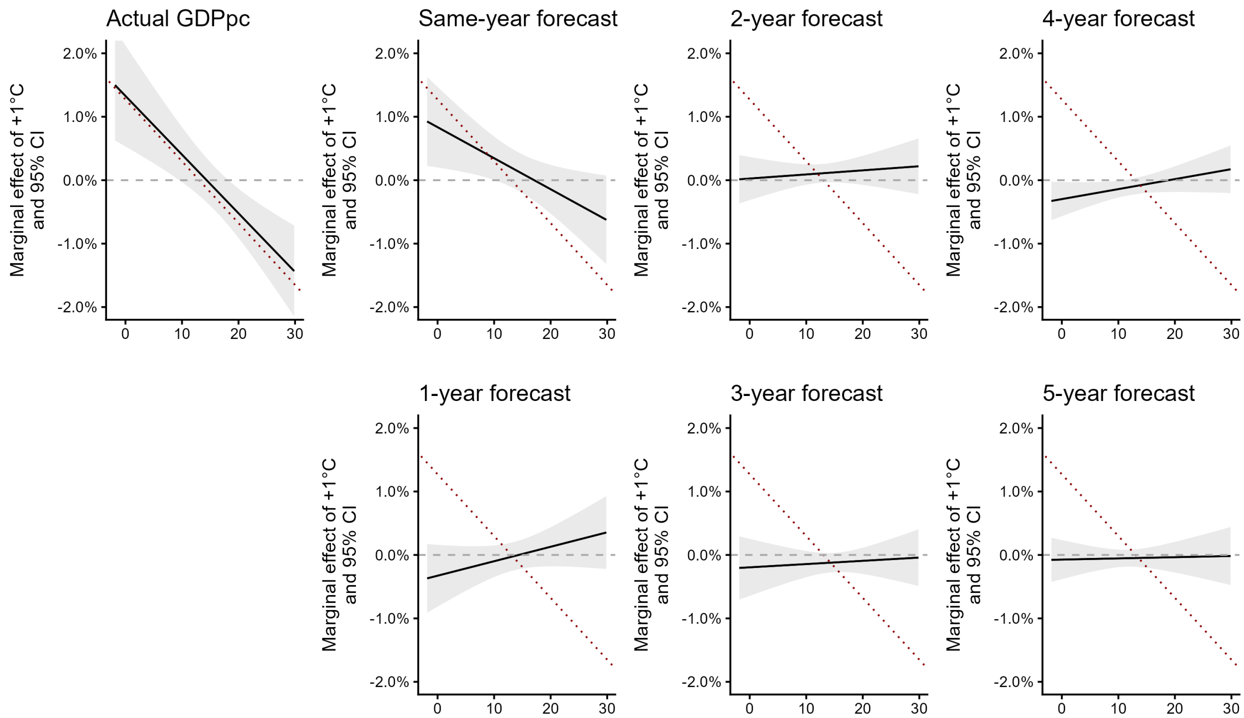


Figure 2: Marginal effects of temperature on growth for observed outcomes and IMF forecasts at different horizons under the BHM 2015 damage function. The dotted line represents the central estimate from the original paper.

Figure 2 visualizes this pattern, showing the marginal effect of a one-degree temperature increase. Actual growth exhibits the characteristic inverted-U relationship with temperature documented by Burke et al. (2015). This effect diminishes already substantially in contemporaneous forecasts (the Fall WEO for the ongoing year ("0-year")) and vanishes entirely for one-year-ahead and longer horizons.

We conduct the same placebo test for one of the key regression models estimated by Kahn et al. (2021) with similar results, as illustrated in Figure 3. For temperature deviations below ("dtemp_minus") and above ("dtemp_plus") a country's historical norm, using actual GDP per capita (GDPpc) growth or the WEO's same-year forecast (i.e., made in the fall of the year in question) produces point estimates similar to the original paper.[5] When using other forecasts, the point estimates shift toward zero and are typically not significantly different from zero. Table 4 presents the results of the key coefficients for this specification.

This validation has two important implications. First, it confirms that IMF forecasts—despite their sophistication in capturing economic fundamentals—do not systematically ac-

[5]Differences in the exact results are likely driven by the fact that our sample composition is not identical to that of Kahn et al. (2021), as well as potential differences between the actual GDPpc growth rates in the WEO data and the original paper's data sources.

Dependent Variables:	Actual GDP growth	year-0	year-1	year-2	year-3	year-4	year-5
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
Temp. increase	-1.0986***	-0.8247	-0.2281	-0.0991	0.0472	0.0950	-0.3243**
	(0.3454)	(0.5532)	(0.1683)	(0.1437)	(0.1805)	(0.1517)	(0.1478)
Temp. decrease	-1.3635**	-1.1368	-0.5263	-0.3107	0.3557	0.2202	-0.4419
	(0.6450)	(1.1619)	(0.3805)	(0.3597)	(0.3908)	(0.3192)	(0.3989)
Fixed-effects							
Country	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics							
Observations	3,398	3,122	3,084	2,816	2,772	2,502	2,452
R^2	0.44354	0.34030	0.32553	0.35008	0.27510	0.30466	0.28779

Standard errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 4: Regression Results: Real GDP Growth and Forecasts, Kahn et al. (2021) specification.

count for climate shocks, making them appropriate baseline counterfactual for our damage function evaluation. Second, it supports the notion that the temperature-growth relationships identified in the empirical literature reflect genuine climate impacts rather than omitted variable bias from correlated country-specific trends.

5 Comparison of impact specifications using forecast errors

We now evaluate the nine different impact functions and compare their respective improvements in MAE (and RMSE) of the baseline forecast over different evaluation windows. Our main results demonstrate that climate damage functions reduce forecast errors more than a random shock (modeled as a Bernoulli distribution of a $+/-1\%$ shock with $p = 0.5$) in annual growth rates. Figure 4 summarizes performance across all five evaluation windows (2011-2015 through 2015-2019), with MAE improvements for a specific window/base year ($\Delta MAE^{\tau s}$) and average MAE improvements across all five windows (ΔMAE^{s}) shown as triangles and dots, respectively. Most climate damage functions reduce the IMF forecast errors by 0.1–0.4 percentage points, while one specification (BK 2026) increases the error considerably. To put these magnitudes in perspective, average annual climate impacts over the five-year evaluation window amount to approximately 1–3% of the average annual growth rate of 1.7% over the full sample (based on Figure S2). We also present additional benchmarks. First, we introduce an alternative 'debiased' forecast that removes the average WEO

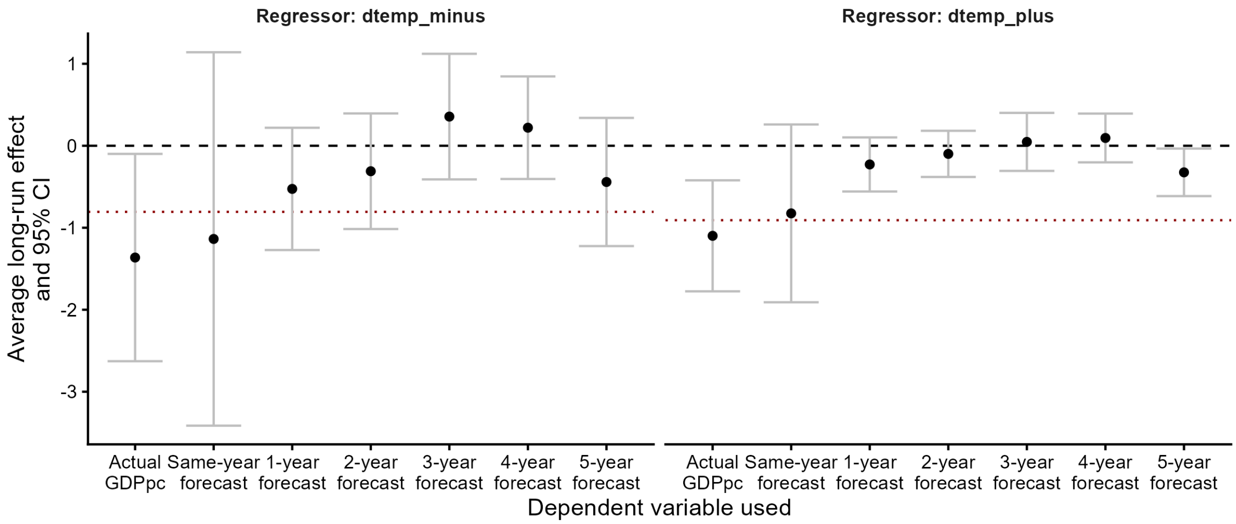


Figure 3: Marginal effects of temperature on GDP growth for observed outcomes and IMF forecasts at different horizons under the KMNPRY (2021) damage function, using a 30-year moving-average specification for historical norms and excluding precipitation controls (Table 1, specification 2, HPJ-FE model in the original paper). The dotted line represents the central estimate from the original paper.

forecast error of −0.88 pp (Table 1). Correcting for this bias yields an improvement of 0.56 pp. Second, we introduce a 'perfect' forecast that matches actual outcomes; such a forecast would improve the baseline forecast without climate impacts by 6.25 pp on average. Thus, most damage functions achieve roughly 1–6% of the gain that would be attainable under a perfect forecast.

Interestingly, the share of country-year observations in which MAE improves (rather than deteriorates) is highest for damage functions that, by design, predict negative impacts for all countries, either with larger magnitudes (HS 2025, COACCH) or smaller magnitudes (DICE 2023, KMNPRY 2021). By contrast, the sign of climate impacts implied by KW 2020 and subsequent work (KLW 2022), as well as by BHM 2015, can vary from year to year; however, this greater flexibility in tracking climate impacts does not translate into a higher share of country-years with MAE improvements.

Since we performed the analysis over multiple five-year windows, we can also compare performance across windows, as shown in Figure 5. Several key patterns emerge. Only four damage functions (COACCH, DICE 2023, HS 2025, KMNPRY 2021) would have improved WEO forecast performance over the 2011-2015 period. By contrast, for the later periods 2014-2018 and 2015-2019, improvements are more widespread across damage functions and larger in magnitude. The HS 2025 meta-study with country-level downscaling and the COACCH damage function, a bottom-up model that is region-specific but driven by global mean surface temperature, show the most consistent performance across evaluation

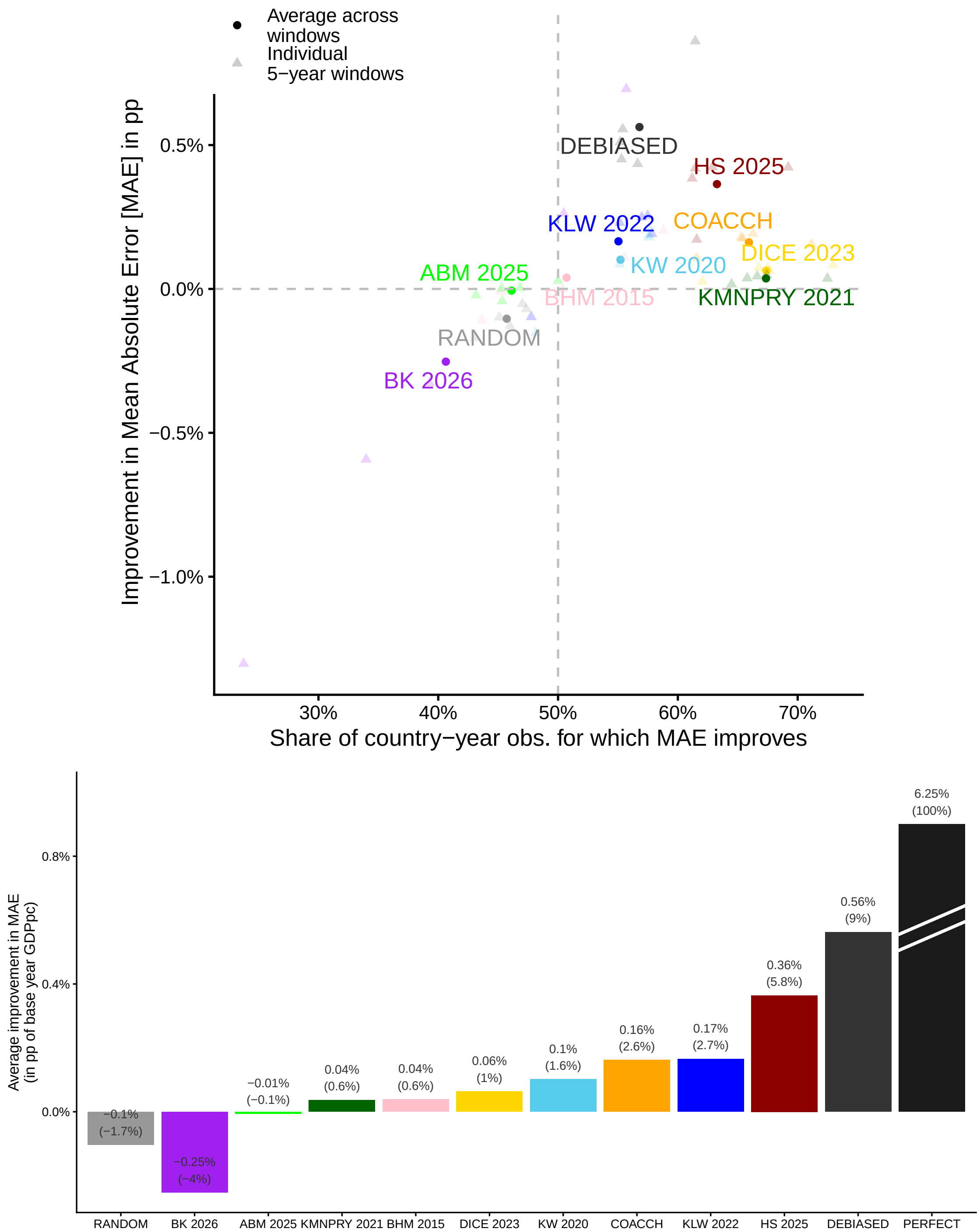


Figure 4: The top panel plots, on the vertical axis, MAE performance of each damage function. MAEs are expressed as a percentage of base-year GDP per capita, and improvement in MAE is measured in percentage points. The horizontal axis shows the share of country-year observations for which MAE improves. The bottom panel reports the average improvement in MAE by damage function, with values in parentheses expressed as a percentage of the average forecast error, which amounts to 6.25 percent of base-year GDP per capita. A "PERFECT" correction would eliminate 100 percent of the forecast error.

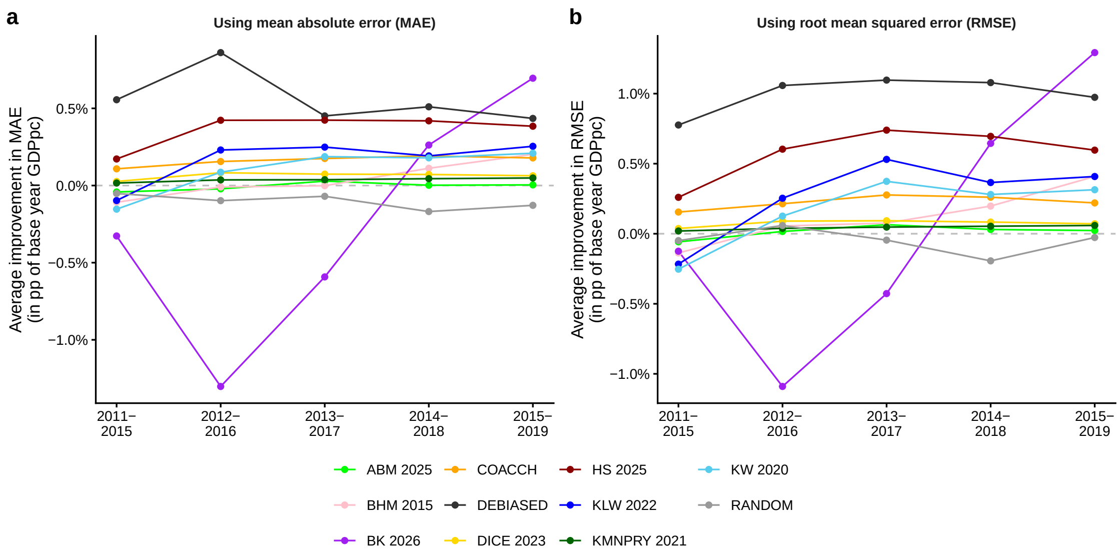


Figure 5: Comparison of MAE- and RMSE-based performance over the different time horizons.

windows. Econometric models, by contrast, show more variable performance, performing well in some periods but poorly in others. A clear outlier is the BK 2026 specification, which deteriorates forecast performance in three out of five windows. Overall, we confirm that RMSE-based performance tends to be more favorable than MAE-based performance to damage functions that imply larger global GDP losses.

Our five-year moving window is designed to be short enough to exclude periods with major economic shocks, yet long enough to capture persistent climate shocks and their cumulative effects on growth. In addition, we compute the improvement in MAE for each individual year from 2011 until 2019. We also assess whether model performance differs when evaluating the cumulative five-year forecast error rather than the average five-year error used in our main results. The results are presented in Figure S2 in the Appendix. We find an average annual WEO GDP forecast MAE of 2.5 pp between 2011 and 2019, which is close to the average error of 2.8% observed over the full sample from 1990 to 2023. Over each five-year horizon, however, this error accumulates to an average of approximately 10.5% of base-year GDP. When focusing on annual growth rate errors, KW 2020 and KLW 2022 perform notably worse, while the remaining models display broadly similar improvements. For the cumulative five-year forecast error, the improvement from incorporating climate damages is broadly consistent with the average results discussed above, with the exception of BK 2026, which now exhibits an improvement.

We find variation in performance across countries (Figures S3 and S4 in the Appendix). High-impact damage functions (KLW 2022, BK 2026, BHM 2015) show more dispersed performance, significantly reducing MAE for some countries while substantially increasing

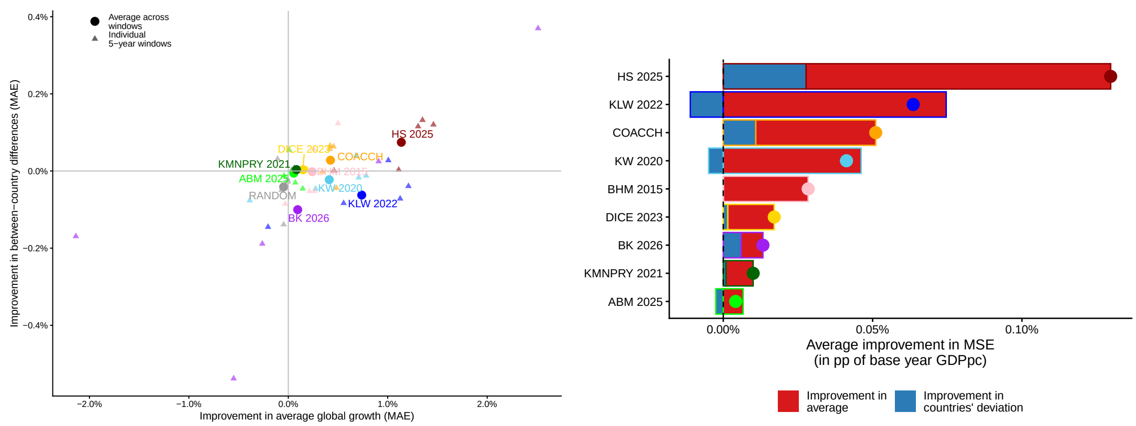


Figure 6: Performance of forecast improvement at global average and country-level deviation, based on the mean-squared error (MSE) (left panel) and aggregation at average with decomposition (right panel).

it for others. In contrast, smaller-impact functions (DICE, ABM, KMNPRY) improve MAE more modestly but for a larger share of countries.

Given that WEO forecasts are, on average, overly optimistic during the 2010s, this raises the question of whether the forecast improvements from incorporating climate impacts lead to lower forecast error across the board or instead arise from highly heterogeneous improvements across countries. Figure 6 depicts the decomposition of MSE and MAE into the two components defined in Section 3.2: improvements in the average forecast error across countries and improvements in explaining between-country differences in forecast errors. The results show that forecast improvements are driven primarily by reductions in the average error across countries, rather than by improved country-specific performance. Indeed, several damage functions reduce forecast performance along the latter dimension, an effect that is more than offset by improvements in average forecast errors across countries. Overall, for all damage functions, forecast improvements are driven largely by the global component rather than by country-specific outcomes.

Finally, some years and specifications stand out by showing negative improvements in the global growth dimension, meaning that they worsen forecasts of average global growth, even while potentially improving cross-country differentiation. This pattern, combined with the deteriorating MAE performance documented earlier, raises concerns, at least about the near-term predictions of these functions.

6 A country-level reality check for damage function predictions

Beyond aggregate performance metrics, our approach provides a valuable reality check on damage function predictions using country-level evidence. Figure 7 illustrates GDP per capita projections for 2015-2019 based on the 2014 IMF WEO forecasts, with and without climate impacts, compared with observed GDP per-capita growth over the same period. Over this period, annual global mean surface temperature increased from approximately 0.9°C above pre-industrial levels in 2014 to 1.1°C in 2019. Damage functions that predict substantially lower growth than observed raise questions about their validity for short-term climate risks assessments, as they can be reconciled with data only by assuming much higher counterfactual growth than implied by professional forecasts, which are typically positively biased.

For some industrialized economies, such as Canada or Japan, as well as middle-income and low-income countries like Egypt, Eritrea and Madagascar, the BK 2026 damage function predicts significant reductions in GDP per capita that are at odds with observed growth patterns during this period. The implied economic damage would have been highly salient and widely discussed had it occurred—representing contractions or stagnation in countries that actually experienced positive growth, in some cases even exceeding forecasts. This discrepancy suggests that the function may overstate near-term climate impacts and should therefore be interpreted with caution.

Notably, this inconsistency does not apply to the BHM 2015 damage function, which predicts smaller annual impacts but more pronounced accumulation over longer time horizons. This contrast highlights that the functional form of damage accumulation—whether impacts are predominantly contemporaneous or persistent—matters substantially for validation against observed outcomes. Interestingly, for several countries including Nigeria, Mexico, Brazil, and Russia, actual economic growth was considerably lower than predicted even by the high-impact functions, although this likely reflects non-climate factors such as the 2014-2016 commodity price collapse and country-specific political-economic challenges.

An interesting pattern emerges when comparing damage functions that use country-level annual temperature variations (particularly KW 2020, KLW 2022) with those that use global temperatures indicators (COACCH, DICE 2023) or abstract from year-to-year temperature volatility when projecting impacts (KMNPRY 2021). In some cases, particularly in high-income countries (e.g., Germany, France and Japan), functions based on annual temperature fluctuations generate growth projections that are considerably more volatile than the observed trajectory of GDP per capita.

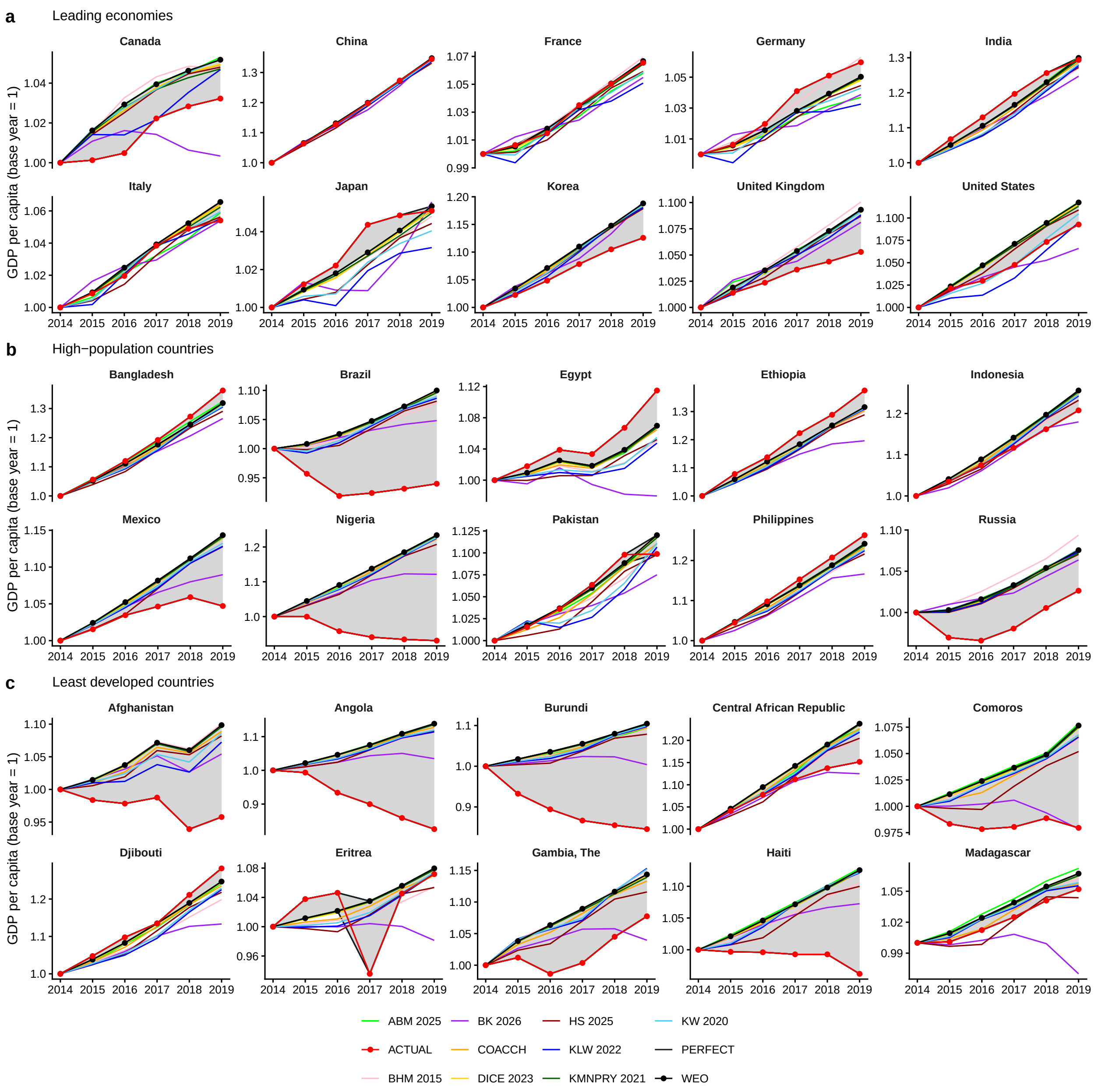


Figure 7: Projection and climate impacts for selected example countries. GDP per capita in 2014 is normalized to 1 in all cases.

7 Reconstructing climate impacts for Europe using a weighted damage function

While some meta-studies aim to calibrate meta damage functions using statistical averages of selected damage functions (Howard and Sterner, 2017, 2025), we instead use performance metrics to construct alternative weighted averages of individual damage functions. Specifically, we build a performance-weighted ensemble of damage functions based on their average forecast improvement in MAE relative to a baseline model. For each damage function s, we compute the average forecast improvement across the five validation windows (2010–2014):

$$R_s = \frac{1}{5} \sum_{\tau=2010}^{2014} \frac{\text{MAE}^{\tau}_{\text{baseline}} - \text{MAE}^{\tau}_{s}}{\text{MAE}^{\tau}_{\text{baseline}}}. \tag{20}$$

All functions with $R_s \leq 0$ (i.e., worse than the baseline) are excluded from the ensemble. We also exclude functions that show a clear deterioration in the country-deviation dimension (see Figure 6).[6] For the remaining functions $S^* = \{s : R_s > 0\}$, we define weights based solely on their relative forecast improvement:

$$w_s = \frac{R_s}{\sum_{j \in S^*} R_j} \tag{21}$$

This yields the following weights for the retained damage specifications ($R_s > 0$):

Name	HS 2025	COACCH	DICE 2023	BHM 2015	KMNPRY 2021
Weight (w_s)	0.55	0.24	0.10	0.06	0.06

Table 5: Performance-based weights for damage function ensemble. Only specifications that improve forecast accuracy over the baseline ($R_s > 0$) and do not show clear signs of deterioration in the country-deviation dimension are included.

Using these weights, we extend the forecast horizon from 2010 to 2023, the most recent year covered by our climate data source (Gortan et al., 2024). Because forecasts are available only for five years at a time, we construct a counterfactual series by chaining successive forecasts over the full horizon. Specifically, we use the annual one-year-ahead growth forecast (i.e., forecast of τ for year $\tau + 1$) to compute a GDP per capita series $\left\{\hat{y}_{it}^{\tau=t-1}\right\}$, obtained by chaining growth forecasts as $\left\{\hat{y}_{it}^{\tau=t-1}\right\} = y_{i0} \prod_{t=1}^{T}(1 + \hat{g}_{it}^{\tau=t-1})$. This approach

[6] For this screening criterion, we use the MSE rather than the MAE (i.e., results from the right panel of Figure 6) because, unlike the MAE, the two disaggregated components sum to the total score improvement for the MSE, making it more internally consistent.

generates a GDP per capita projection under the assumption that, in each year, the one-year-ahead forecast is realized instead of observed growth. Although forecast errors accumulate over time and forecasts are issued at different times, the resulting series provides a quasi-counterfactual path of economic growth that does not condition on actual realizations. By chaining these forecasts, we construct a counterfactual GDP trajectory through 2023. We use this projected GDP per capita series as the reference path to which we apply the estimated climate-impact factors, and we report the implied climate impacts. Combining the individual damage function estimates with the performance-weighted ensemble estimate for 2023, we focus on Europe to assess the cumulative impact of observed warming from 2010 to 2023 on European GDP

$$\Delta_{\mathrm{EU}}^{2023} = \sum_{s \in S^*} w_s \cdot \Delta_s^{2023} = -1.0\%. \tag{22}$$

This approach provides a more robust estimate of current climate impacts for Europe, and allows comparison with bottom-up studies such as the European Climate Risk Assessment (EUCRA) or similar work (Usman et al., 2025), thereby facilitating a comparison between sectoral model-based estimates and climate-econometric macroeconomic estimates. The reported uncertainty range reflects only part of the structural uncertainty arising from differences in functional forms and estimation approaches in the climate-economy literature. In particular, these intervals capture cross-model variation but do not incorporate parameter estimation uncertainty within individual damage functions. A comprehensive uncertainty assessment that considers parametric uncertainty jointly with model uncertainty is beyond the scope of this paper and is therefore left for future research.

The country-level estimates (Figure 8) reveal substantial heterogeneity across EU27 member states (in the Appendix, Figure S6 shows all damage functions for these countries, while Figure S5 shows the weighted average based on total MAE improvement). The geographic distribution of impacts shows a pronounced north-south gradient, with Southern European countries experiencing the most negative impacts and Northern European countries much smaller losses. This pattern reflects both higher baseline temperatures in Southern Europe and slower temperature increases in Northern Europe over the 2010-2023 period. The large positive impacts estimated for Northern countries under the BHM 2015 specification are at odds with results from other damage functions, although this specification receives a lower weight due to its MAE-based performance.

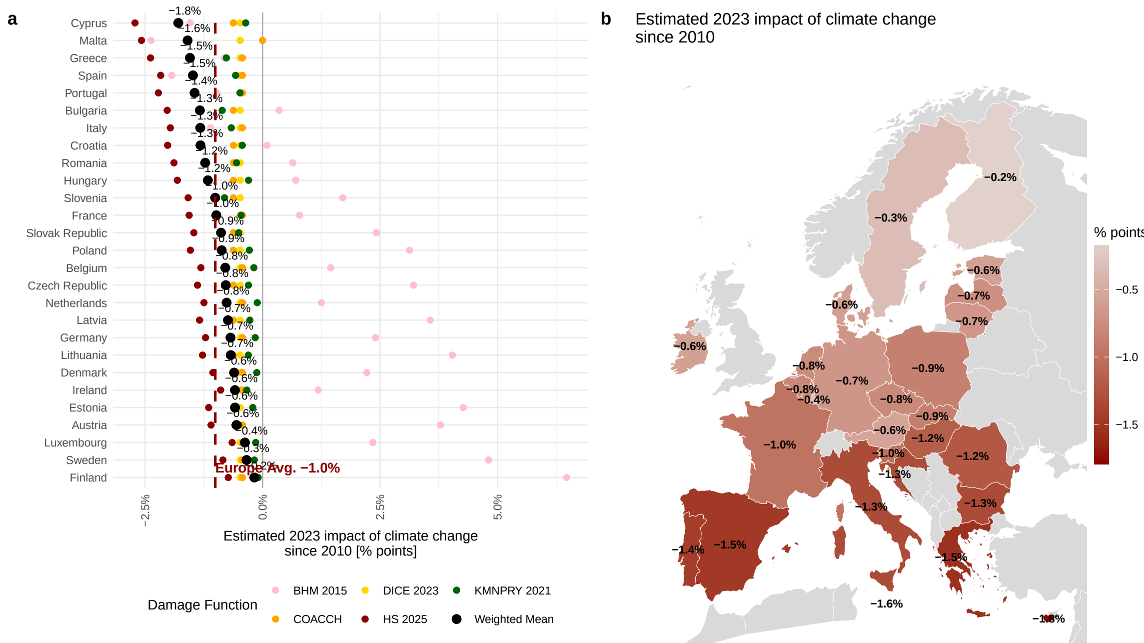


Figure 8: Weighted ensemble estimates of climate change impacts on GDP per capita from 2010 to 2023. Panel (a) shows country-level estimates by individual damage function (colored points) and weighted mean values (black dots) for EU27 countries. Panel (b) displays the geographic distribution of impacts, with darker shades of red indicating larger negative impacts. Grey shading indicates non-EU countries.

8 Discussion

Our results offer several important insights for the climate economics literature and for the quantification of the short-term economic impacts from climate change.

First, our methodology offers a novel approach to validating climate damage functions that complements existing methods. Unlike traditional cross-validation approaches that struggle with the low explanatory power of climate-growth regressions, our approach leverages professional forecasts to provide a more realistic assessment of damage function performance in explaining observed economic growth. Crucially, WEO growth forecasts account for neither future weather shocks nor their expected impacts, as confirmed by our empirical tests. This supports our use of these forecasts as counterfactual growth trajectories and allows us to evaluate damage functions while circumventing concerns about omitted drivers of growth correlated with climate variables and the resulting omitted variable bias. Nevertheless, WEO forecasts have other limitations, most notably an optimistic bias. Annual cross-country averages of forecast errors are positive in most years (Artis, 1988; Celasun et al., 2021). Importantly, this bias dates back to the 1970s, before the world started to experience a marked increase in global temperatures. Therefore, evidence based on annual correlations between average forecast errors across countries and average climate impacts across countries should be interpreted with caution. We cannot rule out that the systematic negative impacts of global warming predicted by some damage functions coincide, by chance, with the systematic optimism of WEO forecasts. On the contrary, evidence based on the correlation between country-specific forecast errors and country-specific climate impacts relies on many countries' time series and should be more robust.[7]

Second, our findings suggest important differences among some widely used and discussed damage functions based on their consistency with observed economic outcomes. Functions that predict more gradual, cumulative impacts (such as Kahn et al. (2021) and van der Wijst et al. (2023)) generate GDP losses that are more consistent with observed data than functions predicting large responses to year-to-year climate shocks. This has important implications for long-term climate impact projections. While our method cannot be used to assess the general validity of competing damage functions, it suggests that those predicting gradually increasing damages may be more safely used to improve macroeconomic projections.

Third, our results raise important questions about the validity of some implications of the most severe damage function estimates in the literature. In particular, damage functions such as that of Bilal and Känzig (2026), which imply very high SCC exceeding $1,000 per ton of CO_2, predict economic contractions in several economies during 2014-2019 that are, in many cases, difficult to reconcile with observed outcomes. This discrepancy suggests these estimates may overstate climate impacts, at least in the short term. Our results do not imply these damage functions are necessarily incorrect about long-term climate impacts—small biases in annual sensitivity can compound substantially over decades. Rather, our findings suggest their estimated short-term temperature sensitivity may be upwardly biased, with implications both for near-term impact projections and for the credibility of long-term extrapolations based on these parameters.

However, our analysis also has important limitations. First and most importantly, we cannot definitively decompose forecast errors into climate-related and non-climate components—other factors may coincidentally generate errors with the same sign as predicted climate impacts. However, the magnitude of the observed improvements aligns reasonably well with ex-ante expectations given the modest warming during the evaluation period. Nevertheless, we emphasize that our approach evaluates whether damage functions can help explain forecast errors, not whether they perfectly isolate causal climate effects from all other sources of error.

[7] When evaluating damage functions, the country-specific results remove the global average forecast error, implying that remaining country-level average forecast errors are equal to zero on average across countries. Therefore, a damage function that mostly predicts negative impacts might perform better for the countries with positive forecast error averages but will underperform in all the other countries. On balance, this damage function does not benefit anymore from predicting systematic negative impacts.

Second, our validation period (2011-2019) covers relatively limited warming (approximately 0.2°C globally), so our results may not generalize to longer horizons over which temperatures rise substantially. However, this limitation arguably strengthens rather than weakens the policy relevance of our findings. Our analysis validates precisely the near-term impacts most relevant for current policy decisions and investment planning. Damage functions that fail to predict plausible impacts during a period of modest warming raise concerns about their reliability for near-term projections, even if their long-term sensitivities might be correct. For long-term impact assessment, our approach therefore complements, rather than replaces, other validation methods such as paleoclimate evidence or process-based modeling.

Third, our approach cannot directly validate the functional form governing long-term damage accumulation, which drives much of the variation in impact and social cost of carbon estimates across studies. While we observe differences between functions predicting temporary versus persistent impacts, we cannot fully adjudicate between competing hypotheses about how damages accumulate over multi-decade timescales. This remains an important area for future research (Piontek et al., 2019; Nath et al., 2024).

Fourth, by excluding crisis years (2008-2010, 2020-2023), we avoid contamination from extreme non-climate shocks, but we may also miss periods in which climate impacts might be more salient or forecast errors larger. This conservative choice strengthens the validity of our climate-naïve counterfactual assumption, but it may understate the potential for climate impacts to explain forecast errors during more volatile periods.

Fifth, a critical limitation of existing damage functions is that most econometric estimates capture only short-run, localized responses to weather shocks rather than economy-wide adaptation to persistent climate change, which we also cannot address in this work. Recent empirical work documents substantial heterogeneity in how economies adapt across timescales: immediate responses to weather shocks, such as labor reallocation and input substitution, differ markedly from longer-run adaptations involving capital stock adjustments, technological adoption, and sectoral reallocation (Lemoine et al., 2025; Carleton et al., 2022; Burke and Emerick, 2016).

Finally, our analysis focuses on impact on GDP growth and does not include other dimensions of climate damage, notably non-market and non-use values. These include impacts on health, terrestrial and maritime ecosystems, direct effects on human well-being, among others. Our findings about GDP impacts should therefore be interpreted within the broader context of multidimensional climate risks.

9 Conclusion

This paper introduces a novel approach to validating climate damage functions by leveraging professional economic forecasts as a climate-naïve counterfactual. Our methodology provides both backward-looking validation of econometric identification strategies and a forward-looking assessment of damage function performance in explaining observed economic outcomes. Our key findings demonstrate that climate damage functions can explain meaningful but modest portions of IMF forecast errors. The best-performing functions reduce mean absolute errors by 0.1–0.4 percentage points, representing approximately 1-6% of baseline forecast error, or at most 3% of average annual growth. This indicates that climate variables contain genuine information about economic outcomes beyond what professional forecasters capture using conventional economic indicators, while also highlighting that these effects remain modest relative to total year-to-year economic variation during periods of moderate warming. These results should not be extrapolated far into the future, as climate impacts may intensify if greenhouse gas concentrations continue to rise unabated.

However, substantial heterogeneity exists across damage functions. Bottom-up models such as COACCH, as well as specifications based on smoothed climate variation, show more consistent performance than many functions estimated using annual climate fluctuations. Some damage functions perform well primarily because they predict gradual GDP losses across countries that coincidentally offset the WEO's systematic positive bias. In contrast, a smaller subset of functions is also able to reduce the cross-country dispersion of forecast errors. Most importantly, damage functions that generate very high social costs of carbon imply economic contractions in many economies during 2014-2019 that are difficult to reconcile with the counterfactual scenarios produced by professional forecasters, raising concerns about their suitability for near-term impact analysis.

For policymakers, our results validate concerns about climate-economy linkages and support the integration of climate factors into economic forecasting. At the same time, they counsel caution in relying on the most extreme damage estimates for near-term policy analysis. Damage functions predicting social costs above $1,000 per ton appear inconsistent, at least, with observed near-term impacts for several countries. More moderate damage functions that show greater consistency with observed data may therefore provide more reliable foundations for policy analysis, particularly over near- to medium-term planning horizons.

Despite its limitations—including modest warming during the validation period and an inability to fully adjudicate long-term damage accumulation—our approach represents an important step toward more rigorous validation of climate damage functions. By comparing predicted impacts with observed outcomes using professional forecasts as benchmarks, we provide new evidence on which specifications are most consistent with historical experience, thereby complementing other validation methods in the ongoing effort to quantify the economic impacts of climate change.

References

Akyapı, Berkay, Matthieu Bellon, and Emanuele Massetti, "Estimating Macrofiscal Effects of Climate Shocks from Billions of Geospatial Weather Observations," *American Economic Journal: Macroeconomics*, July 2025, *17* (3), 114–159.

Aligishiev, Zamid, Emanuele Massetti, and Matthieu Bellon, *Macro-fiscal implications of adaptation to climate change*, International Monetary Fund, 2022.

Artis, Michael J, "How Accurate Is the World Economic Outlook? A Post Mortem on Short-Term Forecasting at the International Monetary Fund," in "World Econonomic and Financial Surveys, 1988," International Monetary Fund, 1988.

Barrage, Lint and William Nordhaus, "Policies, projections, and the social cost of carbon: Results from the DICE-2023 model," *Proceedings of the National Academy of Sciences of the United States of America*, 2024, *121* (13), e2312030121.

Bilal, Adrien and Diego Känzig, "Does Unilateral Decarbonization Pay For Itself?," Technical Report, National Bureau of Economic Research, Cambridge, MA 2025.

_ and Diego R Känzig, "The Macroeconomic Impact of Climate Change: Global Versus Local Temperature*," *The Quarterly Journal of Economics*, February 2026, p. qjag011.

Burke, Marshall and Kyle Emerick, "Adaptation to Climate Change: Evidence from US Agriculture," *American Economic Journal: Economic Policy*, August 2016, *8* (3), 106–140.

_ , Solomon M. Hsiang, and Edward Miguel, "Global non-linear effect of temperature on economic production," *Nature*, November 2015, *527* (7577), 235–239.

Callahan, Christopher W. and Justin S. Mankin, "Globally unequal effect of extreme heat on economic growth," *Science Advances*, 2022, *8* (43), eadd3726.

Carleton, Tamma, Amir Jina, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, Robert E Kopp, Kelly E McCusker, Ishan Nath, James Rising, Ashwin Rode, Hee Kwon Seo, Arvid Vi-

aene, Jiacan Yuan, and Alice Tianbo Zhang, "Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits*," *The Quarterly Journal of Economics*, April 2022, p. qjac020.

Celasun, Oya, Jungjin Lee, Mico Mrkaic, and Allan Timmermann, "An evaluation of world economic outlook growth forecasts, 2004–17," Technical Report 2021/216, International Monetary Fund 2021.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken, "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, July 2012, *4* (3), 66–95.

Desbordes, Rodolphe and Markus Eberhardt, "Climate change and economic prosperity: Evidence from a flexible damage function," *Journal of Environmental Economics and Management*, 2024, *125*, 102974.

Dreher, Axel, Silvia Marchesi, and James Raymond Vreeland, "The political economy of IMF forecasts," *Public Choice*, October 2008, *137* (1), 145–171.

Drouet, Laurent, Valentina Bosetti, Simone A. Padoan, Lara Aleluia Reis, Christoph Bertram, Francesco Dalla Longa, Jacques Després, Johannes Emmerling, Florian Fosse, Kostas Fragkiadakis, Stefan Frank, Oliver Fricko, Shinichiro Fujimori, Mathijs Harmsen, Volker Krey, Ken Oshiro, Larissa P. Nogueira, Leonidas Paroussos, Franziska Piontek, Keywan Riahi, Pedro R. R. Rochedo, Roberto Schaeffer, Jun'ya Takakura, Kaj-Ivar van der Wijst, Bob van der Zwaan, Detlef van Vuuren, Zoi Vrontisi, Matthias Weitzel, Behnam Zakeri, and Massimo Tavoni, "Net zero-emission pathways reduce the physical and economic risks of climate change," *Nature Climate Change*, November 2021, *11* (12), 1070–1076.

Eicher, Theo S. and Yuan Gao Rollinson, "The accuracy of IMF crises nowcasts," *International Journal of Forecasting*, January 2023, *39* (1), 431–449.

Emmerling, Johannes, Pietro Andreoni, and Massimo Tavoni, "Global inequality consequences of climate policies when accounting for avoided climate impacts," *Cell Reports Sustainability*, January 2024, *1* (1), 100008. Publisher: Elsevier.

Gatti, Roberta, Daniel Lederman, Asif M. Islam, Ha Nguyen, Rana Lotfi, and Mennatallah Emam Mousa, "Data transparency and GDP growth forecast errors," *Journal of International Money and Finance*, February 2024, *140*, 102991.

Genberg, Hans, Andrew Martinez, and Michael Salemi, "The imf/weo forecast process," *Independent Evaluation Office of the International Monetary Fund Background Paper No. BP/14/03, Washington, DC.*, 2014.

_ , _ , **Carlos de Resende, Francesco Luna, and Franz Loyola**, "IMF Forecasts: Process, Quality, and Country Perspectives," *Independent Evaluation Office of the International Monetary Fund Evaluation Report, Washington, DC.*, 2014.

Gortan, Marco, Lorenzo Testa, Giorgio Fagiolo, and Francesco Lamperti, "A unified dataset for pre-processed climate indicators weighted by gridded economic activity," *Scientific data*, 2024, *11* (1), 533.

Hamilton, James D., "Why You Should Never Use the Hodrick-Prescott Filter," *The Review of Economics and Statistics*, December 2018, *100* (5), 831–843.

Howard, Peter H. and Thomas Sterner, "Few and Not So Far Between: A Meta-analysis of Climate Damage Estimates," *Environmental and Resource Economics*, September 2017, *68* (1), 197–225.

_ **and** _ , "Methodology Matters: A Careful Meta-Analysis of Climate Damages," *Environmental and Resource Economics*, 2025.

IMF, "Small States' Resilience to Natural Disasters and Climate Change–Role for the IMF," 2016.

IPCC, *Climate Change 2022: Impacts, Adaptation and Vulnerability* Summary for Policymakers, Cambridge, UK and New York, USA: Cambridge University Press, 2022. Type: Book.

Kahn, Matthew E., Kamiar Mohaddes, Ryan N. C. Ng, M. Hashem Pesaran, Mehdi Raissi, and Jui-Chung Yang, "Long-term macroeconomic effects of climate change: A cross-country analysis," *Energy Economics*, December 2021, *104*, 105624.

Kalkuhl, Matthias and Leonie Wenz, "The impact of climate conditions on economic production. Evidence from a global panel of regions," *Journal of Environmental Economics and Management*, September 2020, *103*, 102360.

Kotz, Maximilian, Anders Levermann, and Leonie Wenz, "The effect of rainfall changes on economic production," *Nature*, January 2022, *601* (7892), 223–227. Number: 7892 Publisher: Nature Publishing Group.

_ , _ , and _ , "The economic commitment of climate change," *Nature*, April 2024, *628* (8008), 551–557. Publisher: Nature Publishing Group.

_ , Leonie Wenz, Annika Stechemesser, Matthias Kalkuhl, and Anders Levermann, "Day-to-day temperature variability reduces economic growth," *Nature Climate Change*, April 2021, *11* (4), 319–325. Number: 4 Publisher: Nature Publishing Group.

Lemoine, Derek, Catherine Hausman, and Jeffrey G. Shrader, "A Guide to Climate Damages," Technical Report w34348, National Bureau of Economic Research October 2025.

Mohaddes, Kamiar and Mehdi Raissi, "Rising temperatures, melting incomes: Country-specific macroeconomic effects of climate scenarios," *PLOS Climate*, 2025, *4* (9), e0000621.

Morris, Jennifer, Steven K. Rose, John Reilly, Angelo Gurgel, Sergey Paltsev, and C. Adam Schlosser, "Reconciling widely varying estimates of the global economic impacts from climate change," *Nature Climate Change*, February 2025, *15* (2), 124–127. Publisher: Nature Publishing Group.

Nath, Ishan B., Valerie A. Ramey, and Peter J. Klenow, "How Much Will Global Warming Cool Global Growth?," July 2024.

Newell, Richard G., Brian C. Prest, and Steven E. Sexton, "The GDP-Temperature relationship: Implications for climate change damages," *Journal of Environmental Economics and Management*, March 2021, p. 102445.

OECD, "OECD global long-run economic scenarios: 2025 update," Technical Report 36, Paris 2025.

Piontek, Franziska, Laurent Drouet, Johannes Emmerling, Tom Kompas, Aurélie Méjean, Christian Otto, James Rising, Bjoern Soergel, Nicolas Taconet, and Massimo Tavoni, "Integrated perspective on translating biophysical to economic impacts of climate change," *Nature Climate Change*, June 2021.

_ , Matthias Kalkuhl, Elmar Kriegler, Anselm Schultes, Marian Leimbach, Ottmar Edenhofer, and Nico Bauer, "Economic Growth Effects of Alternative Climate Change Impact Channels in Economic Modeling," *Environmental and Resource Economics*, August 2019, *73* (4), 1357–1385.

Ranger, Nicola Ann, Olivier Mahul, and Irene Monasterolo, "Assessing financial risks from physical climate shocks," Technical Report 168725, The World Bank, Washington, D.C 2022.

Rennert, Kevin, Frank Errickson, Brian C. Prest, Lisa Rennels, Richard G. Newell, William Pizer, Cora Kingdon, Jordan Wingenroth, Roger Cooke, Bryan Parthum, David Smith, Kevin Cromar, Delavane Diaz, Frances C. Moore, Ulrich K. Müller, Richard J. Plevin, Adrian E. Raftery, Hana Ševčíková, Hannah Sheets, James H. Stock, Tammy Tan, Mark Watson, Tony E. Wong, and David Anthoff, "Comprehensive Evidence Implies a Higher Social Cost of CO2," *Nature*, September 2022, pp. 1–3. Publisher: Nature Publishing Group.

Ricke, Katharine, Laurent Drouet, Ken Caldeira, and Massimo Tavoni, "Country-level social cost of carbon," *Nature Climate Change*, October 2018, *8* (10), 895–900.

Rohde, Robert A. and Zeke Hausfather, "The Berkeley Earth Land/Ocean Temperature Record," *Earth System Science Data*, 2020, *12* (4), 3469–3479.

Takakura, Jun'ya, Shinichiro Fujimori, Naota Hanasaki, Tomoko Hasegawa, Yukiko Hirabayashi, Yasushi Honda, Toshichika Iizumi, Naoko Kumano, Chan Park, Zhihong Shen, Kiyoshi Takahashi, Makoto Tamura, Masahiro Tanoue, Koujiro Tsuchida, Hiromune Yokoki, Qian Zhou, Taikan Oki, and Yasuaki Hijioka, "Dependence of economic impacts of climate change on anthropogenically directed pathways," *Nature Climate Change*, October 2019, *9* (10), 737–741.

Tol, Richard S. J., "Social cost of carbon estimates have increased over time," *Nature Climate Change*, June 2023, *13* (6), 532–536. Publisher: Nature Publishing Group.

Tol, Richard S.J., "A meta-analysis of the total economic impact of climate change," *Energy Policy*, 2024, *185*, 113922.

Usman, Sehrish, Guzmán González-Torres Fernández, and Miles Parker, "Going NUTS: The regional impact of extreme climate events over the medium term," *European Economic Review*, September 2025, *178*, 105081.

van der Wijst, Kaj-Ivar, "MIMOSA: Integrated Assessment Model for Cost-Benefit Analysis: GitHub repository. `https://github.com/UtrechtUniversity/mimosa/`," 2024.

van der Wijst, Kaj-Ivar, Francesco Bosello, Shouro Dasgupta, Laurent Drouet, Johannes Emmerling, Andries Hof, Marian Leimbach, Ramiro Parrado,

Franziska Piontek, Gabriele Standardi, and Detlef van Vuuren, "New damage curves and multimodel analysis suggest lower optimal temperature," *Nature Climate Change*, March 2023.

Waidelich, Paul, Fulden Batibeniz, James Rising, Jarmo S. Kikstra, and Sonia I. Seneviratne, "Climate damage projections beyond annual temperature," *Nature Climate Change*, June 2024, *14* (6), 592–599. Publisher: Nature Publishing Group.

Appendix A Additional results

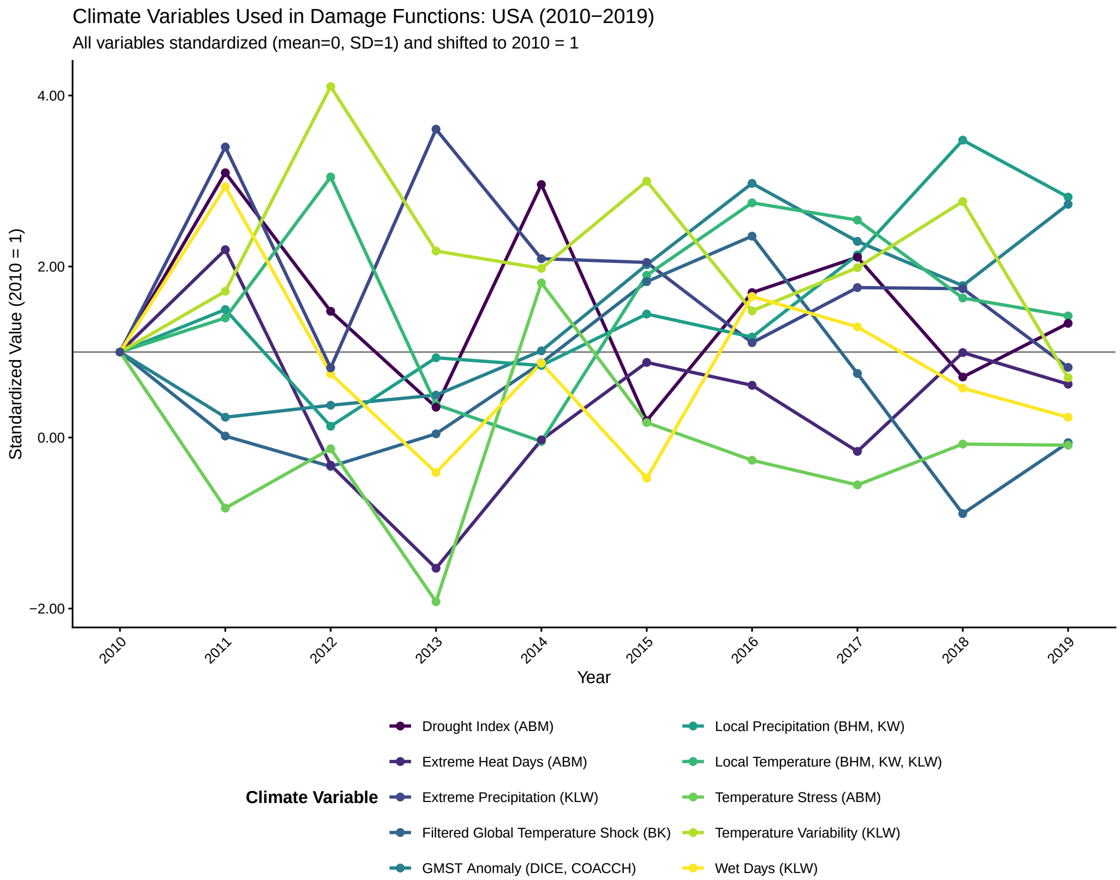
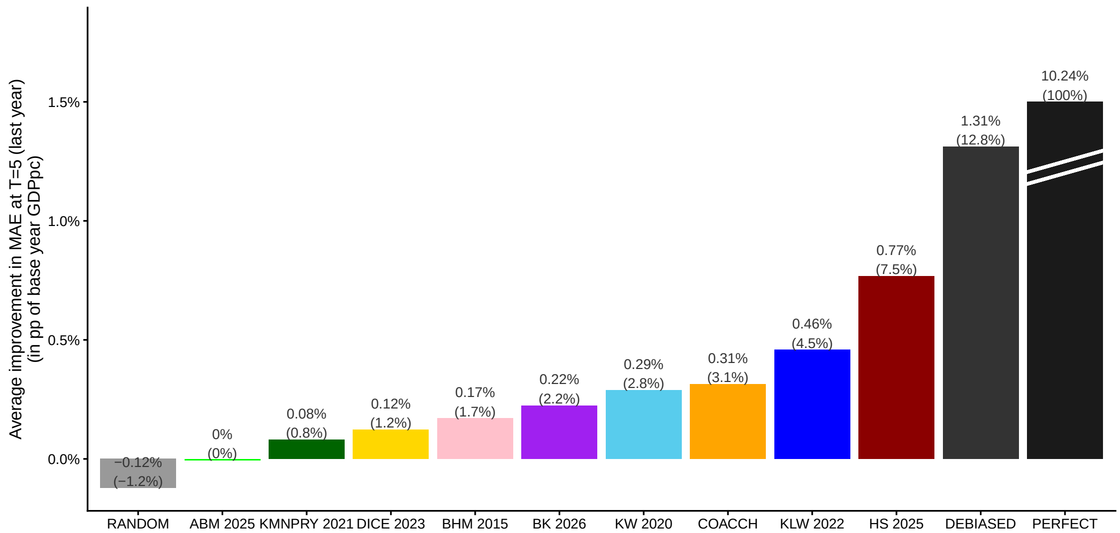


Figure S1: Climate Variables Used in Damage Functions for the United States (2010-2019)
Notes: This figure shows the evolution of all climate variables used by different damage functions in our analysis for the USA from 2010 to 2019. All variables are normalized to 2010 = 1 to enable comparison across different scales and units. Local annual temperature and precipitation are used by BHM 2015, KW 2020, KLW 2022 and KMNPRY 2021. Global temperature is used by BK 2026. Global mean surface temperature (GMST) anomaly relative to 1850-1900 is used by DICE 2023 and COACCH. The ABM 2025 variables include drought index (PDSI < -4), temperature stress (9-12 months), and extreme heat days (days > 35°C).



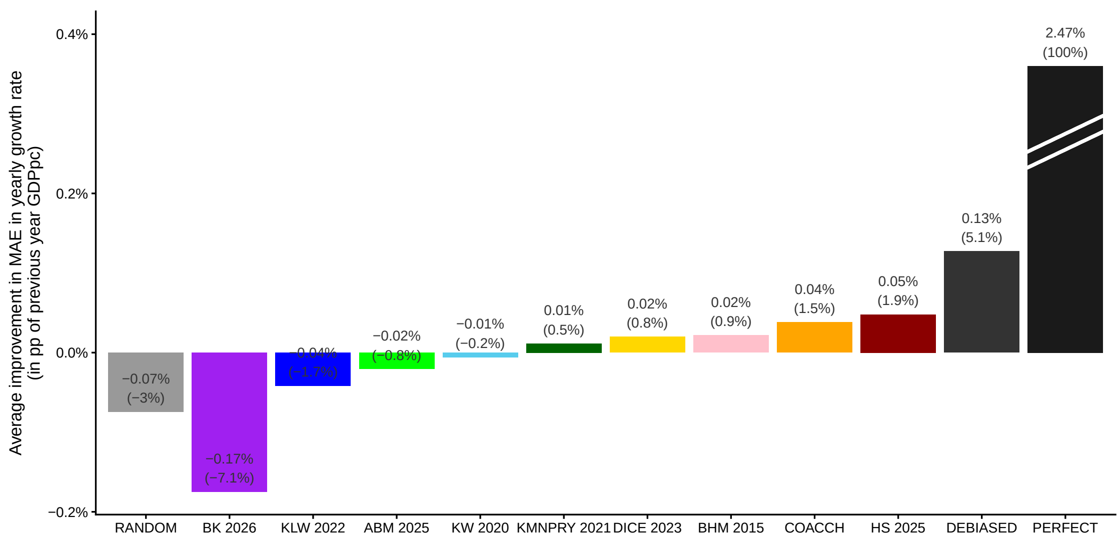


Figure S2: Aggregated performance by climate damage function based on different definitions of the forecast error. The top panel shows the evaluation if only the last time frame (T=5) is used to assess the cumulative effect over five years. The bottom panel on the other hand shows the averaged improvement in year-by-year growth rates, thereby isolating the annual and non-cumulative effect.

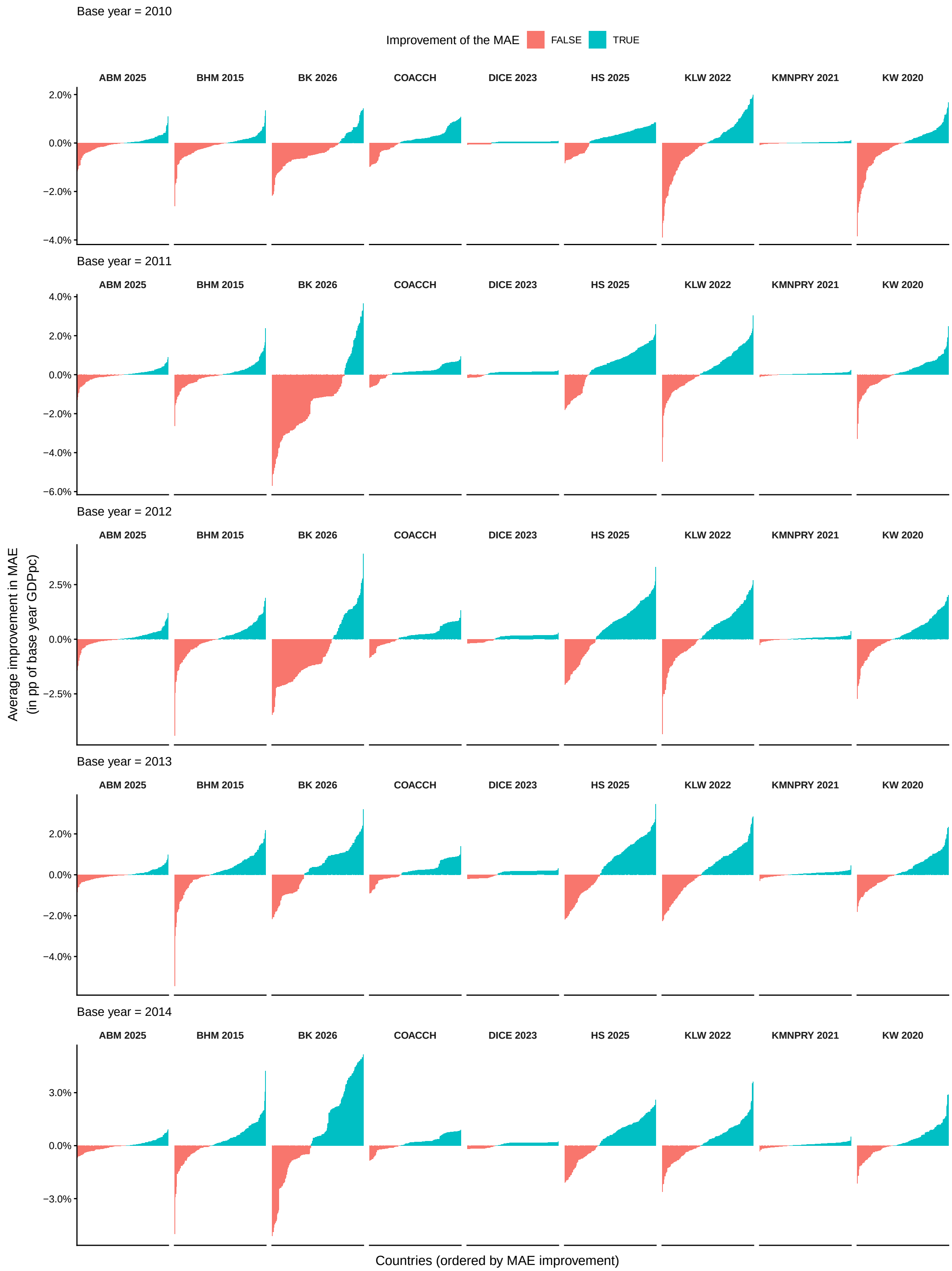


Figure S3: Distribution of MAE improvements by damage function and base year by country. The improvement in MAE in % is ordered from the worst to the best improvement of the country level MAE.

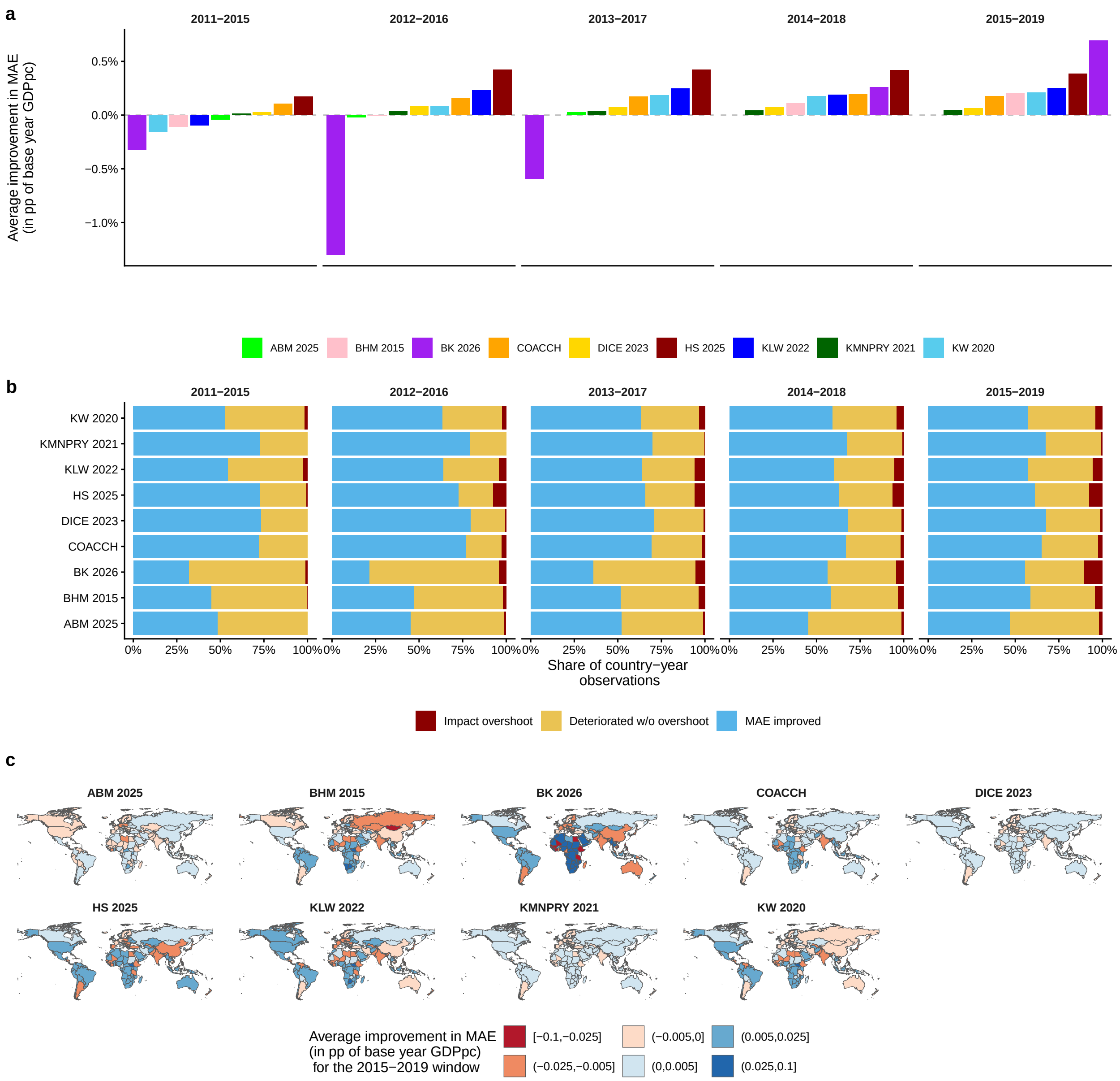


Figure S4: Performance by evaluation window and country

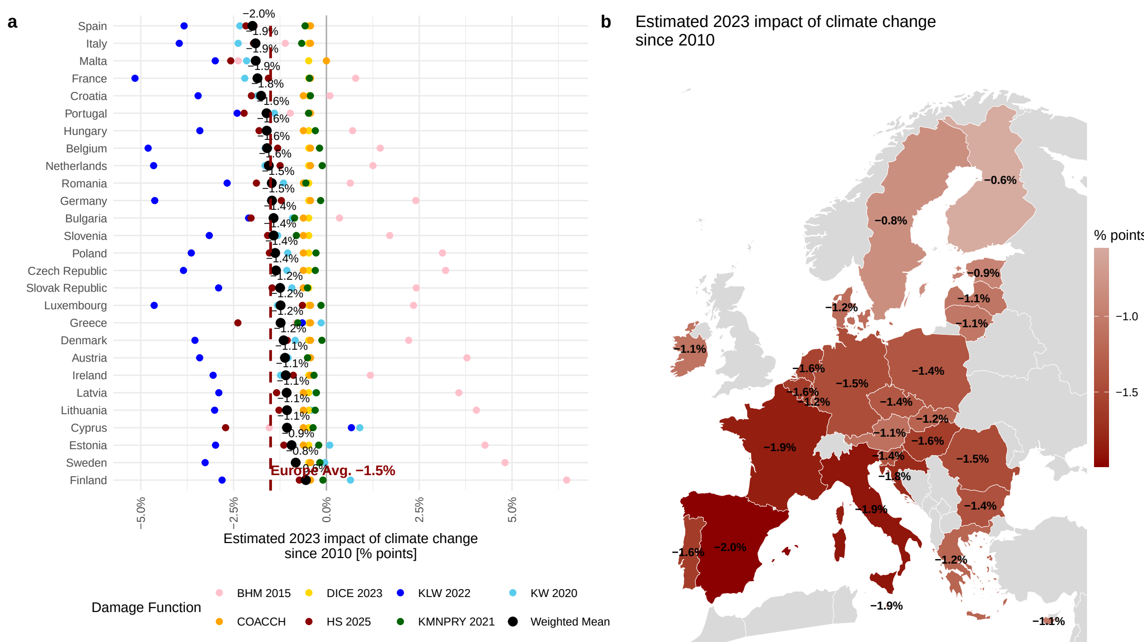


Figure S5: Weighted ensemble estimates of climate change impacts on GDP per capita in 2023, here **weighted based on the total MAE improvement**. Panel (a) shows country-level estimates by individual damage function (colored points) and weighted mean values (black dots) for EU27 countries. Panel (b) displays the geographic distribution of impacts, with shades of red indicating negative impacts and shades of blue indicating positive impacts.

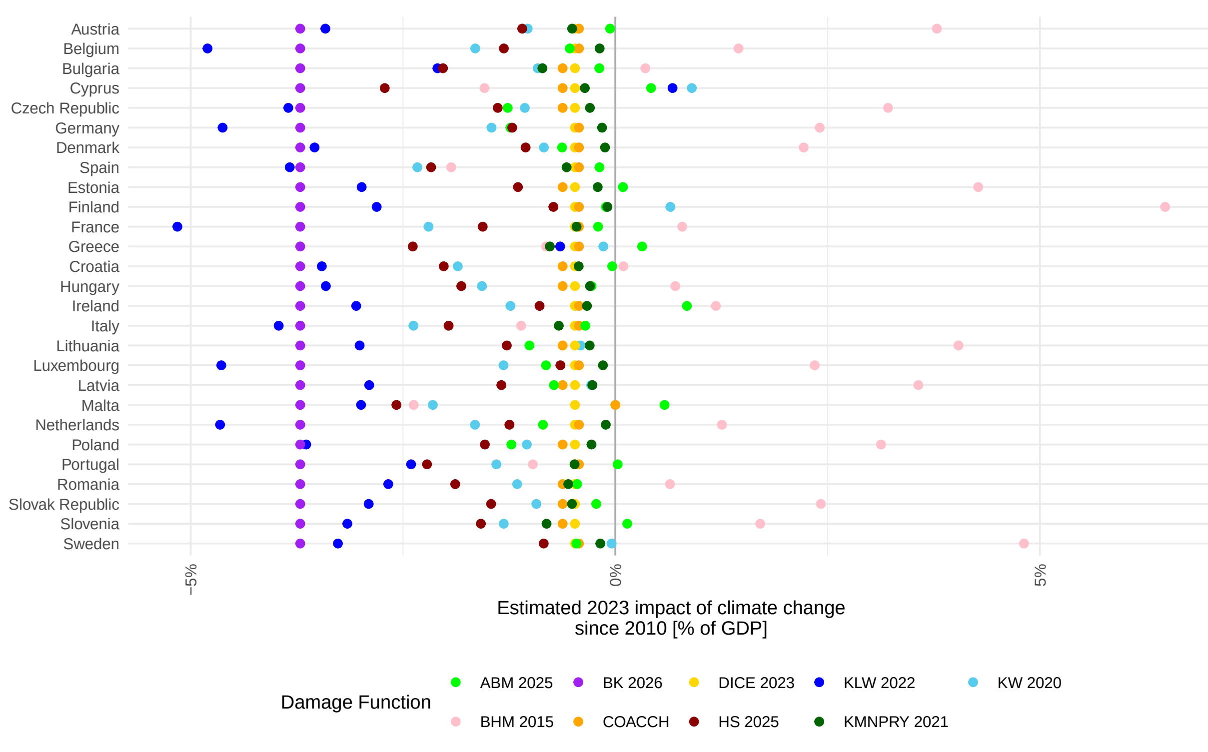


Figure S6: Full ensemble estimates of climate change impacts on GDP per capita in 2023 in Europe

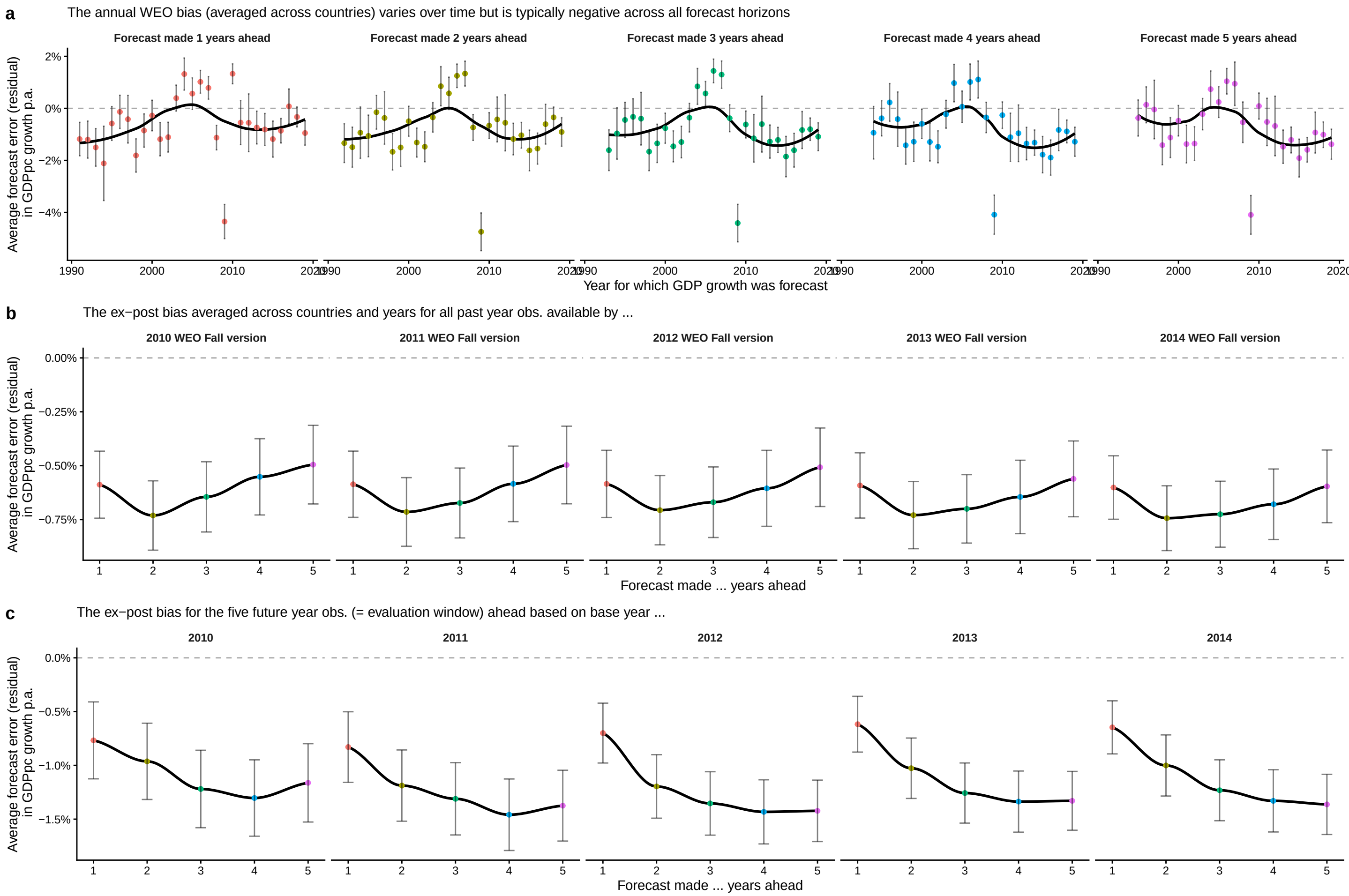


Figure S7: WEO forecast error by time horizon and window

Appendix B Additional information on implementing damage functions

B.1 BHM

BHM estimate the following regression model:

$$ln(y_{it}) - ln(y_{i,t-1}) = \underbrace{\beta_1 T_{it} + \beta_2 T_{it}^2}_{\text{Used for impact projections}} + \beta_3 P_{it} + \beta_4 P_{it}^2 + \gamma_i + \alpha_t + \theta_{i,1} t + \theta_{i,2} t^2 + \epsilon_{it} \quad (23)$$

Their projection approach calculates the impact on annual growth (in logarithmic scale) as the difference in the annual temperature dose-response function relative to average mean temperature over 1980–2010:

$$\delta_{it}^{BHM} = \left(\beta_1 T_{it} + \beta_2 T_{it}^2\right) - \left(\beta_1 T_{i,1980-2010} + \beta_2 T_{i,1980-2010}^2\right) \quad (24)$$

where $\beta_1 = 0.012718353$ and $\beta_2 = -0.00048709$ (see their Extended Data Table 1, col. 1). Adopting this approach would mean that our impact estimates also include the warming that occurred in the 1990s and 2000s. To avoid this, we instead use the 2006–2010 average mean temperature as a baseline such that:

$$\delta_{it}^{BHM} = \left(\beta_1 T_{it} + \beta_2 T_{it}^2\right) - \left(\beta_1 T_{i,2006-2010} + \beta_2 T_{i,2006-2010}^2\right) \quad (25)$$

Here, the choice of baseline aims to be as close as possible to the 2010 start year of our projections while averaging across multiple years to smooth out year-to-year variations.

Note that BHM project damages as follows: $y_{t-1} \times (1 + g_t + \delta_t)$. This implicitly assumes that growth impacts are in % of GDPpc, not in logarithmic scale (which is approximately identical for small impacts but can diverge for more substantial impacts). Therefore, we calculate the cumulative impact since the base year τ as a logarithmic change following Kalkuhl and Wenz (2020):

$$\Delta_{it}^{BHM,ln} = \sum_{s=\tau+1}^{t} \delta_{is}^{BHM} \quad (26)$$

Lastly, as this cumulative impact is still expressed in logarithmic changes, we convert to % of GDPpc:

$$\Delta_{it}^{BHM} = exp(\Delta_{it}^{BHM,ln}) - 1 \quad (27)$$

B.2 COACCH

COACCH estimates climate change-induced GDP losses through region-specific, quadratic reduced-form damage functions that depend on the shift in annual global mean surface temperature vis-à-vis the 1986–2005 average ($\Delta T_{global,t,1986-2005}$):

$$\Delta_{it}^{COACCH} = a_r \left(b_{r,1} \Delta T_{global,t,1986-2005} + b_{r,2} \Delta T^2_{global,t,1986-2005} \right) \tag{28}$$

for each of the 26 regions r in the IMAGE integrated assessment model.[8] Here, a_r is a parameter estimated via quantile regression to capture region-specific different percentiles of damages. For the central damage specification, we use $a_r = 1 \; \forall \; r$.

We map IMAGE regions to the countries in our WEO sample using the matching files from van der Wijst (2024). Then, we use the same region-specific damage function parameters $b_{r,1}$ and $b_{r,2}$ from `https://zenodo.org/records/5546264` for all countries in the respective IMAGE region. Notably, COACCH provides damage function estimates with and without sea level rise. For comparability with the other damage functions, which exclude sea level rise, we use the damage function parameters that do not include sea level rise impacts.

Note that COACCH damage functions return impacts in % of GDP. Therefore, no conversion from logarithmic scale to % of GDP is required. We assume that GDP impacts (as returned by the damage function) are equivalent to GDPpc impacts, which is equivalent to assuming that annual climate change-driven impacts on population sizes are negligible.

Importantly, our methodology only considers impacts that materialize after the year τ in which the respective WEO forecast is made. Therefore, we only consider impacts due to increases in global mean surface temperature that occurred after year τ:

$$\begin{aligned} \Delta_{it}^{COACCH} = & \left(b_{r,1} \Delta T_{global,t,1986-2005} + b_{r,2} \Delta T^2_{global,t,1986-2005} \right) - \\ & \left(b_{r,1} \Delta T_{global,\tau,1986-2005} + b_{r,2} \Delta T^2_{global,\tau,1986-2005} \right) \end{aligned} \tag{29}$$

B.3 KW 2020

KW 2020 estimate the following ADM1-level fixed effects regression model for the specification that ultimately informs their GDP impact projections (see their Table 4, column 5), using the first difference of log-transformed gross regional income per capita as a dependent variable:

[8] Note that COACCH damage functions are also available for other aggregated regions of the REMIND and WITCH model. We use the IMAGE resolution as it has the largest number of regions, allowing for a higher granularity of impacts.

$$
\begin{aligned}
ln(y_{it}) - ln(y_{i,t-1}) = & \underbrace{\beta_1 \Delta T_{it} + \beta_2 \Delta T_{it} T_{i,t-1} + \beta_3 \Delta T_{i,t-1} + \beta_4 \Delta T_{i,t-1} T_{i,t-1}}_{\text{Used for impact projections}} + \beta_5 T_{i,t-1} \\
& \beta_6 \Delta P_{it} + \beta_7 \Delta P_{it} P_{i,t-1} + \beta_8 \Delta P_{i,t-1} + \beta_9 \Delta P_{i,t-1} P_{i,t-1} + \beta_{10} P_{i,t-1} + \\
& \gamma_i + \alpha_t + \theta_{i,1} t + \theta_{i,2} t^2 + \epsilon_{it}
\end{aligned} \tag{30}
$$

where T_{it} and P_{it} denote the ADM1-level annual mean temperature (in $°C$) and annual precipitation (in mm), respectively. Δ denotes the first difference of the respective climate variable. $\theta_{i,1}$ and $\theta_{i,2}$ denote region-specific quadratic time trends.

KW 2020 only use the four terms involving first differences of annual mean temperature in their projections. Following their methodology, we estimate annual impacts due to year-to-year temperature shifts in country i and year t as follows:

$$
\delta_{it}^{KW} = \beta_1 \Delta T_{it} + \beta_2 \Delta T_{it} T_{i,t-1} + \beta_3 \Delta T_{i,t-1} + \beta_4 \Delta T_{i,t-1} T_{i,t-1} \tag{31}
$$

where $\beta_1 = 0.00641$, $\beta_2 = -0.00109$, $\beta_3 = 0.00345$ and $\beta_4 = -0.000718$. Notably, these coefficients are applied directly at the country level. As Emmerling et al. (2024) illustrate, this simplification is valid (and increases the consistency with the other, mostly country-level damage functions), as implementing KW 2020 at ADM1 and country level yield similar results.

These level effects of year-to-year shifts are added up across years, as specified in the Supplementary Information of KW 2020. As for BHM, we convert logarithmic changes to % of gross regional product per capita

$$
\Delta_{it}^{KW} = exp(\sum_{s=\tau+1}^{t} \delta_{is}^{KW}) - 1 \tag{32}
$$

B.4 KLW 2022

KLW 2022 estimate a similar ADM1-level fixed effects regression model, using a slightly modified version of the KW 2020 model that adds additional terms for day-to-day temperature variability, extreme daily rainfall, monthly precipitation deviations, and the number of wet days (all in levels, not in first differences):

$$
\begin{aligned}
\Delta ln(y_{it}) = & \underbrace{\beta_1^T \Delta T_{i,t} + \beta_2^T \Delta T_{i,t-1} + \beta_3^T \Delta T_{i,t} T_{i,t} + \beta_4^T \Delta T_{i,t-1} T_{i,t-1}}_{\text{Used for impact projections}} + \\
& \beta_1^P P_{i,t} + \beta_2^P P_{i,t}^2 + \\
& \beta^{\tilde{T}} \tilde{T} + \beta_1^{RD} RD_{x,t} + \beta_2^{RD} RD_{x,t}^2 + \beta_1^{RM} RM_{i,t} + \beta_2^{RM} RM_{i,t}^2 + \\
& \beta_1^{\hat{RD}} \hat{RD}_{x,t} + \beta_2^{\hat{RD}} \hat{RD}_{x,t} T_{i,t} \\
& \alpha_i + \delta_t + \epsilon_{i,t}
\end{aligned} \tag{33}
$$

Here, $\tilde{T}$ is the day-to-day temperature variability, $RD_{x,t}$ is the annual number of wet days with daily precipitation of at least 1 mm, $\hat{RD}_{x,t}$ is the extreme daily rainfall and $RM_{i,t}$ is the standardized monthly rainfall deviation. For specific variable definitions, we refer to Kotz et al. (2022).

Notably, KLW 2022 do not provide impact projections themselves, but the same author team projects impacts in the (now-retracted) Kotz et al. (2024) by applying the KW 2020 methodology to an updated regression model that accounts for impact persistence. Therefore, we apply the KW 2020 methodology to the KLW 2022 model, as done by Waidelich et al. (2024). Specifically, we calculate annual impacts as

$$
\delta_{it}^{KLW22} = \beta_1^T \Delta T_{it} + \beta_2^T \Delta T_{it} T_{i,t} + \beta_3^T \Delta T_{i,t-1} + \beta_4^T \Delta T_{i,t-1} T_{i,t-1} \tag{34}
$$

where $\beta_1^T = 0.000960192$, $\beta_2^T = -0.002273751$, $\beta_3^T = -0.0010594136$ and $\beta_4^T = -0.0006545396$.

Given that impact projections for the remaining climate indicators exhibit low signal-to-noise ratios due to substantial internal variability and total damages remain dominated by annual mean temperature (Waidelich et al., 2024),[9] we only use annual temperature to project impacts.

Again, we cumulate annual impacts and convert from logarithmic changes to % of GDPpc:

$$
\Delta_{it}^{KLW22} = exp(\sum_{s=\tau+1}^{t} \delta_{is}^{KLW22}) - 1 \tag{35}
$$

Notably, this implementation rests on the same assumption about implementing the ADM1-level estimates directly at the country level (Emmerling et al., 2024).

[9]The dominance of annual temperature impacts for overall results was also supported by the now-retracted Kotz et al. (2024).

B.5 ABM 2025

ABM 2025 regress country-level GDPpc growth (in logarithmic changes) on three different climate variables selected via LASSO, namely the share of days with daily maximum temperature above 35°C, the share of days with daily mean temperature between 9–12 °C, and the share of days that experienced a severe drought (measured through a PDSI < -4):

$$\Delta ln(y_{it}) = \beta_1 \Delta \text{Tx35}_{i,t} + \beta_2 \Delta \text{T9to12}_{i,t} + \beta_3 \Delta \text{PDSI}_{i,t} + \alpha_t \tag{36}$$

Importantly, their independent variables are transformed into z-scores by dividing by the in-sample standard deviation for the respective climate variable.

Therefore, we calculate the annual impact as

$$\delta_{it}^{ABM} = \beta_1 \Delta \text{Tx35}_{i,t} + \beta_2 \Delta \text{T9to12}_{i,t} + \beta_3 \Delta \text{PDSI}_{i,t} \tag{37}$$

where $\beta_1 = -0.00168$, $\beta_2 = 0.00144$ and $\beta_3 = -0.00225$ (see their Table 1, column A).

Lastly, we calculate the cumulative impact due to climatic shifts since the base year τ by cumulating and converting from logarithmic changes to % of GDPpc:

$$\Delta_{it}^{ABM} = exp(\sum_{s=\tau+1}^{t} \delta_{is}^{ABM}) - 1 \tag{38}$$

B.6 BK 2026

We take the regional impulse response function coefficients of global temperature shocks from the GitHub replication package of Bilal and Känzig (2026). Notably, these regional impacts are estimated based on global temperature shocks, defined as somewhat persistent deviations from underlying trends and estimated through the procedure developed by Hamilton (2018). We match the BK 2026 regions to individual countries within the macro-regions and continents (see Table S1) and then apply the respective region-specific β estimates uniformly to all countries within a given region.

We estimate annual impacts for the respective region r based on the evolution of global temperature shocks over the 10 previous years:

$$\delta_{it}^{BK} = \sum_{l=0}^{10} \beta_{r,l} T_{global,t-l}^{shock} \tag{39}$$

BK Region	Countries
CentralEastAsia	CHN, MNG, TWN, HKG, MAC, PRK, KOR, JPN
Europe	AUT, BEL, BGR, HRV, CYP, CZE, DNK, EST, FIN, FRA, DEU, GRC, HUN, IRL, ITA, LVA, LTU, LUX, MLT, NLD, POL, PRT, ROU, SVK, SVN, ESP, SWE, GBR, RUS, CHE, TUR, NOR, UKR, SRB, ALB, BIH, MKD, MNE, BLR, MDA, ISL, ARM, AZE, GEO, KAZ, UZB, TKM, TJK, KGZ, LIE, MCO, SMR, AND, VAT, FRO, GIB, GGY, IMN, JEY
LatinAmerica	MEX, ARG, BRA, CHL, COL, PER, VEN, URY, PRY, BOL, ECU, GUY, SUR, GUF, PAN, CRI, NIC, HND, SLV, GTM, BLZ, CUB, HTI, DOM, JAM, TTO, BHS, BRB, ATG, KNA, LCA, GRD, VCT, DMA, ABW, CUW, SXM, BES, PRI, VIR, VGB, CYM, TCA, AIA, MSR, BLM, MAF, GLP, MTQ, FLK
MiddleEastNorthAfrica	EGY, LBY, TUN, DZA, MAR, SAU, ARE, QAT, KWT, BHR, OMN, YEM, IRQ, IRN, SYR, LBN, JOR, ISR, PSE
NorthAmerica	USA, CAN, BMU, GRL, SPM
Oceania	AUS, NZL, PNG, FJI, SLB, VUT, NCL, PYF, WSM, TON, KIR, MHL, FSM, NRU, PLW, TUV, COK, NIU, TKL, ASM, GUM, MNP, PCN, WLF, NFK, CXR, CCK
SouthAsia	IND, PAK, BGD, NPL, BTN, LKA, MDV, AFG
SoutheastAsia	IDN, MYS, PHL, SGP, THA, VNM, MMR, KHM, LAO, BRN, TLS
SubSaharanAfrica	NGA, ETH, ZAF, KEN, TZA, UGA, GHA, MOZ, AGO, CIV, MDG, CMR, NER, BFA, MLI, MWI, ZMB, SEN, ZWE, RWA, SOM, BDI, BEN, TGO, SLE, LBR, CAF, MRT, ERI, GMB, BWA, NAM, GAB, LSO, GIN, COG, SWZ, DJI, COM, CPV, TCD, COD, GNQ, GNB, MUS, STP, SYC, SSD, SDN

Table S1: BK regions and ISO3 codes

Here T^{shock}_{global} is the year-specific shock in global mean surface temperature taken from the BK replication package (variable "gtmp_noaa_aw_dtfe2s"), which they calculate based on global temperature anomaly data from the NOAA National Centers for Environmental Information.

Then we convert to % of GDPpc:

$$\Delta^{BK}_{it} = exp(\delta^{BK}_{it}) - 1 \tag{40}$$

Importantly, this approach produces short-term results consistent with the structural model results in BK 2026 (and directly rooted in their empirical results). However, its long-term results would be substantially more conservative, as the effects of global temperature shocks disappear entirely after ten years. We verify this by simulating a global temperature trajectory that mimics the one they display in their Figure 13a and calculating impacts for

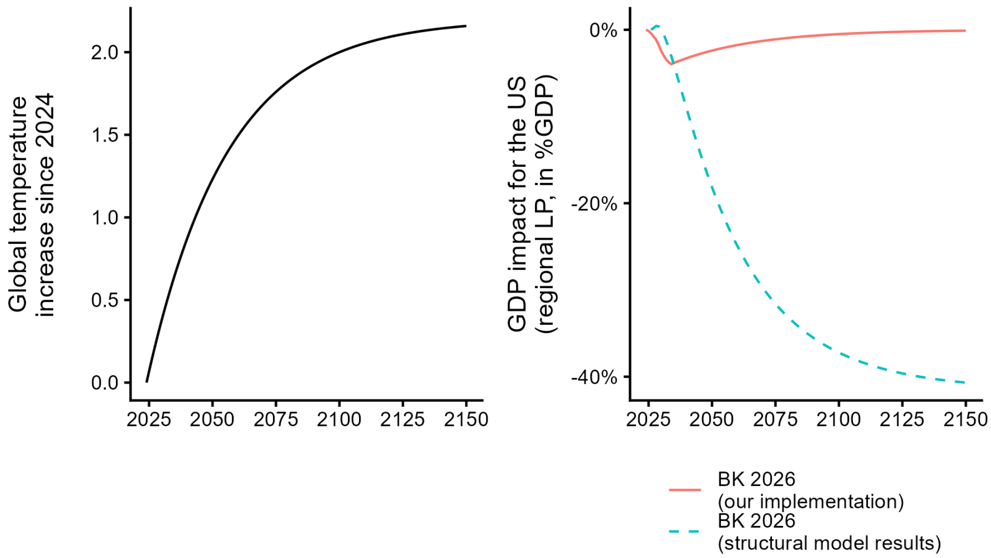


Figure S8: Impact estimates for applying our implementation to a simulated global temperature trajectory mimicking the model results in BK 2026.

the North America region comprising the US as an example region, as shown in Figure S8.

B.7 DICE 2023

DICE 2023 uses a reduced-form quadratic damage function with the shift in global mean surface temperature since the pre-industrial era:

$$\Delta^{DICE}_{global,t} = \beta_1 * \Delta T^2_{global,t,1850-1900} \tag{41}$$

where $\beta_1 = -0.003467$ is derived from a meta-analysis of the extant literature. As the damage estimate is global, we apply the resulting impact uniformly to all countries and, similar to our COACCH implementation, isolate impacts due to warming that occurred after the base year τ:

$$\Delta^{DICE}_{global,t} = \beta_1 * \left(\Delta T^2_{global,t,1850-1900} - \Delta T^2_{global,\tau,1850-1900}\right) \tag{42}$$

Notably, the DICE damage function returns impact in % of output, not in % of GDPpc. As for our COACCH implementation, we assume that the two are equivalent (i.e., assuming that impacts on population size are negligible).

B.8 HS 2025

Howard and Sterner (2025) provide a global damage function based on a meta-analysis of the existing literature, expressed as a share of global GDP relative to a baseline temperature level. We implement the global damage as:

$$\Delta^{HS}_{global,t} = f(\Delta T_{global,t,1850-1900}) - f(\Delta T_{global,\tau,1850-1900}) \tag{43}$$

where $f(\cdot)$ is the meta-analytic damage function from Howard and Sterner (2025) and τ denotes the base year. As in our other implementations, we isolate only the impacts attributable to warming occurring after the base period.

To distribute global damages across countries, we follow OECD (2025), who use the country-level downscaling approach of Tol (2024). This approach allocates global damages to individual countries based on their current climate conditions and GDP per capita, using the 2006–2010 average as the reference period for baseline climate and income characteristics.[10]

B.9 KMNPRY 2021

For their impact projections, KMNPRY 2021 estimate the following regression model using a half-panel jackknife estimator:

$$\Delta ln(y_{i,t}) = \alpha_i + \sum_{j=1}^{4} \gamma_l \Delta ln(y_{i,t-l}) + \sum_{j=0}^{4} \beta_j \, \Delta\tilde{T}_{i,t-j} + \varepsilon_{i,t} \tag{44}$$

Here, $\tilde{T}_{i,t}$ is the absolute deviation of annual temperature from its long-term historical norm (calculated as a moving-average of length m):

$$\tilde{T}_{i,t} = |\ T_{i,t} - \frac{1}{m}\sum_{l=1}^{m} T_{i,t-l}\ | \tag{45}$$

Among the damage functions we consider here, the KMNPRY 2021 function is somewhat unique in that it projects impacts based on differences in (linear) trends in annual mean temperature between a scenario under climate change and historically observed trends, rather than projecting impacts based on year-specific temperature (whether in levels, shifts, or shocks). Specifically, GDP per-capita losses for a given year and RCP-SSP scenario due to warming since an initial year $t = 0$ are calculated as

[10]The global damage function is based on Howard and Sterner (2025) while the country-level downscaling is following OECD (2025), who use the results of Tol (2024) to capture the current climate and GDP per capita to distribute global damages across countries.

$$\sum_{j=1}^{t}(g_{1,j} - g_0) \times \hat{\psi}(t-j) \tag{46}$$

where

$$g_0 = \mu_0\Big[\Phi\Big(\tfrac{\mu_0}{\omega_0}\Big) - \Phi\Big(-\tfrac{\mu_0}{\omega_0}\Big)\Big] + 2\omega_0 \cdot \phi\Big(\tfrac{\mu_0}{\omega_0}\Big) \tag{47}$$

Here, $\Phi(\cdot)$ and $\phi(\cdot)$ denote the standard normal cumulative distribution function and the standard normal density function, respectively. g_0 is the expected absolute deviation of temperature from its historical norm and μ_0 is the mean deviation implied by the historical trend, estimated as

$$\mu_0 = \frac{m+1}{2}\beta_0 \tag{48}$$

where β_0 is the country-specific time trend in annual mean temperature over 1960–2014 (i.e., their sample period).

ω_0 in Equation 47 denotes its standard deviation, calculated as

$$\Big(\sigma_0^2 \times (1 + \frac{1}{m})\Big)^{0.5} \tag{49}$$

where ω_0^2 is the estimated volatility of annual mean temperature around its trend throughout 1960–2014.

$g_{1,j}$ in Equation 46 is calculated by replacing μ_0 in Equation 47 with

$$\mu_{1,j} = \frac{m+1}{2} \times (\beta_0 + j \times d_{RCP}) \tag{50}$$

where d_{RCP} is the mean increment in a country's trend rise under climate change relative to the trend observed in the sample period (see Equations 10–12 in KMNPRY 2021).

Lastly, $\hat{\psi}(t-j)$ denotes the impulse-response function derived for the regression coefficients for γ and β in Equation 44 :

$$\hat{\psi}(L) = \gamma(L)^{-1}\beta(L) \tag{51}$$

where $\gamma(L) = 1 - \sum_{l=1}^{4}\gamma_l L^l$ and $\beta(L) = \sum_{l=0}^{4}\beta_l L^l$.

For all our results, we use results for the $m = 30$ historical norm. We take country-specific estimates for the historical β_0 and ω_0 (displayed in their Table A.7) and the central estimates for all $\hat{\psi}(t-j)\ \forall\ t-j = 0, 1, ..., 100$ (as illustrated in KMNPRY 2021, Figure 5) from the KMNPRY 2021 R replication package on GitHub (`https://github.com/scentorrino/KMNPRY.2021`), Version 1.0.1, for maximum precision.

Notably, KMNPRY 2021 compare the historical trends over 1960–2014 against expected trends over 2015–2100 and provide $d_{RCP2.6}$ and $d_{RCP8.5}$. Mohaddes and Raissi (2025) expand this methodology to additional scenarios, including RCP7.0. Applying their methodology to our use case means shortening the baseline period from 2014 to 2010 and changing the projection period from 2015–2100 to 2011–2019.

To inform this adjustment, we use Berkeley Earth data on global mean surface temperature anomalies to assess warming trends over these different periods. Estimating a linear time trend in GMST yields $0.047°C$ p.a. for 2011–2019, $0.017°C$ p.a. for 1960–2014 and $0.017°C$ p.a. for 1960–2010. Based on these results, we note that 2011-2019 warming followed a trend that, if extrapolated into the future, aligns most closely with SSP5-RCP8.5 results as per the IPCC AR6 Working Group I report. We also note that, at the global level, temperature trends over 1960–2010 are extremely similar to the KMNPRY 2021 baseline period of 1960–2014. Therefore, we estimate 2011–2019 impacts by using the original RCP8.5 estimates for d_{RCP} and the original μ_0 and ω_0 estimates without adjustments from the GitHub replication package. Compared to an alternative of estimating country-specific historical trends and volatility over 2011–2019, relying on (short-term) RCP8.5 impacts ensures maximum consistency with the paper's original results, consistent with the strong warming surge observed in the 2010s.

Appendix C Sample

The WEO data we use covers a total of $N = 196$ economies. Of these, two (i.e., Macao SAR and Hong Kong SAR) are not included in the climate data by Gortan et al. (2024). An additional 27 economies are not included in our forecast error evaluation because we cannot implement at least one damage function for them (and choose to omit them to ensure a consistent evaluation sample across damage functions). This is primarily because for 13 economies in question, we cannot match them to a COACCH region. In addition, some economies are not included in the KMNPRY 2021 sample, meaning that we lack the original parameters (e.g., the 1960–2014 temperature trend or the projected warming increment under RCP8.5) to implement damages for these countries.

All economies with either missing climate data or impact implementation are listed in Table S2. Economies with missing climate data are not included in our placebo regressions. Economies with missing impacts are not included in our forecast performance evaluation.

Table S2: WEO economies excluded from the forecast evaluation

ISO3	Country	Climate data missing	Impacts missing
ABW	Aruba	FALSE	TRUE
AND	Andorra	FALSE	TRUE
ATG	Antigua and Barbuda	FALSE	TRUE
BHR	Bahrain	FALSE	TRUE
BRB	Barbados	FALSE	TRUE
DMA	Dominica	FALSE	TRUE
FSM	Micronesia	FALSE	TRUE
GRD	Grenada	FALSE	TRUE
HKG	Hong Kong SAR	TRUE	TRUE
KIR	Kiribati	FALSE	TRUE
KNA	St. Kitts and Nevis	FALSE	TRUE
KOS	Kosovo	FALSE	TRUE
LCA	St. Lucia	FALSE	TRUE
MAC	Macao SAR	TRUE	TRUE
MDV	Maldives	FALSE	TRUE
MHL	Marshall Islands	FALSE	TRUE
NRU	Nauru	FALSE	TRUE
PLW	Palau	FALSE	TRUE
SGP	Singapore	FALSE	TRUE
SMR	San Marino	FALSE	TRUE
SYC	Seychelles	FALSE	TRUE
TLS	Timor-Leste	FALSE	TRUE
TON	Tonga	FALSE	TRUE
TUV	Tuvalu	FALSE	TRUE
TWN	Taiwan Province of China	FALSE	TRUE
WBG	West Bank and Gaza	FALSE	TRUE

PUBLICATIONS

When Forecasts Meet Reality: Assessing Climate Damage Functions

Working Paper No. WP/2026/085