Empirically Estimating the Impact of Weather on Agriculture

Oscar Barriga-Cabanillas
University of California, Davis

Leah E.M. Bevis
Ohio State University

Anna L. Josephson
Purdue University

Jeffrey D. Michler
University of Saskatchewan

Aleksandr Michuda
University of California, Davis

Emilia Tjernström
University of Wisconsin

AAEA Annual Meeting, Chicago, 30 July - 1 August, 2017

Climate-Economy Literature

- Estimating the effect of weather on crop output has long been a concern in agricultural and applied economics (Wallace, 1920; Fisher, 1925; Stallings, 1961; Shaw, 1964; Oury, 1965).
- Publicly accessible remote sensing data has spurred renewed interest in understanding how weather effects economic behavior (Dell et al., 2014; Donaldson & Storeygard, 2016).
  - One branch focuses on using weather as a "natural experiment" for identifying causal effects.
  - Another branch focuses on quantifying the effects of climate on economic outcomes.

The Physical-Biological Nexus

- At least since the Morrow Plots were established (1876), crop science has accounted for weather's effects on crop output.
- Economists have little to add in understanding the Physical-Biological nexus.
- Where economists may have a comparative advantage is in understanding how human behavior can modify the bio-physical nexus.

The Physical-Biological-Economic Nexus

- *A priori*, there is no reason to believe that the quantitative effect of the physical world (rain, temperature, wind speed) on the biological world (crop growth) will not be different once the economic world is taken into consideration.
  - Essentially, the difference between using experimental field data and observational household data.
- The relevant unit of observation may be different.
- The relevant weather data may be different.
- The relevant inputs may be different.
- The functional form may be different.

The Long-Term Goal

- Define a set of relevant weather statistics and econometric “best practices” for estimating agricultural production using observational data from developing countries.

The Long-Term Goal

- Define a set of relevant weather statistics and econometric "best practices" for estimating agricultural production using observational data from developing countries.

We want to be able to respond to critiques like, "Strictly speaking, what matters is agricultural drought - the sufficiency of water availability for plant growth - not meteorological drought - commonly based on average annual rainfall. Unless you can measure rainfall during the most critical period of crop flowering, for maize that this occurs at roughly 1/2 the # of maturity days, and the period to maturity is 110 days, then you cannot measure the relevant drought."

The Goal for Today

- Begin to compare across statistics for each country, in order to:
  - Compare the effectiveness of each statistic
  - Compare which statistics are most effective for which countries.

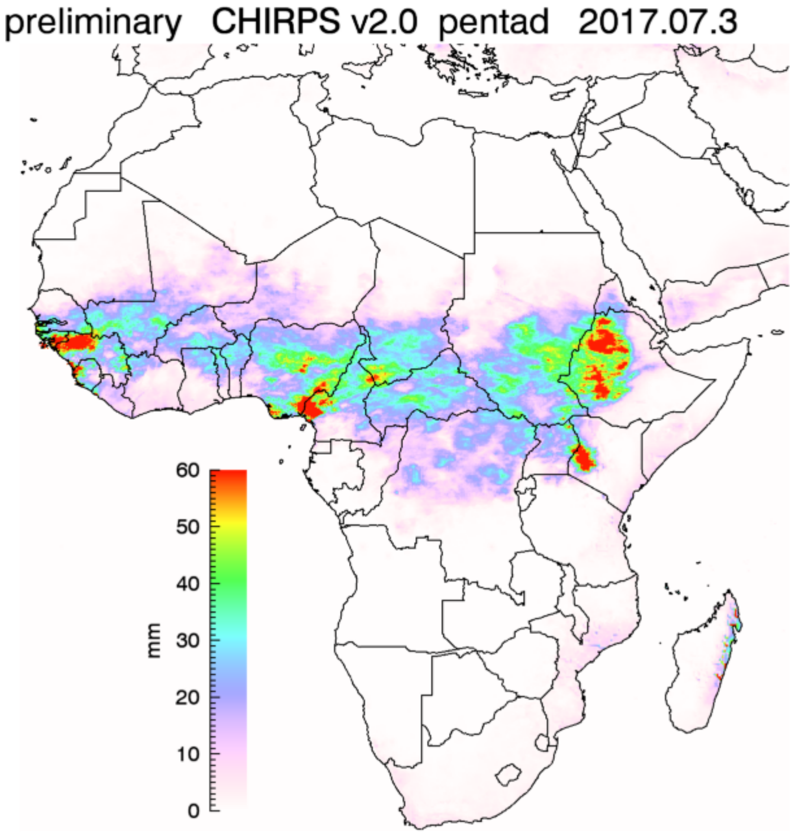
The Goal for Today

- Begin to compare across statistics for each country, in order to:
  - Compare the effectiveness of each statistic
  - Compare which statistics are most effective for which countries.

We want to be able to say, “It turns out that for most countries, all statistics provide basically the same predictive ability, though for COUNTRY XX, it was actually far better to use TEMP STAT YY than the other temperature stats.”

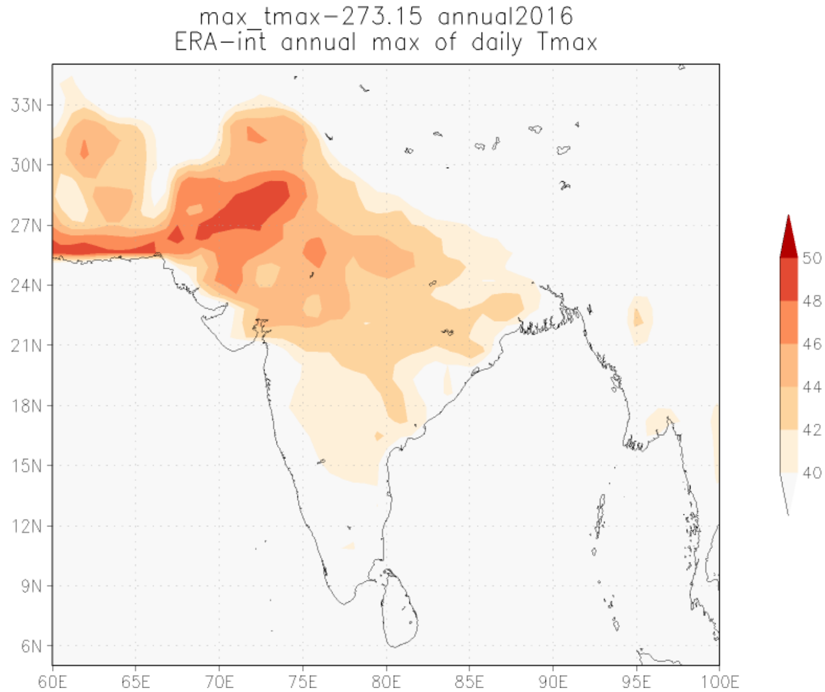
CHIRPS Data

- Climate Hazards Group InfraRed Precipitation with Station data
- 30+ year quasi-global rainfall dataset
- 1981 to the near present at up to daily frequency
- Incorporates satellite imagery at 0.05° with station data to create gridded rainfall time series



ECMWF Data

- European Centre for Medium-Range Weather Forecasts
- Provides accurate medium-range global weather forecasts out to 15 days and seasonal forecasts out to 12 months
- 1975 to the near present at up to daily frequency
- Up to horizontal resolution of T1279 spectral wave truncation, and .01 hPa vertical resolution



Setup

- Developed Stata command to run on different datasets.
- Generates the same set of independent variables.
  - Creates consistency in implementation.
- Use nationally defined main rainy season for rainfall data.
- Use FAO defined crop growing season for temperature data.

Weather Variables

- Various moments and descriptive statistics:
  - Mean, median, standard deviation, skewness, min, max,
  - As well as deviations of annual statistics from long-run as z-scores.

$$z_{jt} = \left| \frac{r_{jt} - \bar{r}_j}{\sigma_{r_j}} \right|$$

Rain-specific Variables

- Total (cumulative) seasonal rainfall.
- Number and percentage of days without rain in a year (and as deviations from long-run in z-scores).
    - Important for capturing drought effects.

Temperature-specific Variables

- Growing Degree Days: number and percentage of days within some lower and upper bound in a given year (and as deviations from long-run in z-scores).

$$GDD = \mathcal{I}\left(T_l \leq \frac{T_{max} + T_{min}}{2} \leq T_h\right)$$

- As in Schlenker & Roberts (2009), number and percentage of days in a set number of quantile bins (per-bin deviations from long-run in z-scores).

Hurdles and Lessons Learned

- Standardization can be difficult across data sets and across crops within data sets.
  - Do you use the same definition of rainy season and growing season? Or local, crop specific definitions?
  - How do you deal with different levels of aggregation and accuracy in GPS data?
  - Do you use the same set of inputs or whatever inputs are available?
  - The issue of measurement error, both in satellite data and production data, still is not very well understood.
- Computational power is increasingly important.
  - Large gains from using multi-processing/ parallelization in generating variables.

Data Sets

- **Bangladesh & India: 7 years, 5 crops, 18,889 observations**
- **India: 25 years, 13 crops, 16,823 observations**
- **Kenya: 3 years, 1 crop, 3,459 observations**
- Zimbabwe: 4 years, 5 crops, 7,837 observations
- **Ethiopia: 6 years, 10 crops, 6,090 observations**
- Uganda: 2 years, 6 crops, 6,704 observations
- Nicaragua: 3 years, 5 crops, 4,800 observations

Cumulative Seasonal Rainfall

	CHIRPS	ECMWF	Obs.	Farms	Years
Bangladesh & India	1,056	1,076	6,316	787	5
	(319.4)	(283.8)			
India	597	596	705.1	244.6	18
	(248.2)	(248.2)			
Ethiopia	504.1	485.4	3,521	938	4
	(169.3)	(358.8)			
Kenya	127.8	487.07	3,459	1,153	3
	(215.8)	(845.2)			

Generic Specification

The basic model specification, which comes from Deschênes and Greenstone (2007), is

$$Y_{ht} = \alpha_h + \gamma_t + \mathbf{X}_{ht}\pi + \sum_i \beta_i f_i(W_{iht}) + u_{ht}$$

- Y_{ht} is yield (kg/ha) of a crop by household h in year t.
- α_h is a household fixed effect.
- γ_t is a year fixed effect.
- $\mathbf{X}_{ht}$ represents a vector of input variables, specific to the data set.
- $f_i(W_{iht})$ represents the functional form (f_i) of our weather variables of interest (W_{iht}), discussed in more detail below, where i represents a particular measurement of weather.
- u_{ht} is an error term.

Variables of Interest

- **Mean** - first moment of daily rain/temp over the season
- **Z-score** - of the mean
- **Median** - median daily rain/temp over the season
- **Standard deviation** - second moment of daily rain/temp over the season
- **Skew** - third moment of daily rain/temp over the season
- **Maximum** - the greatest value of daily rain/temp during the season
- **Total** - cumulative daily rain over the season
- **z-score** - of the total
- **No rain days** - number of days in the season without rain
- **No rain share** - percentage of days in the season without rain
- **GDD** - number of days where temp was between 8° to 32° C
- **Temperature bins** - the percentage of days that fall in each temperature quintile during the season

Linear Specification

- Base case, we include each of the individual climate / weather measures, with simple households and year fixed effects.

$$Y_{ht} = \alpha_h + \gamma_t + \beta W_{ht} + u_{ht}$$

- Add various input variables, as well as the climate / weather measures, with household and year fixed effects.

$$Y_{ht} = \alpha_h + \gamma_t + \mathbf{X}_{ht}\pi + \beta W_{ht} + u_{ht}$$

Quadratic Specification

- Base case, we include each of the individual climate / weather measures and their squares, with simple households and year fixed effects.

$$Y_{ht} = \alpha_h + \gamma_t + \beta_1 W_{ht} + \beta_2 W_{ht}^2 + u_{ht}$$

- Add various input variables, as well as the climate / weather measures and their squares, with household and year fixed effects.

$$Y_{ht} = \alpha_h + \gamma_t + \mathbf{X}_{ht}\pi + \beta_1 W_{ht} + \beta_2 W_{ht}^2 + u_{ht}$$

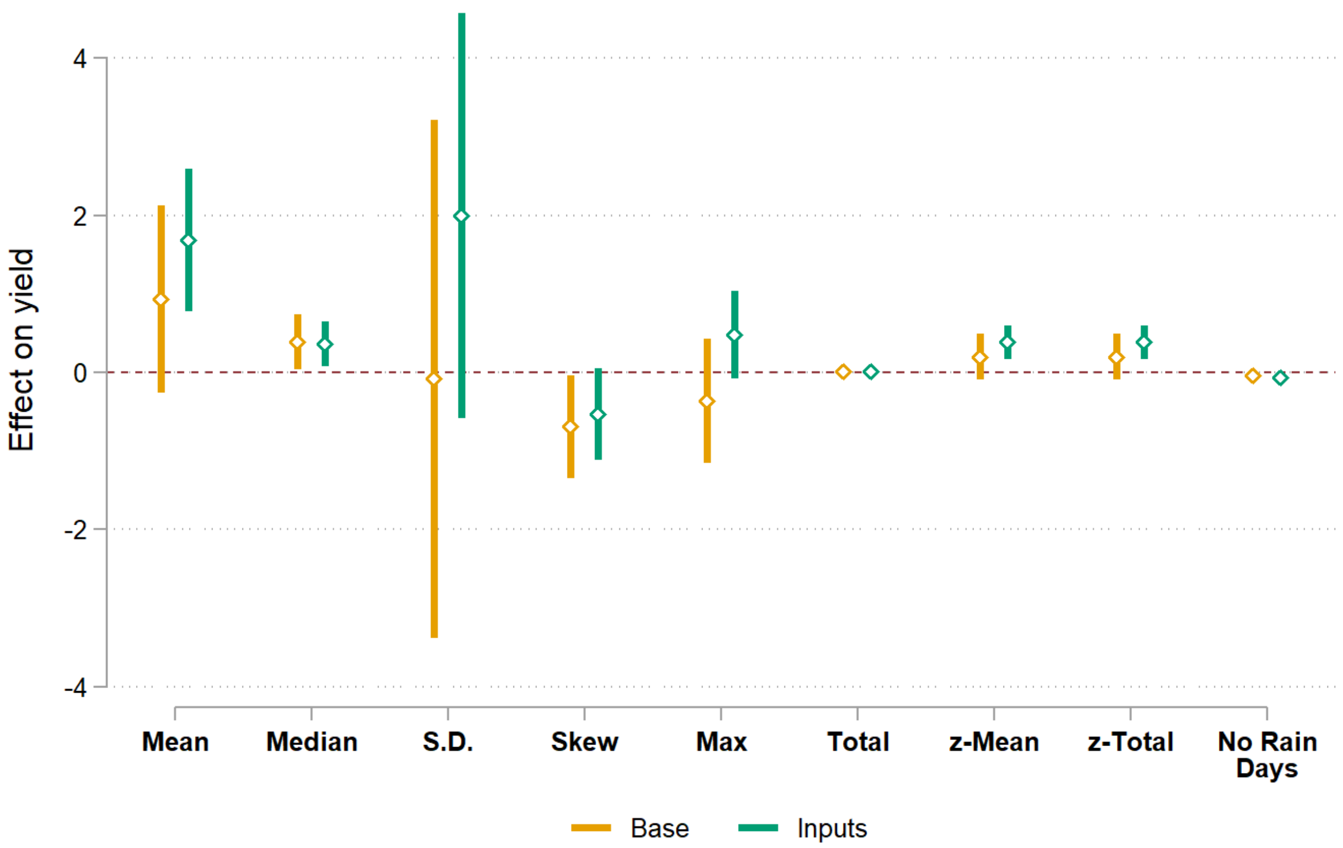
Combined Specification

We also combine various measures of rainfall, temperature, and rainfall with temperature.

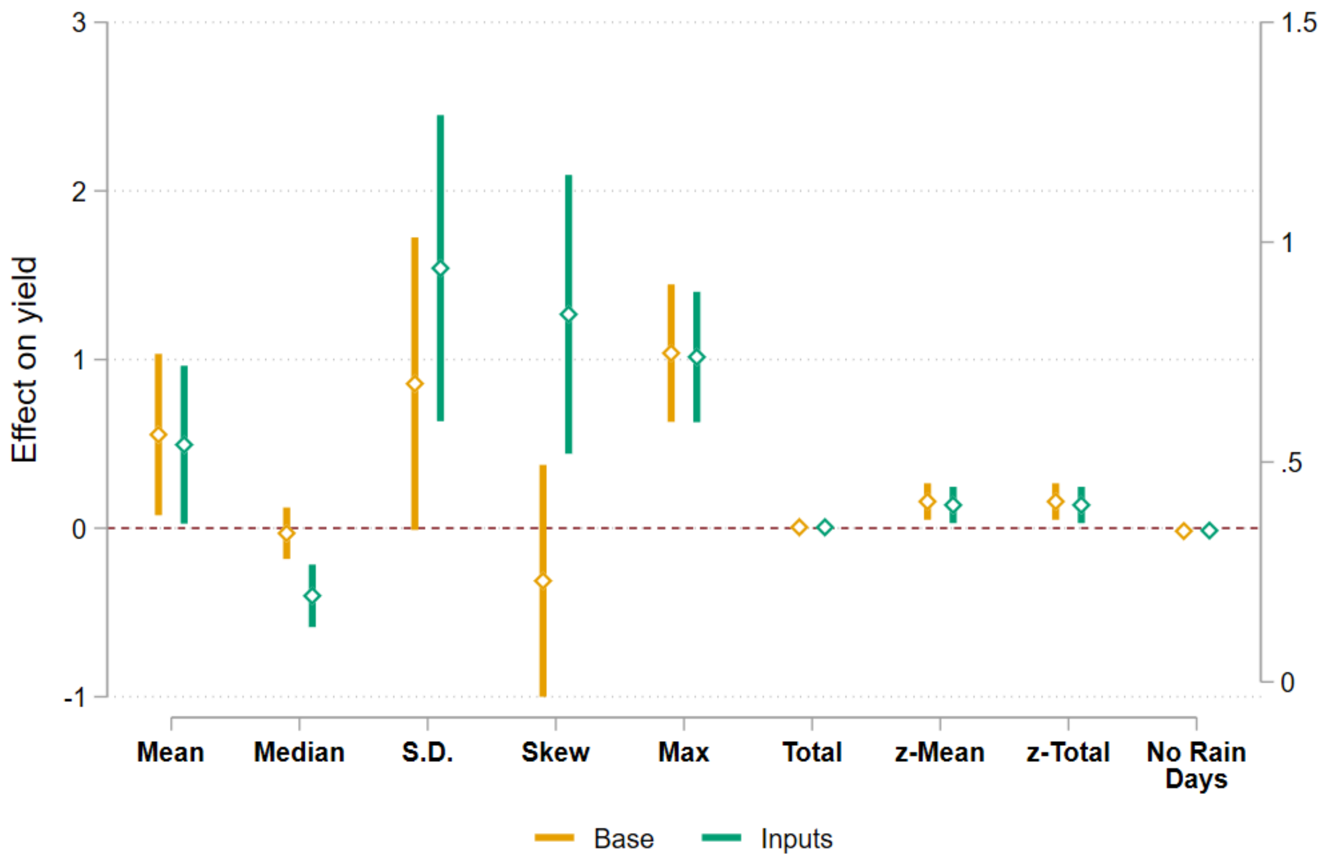
- Mean and standard deviation of both rain and temp
- Mean, standard deviation, and skew of both rain and temp
- Median of rain and temp
- Maximum rain and temp
- Z-score of mean rain and mean temp
- Number of no rain days and GDD

Bangladesh & India Rice - CHIRPS

India Sorghum - CHIRPS

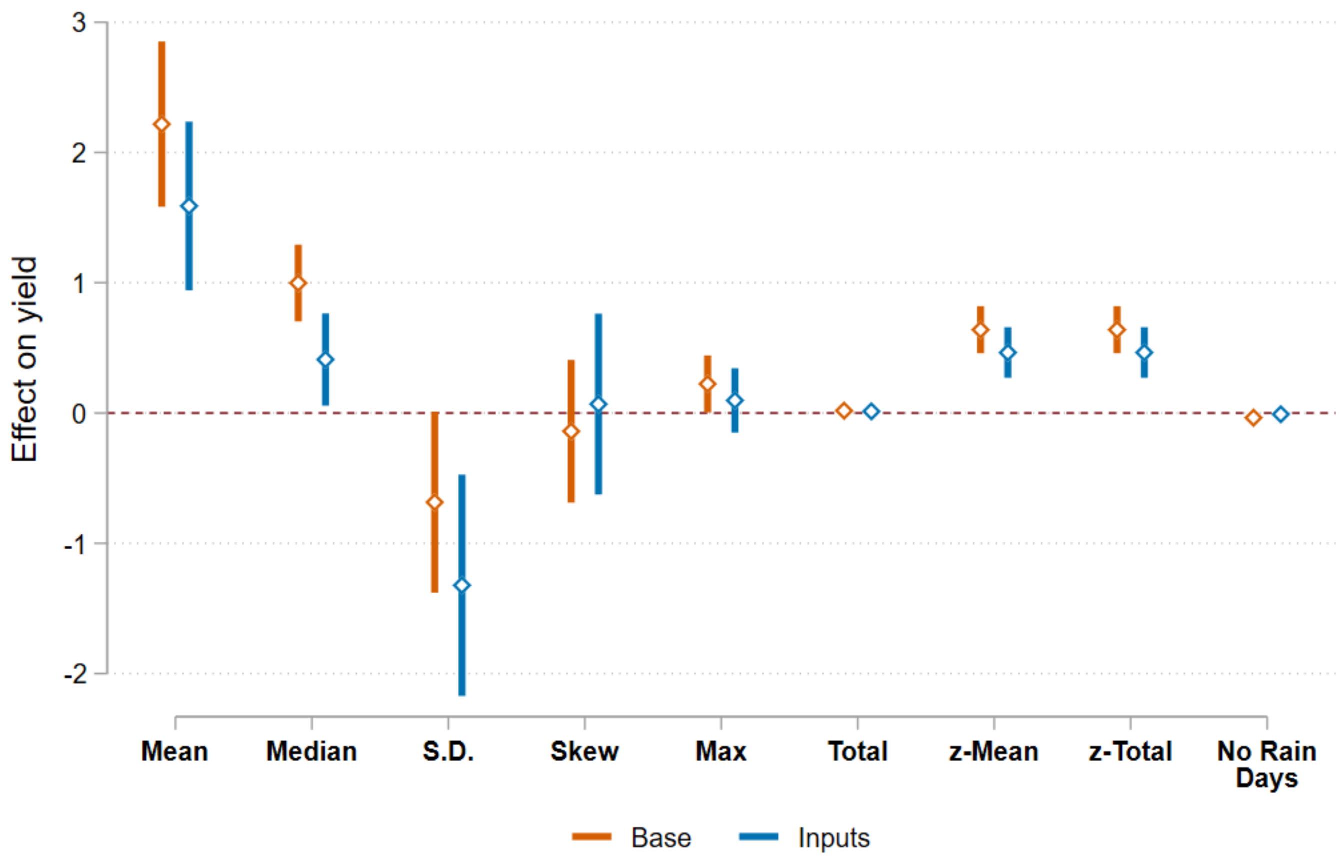


Ethiopia Tef - CHIRPS

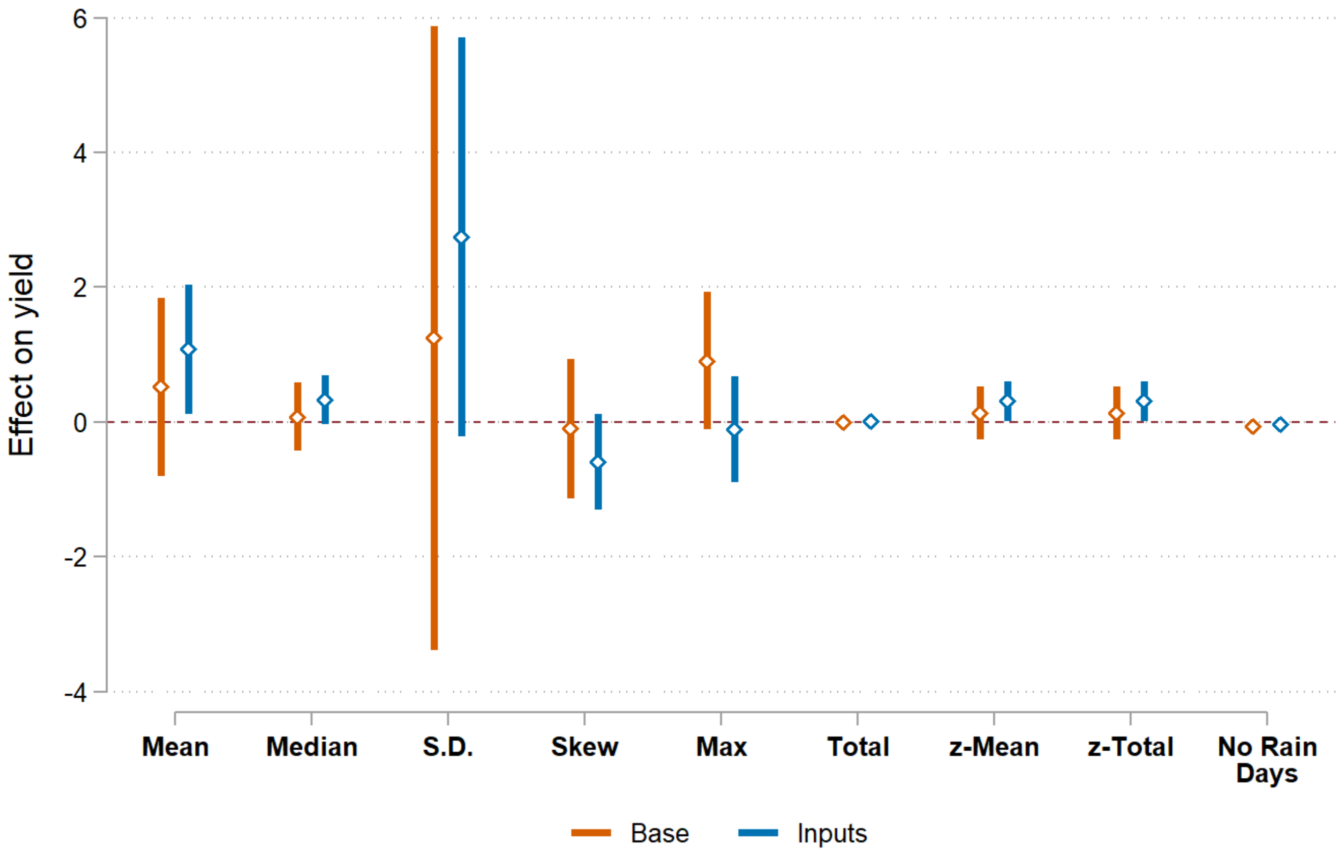


Kenya Maize - CHIRPS

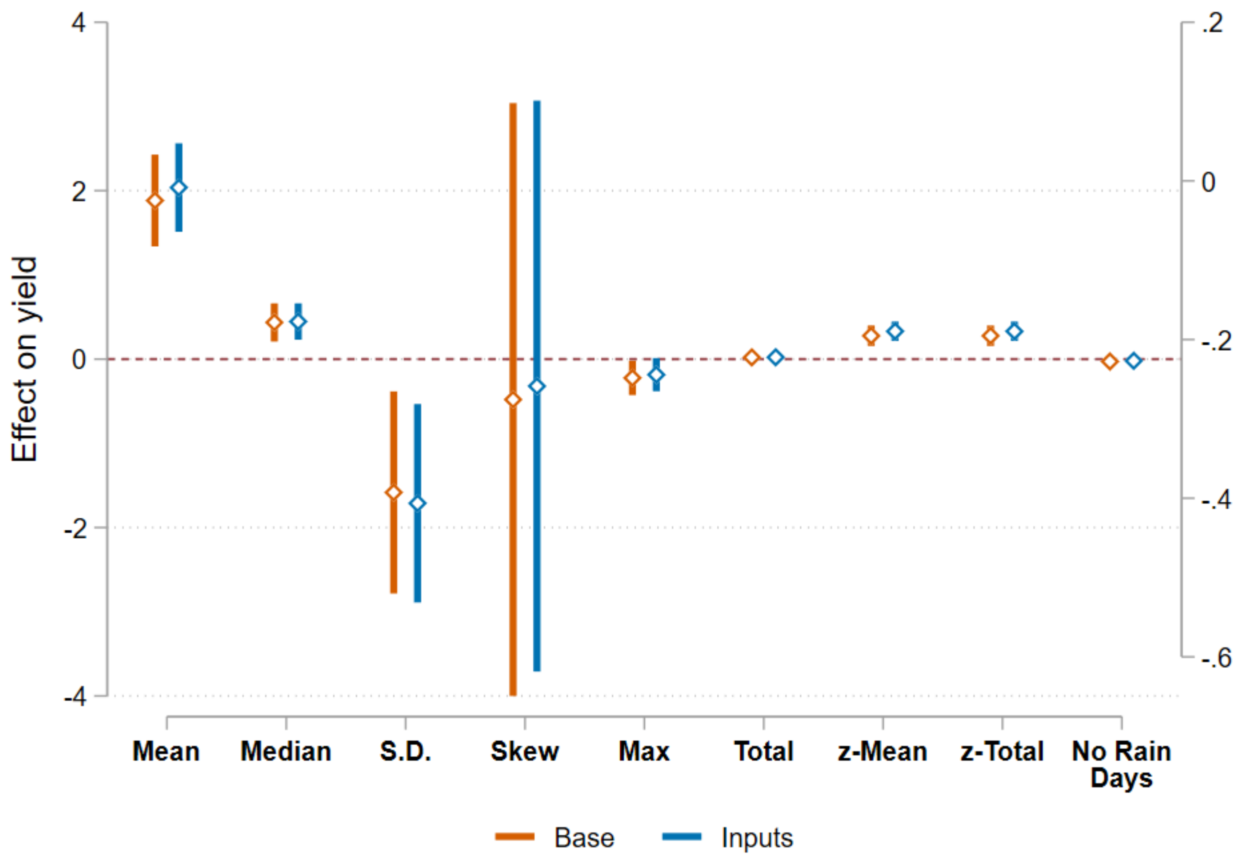
Bangladesh & India Rice - ECMWF



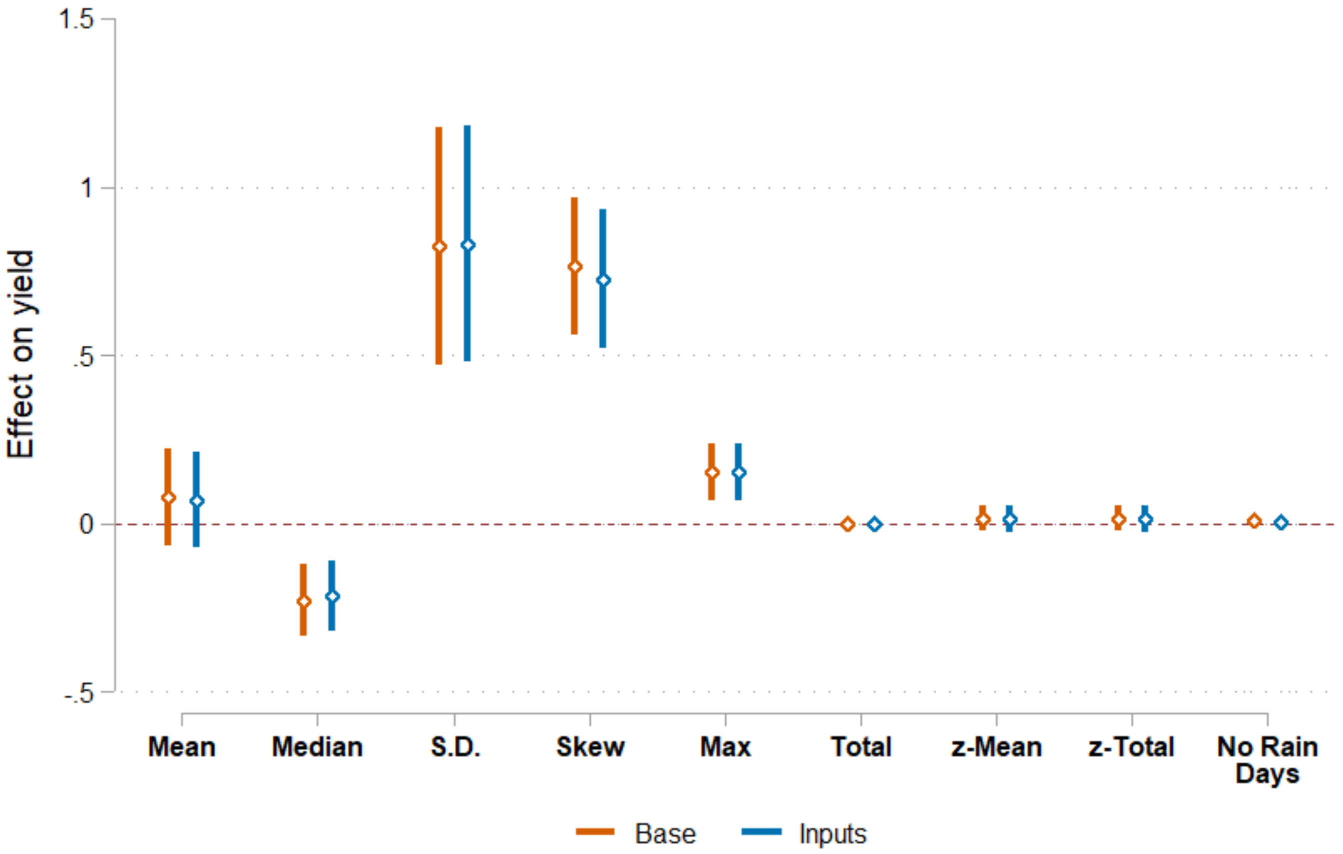
India Sorghum - ECMWF



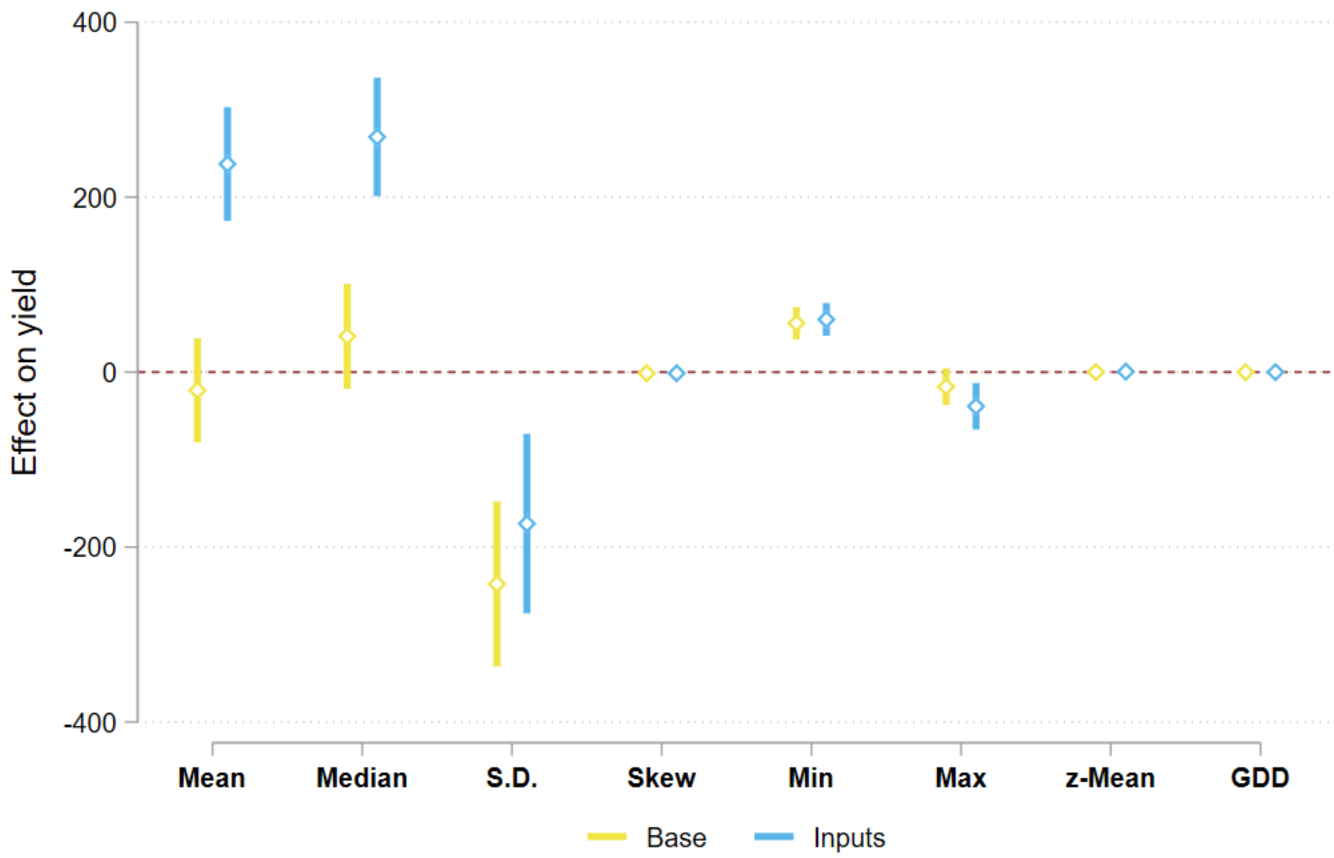
Ethiopia Tef - ECMWF



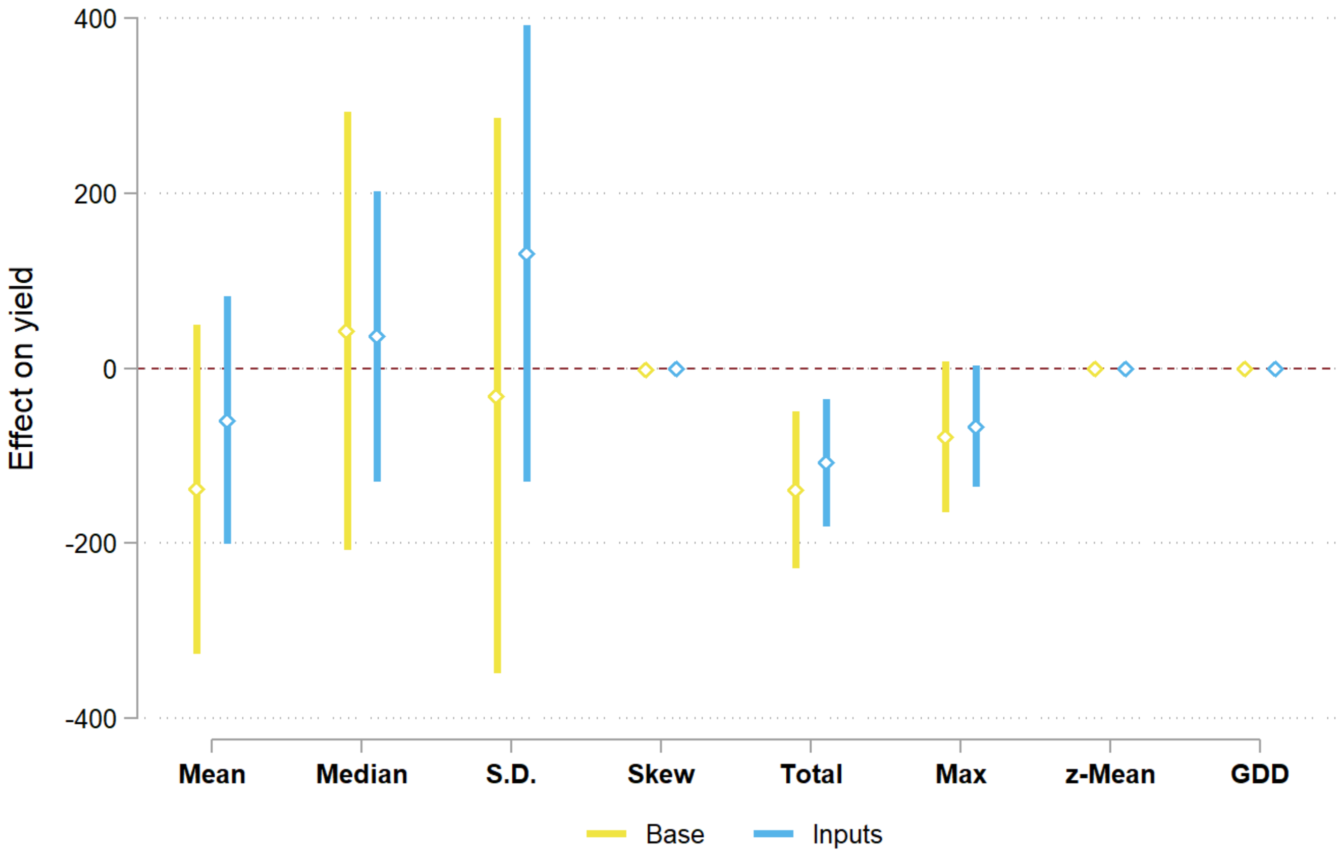
Kenya Maize - ECMWF



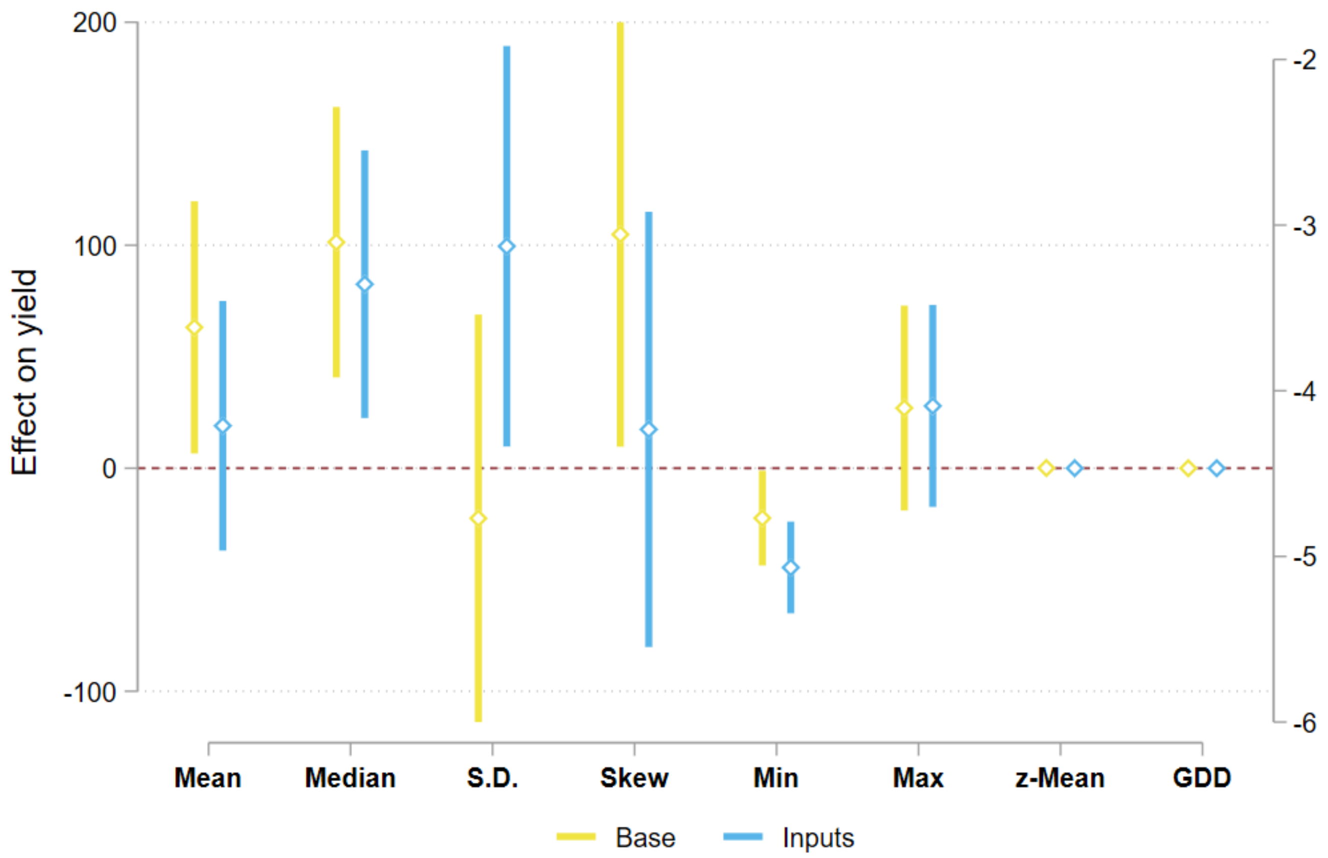
Bangladesh & India Rice - Temperature



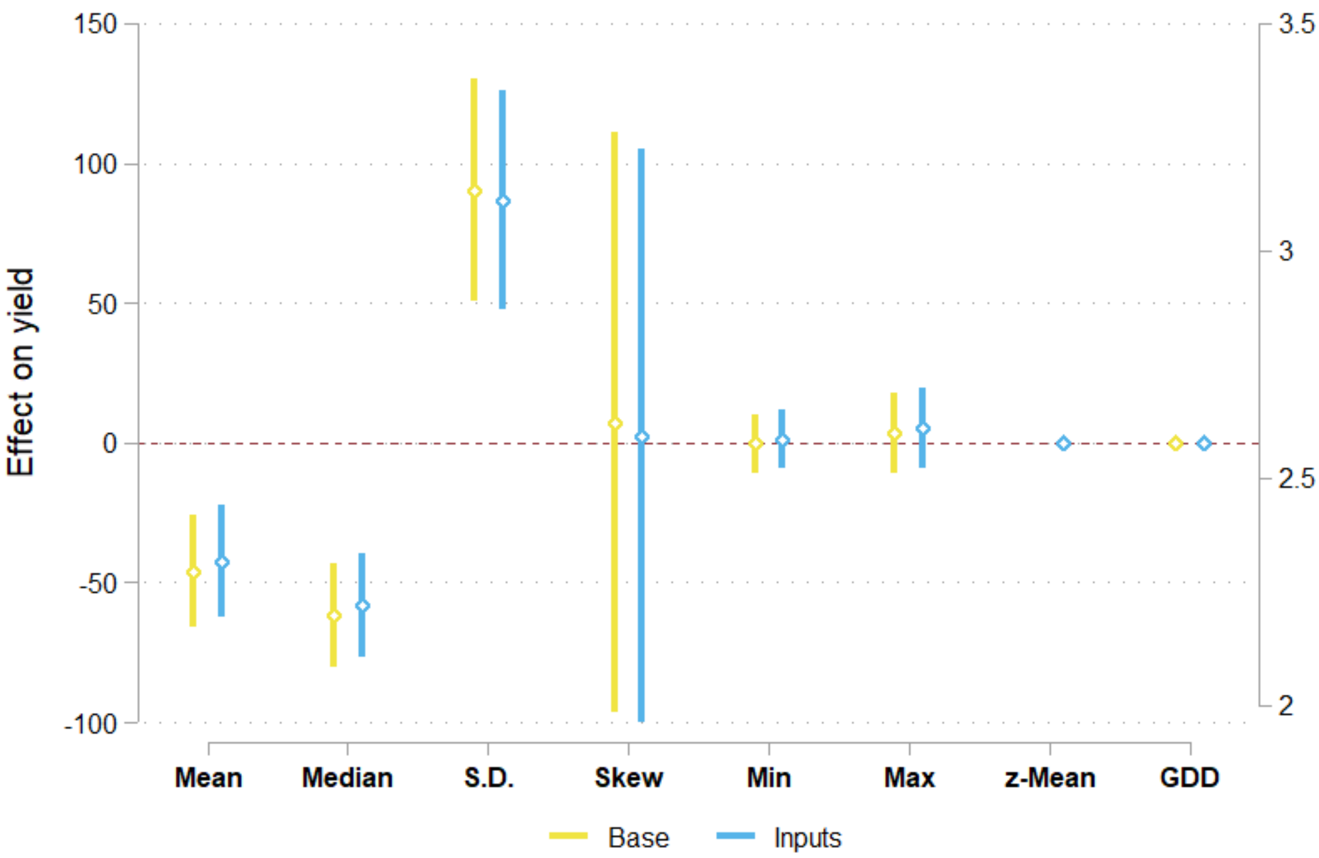
India Sorghum - Temperature



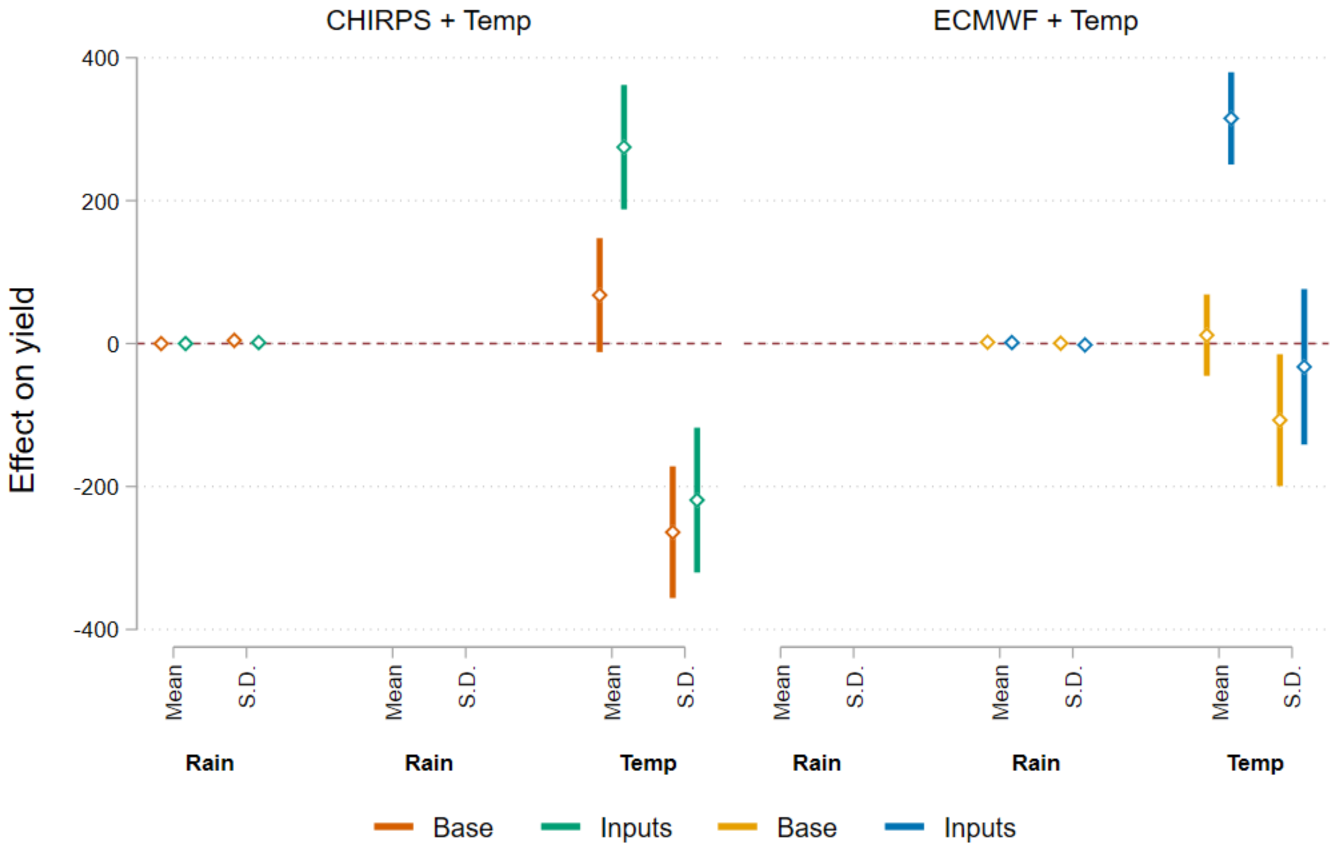
Ethiopia Tef - Temperature



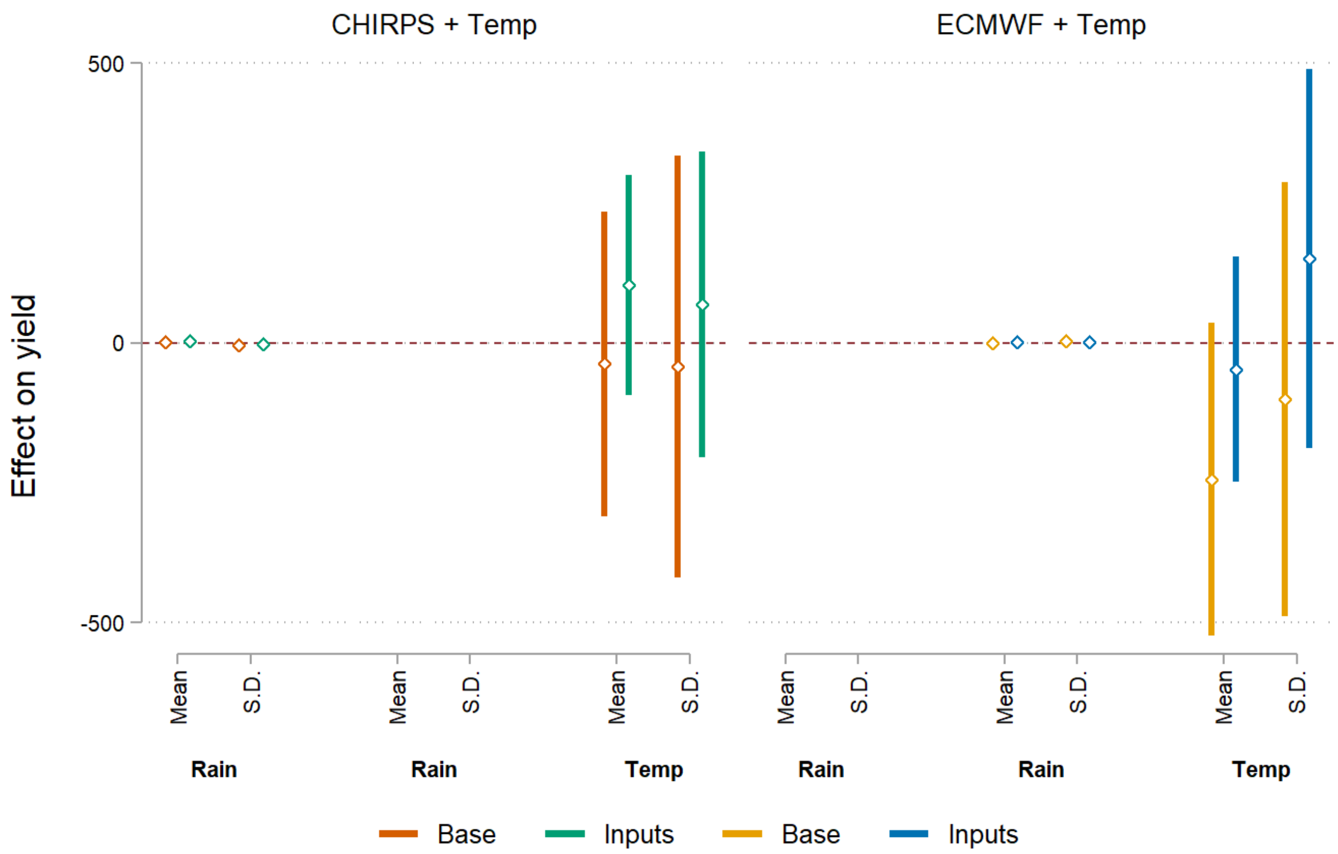
Kenya Maize - Temperature



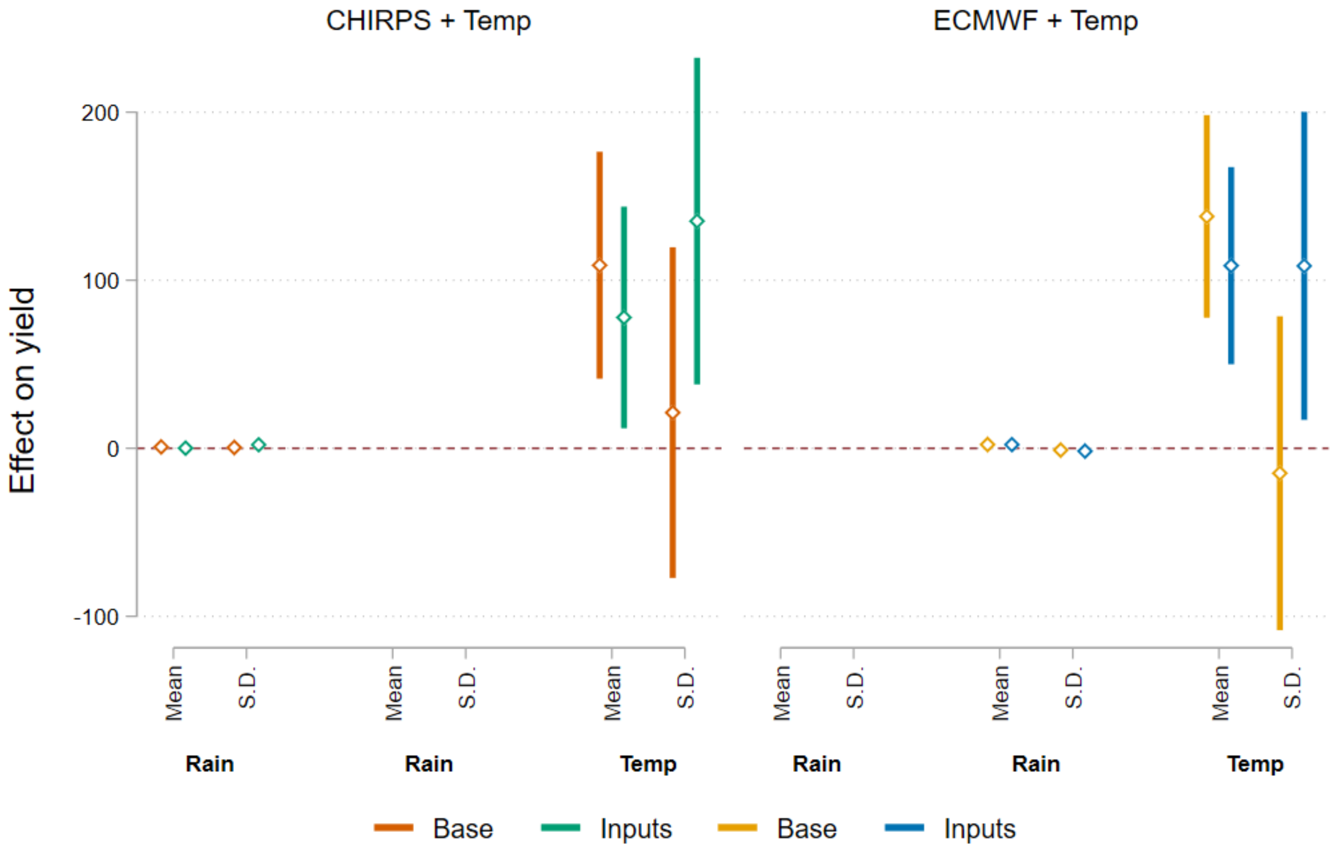
Bangladesh & India Rice - Mean + S.D.



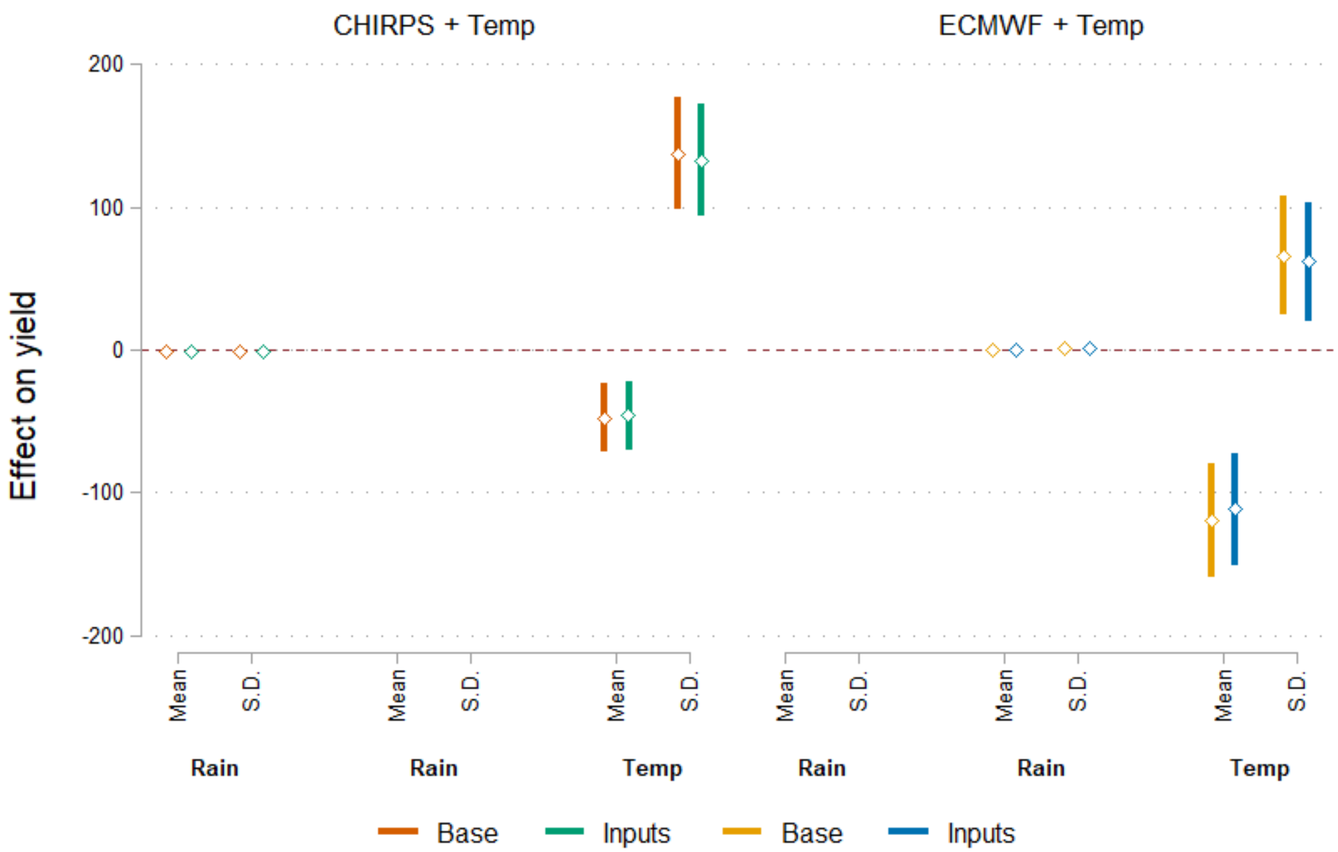
India Sorghum - Mean + S.D.



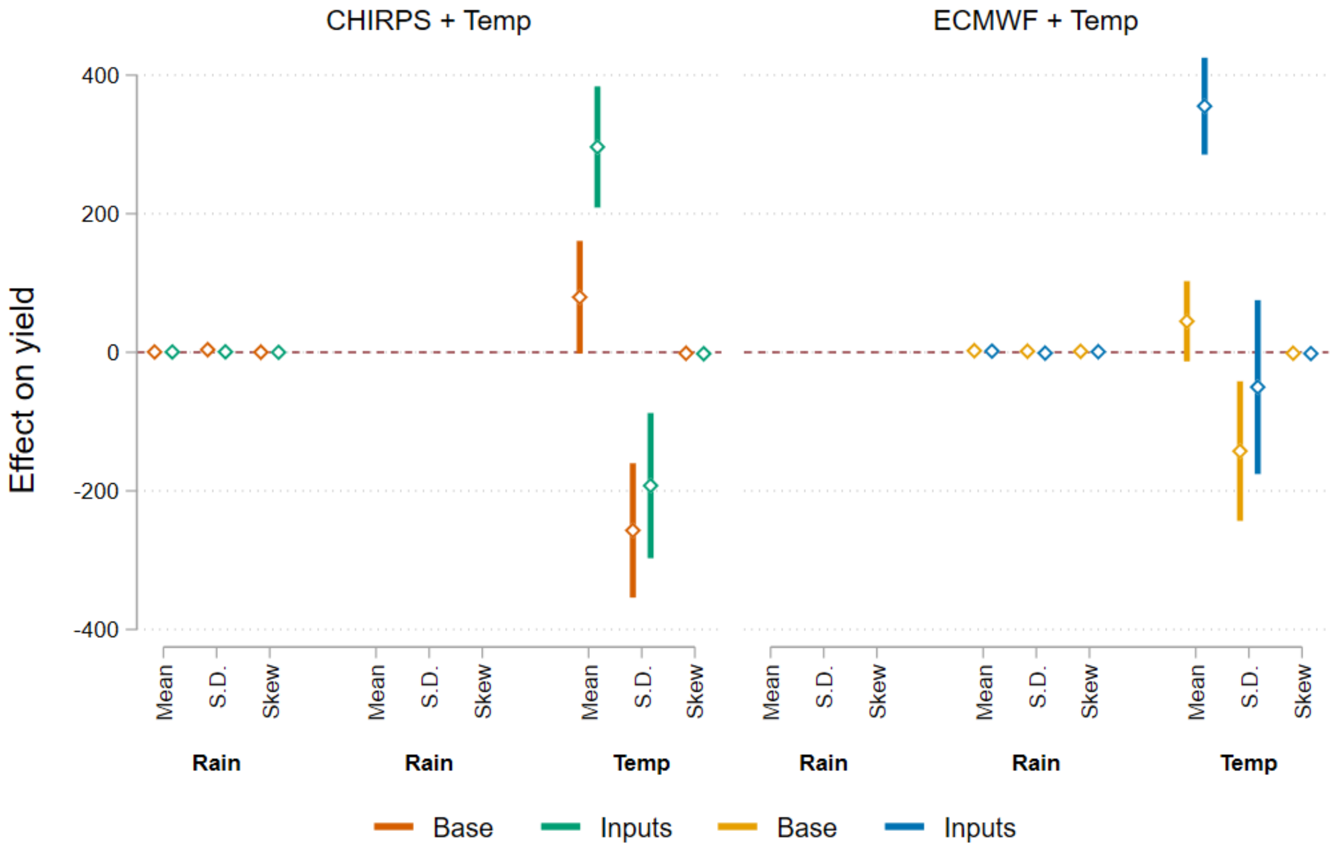
Ethiopia Tef - Mean + S.D.



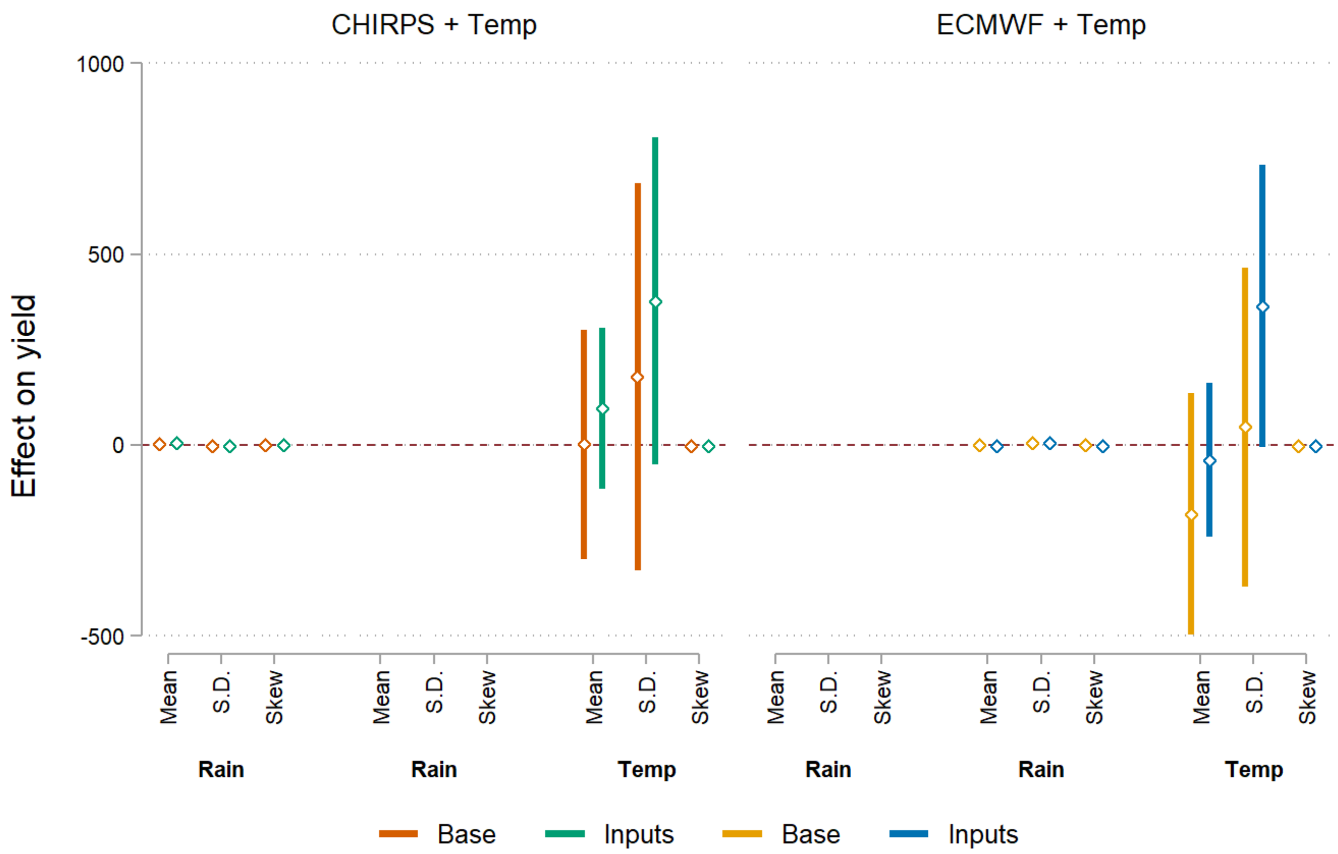
Kenya Maize - Mean + S.D.



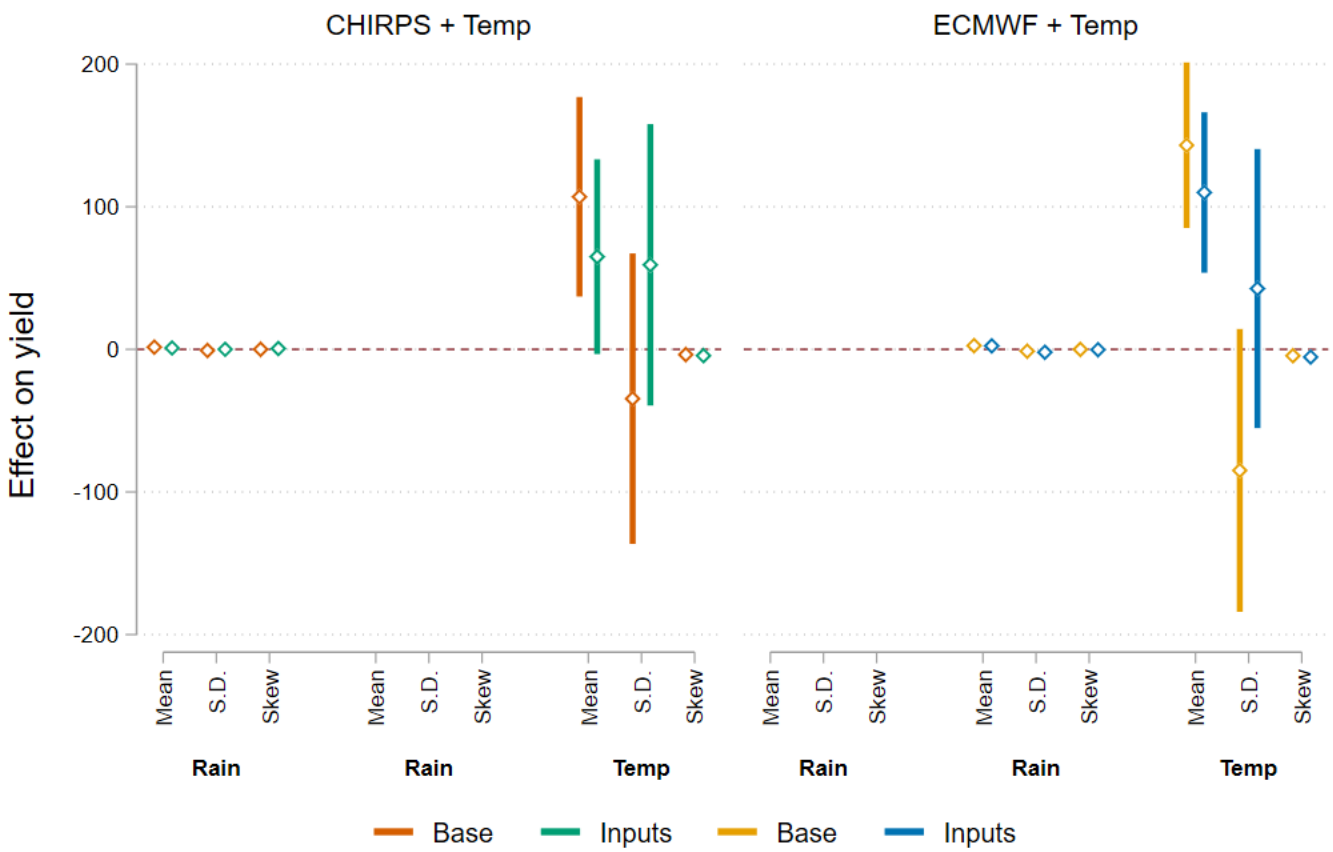
Bangladesh & India Rice - Mean + S.D. + Skew



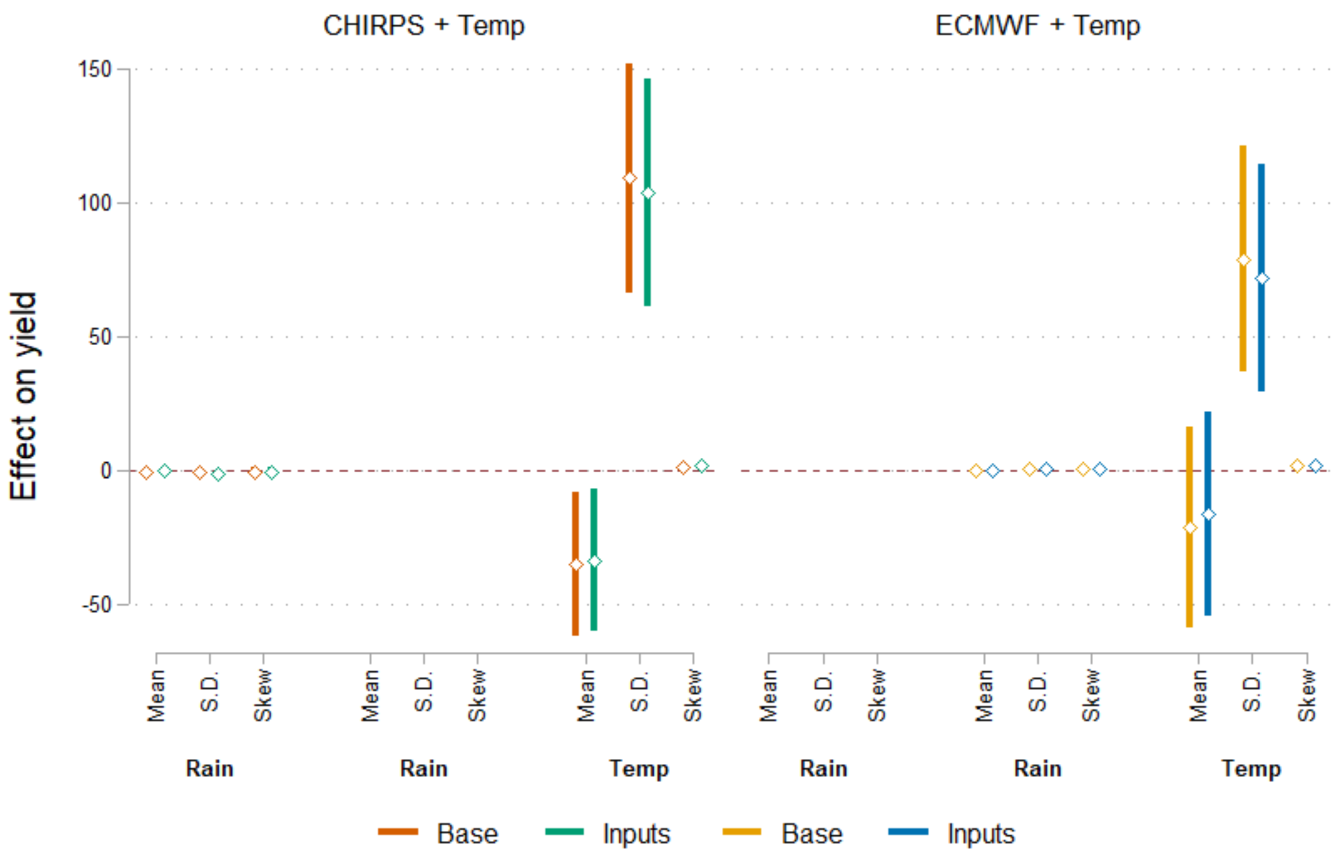
India Sorghum - Mean + S.D. + Skew



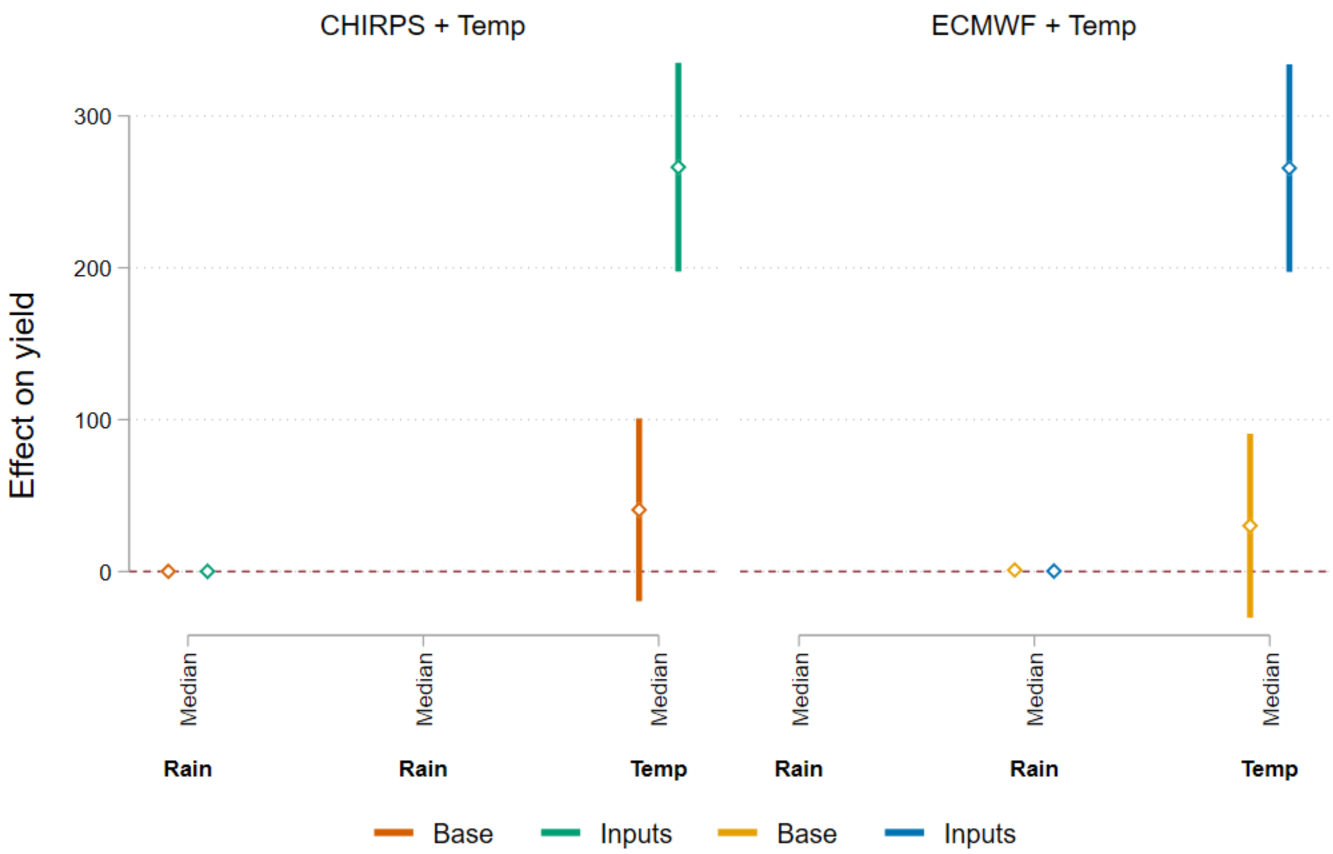
Ethiopia Tef - Mean + S.D. + Skew



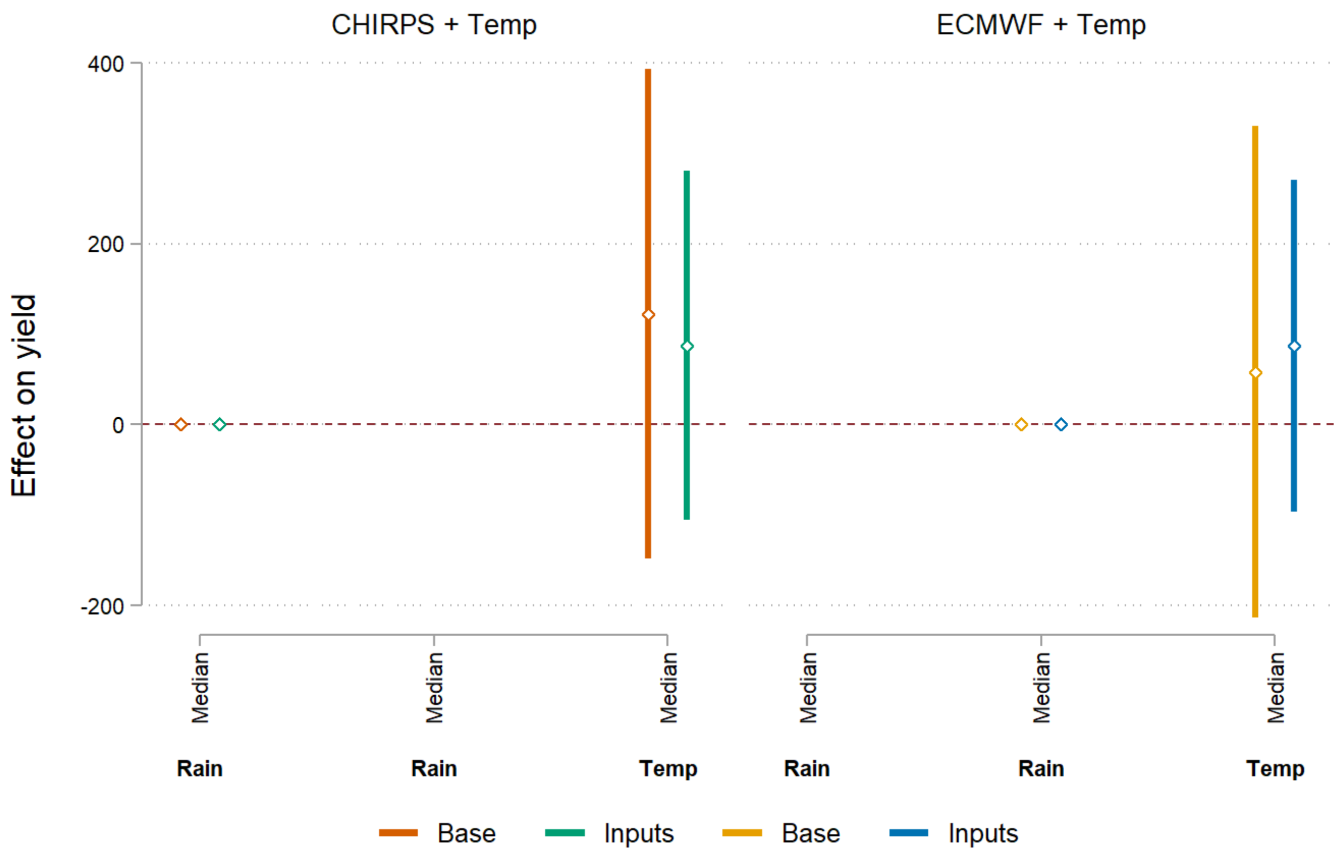
Kenya Maize - Mean + S.D. + Skew



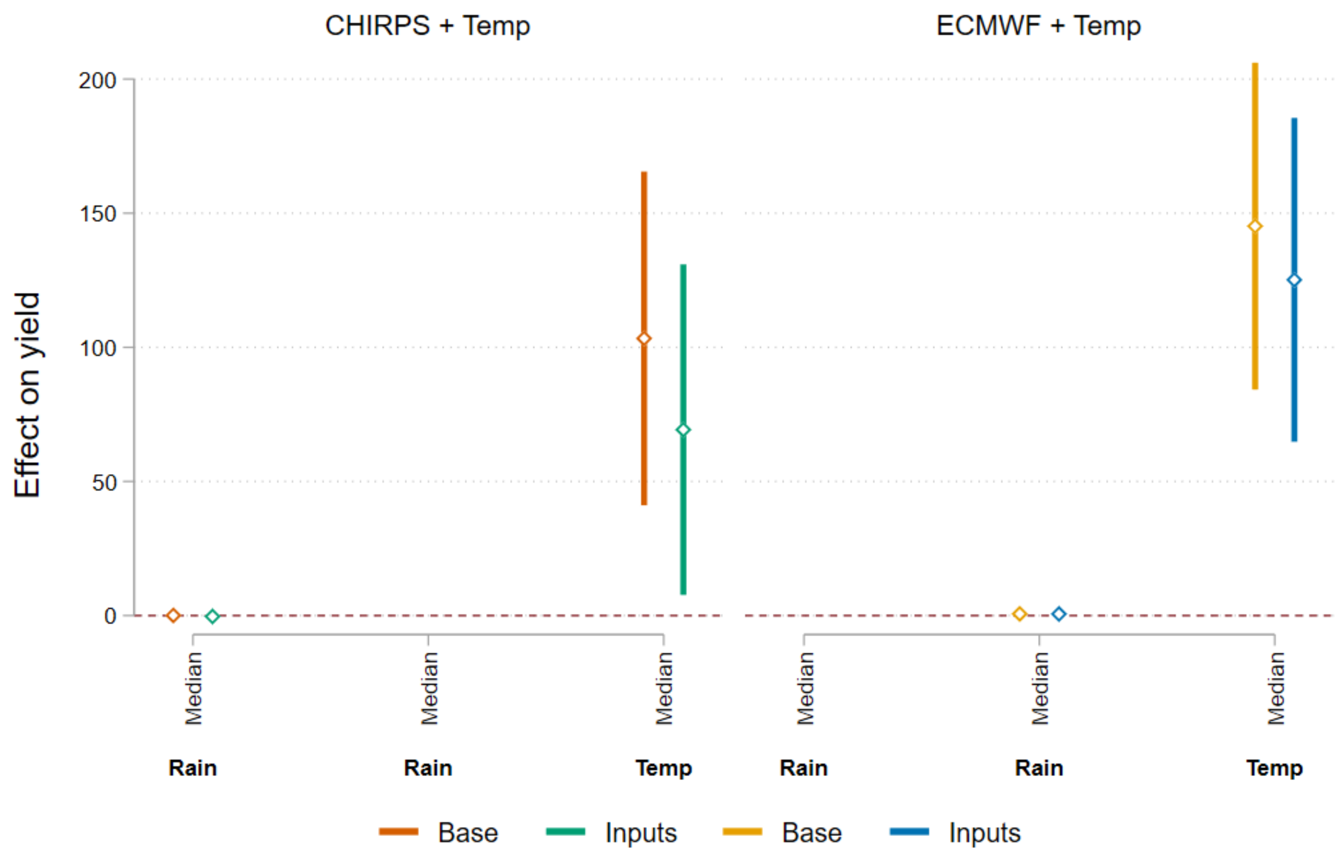
Bangladesh & India Rice - Median



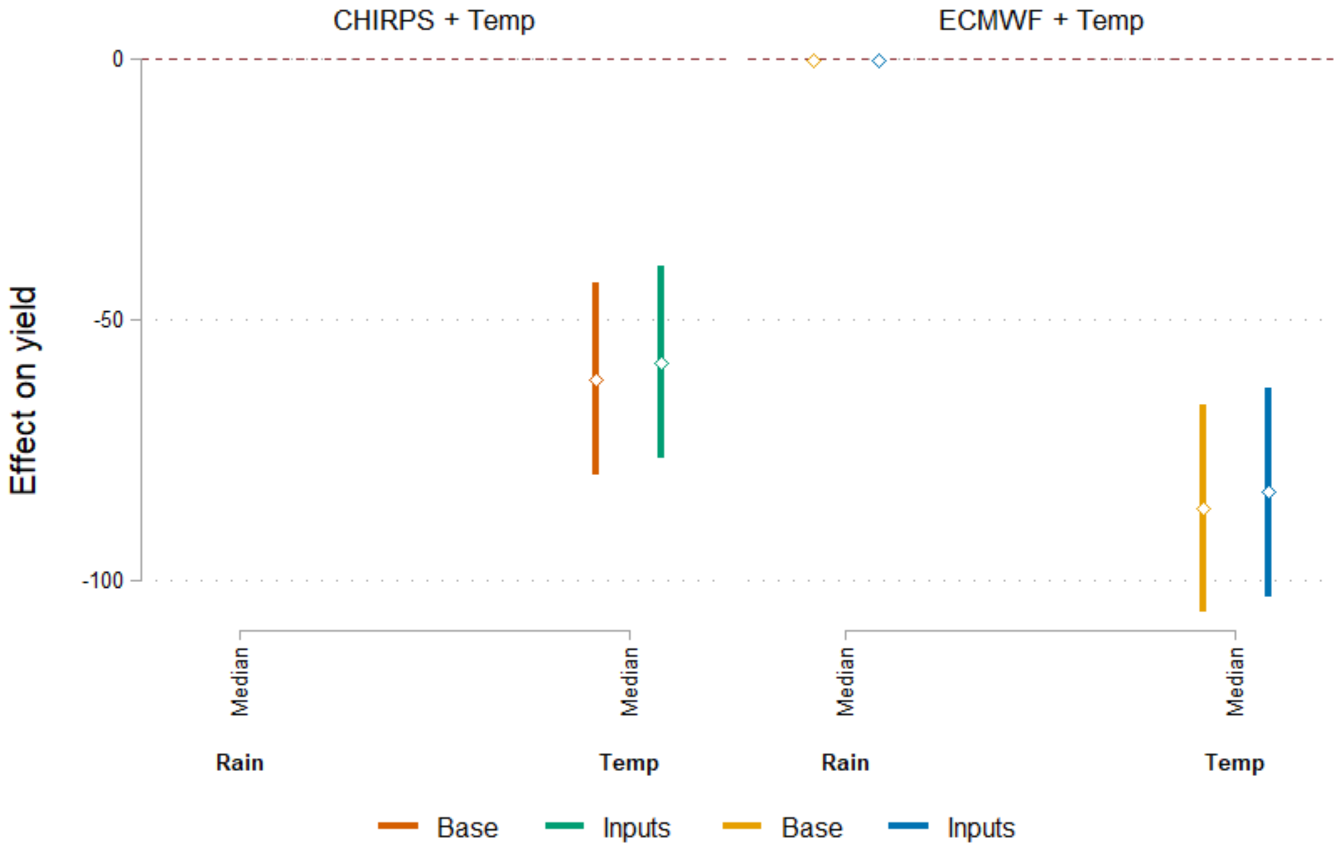
India Sorghum - Median



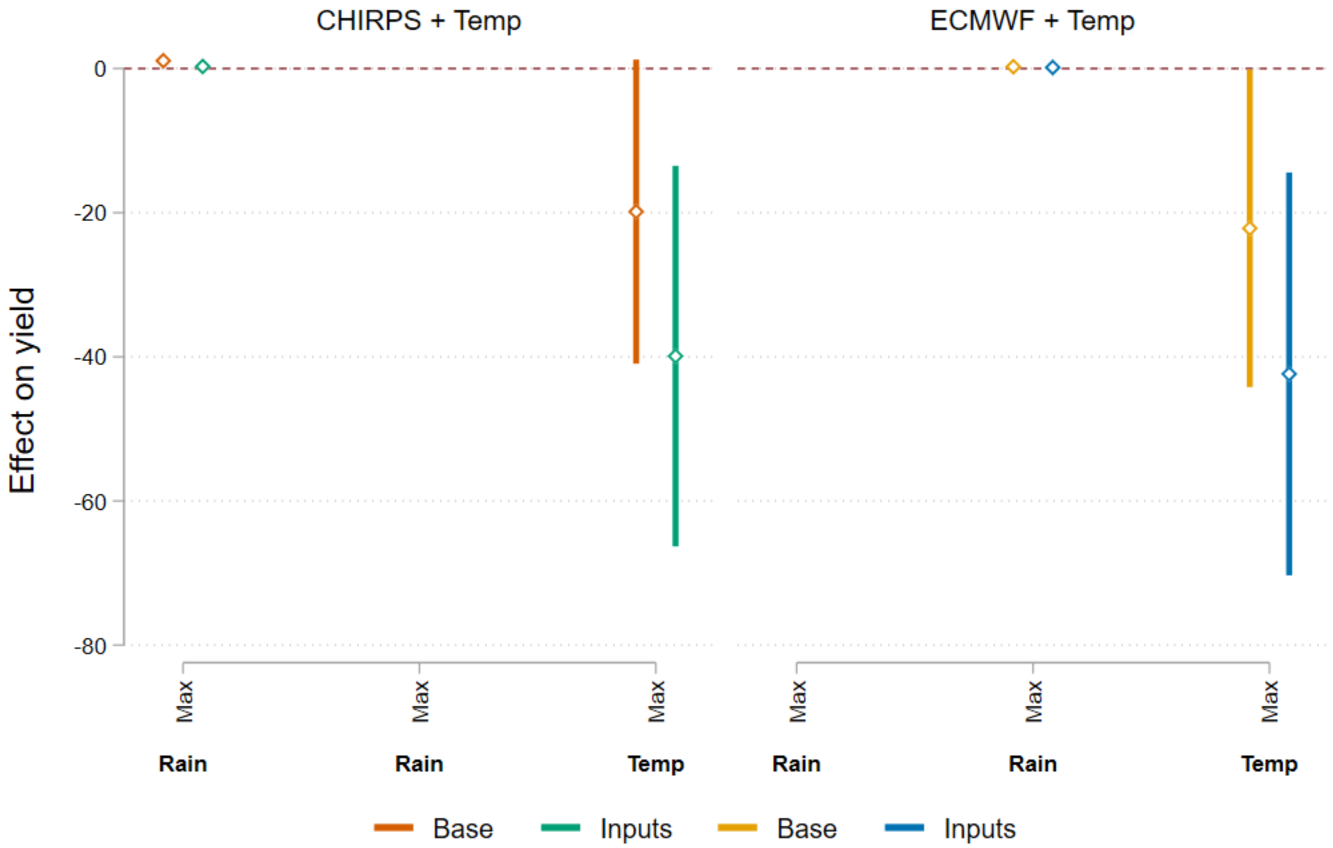
Ethiopia Tef - Median



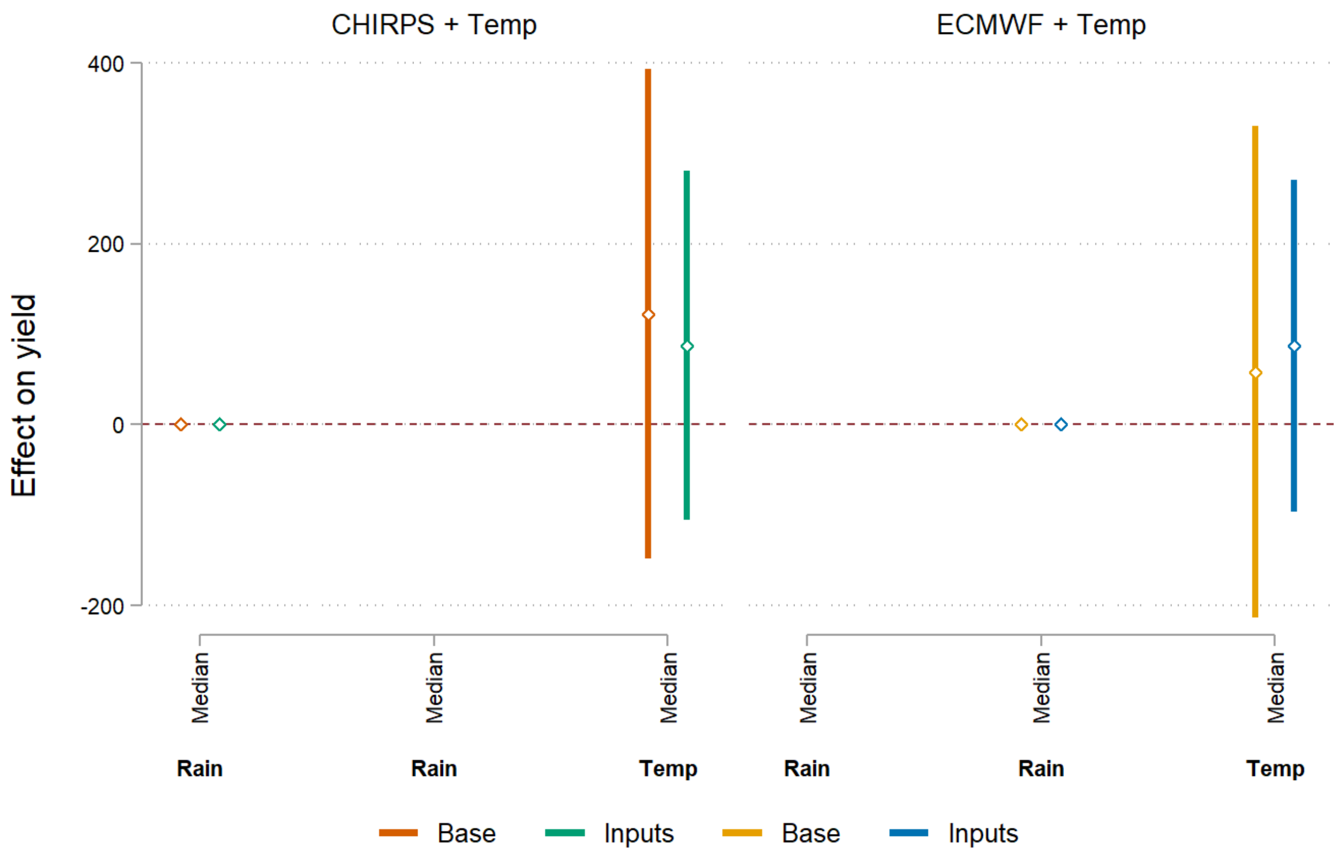
Kenya Maize - Median



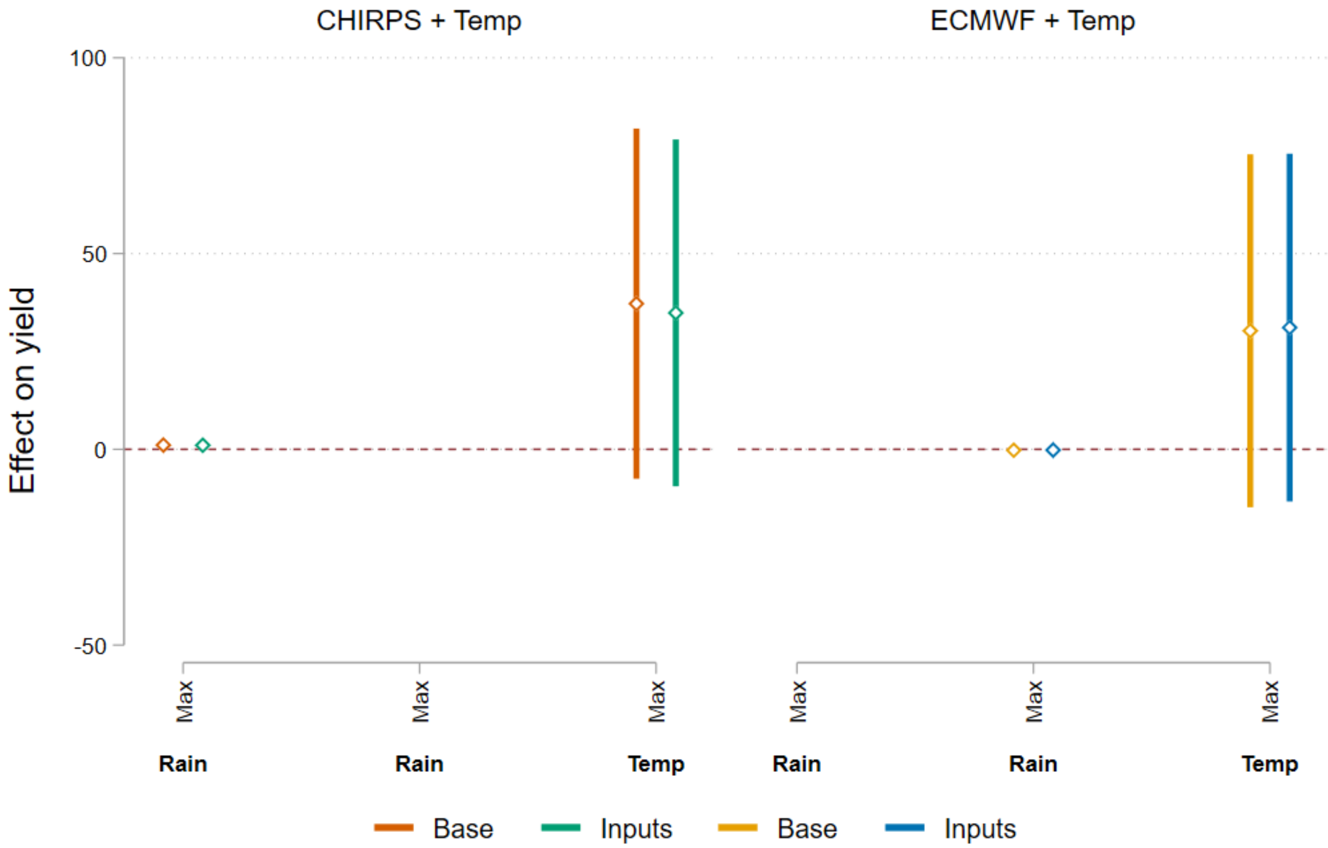
Bangladesh & India Rice - Maximum



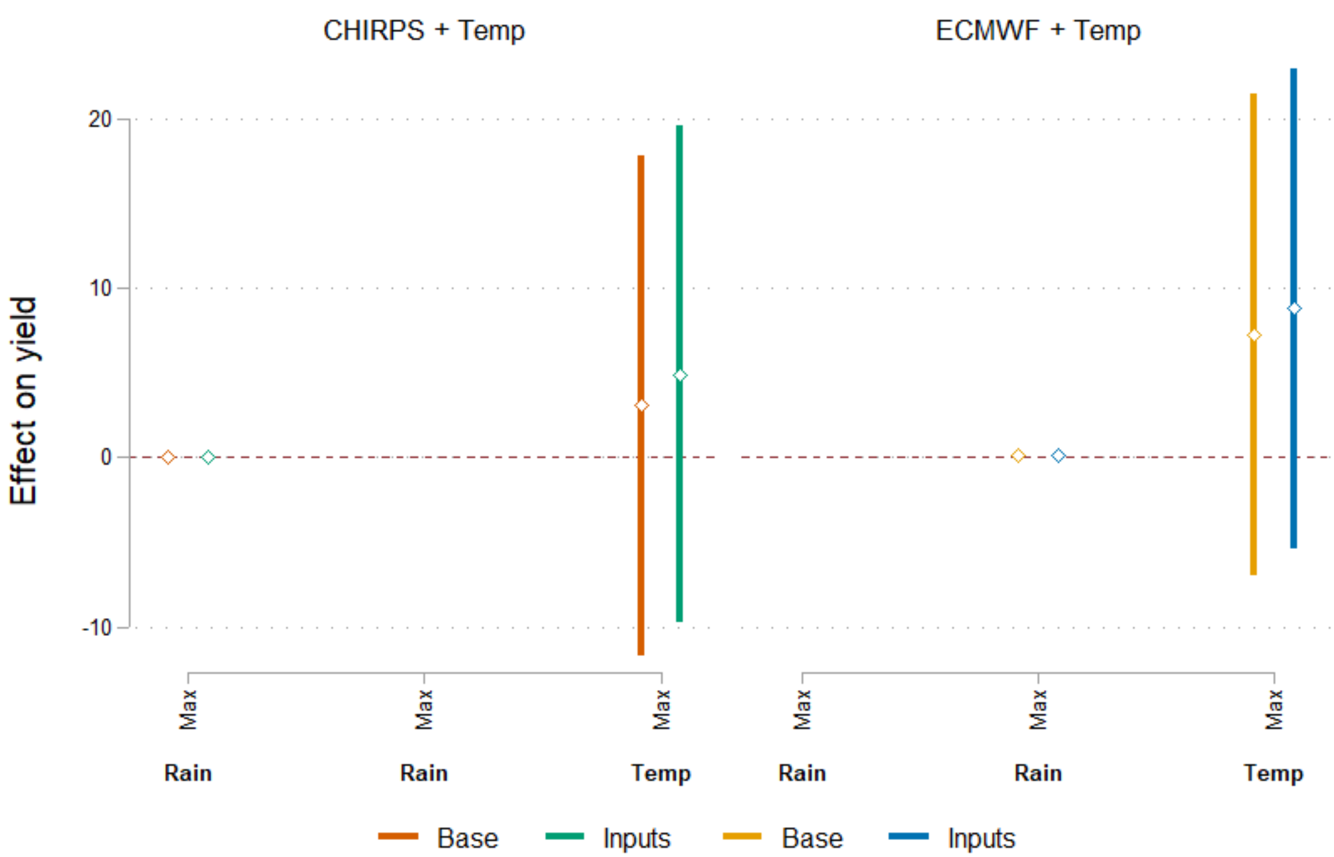
India Sorghum - Maximum



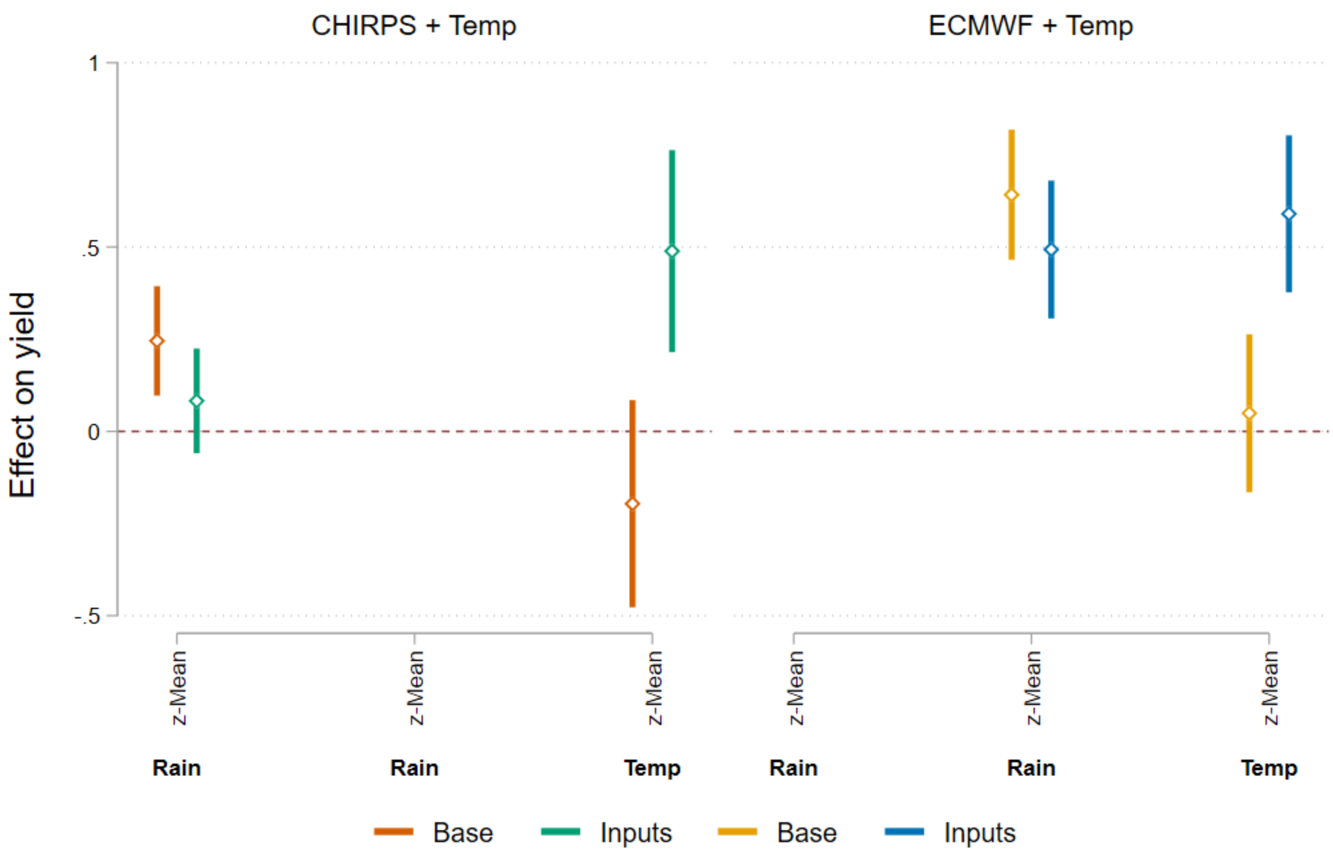
Ethiopia Tef - Maximum



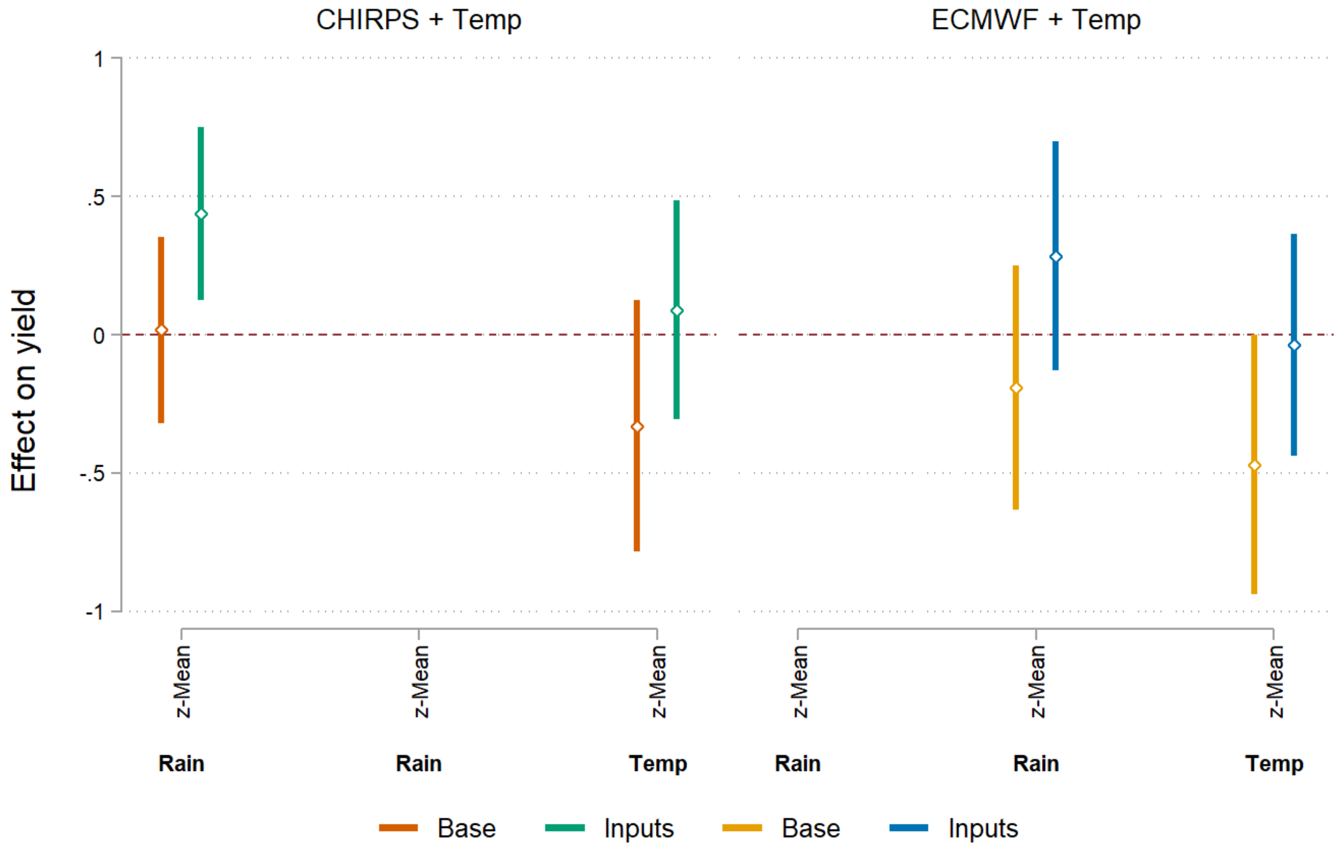
Kenya Maize - Maximum



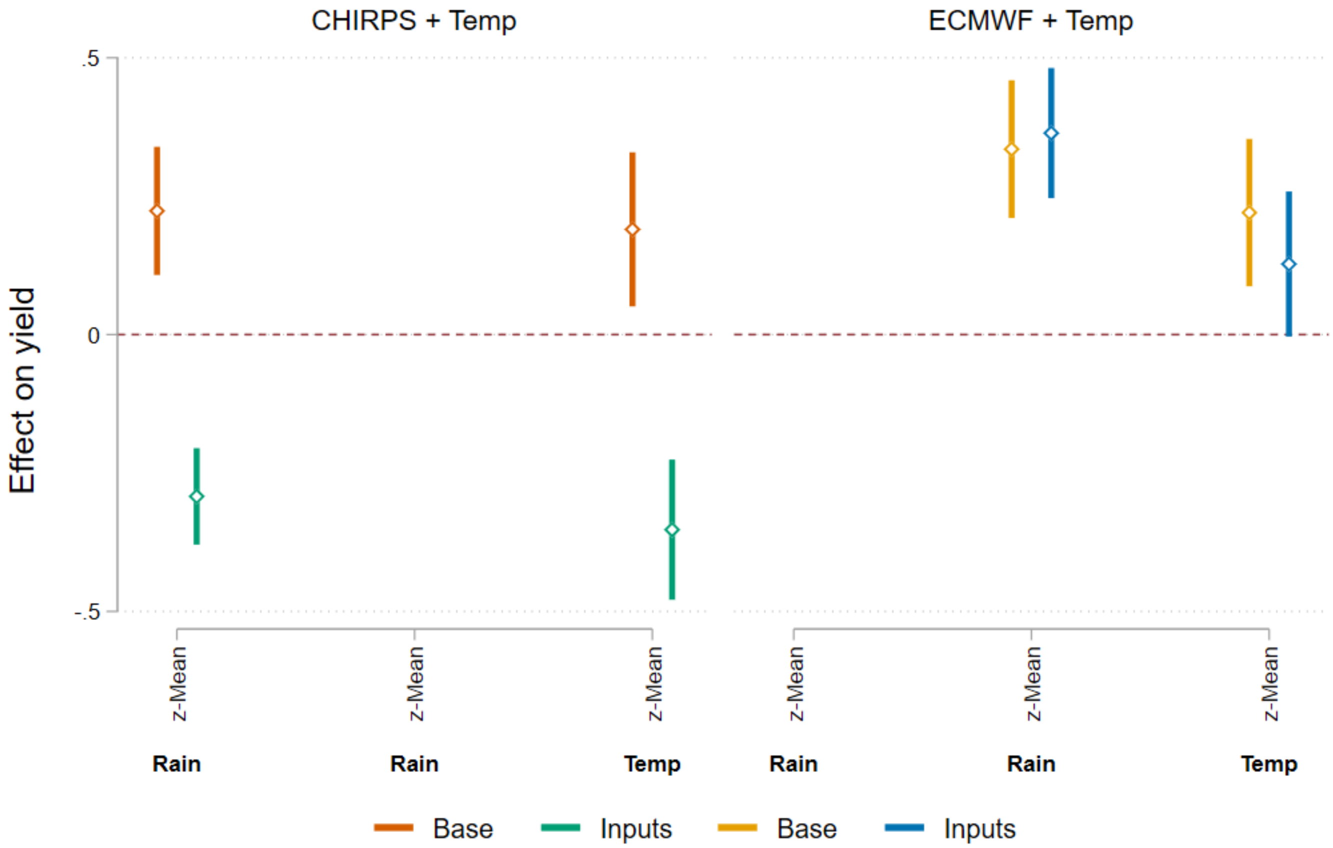
Bangladesh & India Rice - z-Score



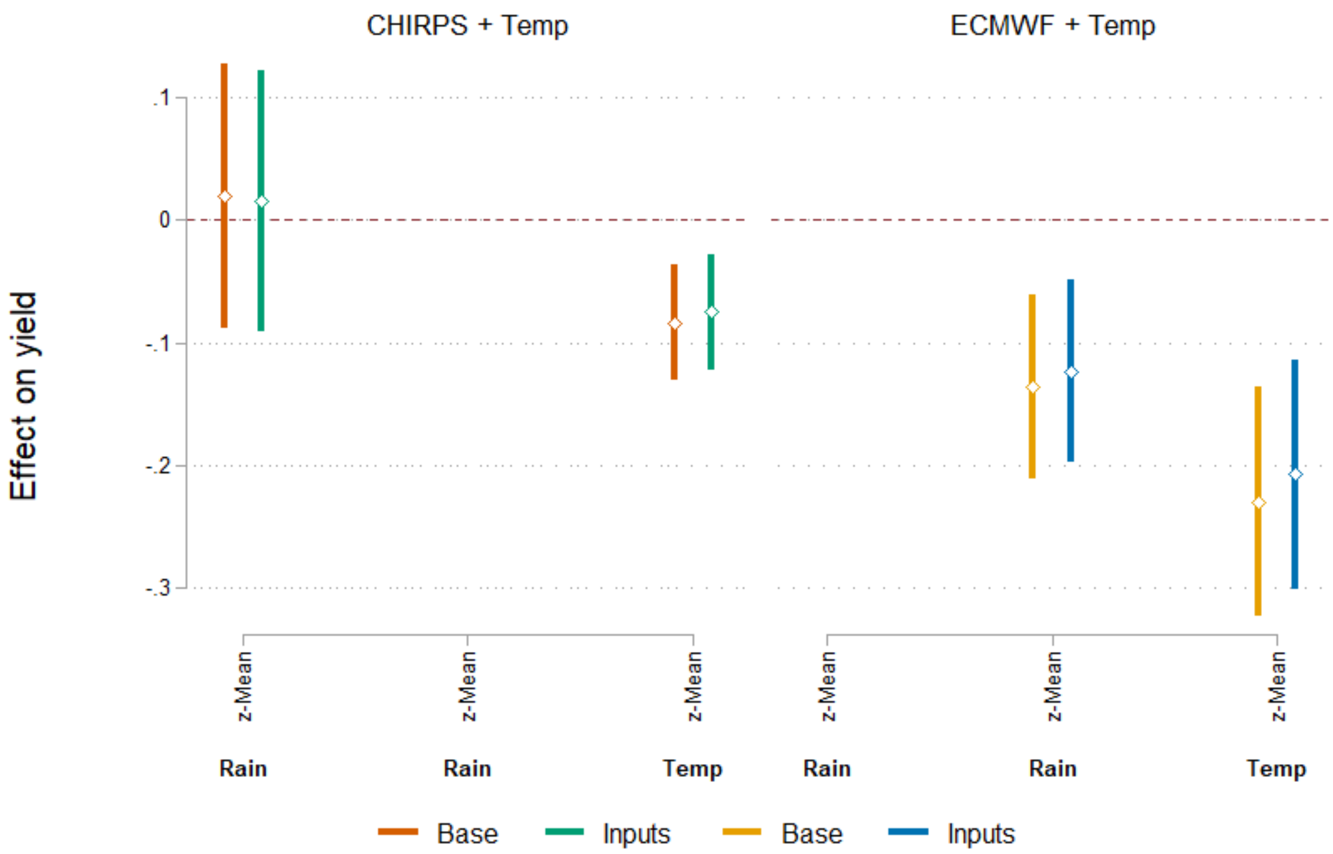
India Sorghum - z-Score



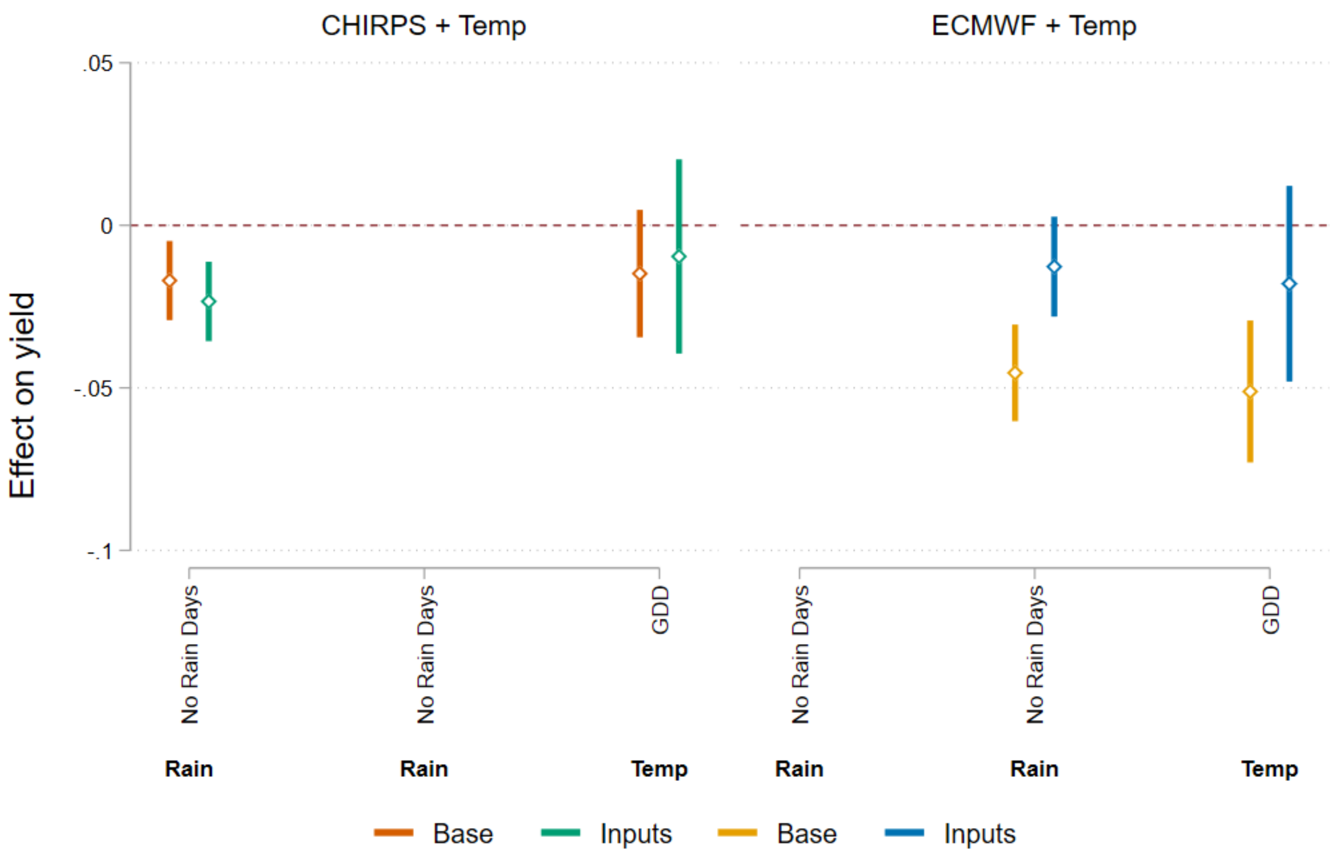
Ethiopia Tef - z-Score



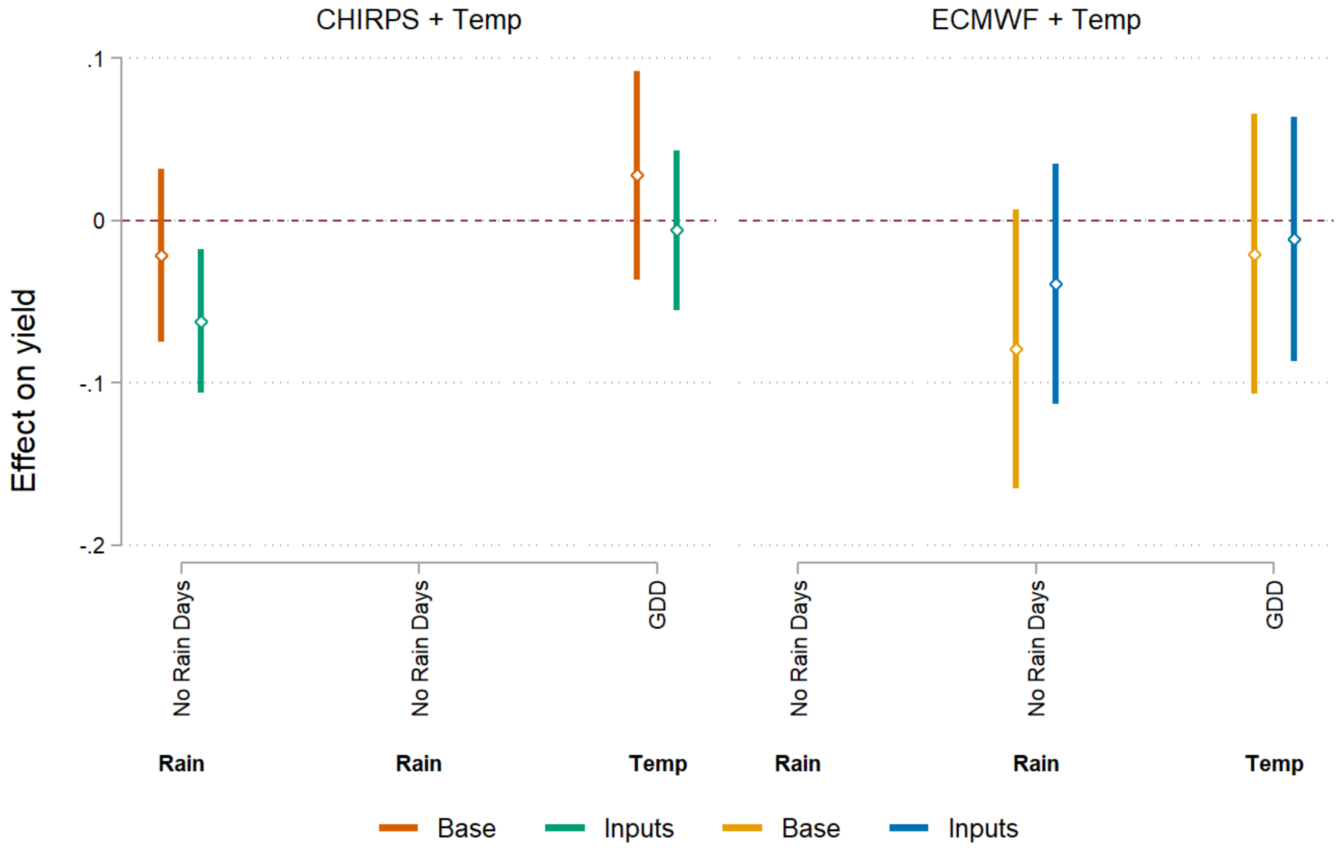
Kenya Maize - z-Score



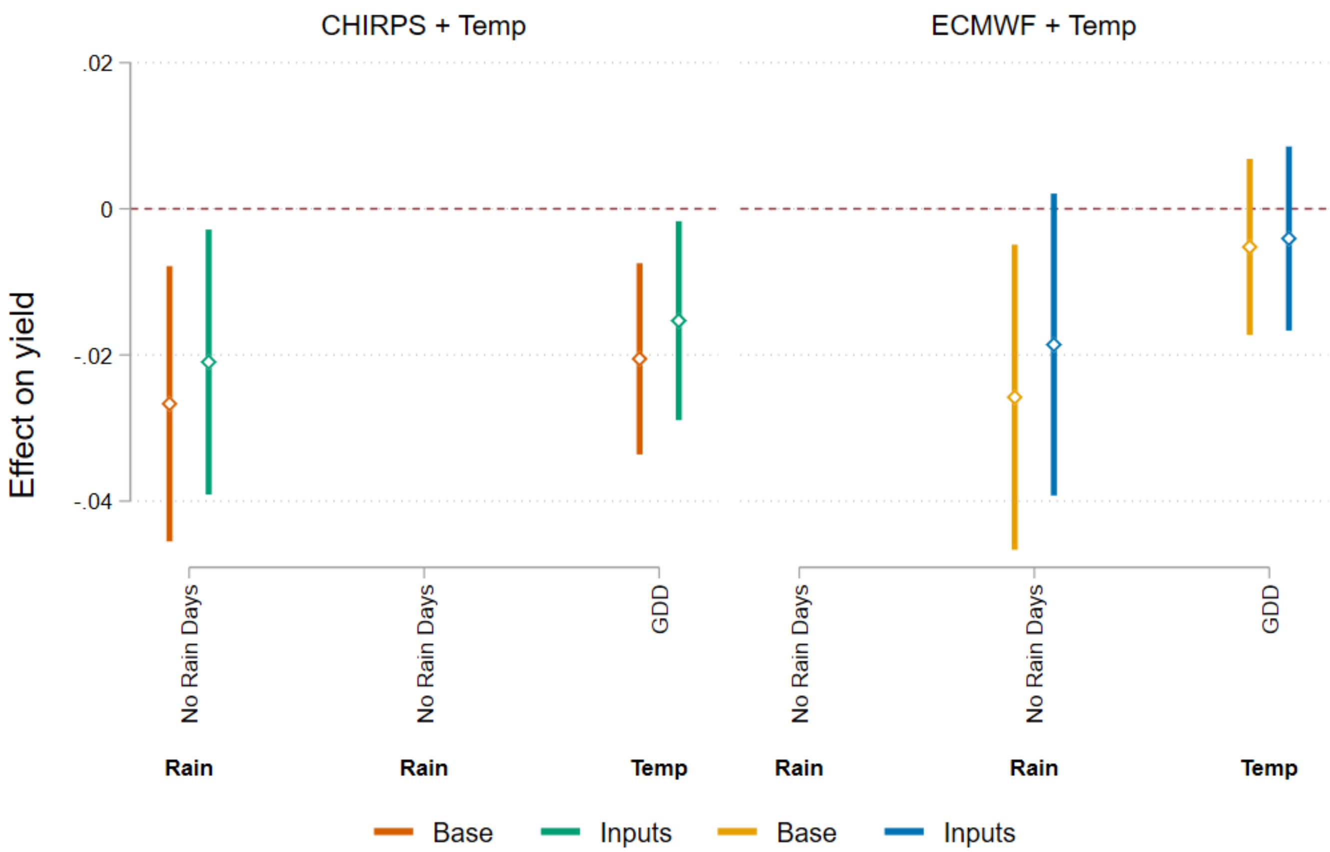
Bangladesh & India Rice - No Rain Days + GDD



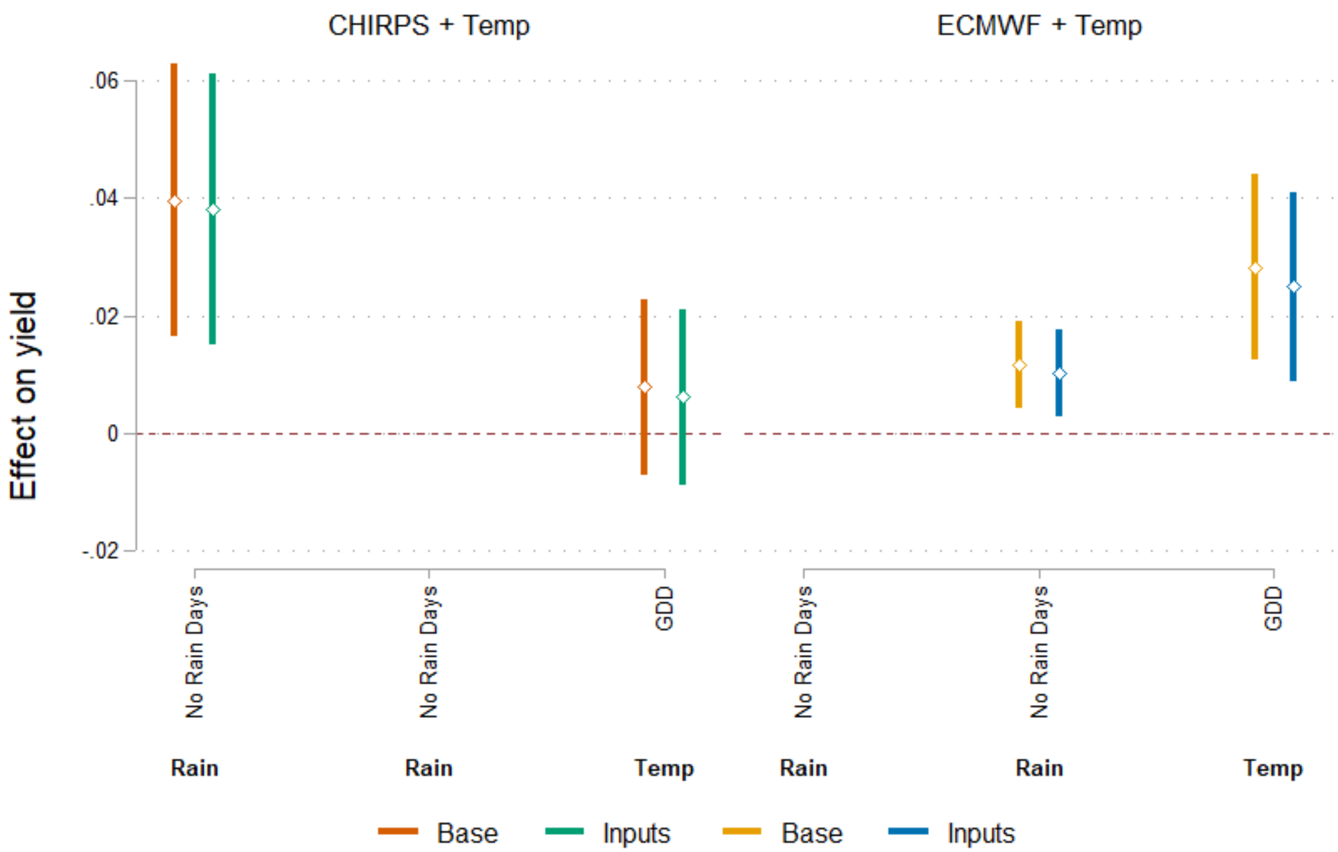
India Sorghum - No Rain Days + GDD



Ethiopia Tef - No Rain Days + GDD



Kenya Maize - No Rain Days + GDD



What patterns can we identify?

- If using moments of the rainfall distribution, the source of the data matters.
- There is some consistency in the direction of effect that rainfall has on yields.
- Max, total, z-score, and number of no rain days do not seem to matter.
- More than rainfall, there is no consistent pattern in temperature across countries.
- z-score and GDD do not seem to matter.
- Combining temperature and rainfall makes rainfall not significant, except when combining the number of no rain days with GDD.

Next Steps

- Check more variables (quadratic, SPI, temperature bins)
- Interaction terms, with weather variables and with inputs
- Multi-crop regressions with flexible functional forms
- Check different levels of aggregation
- Soil quality measures
- More countries!!!