Appendix A. Supplemental Information

A.1. Literature Review

Since the adoption of the Paris Agreement and the emergence of Nationally Determined Contributions (NDCs) and Long-term Strategies (LTS), model-based research has actively explored the feasibility, implications, and opportunities surrounding these policies and other emissions reduction pathways. Many of these studies focus on the policy implementation while relying on business-as-usual assumptions in other areas of the modeling framework. (Iyer et al., 2015b) examine the NDCs in 2015 and the energy-economic implications across policy scenarios which vary the timing of mitigation actions. (Fawcett et al., 2015) also assess these NDC pledges by computing probabilistic temperature outcomes with a global climate model based on several scenarios constructed with an integrated assessment model. (Ou et al., 2021) then evaluate the updated 2020 NDC pledges using additional simulations, emphasizing that additional ambition is needed to achieve long-term goals. These studies use a limited number of scenarios in determining emissions trajectories, trading off the evaluation of uncertainty with finely-tuned scenario pathways. (Gambhir et al., 2022) approach emissions mitigation using several temperature target scenarios as well as an NDC scenario to identify transition risk metrics within an integrated assessment framework. The authors find that different types of risks emerge as being most sensitive to the future temperature pathway on different timescales. (Binsted et al., 2020) used NDC scenarios to quantify the economic implications of stranded assets under the Paris Agreement, finding significant cost burdens associated with the policies. (Santos Da Silva et al., 2019) model two NDC scenarios using an integrated assessment framework in which one scenario does not have access to CCS technologies, and evaluates resulting food-energy-water nexus outcomes.

There exists also a broad literature of uncertainty and sensitivity analysis centered around climate mitigation modeling research. However, many of these studies evaluate only a few deeply uncertain factors in their simulations, often only implemented individually rather than through a factorial ensemble. (Iyer et al., 2015a) explore varying the cost of financing clean energy projects in the electric power sector across regions due to investment risk and variations in institutional quality under a generic 50% emissions reduction policy. This study found that these disparities in investment risks significantly affected the total costs of mitigation, and that more industrialized regions take on a greater share of the mitigation requirements. (Kanyako and Baker, 2021) perform an uncertainty analysis on wind energy costs for a carbon tax and a 1.5° scenario, exploring impacts on wind generation share across a distribution of cost forecasts. (Ou et al., 2018) compare two low-carbon pathways (each comprised of several technology assumptions) in the US under two different mid-century emissions reductions targets, evaluated with water consumption and air pollution metrics. (Moksnes et al., 2019) prepare an ensemble of 324 scenarios varying six uncertain factors related to energy systems (including a simple CO_2 target) and perform scenario discovery on the resulting cost and capacity mix outcomes.

Several studies use an ensemble of model realizations in climate mitigation contexts. McJeon et al., 2011 uses a large, 768-member ensemble and scenario discovery to explore the impacts of technology assumptions on stabilization costs under two temperature stabilization scenarios. Groves et al., 2020 develops 3,003 realizations of Costa Rica's decarbonization plan to assess the economic value of the plan independent of international pledges. Although many previous modeling efforts have examined impacts of climate mitigation measures and parametric uncertainties on energy-economic outcomes, there remains a gap in evaluating countries' NDC + LTS pledges across a wide range of deeply uncertain factors in a large ensemble framework. This study seeks to confirm the results of prior research in a robust NDC- + LTS-consistent mitigation context, as well as examine interactive effects of previously independent sensitivity factors in a large ensemble of model realizations.

Table S1: Non-exhaustive list of existing work.

Authors	Short Description	Approach to Uncertainty
McJeon et al., 2011	768-member large ensemble of GCAM runs exploring impacts of technology assumptions on stabilization costs	Scenario discovery, reporting density and coverage statistics on extreme outcomes
Fawcett et al., 2015	600-member temperature projection ensemble applied to several GCAM Paris Agreement scenarios	Temperature outcomes presented probabilistically
Isley et al., 2015	XLRM framework generating 6,000 combinations of uncertain parameters and 6 policies in agent-based model	Exploratory modeling to explore decarbonization rates and policy choices
Iyer et al., 2015b	Four GCAM scenarios varying model assumptions to explore Paris Agreement implications on 2°C	Using a small number of detailed representative scenarios to assess implications of INDCs
McFarland et al., 2015	Set of temperature projections applied to GCAM-USA, ReEDS, IPM to look at electricity supply/demand	Multi-model comparison
Wilkerson et al., 2015	Carbon price scenarios applied to GCAM, MERGE, and EPPA	Multi-model comparison
Kober et al., 2016	Climate policies centered on Latin America, using GCAM, POLES, TIAM-ECN, and TIAM-WORLD	Multi-model comparison
Lucena et al., 2016	Five scenarios of Brazil's energy mix using EPPA, GCAM, MESSAGE-Brazil, Phoenix, POLES, and TIAM-ECN	Multi-model comparison
Van Der Zwaan et al., 2016	Five scenarios of energy technology deployment in Latin America using EPPA, GCAM, Phoenix, POLES, TIAM-ECN, and TIAM-WORLD	Multi-model comparison
Pietzcker et al., 2017	Integration of wind and solar in IAMs using AIM/CGE, IMAGE, MESSAGE, POLES, REMIND, and WITCH	Multi-model comparison
Kriegler et al., 2018	Strengthening short-term goals to meet Paris Agreement with 13 scenarios across three policy dimensions using REMIND-MAgPIE	Constructing representative scenarios with detailed sectoral assumptions to assess policy impacts
Lamontagne et al., 2018	33,750-member ensemble of GCAM runs splitting SSP assumptions into individually sampled elements	Scenario discovery using CART
Arango-Aramburo et al., 2019	Climate-impacted hydropower in Colombia using two GCMs, two RCPs, and 4 IAMs: GCAM, TIAM-ECN, MEG4C, Phoenix	Multi-model comparison
Lamontagne et al., 2019	5,200,000-member ensemble using DICE, sampling 24 uncertain factors and growth rate of global abatement	Time-varying sensitivity analysis
Moksnes et al., 2019	324-member ensemble using OSeMOSYS-SAMBA to explore South American electricity infrastructure	Scenario discovery using a Gaussian mixture model and PRIM
Binsted et al., 2020	Four global GHG mitigation scenarios using GCAM to explore stranded assets in Latin America	Used 36 sensitivity scenarios to perform sensitivity analysis

Burleyson et al., 2020	Four scenarios each run using GCAM-USA and BEND to explore US buildings electricity consumption	Two-model comparison
Groves et al., 2020	3,003-member ensemble varying over 300 uncertainties to explore Costa Rica's national decarbonization plan	Scenario discovery using PRIM to identify vulnerabilities
Dolan et al., 2021	3,000-member ensemble of GCAM runs varying seven dimensions of uncertainties to explore impacts of water scarcity	Scenario discovery using CART
Kanyako and Baker, 2021	1,000-member ensemble of GCAM runs with technology costs sampled from expert elicitation data	Uncertainty propagation from expert elicitation data
Ou et al., 2021	Five emissions scenarios using GCAM coupled with simple climate model MAGICC	Probabilistic temperature outcomes using detailed emissions scenarios
Solano-Rodríguez et al., 2021	XLRM framework generating 480 alternatives for oil production in Latin America using BUEGO	Latin hypercube sampling to generate ensemble of alternatives
Birnbaum et al., 2022	3,000-member ensemble of GCAM runs exploring water scarcity in Latin America	Scenario discovery using CART
Gambhir et al., 2022	11 scenarios of temperature outcomes and socioeconomic/technological choices for 2°C pathways using GCAM	Comparison of risk metrics across detailed representative scenarios
Browning et al., 2023	Using three scenarios to analyze net-zero by 2050 in the US across 16 models	Multi-model (and multi-modeling team) comparison of detailed representative scenarios
Huang et al., 2023	28,706-member ensemble of GCAM runs coupled with TM5-FASST to explore air quality implications from climate mitigation under uncertainty	Large ensemble scenario analysis and model coupling
van de Ven et al., 2023	Three scenarios of climate action applied to GCAM-PR, GEMINI-E3, MUSE, and TIAM-Grantham	Multi-model comparison to explore feasibility of climate ambition
Woodard et al., 2023	3,989-member ensemble of GCAM runs varying 12 uncertainties chosen from expert elicitation	Scenario discovery using CART

51

A.2. Computing Metrics from GCAM Ensemble

Table S2: Descriptions of each metric and how each is calculated from GCAM outputs.

Metric	Short Description
Electricity Price	Marginal levelized cost of new generation (analogous to wholesale electricity costs). When aggregated from several regions, a weighted average based on total regional electricity generation is applied. Queried directly from GCAM outputs.
Electricity Share in Final Energy	Also termed "Electrification Rate", the proportion of total final energy delivered to end use sectors as electricity in each region. When aggregated from several regions, a weighted average based on total regional final energy is applied. Total final energy is queried directly from GCAM outputs, from which the proportion of electricity can be computed.
Stranded Assets	The cumulative costs of premature retirement of electric generating capacity over time in each region. Can be split by technology. Premature retirement refers to a generating unit being forced offline before the end of its economic life (e.g., due to mitigation policy constraining emissions or increasing costs to inefficient levels). Results from individual regions can be summed. Stranded assets are computed from GCAM outputs using the "plutus" R package (Zhao et al., 2021).
Capacity Investments	The cumulative capital costs of new electric generating capacity over time in each region. This metric gives one angle of a policy's economic impacts, and can be split by technology. Capacity investments are computed from GCAM outputs using the "plutus" R package (Zhao et al., 2021).
Energy Burden	An aggregated metric of distributional energy justice, computed as a residential energy burden by dividing per capita residential energy expenditures by per capita GDP. From GCAM outputs, residential energy expenditures are computed using residential building service costs (which includes levelized installed costs of service equipment in addition to fuel costs) and final energy consumption in residential sectors. Population and GDP are exogenous inputs to GCAM. This metric does not include transport service costs.
Level of CO_2 Removal	The quantity (mass of CO_2) removed from the atmosphere via Bioenergy with CCS (BECCS) and Direct Air Capture (DAC). Results from individual regions can be summed. Queried directly from GCAM outputs.
Land Use Change Emissions	The net quantity (mass of CO_2) of land use change emissions, representing regional and global carbon stocks. Results from individual regions can be summed. Queried directly from GCAM outputs.

A.3. Supplemental Figures

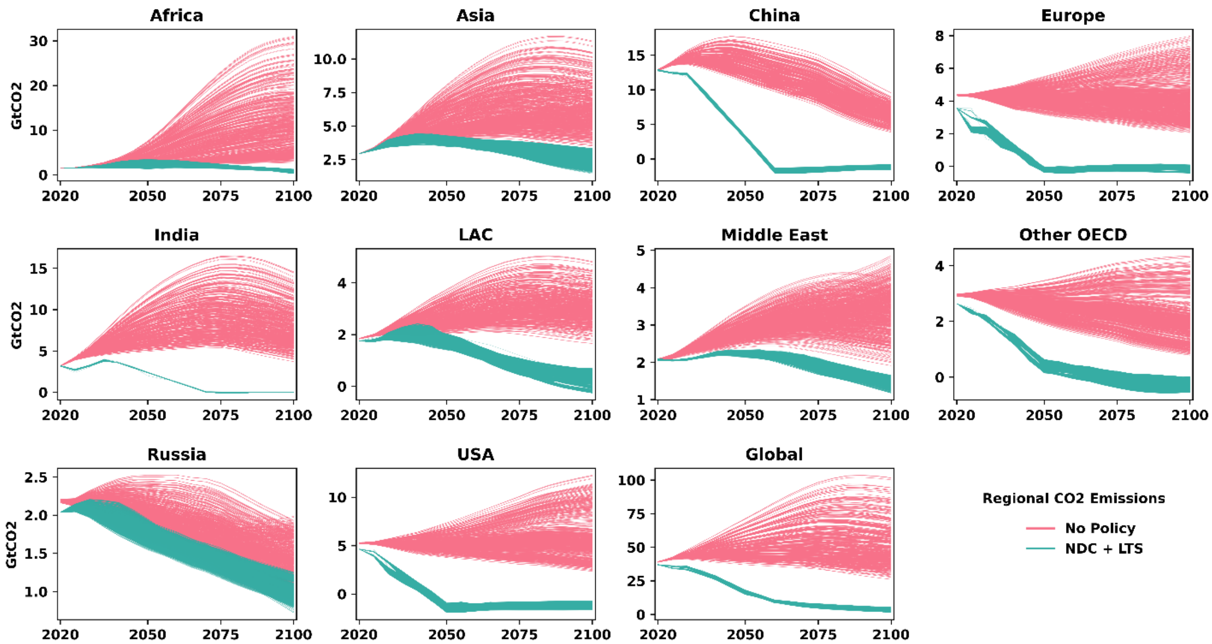

Figure S1: CO_2 emissions trajectories across regions and globally, split by climate pledge policy sensitivity. "Other OECD" includes Canada, Japan, South Korea, Australia, and New Zealand. “Asia” includes Pakistan, Indonesia, Central Asia, South Asia, and Southeast Asia. “LAC” refers to Latin America and the Caribbean.

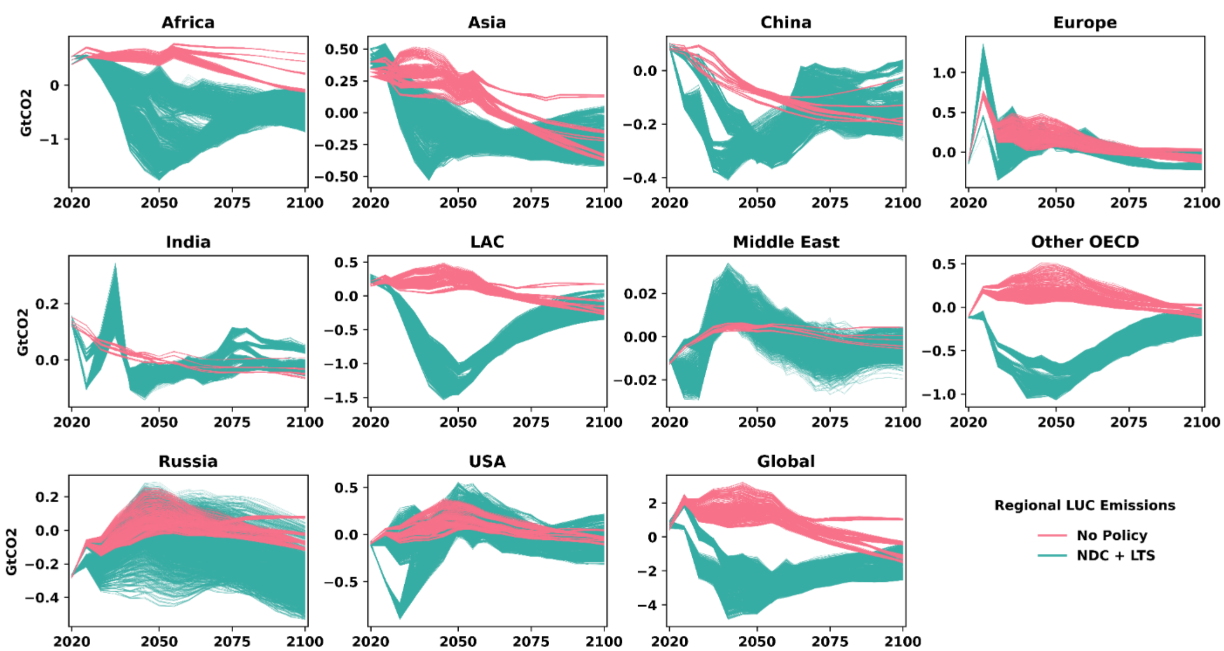

Figure S2: Land use change emissions trajectories across regions and globally, split by climate pledge policy sensitivity. "Other OECD" includes Canada, Japan, South Korea, Australia, and New Zealand. “Asia” includes Pakistan, Indonesia, Central Asia, South Asia, and Southeast Asia. “LAC” refers to Latin America and the Caribbean.

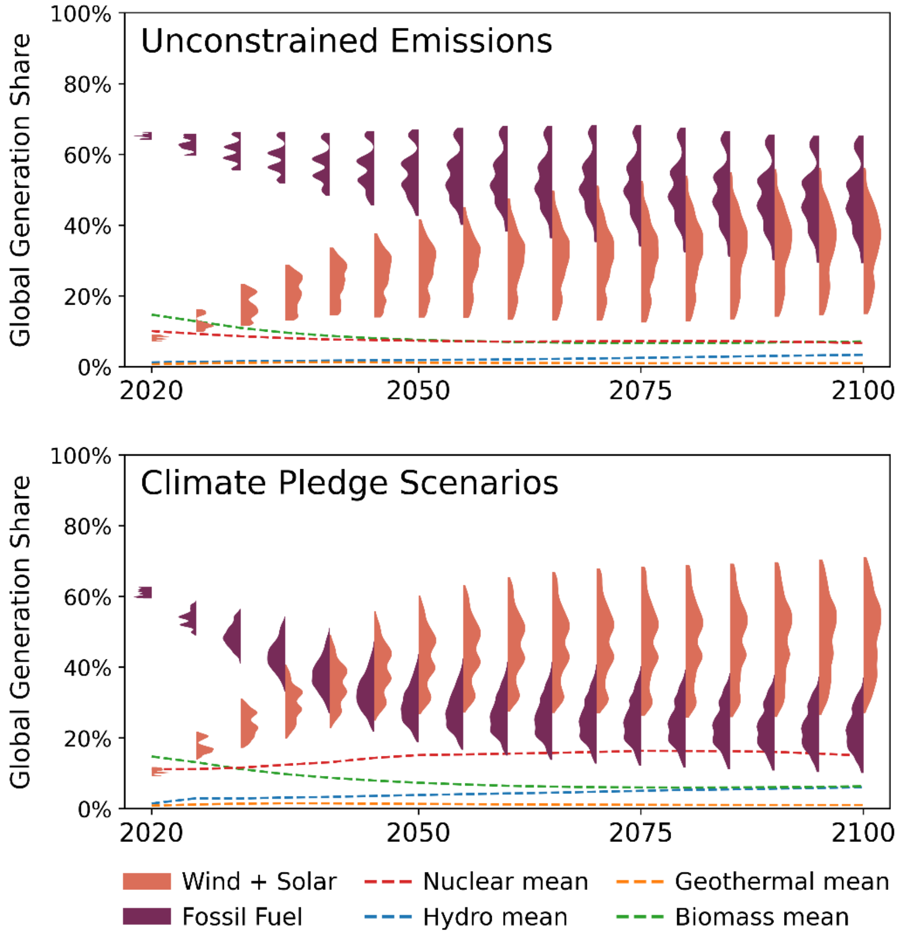

Figure S3: Evolution of the electricity generation mix as a split violin plot for No Policy cases (top) and climate pledge scenarios (bottom). Fossil fuels remain dominant in the No Policy case, although renewables still increase over time. In the NDC + LTS case, wind and solar trade places with fossil generation to become the leading producer of electricity. Fossil generation does not go to zero, partially because not every country has committed to NDC/LTS pledges, but also because of the significant amount of CO_2 removal technologies employed in the model. Variability for other generation types is relatively small; these are shown instead as dotted lines representing the mean.

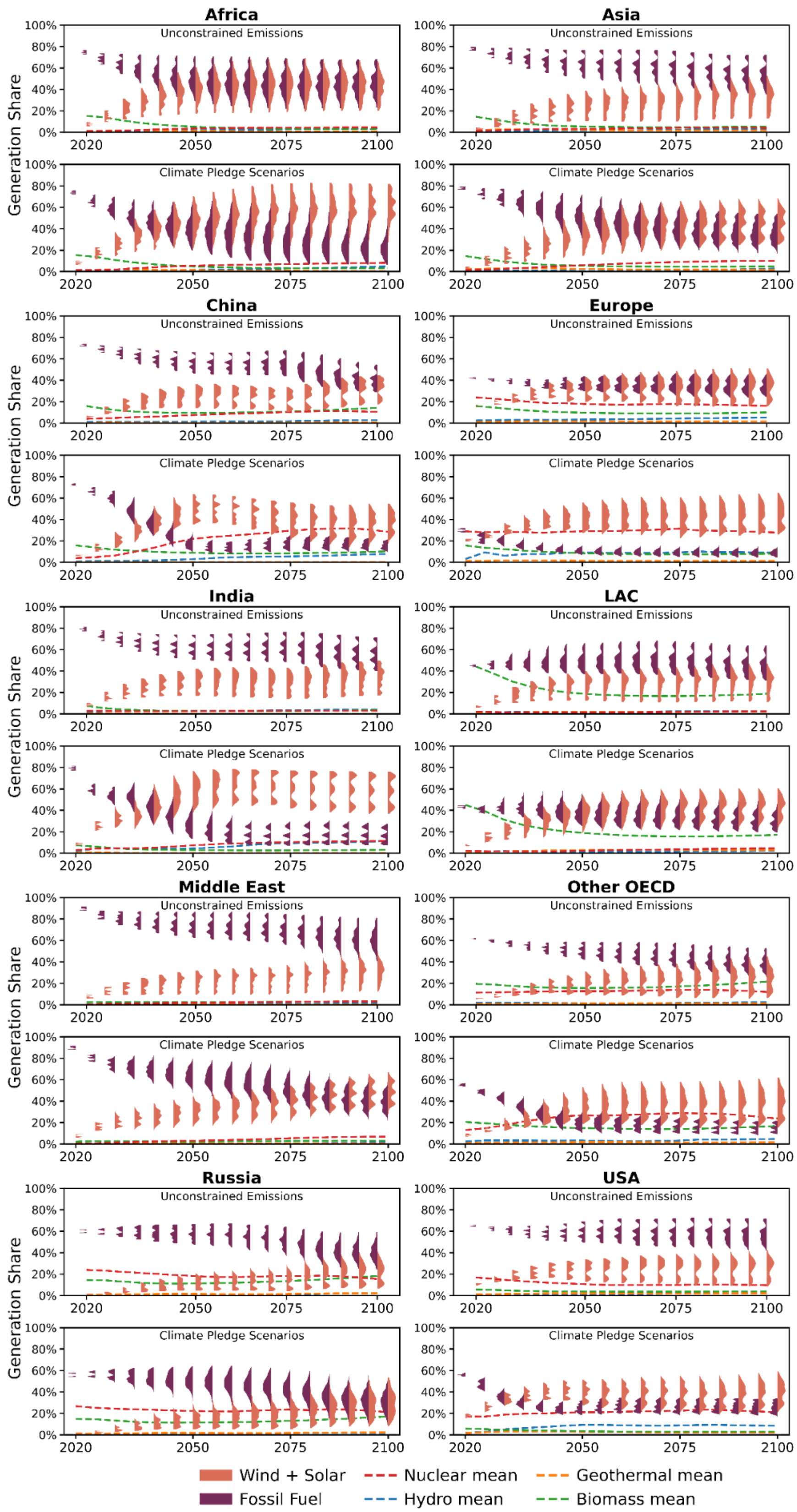

Figure S4: Generation share violin plots similar to Figure S3, split out into ten aggregated global regions.

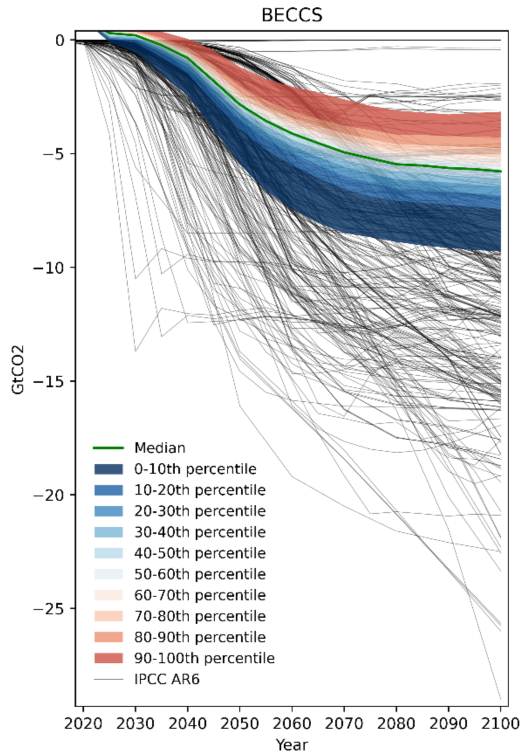

Figure S5: Bioenergy with CCS (BECCS) for climate pledge scenarios as percentiles. Negative values represent CO_2 being removed. Black lines show scenarios from IPCC AR6 (Riahi, 2022).

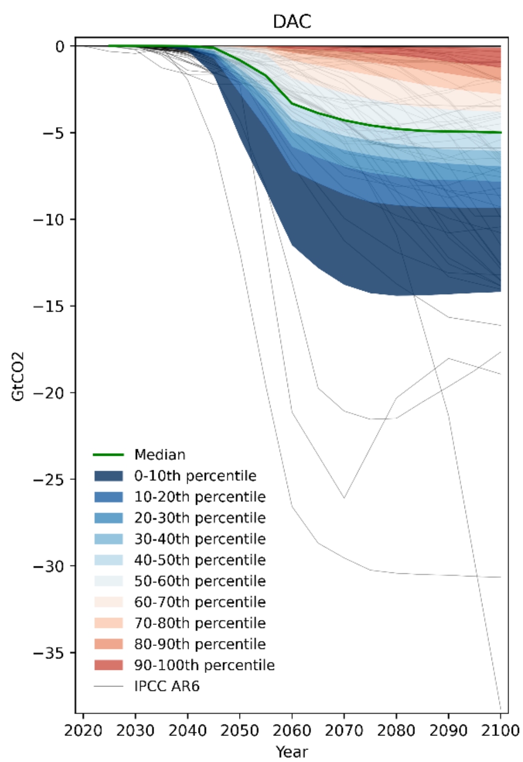

Figure S6: Direct Air Capture (DAC) for climate pledge scenarios as percentiles. Negative values represent CO_2 being removed. Black lines show scenarios from IPCC AR6 (Riahi, 2022).

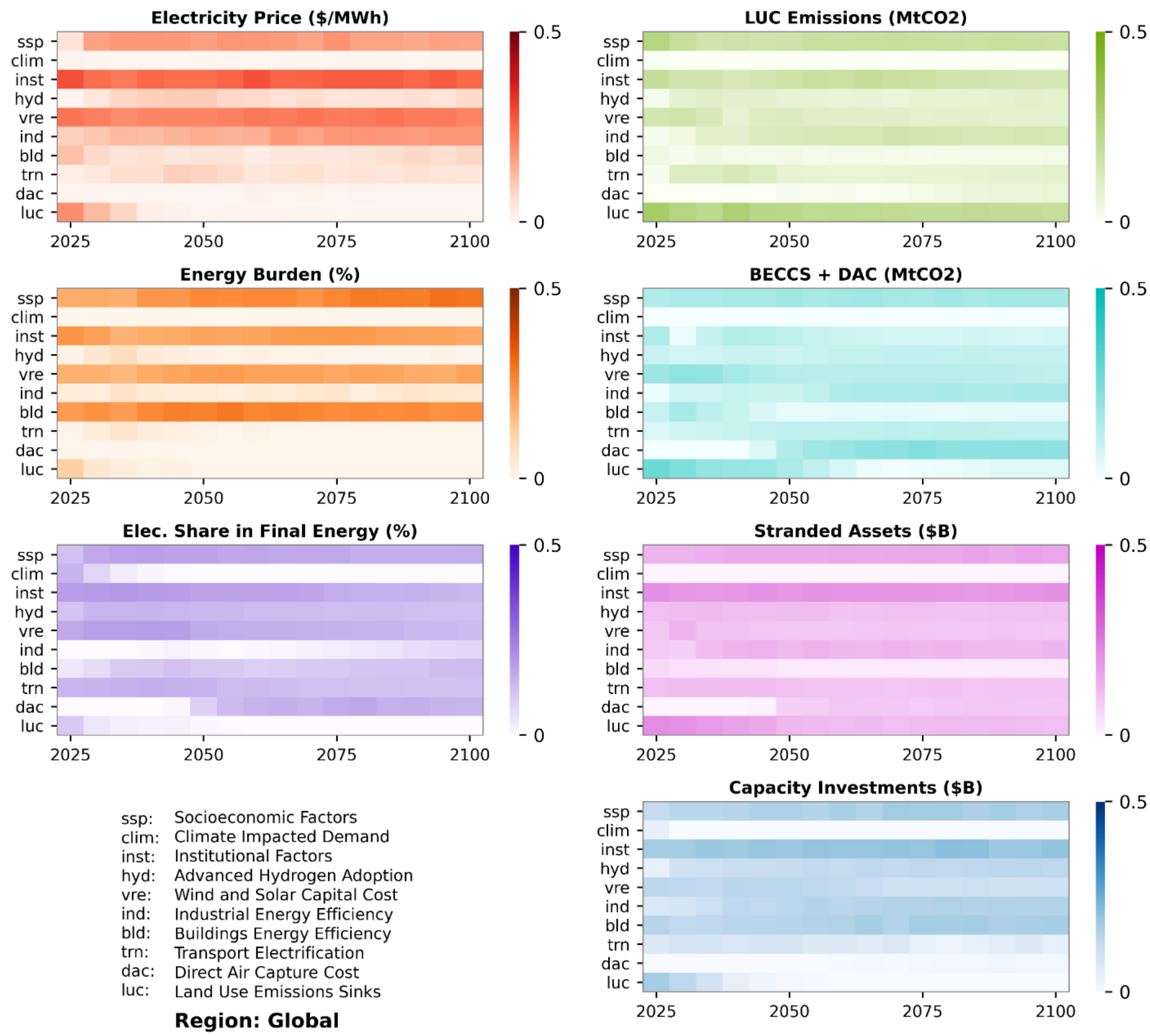

Figure S7: Feature importance analysis for seven representative metrics across the 3,840 simulations implementing national climate pledges. Each panel is presented as a heatmap quantifying relative influence by the scenario sensitivities in each row on each output metric over time. A higher score (darker color) indicates higher influence in the random forest model from the inclusion of each feature (listed in bottom left of figure). Because only NDC + LTS scenarios are examined here, this sensitivity is not listed. In general, *Socioeconomic Factors* is a relevant driver in nearly all outcome metrics, as it controls the scale of economic activity as well as resource demand. The electricity price panel confirms the critical drivers seen in **Error! Reference source not found.**, while also notable is the increasing potential role of *Industry Energy Efficiency*, which affects industrial sectors including iron & steel, cement, aluminum, chemicals, and fertilizer production. This sensitivity also has an increasing importance in several other economic metrics as well as negative emissions. Feature importance is quantified by the average improvement in mean squared error (MSE) achieved in the random forest model from permuting each feature in out-of-bag samples, scaled to sum to one in each timestep. Feature importance here does not in itself indicate the direction of influence.

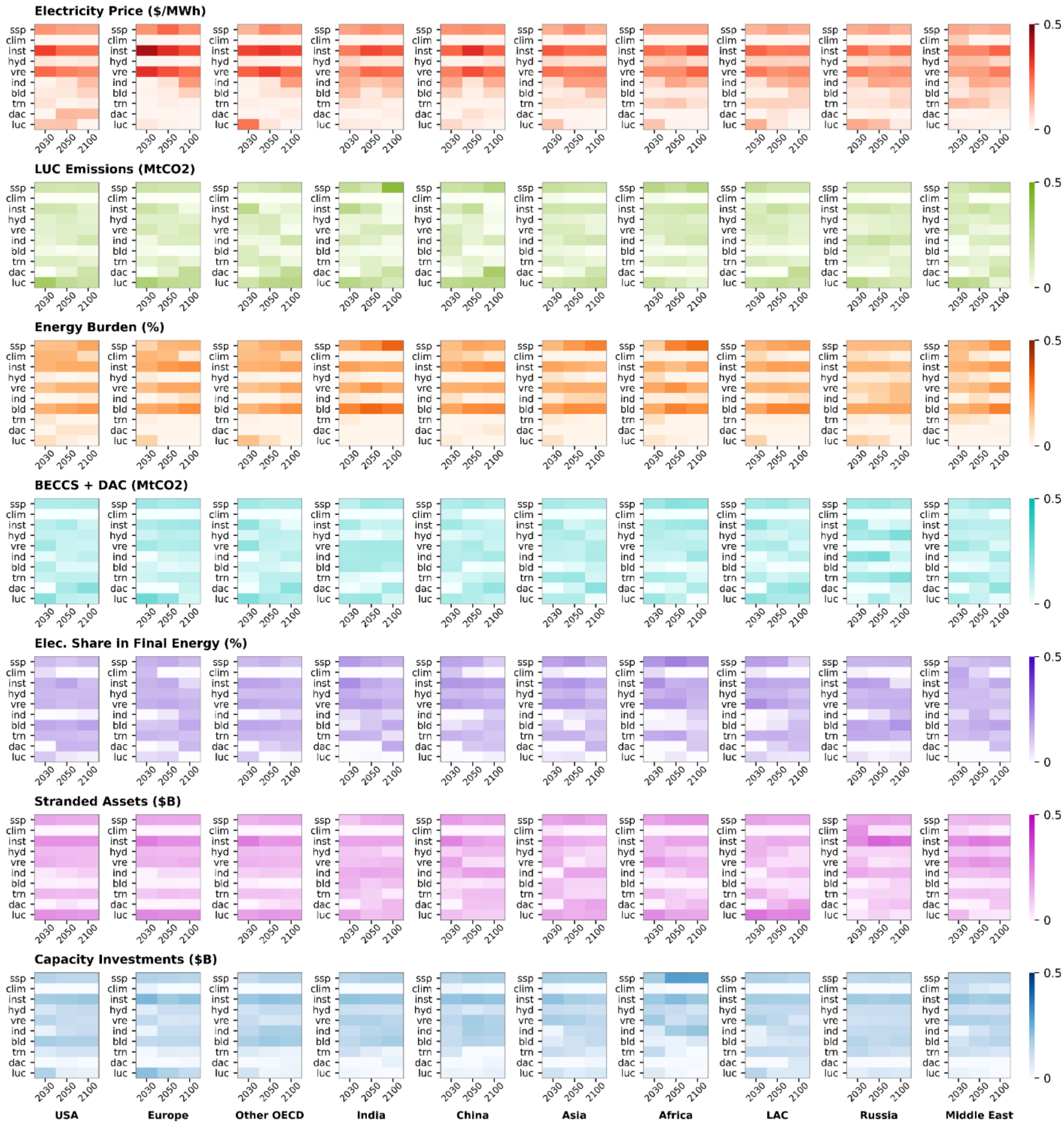

Figure S8: Feature importance analysis for seven representative metrics across the 3,840 simulations implementing national climate pledges, split by region (column) and only showing values for 2030, 2050, and 2100. Each panel is presented as a heatmap quantifying relative influence by the scenario sensitivities in each row on each output metric. A higher score (darker color) indicates higher influence in the random forest model from the inclusion of each feature. Because only NDC + LTS scenarios are examined here, this sensitivity is not listed.

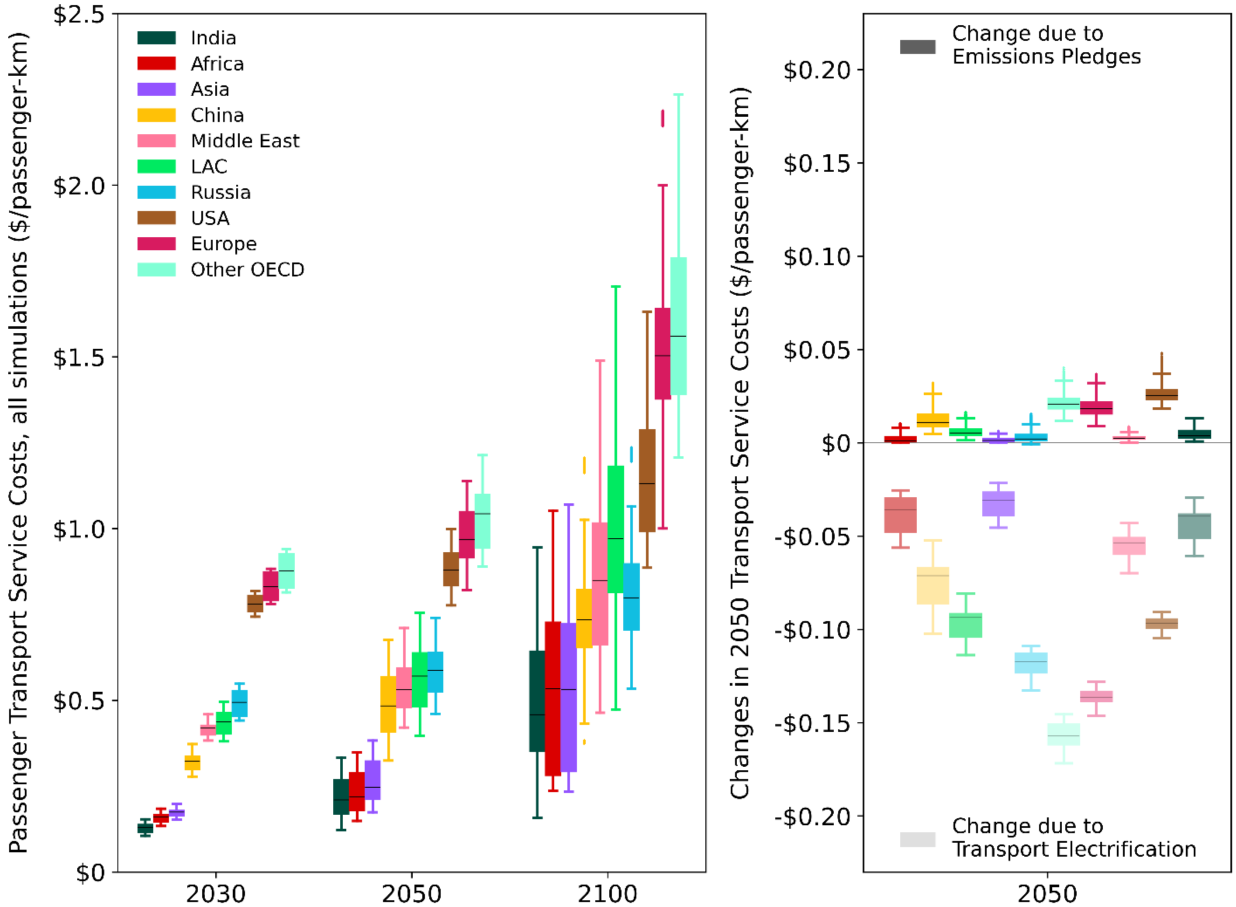

Figure S9: (left) Cost of transport services in the passenger transport sector for aggregated global regions in three model periods, showing all 5,760 simulations; **(right)** Change in passenger transport service costs caused by two scenario sensitivities (climate pledges and *Electrification of Transport*) for each model configuration, computed as the difference between pairs of realizations which differ only by inclusion/exclusion of these two scenario levers. Developed regions tend to experience the highest costs, a trend which does not change over time. Passenger transport service costs increase over time across regions, but total expenditures remain relatively stable when scaled by GDP. "Other OECD" includes Canada, Japan, South Korea, Australia, and New Zealand. “Asia” includes Pakistan, Indonesia, Central Asia, South Asia, and Southeast Asia. “LAC” refers to Latin America and the Caribbean.

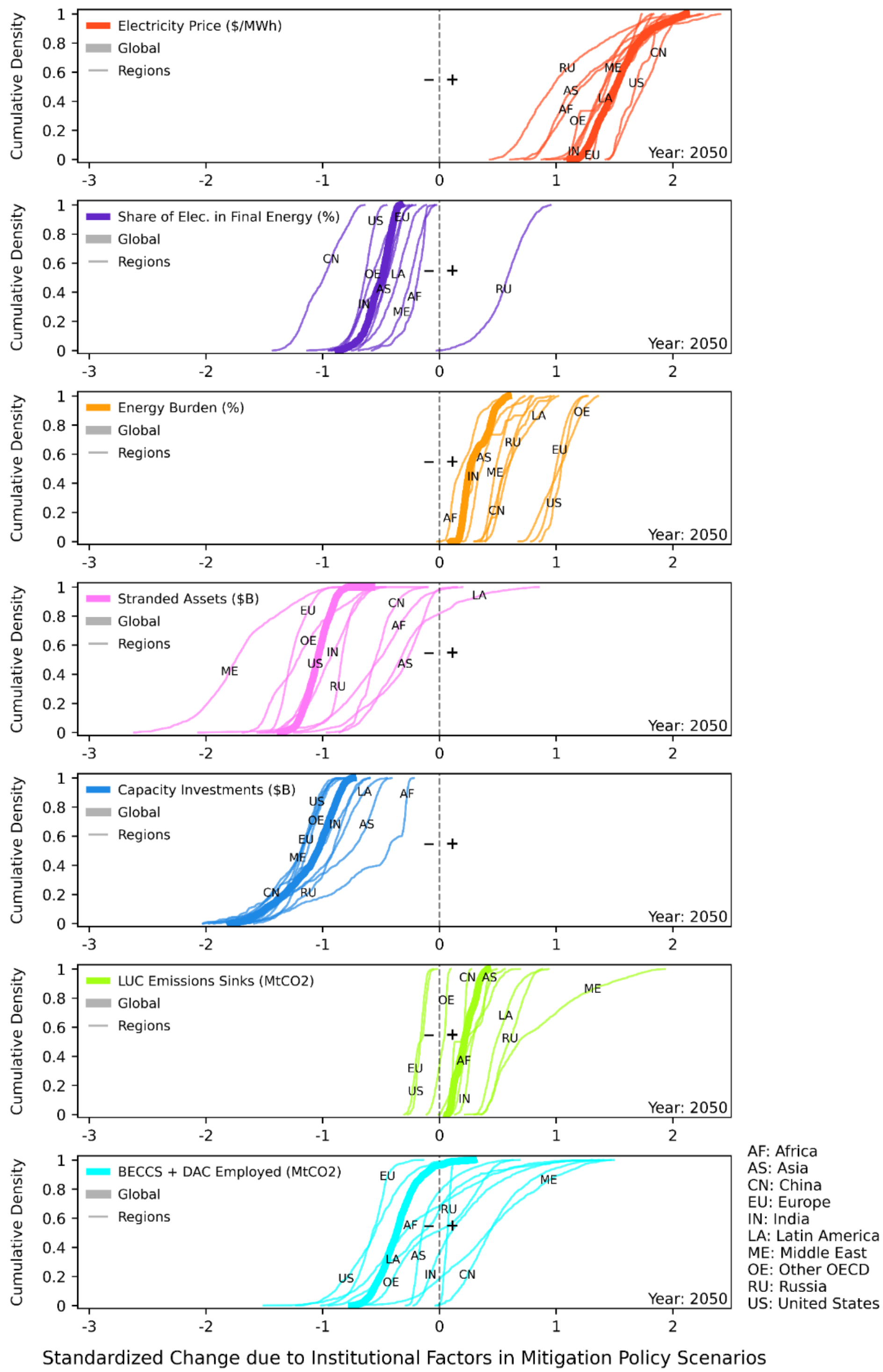

Figure S10: CDF plot showing standardized changes in the values of select metrics when investment costs are regionally and technologically differentiated in each scenario configuration (only showing scenarios with NDCs + LTS implemented). A curve lying entirely to the right (left) of zero implies that institutional factors always increase (decrease) that metric. Thicker lines refer to global weighted means, while thinner lines refer to ten aggregated global regions (legend at bottom right). Note that a steep CDF curve here suggests that varying this sensitivity results in a very consistent change in the outcome; it does not represent the underlying variability of the outcome itself.

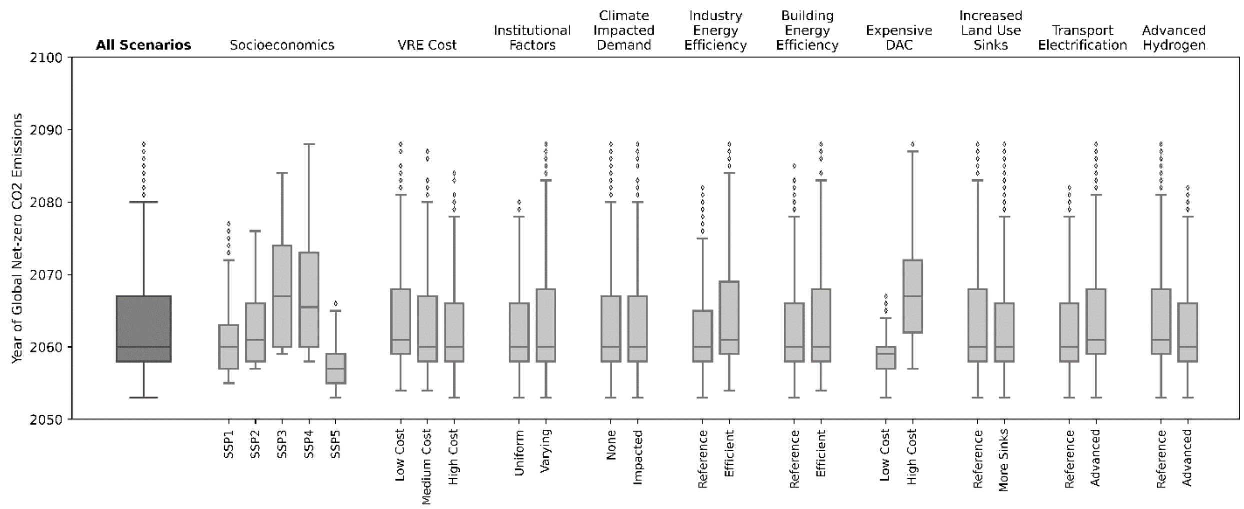

Figure S11: Year in which global net-zero CO_2 emissions is achieved across all realizations with national emissions pledges, split by scenario sensitivity. Visually, *Socioeconomic Factors* and *Direct Air Capture Cost* show the greatest variability, followed by *Industry Energy Efficiency* and *Cost of Wind and Solar* (VRE Cost). Net-zero year is determined by linear interpolation between GCAM's five-year timesteps.

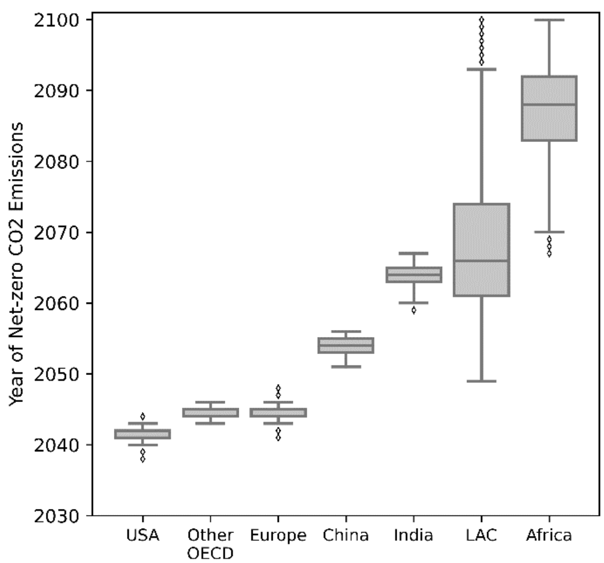

Figure S12: Year in which net-zero CO_2 emissions is achieved across aggregate regions, for all realizations with national emissions pledges. Russia, Asia, and Middle East do not reach net-zero in any simulation due to one or more countries within each region not reaching net-zero. For LAC, 93 realizations out of 3,840 do not reach net-zero by 2100. For Africa, 103 realizations out of 3,840 do not reach net-zero by 2100.