Supplementary Information for

THRESHOLD EFFECTS OF EXTREME WEATHER EVENTS ON CEREAL YIELDS IN INDIA

Authors: Francisco Fontes[1], **Ashley Gorst**[2] **and Charles Palmer**[3]

Affiliations: Food and Agriculture Organization of the United Nations (Francisco Fontes); Vivid Economics (Ashley Gorst); Department of Geography and Environment, London School of Economics and Political Science (Charles Palmer)

[1] Corresponding author. Email address: frapfontes@gmail.com
[2] ashley.gorst@vivideconomics.com
[3] C.palmer1@lse.ac.uk

Data and methods – additional material

1. Data

The SPEI index is sourced from Vicente-Serrano et al. (2010) and can be accessed using the following URL: https://spei.csic.es/database.html

As explained by Vicente-Serrano et al. (2010), the SPEI relies on the concept of a climatic water balance, which is given by precipitation minus evapotranspiration:

$$D_i = P_i - PET_i \qquad (1)$$

Where P_i represents precipitation and uses the TS v. 4.03 dataset. PET_i represents the potential evapotranspiration and is calculated using the FAO-56 Penman–Monteith equation. Finally, D_i represents a measure of water surplus or deficit for month *i*.

This variable is then aggregated at different time scales to obtain the cumulative water balance over a given period. Following this, the variable is then standardized using the log-logistic distribution.

As with all gridded datasets, they depend on data from weather stations, which are likely to change over time. While it is impossible for us to know how potential changes in stations over time has affected the precipitation variables and whether this had any effect on the SPEI, we note that in Harris et al. (2020), the coverage maps for the CRU for India seem to indicate good coverage, especially for the time-periods of data relevant for our analysis.

Figure S.1 plots the distribution of the *kharif* and *rabi* SPEI indices for our sample.

Figure S.2 is constructed by using the average share of the cereal area devoted to a given crop.

Figure S.3 is constructed by using the sum of district total gross cropped area in our subsample for each year. We then set the value in 1966 as 100.

To construct panel (a) of Figure S.4, we use rainfall data collected by the Indian Meteorological Department. The rainfall data are available in gridded format at a resolution of 0.25 degrees by 0.25 degrees (Pai *et al.*, 2014). District-level weather data are then obtained by taking a weighted average of gridded observations from grid cells that fall in each district.

Agro-ecological classification, on the basis of whether they are arid and humid, at the district level, (Figure S.4, panel (b)) is constructed using data in the ICRISAT dataset that subdivides India into 20 agro-ecological zones, which takes into account rainfall, potential evapotranspiration and soil type. These data are originally sourced from the National Bureau of Soil Sciences and Land Utilisation Pattern, ICAR, Nagpur (Gajbhiye and Mandal, 2010).

The agro-climatic regions of 2000 are used for the full analyzed period. To allow for a large number of observations in each sub-sample, we simplify this classification by merging all arid and semi-arid areas together irrespective of soil type (see red and orange districts in Figure S.4 (b)). We also merge humid, sub-humid and perhumid areas (see districts in yellow and green in Figure S.4 (b)).

2. Empirical approach

Significance of thresholds

See Table S.4 for p-values of the different threshold models and Table S.5 for the location of thresholds in all estimated specifications. Note that, as described in Wang (2015), the threshold tests are sequential. In other words, when testing for the one-threshold model the null hypothesis is the linear model with the alternative being the single-threshold model. When testing for a double-threshold model, the null hypothesis becomes the single-threshold model and the double-threshold model becomes the alternative. In some cases (e.g. barley in the *kharif* season), it is possible that we have a p-value above 0.1 for a single-threshold (i.e. a model with no threshold is preferred) and a p-value below 0.1 for a double-threshold model (i.e. a double-threshold model is preferred to a single-threshold model). In these cases, this means that we opt for the no-threshold model because we fail to reject the null of no threshold in the first place. Throughout our paper, we test for significance at the 10% level to determine the admissibility of a threshold. However, as can be seen from Tables S.4 and S.5, the robustness of almost all of our thresholds is illustrated by the low p-values and the relatively small size of the confidence intervals for most of our estimates.

Multiple threshold variables

One important limitation of the threshold model used in the paper is its inability to account for multiple threshold variables simultaneously. Given the importance of both temperature and rainfall, it could have made sense to estimate thresholds of rainfall conditional on temperature (or vice versa). However, the panel threshold model proposed in Hansen (1999) is only capable of handling one threshold variable. The main justification for this is that, according to Hansen (2000), there is no distributional theory underlying such models. As such, even if we were to estimate such models using, for example, regression trees, we would not be able to calculate the degree of confidence associated with each threshold nor would we be able to construct their confidence intervals. To our knowledge, there have been developments to allow for the incorporation of multiple threshold variables in a time series setting (Chen et al., 2012). However, we are not aware of such developments for non-dynamic or dynamic panel data.

Beyond the direct policy relevance of the SPEI index as a multi-scalar drought index, the fact that it combines rainfall and temperature is an important reason as to why we adopted it as our threshold variable.

Estimated impacts

To obtain the plots in Figures 2-3, we plot the estimated deviations in log yields for the observed range of the SPEI. In the case of one threshold, we calculate:

$$Effect = \begin{cases} SPEI * \beta_1 \ if \ SPEI \leq \gamma \ (i.e. \ SPEI \ below \ T1) \\ SPEI * \beta_2 \ if \ SPEI > \gamma \ (i.e. \ SPEI \ above \ T1) \end{cases} \qquad (2\cancel{1})$$

To make the results more interpretable, the results are interpreted as the estimated deviation in $\ln(y_{it})$ for a given value of the SPEI, when compared to a SPEI=0. It should be noted that the actual value of predicted log yields at SPEI is not 0 (which would imply a yield in levels of 1 for each district). Instead our effect formula in (1) estimates the SPEI-log(yield) relationship conditional on other estimated parameters (district-specific quadratic trends and constant). However, to compare the results with the bins approach, we focus only on the marginal effects. Since, by definition, at SPEI=0, the effect (as calculated above) is always zero. We opt for this approach for two reasons. First, setting the SPEI equal to zero is the most natural value to compare deviations against because it is the centre of the distribution. Second, our bin approach robustness check (see below) is also computed using $\ln(y_{it})$ as a dependent variable and uses the centre of the distribution as a reference category. As such, to make the two approaches broadly comparable, we need to calculate the effect as shown above.

Advantages of threshold model over alternative methods

Compared to other methods, this approach has several advantages. First, it does not impose a global linear relationship between the independent variable and the dependent variable as in the case of a linear regression. Second, it does not impose a symmetrical relationship around an estimated peak as in the case of a quadratic model, nor does it impose a strict functional form as is the norm in the case of higher-order polynomial regressions. In this sense, it allows for an approach that is more data-driven.

Other models, most notably the multivariate adaptive regression splines (MARS) model, have also been used to identify non-linear relationships between two variables (e.g. Zipper et al., 2016). It has an advantage over the threshold regression model in that it allows for a large number of thresholds. In our case, however, this is unlikely to be important, since a three-threshold model (the maximum allowed by our statistical software) is rejected in every case. However, it has several important drawbacks. First, analyses carried out using the MARS model do not compute a confidence interval for the location of the threshold. One advantage of the threshold model is that, by computing and testing the statistical significance of the threshold and its confidence interval, it provides useful information regarding how precisely the threshold is estimated and how confident we are regarding its location for a given sample. Second, confidence intervals are also constructed around the coefficients so that we know how noisy our estimates are for a given range of the threshold variable.

Compared to the bins approach (described in more detail below), our method also has several advantages. First, the bins approach is more subjective that the threshold model as the results will depend a lot on both the choice of a base category and the coarseness of the bins. This means that depending on the base category we choose and how coarse the bins are, we may come up with different thresholds simply based on our choices of these two parameters (base category and extent of bin coarseness).

Second, a key purpose of this paper is to identify the location of the SPEI at which the yield-SPEI relationship is likely to change (i.e. "thresholds"). Doing this with bins is much more challenging. With very small bins, it may be very difficult to identify where the average relationship changes given that estimates using bins are noisier. On the other hand, if we use coarser bins, it becomes easier to see where the threshold might lie (as changes in impacts become clearer), but the range of where the threshold lies could be very large. For example, for rice in the *kharif* season we note that impacts appear to become worse in the bin located at -1.75 (which covers SPEI values between -1.5 and -2). Yet, what we learn from this is that the threshold is likely to be located somewhere between the previous bin and this bin (which covers the -1 to -2 range). This range may be too large to be of any practical relevance. For example, a threshold of -2, would imply an event that happens roughly once every 50-60 years, whereas a threshold of -1 would imply an event that is likely to happen every 4-6 years, depending on the sample.

Third, the bins approach may also be problematic when the SPEI-yield relationship changes somewhere in (or very close to) either the base category or close to the extremes. In our paper, we find some thresholds in or very close to what would typically be considered a "normal" range of the SPEI (-.5 to +.5). Invariably, these thresholds would be difficult to pick up using a bins approach, especially if we set the "normal" range of the SPEI as a base category. With regards to identifying thresholds at extreme values of the SPEI, the bins approach could also be more problematic. Near the extremes, if we use absolute rules (say 0.25 or 0.5 increments of the SPEI) we could end up with few observations, which may then make it impossible to identify thresholds. If we use relative rules (e.g. 2 or 5 percentiles), we may end up with a more accurate, but much bigger range of the SPEI, which might be of limited use for policy applications. With regards to identifying thresholds at extremes, the threshold model improves on the bins approach in that it can identify these thresholds. However, often, when thresholds are estimated close to the extremes, it will also have wider Cis for these cases or no Cis for the threshold location when these are located at the cut-off point of the trimming, which reflects the uncertainty around the threshold.

A final point is that, when computing total losses, the choice of a base category that includes 0 (e.g. -0.5 and +0.5) will lead to an underestimate of total losses (see Figure 6), as many events are not considered. It is true that a base category could be constructed at a higher range of the SPEI, but until now we have not observed researchers comparing SPEI ranges to base categories that do not include 0.

3. Robustness checks

Sensitivity to alternative cutoff points

One potential source of subjectivity in a threshold model is the choice of how much of the threshold variable is to be trimmed at the extremes. In our case, we opted for 1%, but we also test the sensitivity of results to alternative trimming points at 0.5%, 1.5% and 2.5%. Overall, as can be seen in Table S.11, our results remain broadly unaffected.

Sensitivity of estimates to clustering and spatial correlation

Throughout the paper we used standard errors clustered at the district level. The main reason for this is anchored in the fact that when large numbers of clusters are available (e.g. typically above 50), the asymptotic conditions for the clustering are met (Cameron and Miller, 2015). Below this number, there is currently no consensus on the extent to which we can trust clustered standard errors. To test the sensitivity of our estimates to clustering on different variables, we cluster the errors using the year as a clustering variable.

We also cluster using a two-way cluster based on state and year. However, in this case, while we show the results in Table S.12, one of the clusters (State) falls clearly in a grey area and as such, of all the clustering sensitivity checks, this is the one that gives us least confidence. From an estimation perspective, we note that the *xthreg* command on Stata does not allow for two-way clustering of standard errors. As such, for the state-year clustering, we use the estimated thresholds from the threshold regression and reproduce the equation using the high-dimensional fixed effects regression command in Stata *reg2hdfe* by Guimarães (2015). We apply this procedure using the *clus_nway* code by Wolfson and Kleinbaum (2020). Since this procedure could be problematic when the dimension of one of the clusters is small (which is the case for the state dimension), we also consider using a wild bootstrap (using the *boottest* command), a method suggested in Cameron and Miller. We do not report the results of the wild bootstrap as they were almost identical to the results produced by the other procedures.

To test the sensitivity of the estimates to spatial correlation, we use Conley (1999) errors using the *stata* routine procedure proposed by Hsiang (2010). Since, the spatial correlation part of the procedures requires a distance cutoff, we test multiple cutoffs at 50, 100, 250, 500 and 1000 km.

Sensitivity to alternative lags for the kharif season

While using June-September is a common definition of the *kharif* season in India, it is not the only definition. Some researchers prefer June-October. As a result, to ensure that our results are not driven by the definition we use for *kharif* season, we construct two alternative *kharif* SPEI indices, covering June-October and June-November, respectively. As can be seen in Table S.13, the results remain broadly similar.

Sensitivity to the inclusion of controls

Our preferred specification is a reduced-form function where controls are excluded. This is arguably the most common type of specification in the climate literature. To test the robustness of our results, we estimate our results with and without a set of time-varying control variables, X_{it}. We include rural population per hectare of cereal area, total cereal area, fertilizer used and proportion of land under irrigation. However, given that our method requires a balanced panel, we lose many districts. For example, out of 242 districts in the full sample, the inclusion of only four controls leads to the loss of 138 districts. Despite this, while threshold locations and coefficients change a little, the main results (Table S.14) remain very similar. We do not report the barley estimates, however, because for this sample we are left with just eight districts (368 observations) and hence the estimation results are unlikely to be credible.

Application of the bins approach

We also check total impacts and compare these to the estimates from our threshold model using dummy variables to capture the effects of the SPEI on our yield variables at different percentiles of the index (‘bins’ approach). This robustness check adapts a commonly-used method to estimate weather (mostly temperature) impacts on economic outcomes (e.g. Schlenker and Roberts, 2009), which consists of estimating dummy variables for different ranges (bins) of the SPEI. Bins were constructed as follows. The 45th to 55th percentiles of the average seasonal SPEI values is used as the baseline. Then, dummy variables are constructed for values below the 1st percentile, above the 99th percentile, and for every second percentile between the 1st and 99th percentile which did not fall into the baseline category.

To plot the predicted deviations of the SPEI compared to the base category, we start by defining a set of dummy variables that are equal to one if the SPEI is within a given range and zero otherwise. The range is based on the percentiles of the SPEI. Specifically, we generate bins in the following way:

$$\begin{cases} D_1 = 1 \ \ if \ SPEI < p1; 0 \ otherwise \\ D_{1to3} = 1 \ \ if \ p1 \ \le SPEI < p3 \ ; 0 \ otherwise \\ \vdots \\ D_{43to45} = 1 \ \ if \ p43 \ \le SPEI < p45 \ ; 0 \ otherwise \\ D_{45to47} = 0 \ \ if \ \ p45 \ \le SPEI < p47 \\ \vdots \\ D_{53to55} = 0 \ if \ \ p53 \ \le SPEI < p55 \\ D_{55to57} = 1 \ \ if \ p55 \ \le SPEI < p57 \ ; 0 \ otherwise \\ \vdots \\ D_{55to57} = 1 \ \ if \ p97 \ \le SPEI < p99 \ ; 0 \ otherwise \\ D_{99} = 1 \ \ if \ SPEI \ge p99 \ ; 0 \ otherwise \end{cases} \quad (2)$$

In other words, D_1, for example, is equal to one if the SPEI falls below the first percentile and is equal to zero otherwise and D_{1to3} is equal to one if the SPEI falls between the first and third percentile and zero otherwise. Note, we use percentiles to ensure that each bin contains enough observations so that the dummy variables are estimated with a sufficient degree of precision. Most of our crop-specific samples have in excess of 5,000 observations, which means that each dummy contains at least 100 observations (above 50 in the extremes), a sufficient size to ensure that the coefficient is estimated with a minimum degree of accuracy. We also run these regressions using increments of the SPEI (based on 0.25 and 0.5 increments, using the range [-0.5, +0.5] as a base category.

We argue that the bins approach has several advantages as a robustness check against other commonly used methods (i.e. generally imposing a quadratic or cubic relationship). First, it does not impose an *a priori* shape between the SPEI and yield. Second, unlike imposing a quadratic trend, it allows the relationship to be asymmetrical around a given turning point. Third, since the bins are quite small, it also allows us to visually identify whether the "jumps" and thresholds identified by the threshold model seem to be reflected in the data. The drawback, of course, is that the range of observed SPEI values in D_1 will be a lot larger than in the range of D_{43to45}, which means we have less granularity at lower levels of the SPEI. Arguably, since threshold methods require a portion of the data to be trimmed at the extremes (we use 1%), this is not a critical issue. That said, differences in the relationship within the 1st percentile will not be estimated. We set the dummies for SPEI values between the 45th and 55th percentiles as this is the base category against which we can compare the SPEI coefficients. As the SPEI follows a close to normal distribution, in most cases these percentiles will contain the value of zero or close to zero, which make these coefficients comparable to the coefficients obtained by the threshold model.

Once the dummy variables are generated, we estimate the following regression using a fixed-effects regression:

$$ln(y_{it}) = \alpha_{it} + \beta_{i1} dist_i * t + \beta_{i2} dist_i * t^2 + \delta_1 D_1 + \delta_{1to3} D_{1to3} + \cdots + \delta_{43to45} D_{43to45} + \delta_{55to57} D_{55to57} + \cdots + \delta_{97to99} D_{97to99} + \delta_{100} D_{100} + \epsilon_{it} \quad (3)$$

After the regression is estimated, we simply store the set of δ coefficients and plot these.

Analysing potential heterogeneity of thresholds and marginal effects by crop

To test whether thresholds and marginal effects are similar across agro-ecological zones and irrigation status, we run the crop-specific threshold models by AEZ and irrigation status. The results are presented in Tables S.6 to S.10.

For AEZ, we divide each crop sample into an arid and a humid subsample (Figure S.4(b), SI - 1).

To build the irrigation sub-samples, we use the share of the total cropped area for a given crop under irrigation. Our model requires that our sub-samples are temporally consistent (due to the necessity of having a balanced sample). As such, we cannot have districts switching across sub-samples. We therefore calculate for each district the average share grown under irrigated conditions for a given crop and then split the sample based on the median value. This ensures both temporal consistency across the sub-samples as well as a comparable number of observations across both sub-samples.

Finally, it should be noted that the sum of observations in the two irrigation sub-samples may not always be equal to the total number of observations for a given crop. This is because for certain districts irrigation data are missing, implying that these districts are dropped.

4. Procedures for estimating the per ha and total revenue losses

Defining relevant counterfactuals

Beyond the identification of thresholds, to explore the policy relevance of our threshold models, we compare the results from these models to those from different counterfactuals.

We opt to focus only on negative deviations of the SPEI for this part of our analysis for several reasons. First, since floods tend to be more concentrated in space and time, as opposed to droughts, which are more of a creeping phenomenon, our index is more likely to capture droughts better than floods (although it does capture wet years). Second, since for most crops wetter conditions are associated with higher yields (at least up to a certain point), considering all the events above 0 as impacts would be problematic as the revenue losses would actually be revenue gains. Third, although floods have been defined using SPEI or SPI thresholds, arbitrary definitions tend to be used most often for droughts. As such, we opt to focus on the negative range of the SPEI to illustrate our main points about the importance of data-driven, objective thresholds tied to an outcome of interest.

For the per ha losses, since these are evaluated at the location of the threshold, we are constrained by specifications that are linear and return different impacts for SPEI values of T1 and T1-0.1. Here, we use two counterfactuals, namely a log-linear specification and a non-linear specification (quadratic). As we are interested in estimating the losses for the negative range, we imposed a quadratic relationship with a threshold at a SPEI value of 0.

We do not use other non-linear models (e.g. bins) for the per ha effect since the bins may be larger than 0.1 which means that the effect at T1 and at T-0.1 would be in the same bin and thus would not vary.

However, the main point this paper wants to make, rather than which non-linear specification is best, relates to the use of arbitrary thresholds. To illustrate this, in addition to the log-linear and quadratic counterfactuals, for the total costs, we add four additional counterfactuals, some of which were/are actively used in policy and/or research. The four additional counterfactuals are as follows:

1) A 20% negative deviation from long-term average rainfall during June-September. This was the definition for a drought at the district level that was used in India. We note that, for this counterfactual, the comparison is limited to the 1966-2009 period as we do not have the rainfall data for 2010-2011.
2) A threshold arbitrarily set at SPEI=-1 using a log-linear specification – A SPEI of -1 is often considered the trigger for a dry event.
3) A threshold arbitrarily set at SPEI=-1 using the quadratic specification.
4) The bins approach using the percentiles method as explained above.

Defining prices to monetize the yield losses

First, we acknowledge that there is no perfect price for valuing yield losses over a period spanning almost 50 years. We opt to use 2005 national prices in USD because 2005 is a relatively recent year in which relatively few districts were affected by drought, so national prices were less likely to be affected by drought. A fixed year ensures that revenue losses are comparable over time and districts. Two alternatives were considered and discarded for the following reasons:

1. Farm-gate prices – these have the advantage of allowing for district-year specific prices. However, there are at least three potential drawbacks. First, these prices were likely to have been affected by drought in the first place. Second, this would have raised the issue of using real prices and it was not clear whether inflation at the district-level (or even food inflation) would have been an accurate deflator. Third, drought-induced revenue losses would depend a lot more on whether drought occurred (if it hits districts with a higher price elasticity, the cost would be higher).
2. International commodity prices or prices in neighbouring countries – these are not adopted, also for three reasons. First, it is not entirely clear that using international prices would free us from the impact of drought on prices. The only difference is that these prices would most likely have been affected by droughts in large producers rather than in India (e.g. Vietnam, Thailand in the case of rice). Second, the commodities sold in international markets are not the same as the commodities sold by farmers. For instance, in the case of rice, the rice sold in the international market is already milled and differs depending on type. Applying a constant conversion factor and estimating marketing margins is not necessarily a good idea. Finally, we do not want our revenue loss estimates to be overly affected by sudden changes in international prices. For example, in 1986 and 1987 (a particularly bad year in terms of droughts in India), international prices for cereals seem to have decreased substantially. According to the WB pink sheet data, there were decreases ranging from 20-43% compared to 1985 for all cereals except rice (not covered in 1985) and millet (not included in the dataset). Using this option would have masked the very high costs of the 1986 and 1987 droughts.

Procedures for estimating the per ha revenue losses

In terms of the procedure, the calculation of the per ha revenue loss is illustrated for a one threshold case against a log-linear model (the process is almost identical for other counterfactuals. Specifically, we only need to change step 6 and estimate the different counterfactual). We carry out the following steps for each crop sample:

1. After estimating the threshold model, we create a temporary SPEI variable equal to T1 for the full sample.
2. We use the threshold model to calculate the predicted yield at SPEI=T1 using the command *levpredict* on Stata. Note that we cannot simply take the exponent of the predicted ln(y_it) because this would result in a bias.
3. We then create another temporary SPEI variable equal to T1-0.1 and calculate the predicted value at SPEI=T1-0.01 using the command *levpredict* on Stata.

4. We take the difference in the predicted values obtained in steps 2 and 3, which gives us the expected yield difference for a 0.01 change in the SPEI value at the threshold.
5. We then multiply the value obtained in step 4 by the 2005 price per tonne of the cereal we are investigating. Note that to generate the price variable, we use a weighted average of district-level prices where weights are defined by cropped area. Since the prices are expressed in Rupees per quintal (100 kilograms), we first convert the loss into USD by dividing it by the annual exchange rate (44.1 Rupees per USD) from the World Bank WDI database. After, we multiply the USD/per quintal price by 10 to get the USD per tonne price.
6. To compute the "no threshold" counterfactual, we start by estimating a fixed effects model where we assume that SPEI is linearly related to ln y (the no threshold counterfactual):

$$\ln(y_{it}) = \alpha_{it} + \beta_{i1} dist_i * t + \beta_{i2} dist_i * t^2 + \delta SPEI_{it} + \epsilon_{\text{it}}$$

7. We calculate the predicted yield at SPEI=T1 using the regression estimated in step 6.
8. We calculate the predicted value at SPEI=T1-0.11 using the model in step 6.
9. We take the difference between the values in steps 7 and 8 and multiply the difference by the price per tonne.

We then repeat this procedure for each crop to get the results in Figure 4.

Procedure to estimate total revenue losses

The procedure to calculate the total revenue loss bears some similarities with the marginal loss per hectare procedure. However, there are two key differences. Whereas in the marginal cost per ha calculation we impose the location where the effect is estimated, in the total cost calculation, we use observed SPEI values. Secondly, we calculate the costs associated with all events that have SPEI values below zero (i.e. we exclude events above zero for the reasons explained above) and compare these to an SPEI value of 0. The key rationale for this is that because drought is normally defined as a negative departure from normal, it would be hard to justify any point above zero. Also, it is the most natural SPEI value with which to make a comparison of results.

More concretely, to obtain total revenue losses, we start by carrying out the following steps for each crop separately (steps below illustrate the procedure for the log-linear case. For other counterfactuals, we simply need to change the estimated equation in step 7):

1. Using the *levpredict* command, we use the threshold model to predict the predicted yield at the observed SPEI value.

2. We create a temporary SPEI variable equal to zero for the full sample and compute the predicted value at SPEI =0 *levpredict* command. This is essentially the counterfactual value.
3. To generate the predicted loss per hectare, we take the difference between the values obtained in steps 1 and 2.
4. We then multiply the result obtained in step 3 by total cultivated hectares.
5. We multiply the results obtained in 4 by the 2005 price for that specific subsample. This gives us the total losses per district predicted by the threshold model for a given year.
6. We then add the results obtained in 5 for all districts in a given year. This gives us the total losses for a given year under the threshold model.
7. To compute the "no threshold" counterfactual, we start by estimating a fixed effects model where we assume that SPEI is linearly related to ln y (the no threshold counterfactual):
$$\ln(y_{it}) = \alpha_{it} + \beta_{i1} dist_i * t + \beta_{i2} dist_i * t^2 + \delta SPEI_{it} + \epsilon_{\mathrm{it}}$$
8. We repeat steps 1-7 using the linear model.
9. We take the difference between the losses predicted by the threshold model vs. the linear model, which generates an estimate of the size of the threshold effect.

After carrying out steps 1-9 for each crop in each season, we aggregate the revenue losses across all crops and seasons. After doing this we calculate the average difference in revenue lost between the threshold and each counterfactual to obtain Figure 6. We then aggregate the total revenue losses for all crops and seasons by sub-period (1966-1970; 1971-1975, 1976-1980, 1981-1985, 1986-1990, 1991-1995, 1996-2000, 2001-2005, and 2006-2011[4]) for each counterfactual to generate Figures 5(a), 5(b), and Figures S.11 and S.12.

[4] In the case of rainfall, as explained before, we can only carry out this exercise up to 2009, since we did not have access to the dataset for 2010 and 2011.

References

Chen, H., Chong, T. T. L., & Bai, J. (2012). Theory and applications of TAR model with two threshold variables. *Econometric Reviews*, *31*(2), 142-170.

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadhyay, B. (2014). Development of a new high spatial resolution (0.25× 0.25) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam*, *65*(1), 1-18.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of human resources*, *50*(2), 317-372. doi: 10.3368/jhr.50.2.317

Gajbhiye, K. S., & Mandal, C. (2000). Agro-ecological zones, their soil resource and cropping systems. *Status of Farm Mechanization in India, cropping systems, status of farm mechanization in India*, 1-32.

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific data*, *7*(1), 1-18.

Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37), 15594-15598. https://doi.org/10.1073/pnas.0906865106

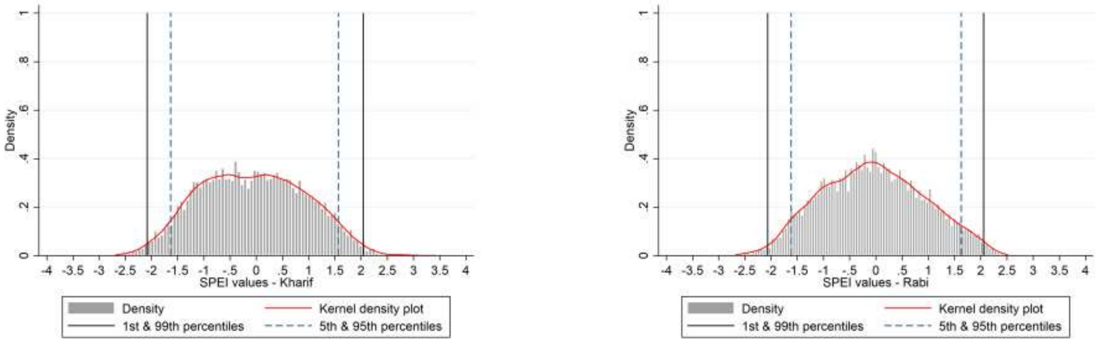
Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, *23*(7), 1696-1718.

Wang, Q. (2015). Fixed-effect panel threshold model using Stata. *The Stata Journal*, *15*(1), 121-134.

Wolfson, P. J., & Kleinbaum, A. (2020). CLUS_NWAY: Stata module to perform Multi-way Clustering for Various Model Specifications.

Guimaraes, P. (2015). REG2HDFE: Stata module to estimate a linear regression model with two high dimensional fixed effects.

Figures and Tables



(a) Distribution June-September SPEI

(b) October-March SPEI

Figure S.1. SPEI distribution over the *kharif* and *rabi* seasons

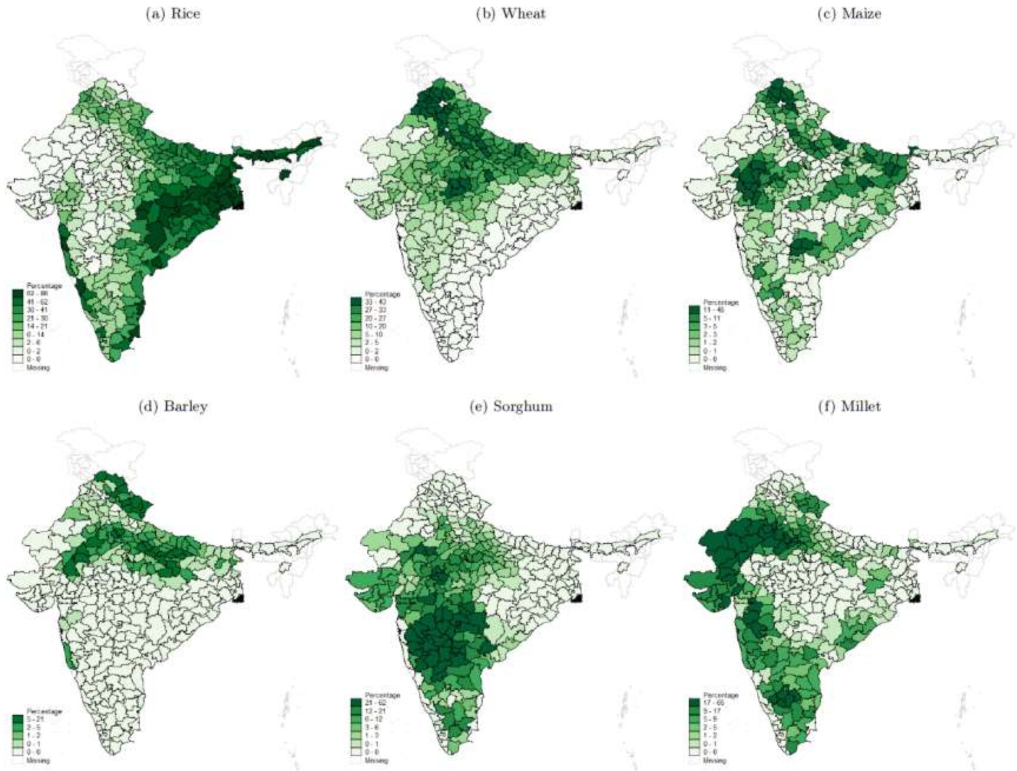


Figure S.2. Area planted by crop

Panels (a)-(f) show the proportion of gross cropped area devoted to each of the six crops used in the analysis. For districts with no data, these areas are shown as white polygons. District boundaries refer to those in 1966.

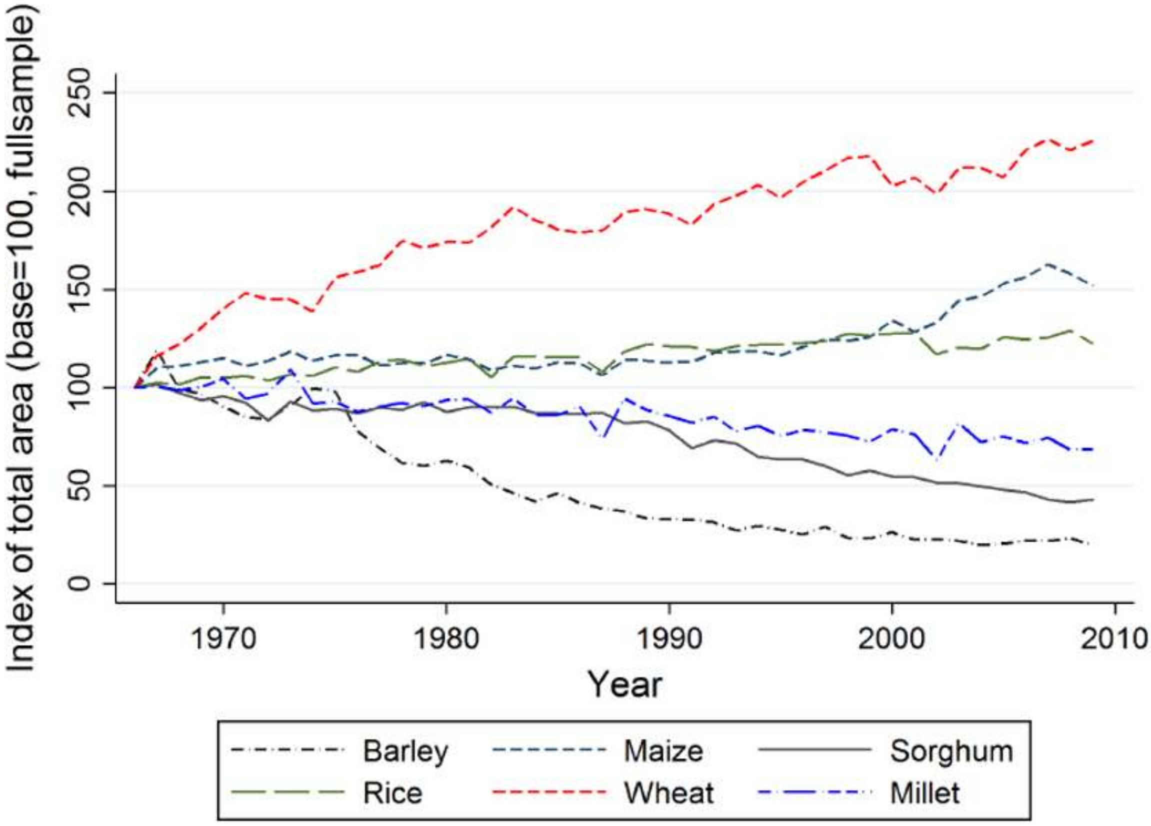


Figure S.3. Total cultivated area by crop

The total cultivated area is indexed at 100 for the year 1966.

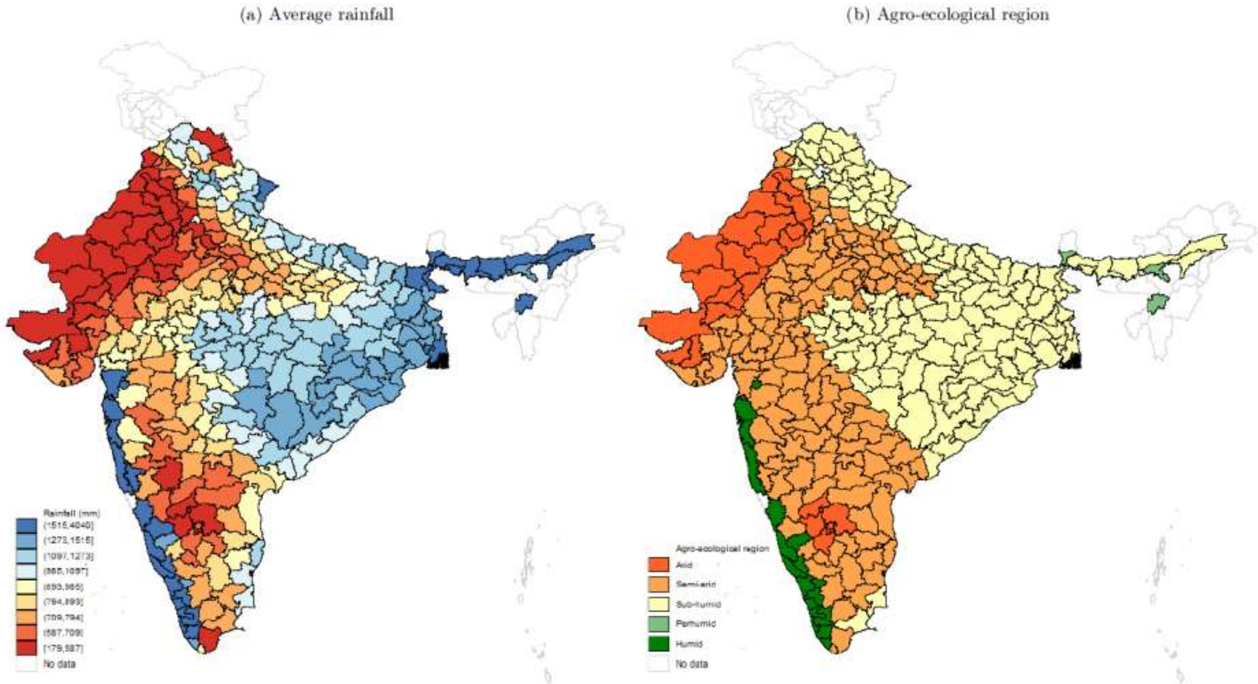
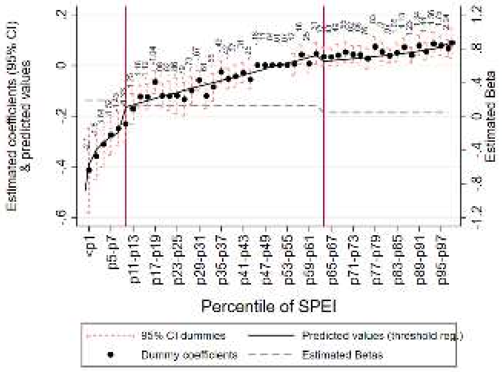
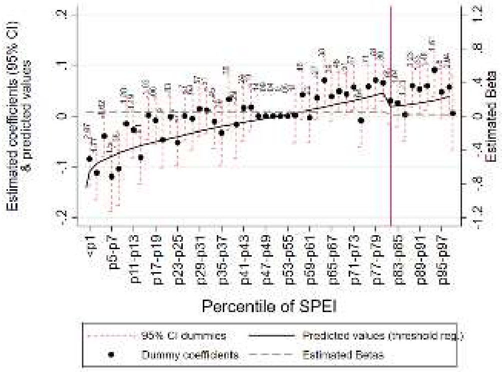


Figure S.4. India rainfall and AEZ classification

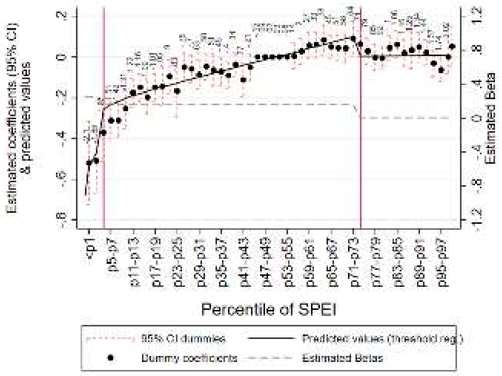
Panel (a) shows district-level average rainfall for the 1957-2009 period. Panel (b) maps the agro-ecological zones. For both, only districts for which we have data are plotted. Districts for which we have no data are shaded in white.



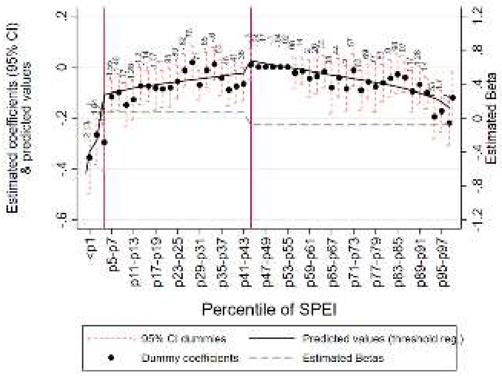
(a) Rice (Kharif)



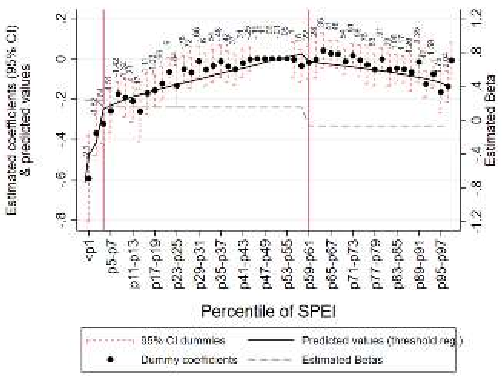
(b) Rice (Rabi)



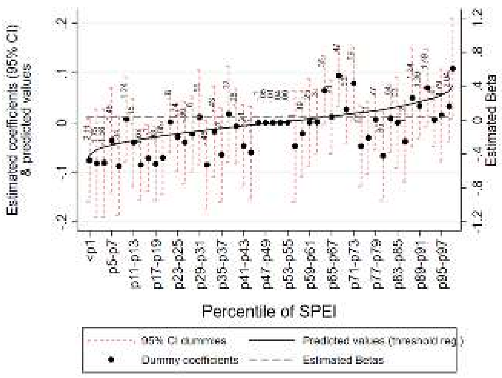
(c) Millet (Kharif)



(d) Maize (Kharif)

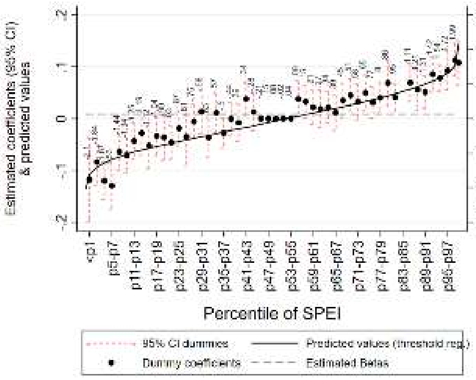


(e) Sorghum - Kharif

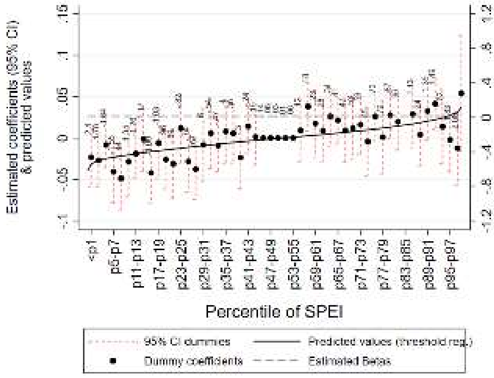


(f) Sorghum (Rabi)

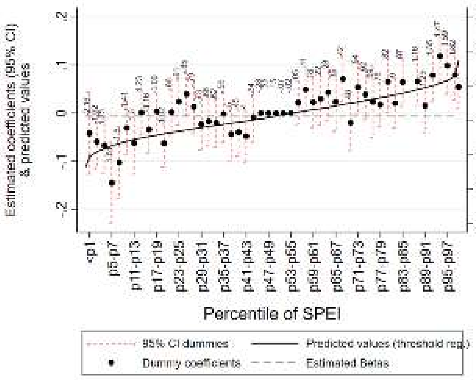
Figure S.5. Plots of dummies (percentiles) vs. threshold regressions (rice, maize, sorghum and millet)



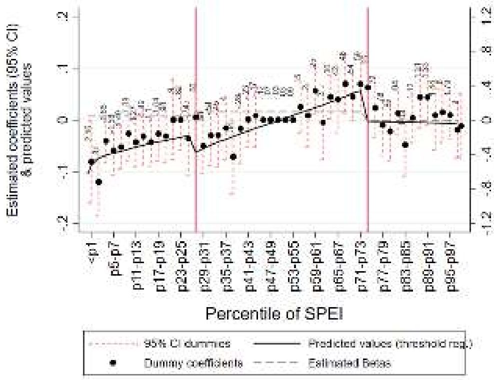
(a) Wheat (Kharif)



(b) Wheat (Rabi)

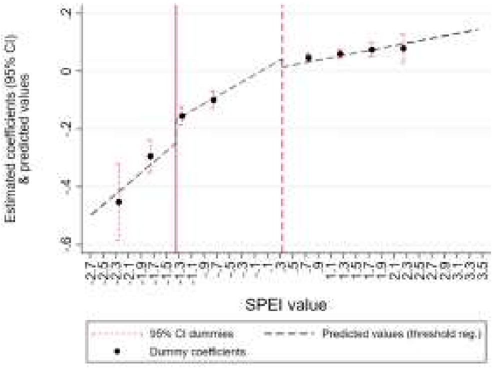


(c) Barley (Kharif)

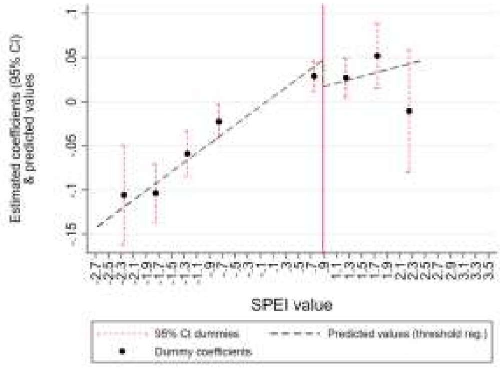


(d) Barley (Rabi)

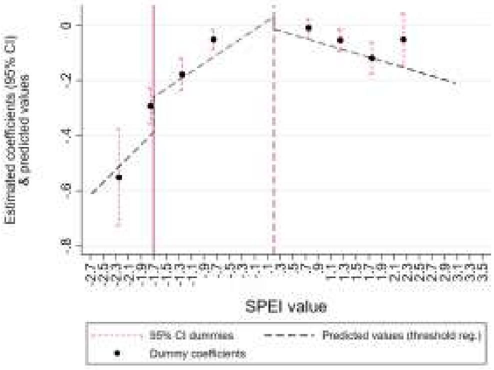
Figure S.6. Plots of dummies (percentiles) vs. threshold regressions (wheat and barley)



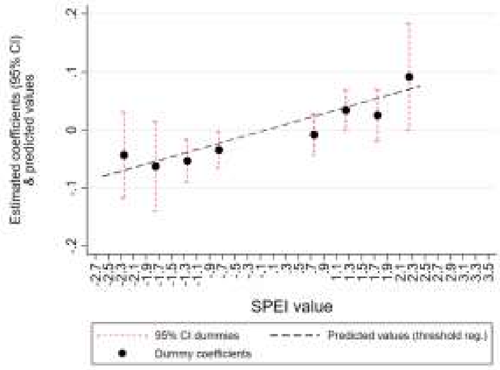
(a) Rice (kharif)



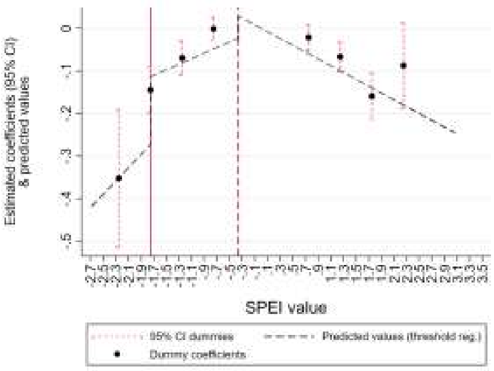
(b) Rice (rabi)



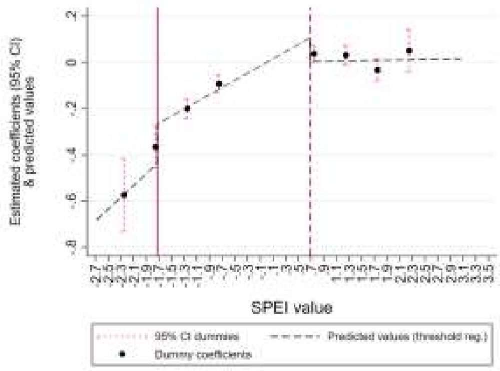
(c) Sorghum (kharif)



(d) Sprghum (rabi)

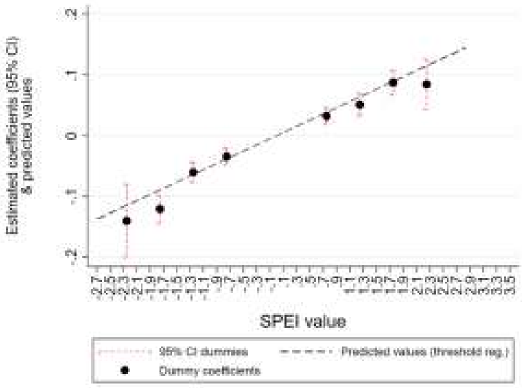


(e) Maize (kharif)

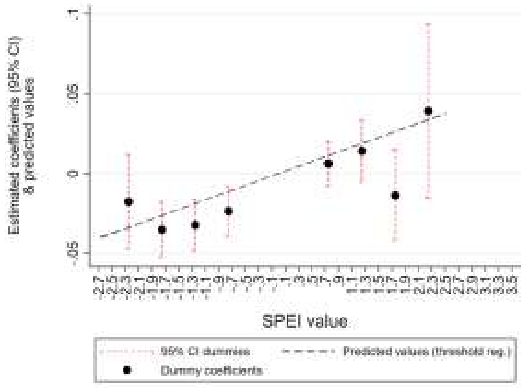


(f) Millet (kharif)

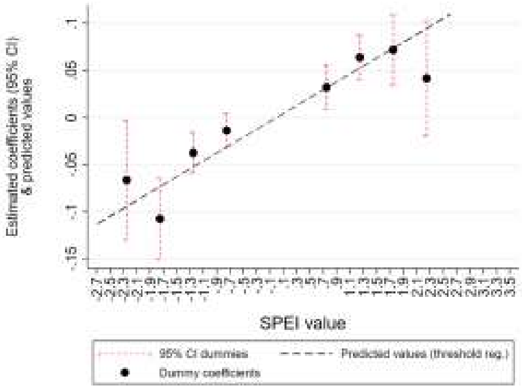
Figure S.7. Plots of dummies (0.5 increments) vs. threshold regressions (rice, sorghum, maize and millet)



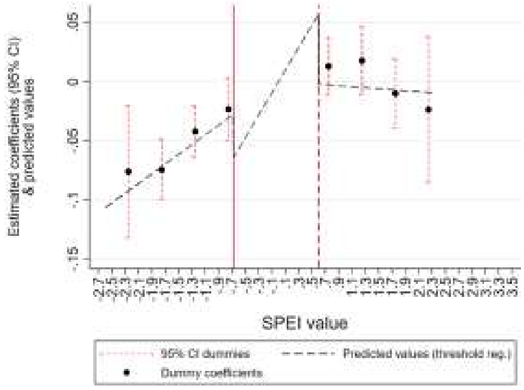
(a) Wheat (kharif)



(b) Wheat (rabi)

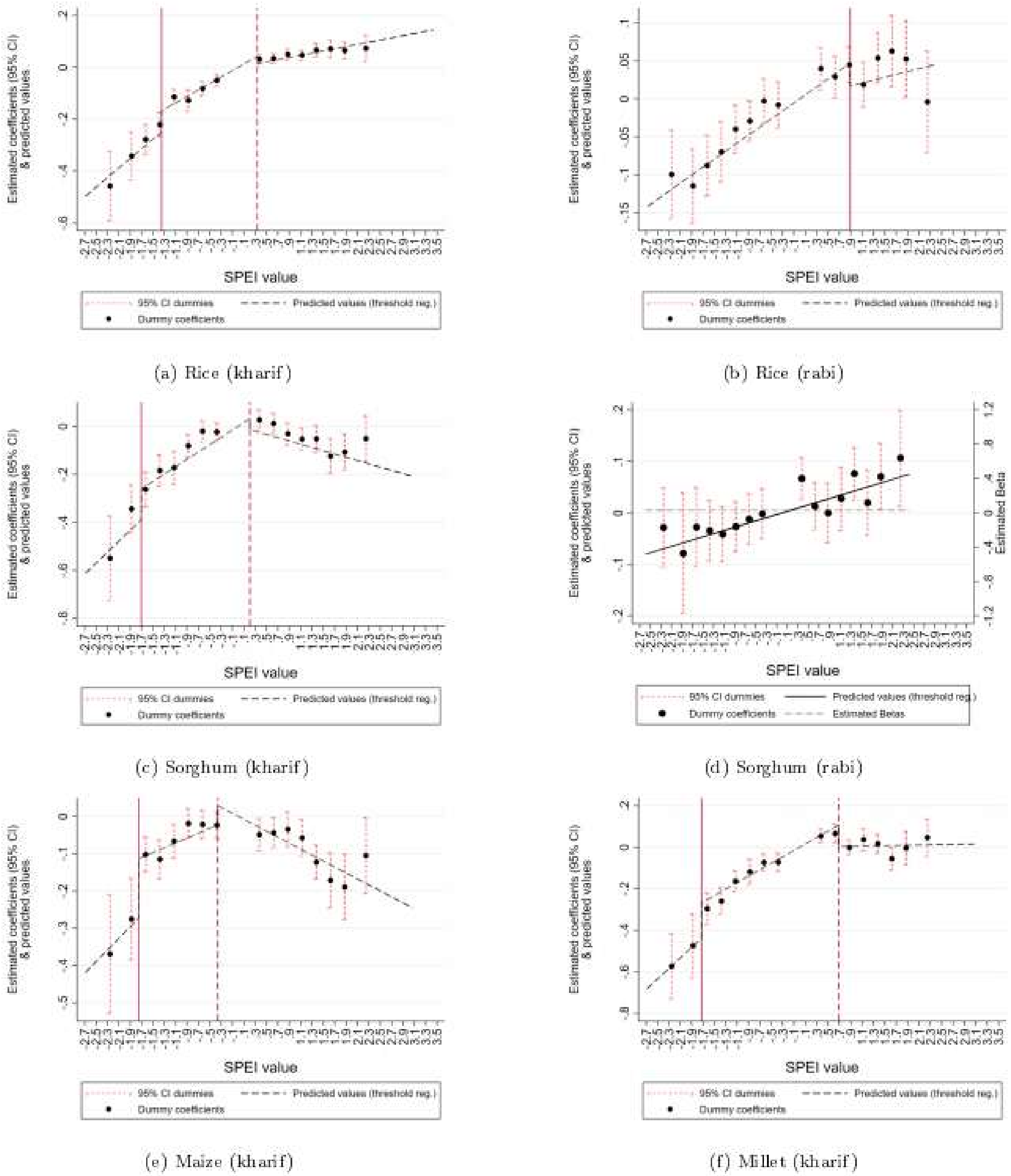


(c) Barley (kharif)



(d) Barley (rabi)

Figure S.8. Plots of dummies (0.5 increments) vs. threshold regressions (wheat and barley)



(a) Rice (kharif)

(b) Rice (rabi)

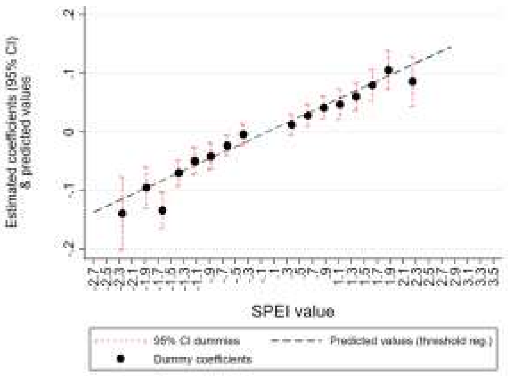
(c) Sorghum (kharif)

(d) Sorghum (rabi)

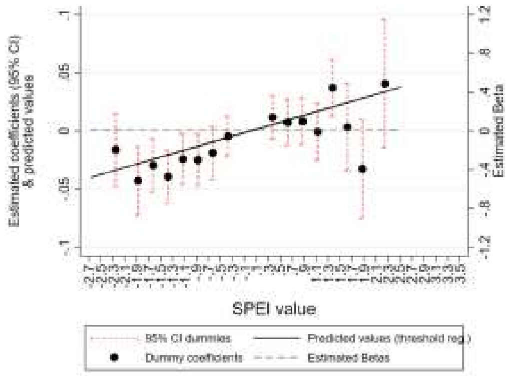
(e) Maize (kharif)

(f) Millet (kharif)

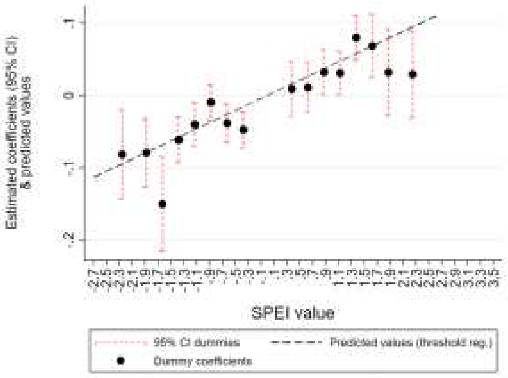
Figure S.9. Plots of dummies (0.25 increments) vs. threshold regressions (rice, sorghum, maize and millet)



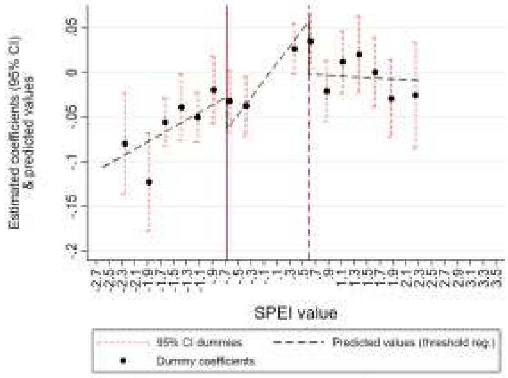
(a) Wheat (kharif)



(b) Wheat (rabi)

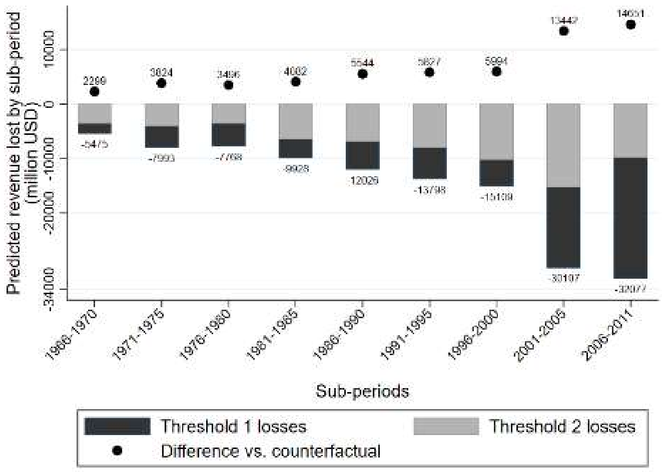


(c) Barley (kharif)

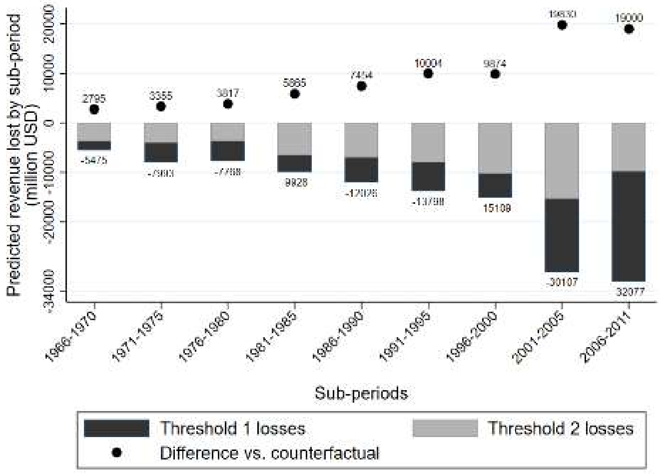


(d) Barley (rabi)

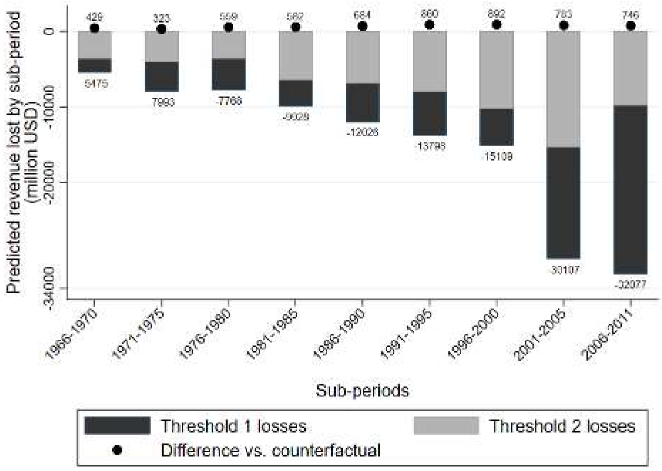
Figure S.10. Plots of dummies (0.25 increments) vs. threshold regressions (wheat and barley)



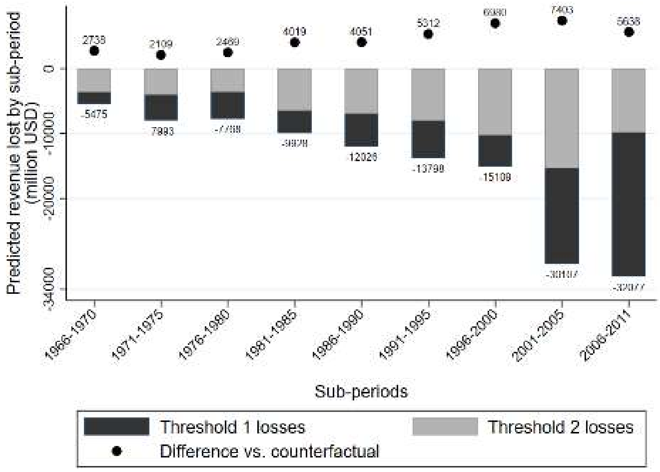
(a) vs. linear



(b) vs. linear (Threshold at $SPEI = -1$)



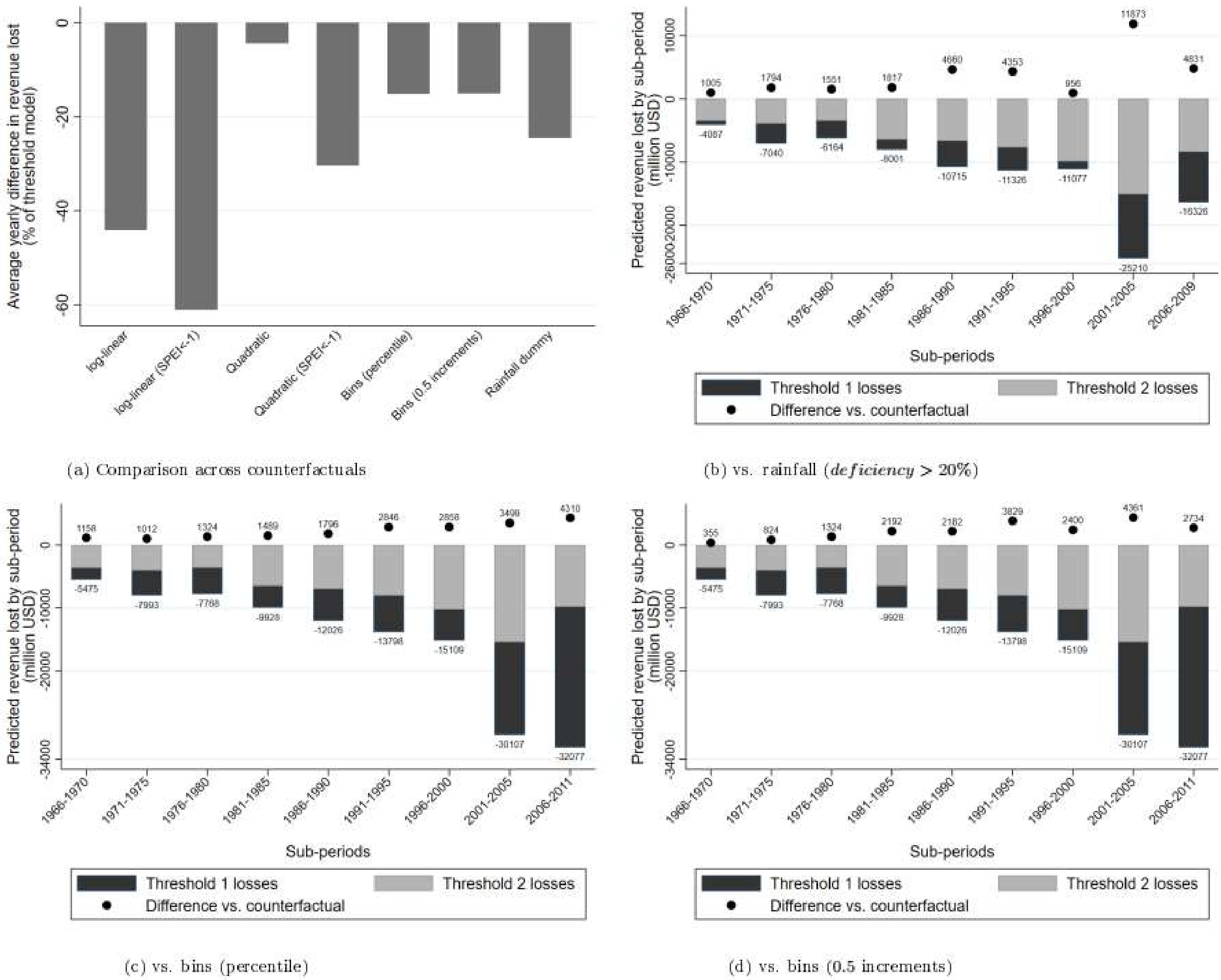
(c) vs. quadratic



(d) vs. quadratic (Threshold at $SPEI = -1$)

Figure S.11. Estimated revenue loss for all crops per sub-period vs. continuous counterfactuals

Note: Total costs per sub-period are estimated by summing the predicted yields given the observed SPEI value vs. predicted yields at SPEI equal 0 for each crop for which a threshold is found. The difference implied by the threshold is estimated by comparing the implied yields under the threshold model given observed SPEI values against the implied yields given observed SPEI values using alternative counterfactuals. The comparison across counterfactuals is obtained by dividing the average difference between the threshold model and a given counterfactual by the average predicted revenue loss using the threshold model. Panel (a) compares the threshold model with the log-linear model and panel (b) imposes an arbitrary threshold at SPEI=-1 for the log-linear model. Panels (c and d) impose a quadratic specification for the negative range of the SPEI (with and without a pre-determined threshold).



(a) Comparison across counterfactuals

(b) vs. rainfall ($deficiency > 20\%$)

(c) vs. bins (percentile)

(d) vs. bins (0.5 increments)

Figure S.12. Comparison across all counterfactuals and estimated revenue loss for all crops per sub-period vs. discrete counterfactuals

Note: Total costs per sub-period are estimated by summing the predicted yields given the observed SPEI value vs. predicted yields at SPEI equal 0 for each crop for which a threshold is found. The difference implied by the threshold is estimated by comparing the implied yields under the threshold model given observed SPEI values against the implied yields given observed SPEI values using alternative counterfactuals. The comparison across counterfactuals is obtained by dividing the average difference between the threshold model and a given counterfactual by the average predicted revenue loss using the threshold model. Panel (a) shows the differences between counterfactuals. Panel (b) compares the threshold model against a simple dummy variable where the pre-determined rainfall threshold is located at 20% negative deviation from rainfall. Panels (g) and (h) compare the threshold model to the bins model (percentiles and increments.

Variable	N	Mean	S. D.	Min.	Max.
Yield					
Cereal yield (t/ha)	11132	1.511	0.843	0.006	5.422
Barley yield (t/ha)	5436	1.488	0.730	0.108	6.000
Maize yield (t/ha)	10055	1.485	0.992	0.006	11.120
Millet yield (t/ha)	9302	0.826	0.462	0.000	4.000
Rice yield (t/ha)	11132	112.989	142.289	0.000	1125.700
Sorghum yield (t/ha)	9464	0.781	0.446	0.001	9.836
Wheat yield (t/ha)	9732	1.719	0.928	0.071	6.129
Area					
Cereal Area (1,000,000 ha)	11132	0.326	0.184	0.000	1.334
Barley (% of total district area)	11009	0.016	0.037	0.000	0.320
Maize (% of total district area)	11132	0.069	0.118	0.000	0.838
Millet (% of total district area)	11132	0.146	0.228	0.000	1.000
Rice (% of total district area)	11132	0.343	0.333	0.000	1.000
Sorghum (% of total district area)	11020	0.165	0.231	0.000	0.929
Wheat (% of total district area)	11131	0.262	0.255	0.000	0.976
Inputs and SPEI					
Proportion of net irrigated area	11125	0.372	0.273	0.000	1.530
Rural population density	10937	3.273	1.796	0.396	13.864
Fertiliser (t/ha)	10937	62.544	63.963	0.000	614.493
Kharif SPEI (4-month)	11132	-0.066	1.000	-2.686	3.432
Rabi SPEI (6-month)	11132	-0.045	0.983	-2.673	2.515

Table S.1. Summary statistics

Note: N refers to the total number of observations. S. D. refers to the standard deviation. Min. and Max. refer to the minimum and maximum values. Rural population density is calculated as the total rural population divided by the gross cropped area. Fertilizer intensity is obtained by dividing total fertilizer used by gross cropped area. The mean *kharif* 4-month SPEI is the SPEI with 4-month lag in September. The *rabi* SPEI is the SPEI with a 6-month lag in March.

Variable	N districts	Time periods	LLC adjusted t-statistics	LLC p-value	IPS t-statistic	IPS p-value
			Barley			
ln barley yield	78	46	-6.064	0.000	-3.631	0.000
			Maize			
ln maize yield	141	46	-21.769	0.000	-22.889	0.000
			Millet			
ln millet yield	136	46	-21.553	0.000	-20.040	0.000
			Rice			
ln rice yield	206	46	-21.028	0.000	-23.038	0.000
			Sorghum			
ln sorghum yield	114	46	-25.274	0.000	-26.02	0.000
			Wheat			
ln wheat yield	181	46	-12.963	0.000	-7.490	0.000

Table S.2. Unit root tests

Notes: This table presents the results of the unit-root tests for the dependent variable of every sub-sample. LLC is the Levin-Liu-Chu panel unit root test and IPS is the Im-Pesaran-Chin panel unit root test. In both cases, the null of all panels having a unit root is rejected at the 1% level.

Sample	β	T1	β	T2	β
Kharif Thresholds					
Barley (N=3588)					
Δ y for 0.1 $\uparrow$ in SPEI					
Maize (N=6486)	0.156***	**-1.746**	0.065***	**-0.358**	-0.080***
Δ y for 0.1 $\uparrow$ in SPEI	*1.688*	[p. 3.9; rp. 25.6]	*0.672*	[p. 43.2; rp. 2.3]	*-0.769*
Millet (N=6256)	0.254***	**-1.724**	0.155***	**0.689**	0.005
Δ y for 0.1 $\uparrow$ in SPEI	*2.892*	[p. 4.2; rp. 23.7]	*1.677*	[p. 74.3; rp. 1.3]	*0.050*
Rice (N=9476)	0.185***	**-1.348**	0.125***	**0.339**	0.042***
Δ y for 0.1 $\uparrow$ in SPEI	*2.032*	[p. 10.7; rp. 9.3]	*1.331*	[p. 63.9; rp. 1.6]	*0.429*
Sorghum (N=5244)	0.228***	**-1.702**	0.153***	**-0.205**	-0.068***
Δ y for 0.1 $\uparrow$ in SPEI	*2.561*	[p. 4.25; rp. 23.6]	*1.653*	[p. 46.4; rp. 2.2]	*-0.657*
Wheat (N=8326)					
Δ y for 0.1 $\uparrow$ in SPEI					
Rabi - Thresholds					
Barley (N=3588)	0.041***	**-0.674**	0.095***	**0.600**	-0.004
Δ y for 0.1 $\uparrow$ in SPEI	*0.419*	[p 32.3; rp. 3.1]	*0.997*	[p 74.7; rp. 1.3]	*-0.040*
Rice (N=9476)	0.053***	**0.890**	0.019***		
Δ y for 0.1 $\uparrow$ in SPEI	*0.544*	[p 80.7; rp. 1.2]	*0.192*		*0.000*
Sorghum (N=5244)					
Δ y for 0.1 $\uparrow$ in SPEI					
Wheat (N=8326)					
Δ y for 0.1 $\uparrow$ in SPEI					

Table S.3. Summary of coefficients (all specifications)

Note: *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. Numbers in bold denote the estimated threshold values of the SPEI. T1 and T2 denote thresholds 1 and 2, respectively. Numbers in italics represent the predicted effect of a 0.01 increase in the SPEI value. For example, for a coefficient of 0.185 means that, for a given event in the rice sample for a value of the SPEI below -1.348, a 0.1 decrease in the index leads to a fall in yield of 2.032 percent. The abbreviations p. and rp. below the value of the identified threshold represent the percentile of the distribution and the associated return period. N refers to the number of observations. The number of districts can be obtained by dividing N by 46.

	Kh -Full	Kh - Low-irrig	Kh - High irrig	Kh -Arid	Kh - Humid	Ra -Full	Ra - Low-irrig	Ra - High irrig	Ra - Arid	Ra - Humid
Rice										
Single	0	0	0	0	0.002	0.01	0.45	0.517	0.007	0.006
Double	0	0.244	0.144	0	0.006	0.497	0.415	0.312	0.146	0.972
Triple	0.938	0.553	0.33	0.431	0.854	0.318	0.522	0.578	0.329	0.388
Wheat										
Single	0.13	0.151	0.297	0.38	0.429	0.184	0.006	0.625	0.893	0.001
Double	0.114	0.456	0.094	0.001	0.002	0.352	0.668	0.001	0.438	0.947
Triple	0.895	0.799	0.982	0.687	0.684	0.715	0.852	0.968	0.969	0.656
Sorghum										
Single	0	0	0	0	0.002	0.833	0.939	0.028	0.375	0.423
Double	0.002	0.583	0.105	0	0.563	0.11	0.268	0.009	0.94	0
Triple	0.474	0.636	0.297	0.499	0.834	0.859	0.847	0.868	0.617	0.898
Millet										
Single	0	0	0	0	0.085	NA	NA	NA	NA	NA
Double	0	0	0.007	0	0.142	NA	NA	NA	NA	NA
Triple	0.745	0.735	0.8	0.878	0.553	NA	NA	NA	NA	NA
Maize										
Single	0	0	0	0	0	NA	NA	NA	NA	NA
Double	0	0.01	0	0	0	NA	NA	NA	NA	NA
Triple	0.738	0.889	0.536	0.368	0.778	NA	NA	NA	NA	NA
Barley										
Single	0.722	NA	NA	NA	NA	0.002	NA	NA	NA	NA
Double	0.044	NA	NA	NA	NA	0.042	NA	NA	NA	NA
Triple	0.989	NA	NA	NA	NA	0.852	NA	NA	NA	NA

Table S.4. P-value for selection of the threshold model

Note: Blank cells indicate that the threshold test did not reject the null of no threshold. NA means that the threshold model was not estimated.

	Kh -Full	Kh - Low-irrig	Kh - High irrig	Kh -Arid	Kh - Humid	Ra -Full	Ra - Low-irrig	Ra - High irrig	Ra - Arid	Ra - Humid
Rice										
γ_1	-1.348 [-1.375, -1.334]	0.583 [0.119; 0.589]	-1.561 [-1.580, 1.534]	-1.759 [-1.790; -1.723]	-1.797 [-1.846, -1.769]	0.890 [0.192,0.909]			0.873 [0.668, 0.883]	-1.873 [-1.900,-1.837]
γ_2	0.339 [0.136,0.347]			0.344 [-0.022, 0.352]	-1.322 [-1.348, -1.769]					
Wheat										
γ_1							1.550 [1.481, 1.577]			1.635 [1.553, 1.669]
γ_2										
Sorghum										
γ_1	-1.702 [-1.800,-1.683]	-1.021 [-1.026, -1.014]	-1.518 [-1.562,-1.518]	-1.775 [-1.826, -1.010]	-0.739 [-0.815, -0.727]			0.580 [0.553, 0.586]		
γ_2	-0.205 [-0.309,-0.196]			-0.205 [-0.276, -0.193]				1.108 [1.024, 1.117]		
Millet										
γ_1	-1.724 [-1.765,-1.702]	-1.730 [-1.767, -1.713]	-1.443 [-1.605, -1.430]	-1.730 [-1.776, -1.709]	0.725 [-0.305, 0.733]	NA NA	NA NA	NA NA	NA NA	NA NA
γ_2	0.689 [0.654,0.699]	0.723 [0.664, 0.731]	0.692 [0.617, 0.690]	0.687 [0.647, 0.700]		NA	NA	NA	NA	NA
Maize										
γ_1	-1.746 [-1.791,-1.721]	-1.743 [-1.786, -1.722]	-1.972 .	-1.764 [-1.791, -1.724]	-0.357 [-0.387, -0.347]	NA NA	NA NA	NA NA	NA NA	NA NA
γ_2	-0.358 [-0.397, -0.347]	-0.359 [-0.467, -0.353]	-0.915 [-0.945, -0.907]	0.553 [0.453, 0.562]	0.825 [0.691, 0.838]	NA	NA	NA	NA	NA
Barley										
γ_1		NA NA	NA NA	NA NA	NA NA	-0.674 [-0.818,-0.665]	NA NA	NA NA	NA NA	NA NA
γ_2		NA	NA	NA	NA	0.600 [0.582,0.609]	NA	NA	NA	NA

Table S.5. Threshold location and confidence intervals

Note**:** Blank cells indicate that the threshold test did not reject the null of no threshold. NA means that the threshold model was not estimated.

Kharif Thresholds					
Sample	β	T1	β	T2	β
Rice - Full sample (N=9476)	0.185***	**-1.348**	0.125***	**0.339**	0.042***
Δ y for 0.1 ↑ in SPEI	*2.032*	[p. 10.7; rp. 9.3]	*1.331*	[p. 63.9; rp. 1.6]	*0.429*
Rice - Low irrigation (N=3496)	0.195***	**0.583**	0.076***		
Δ y for 0.1 ↑ in SPEI	*2.153*	[p. 71.9; rp. 1.4]	*0.790*		
Rice - High irrigation (N=3496)	0.125***	**-1.561**	0.034***		
Δ y for 0.1 ↑ in SPEI	*2.399*	[p. 6.2; rp. 16.0]	*0.346*		
Rice - Arid (N=5060)	0.331***	**-1.759**	0.172***	**0.344**	0.035***
Δ y for 0.1 ↑ in SPEI	*3.924*	[p. 1.6; rp. 56.6]	*1.877*	[p. 55.7; rp. 1.20]	*0.356*
Rice - Humid (N=4416)	0.064***	**-1.797**	0.137***	**-1.322**	0.067***
Δ y for 0.1 ↑ in SPEI	*0.066*	[p. 2.7; rp. 37.4]	*0.147*	[p. 10.1; rp. 9.9]	*0.069*
Rabi - Thresholds					
Rice -Full sample (N=9476)	0.053***	**0.890**	0.019***		
Δ y for 0.1 ↑ in SPEI	*0.544*	[p 80.7; rp. 1.2]	*0.192*		
Rice - Low irrigation (N=3496)					
Δ y for 0.1 ↑ in SPEI					
Rice - High irrigation (N=3496)					
Δ y for 0.1 ↑ in SPEI					
Rice - Arid (N=5060)	0.062***	**0.873**	0.006		
Δ y for 0.1 ↑ in SPEI	*0.640*	[p. 80.2; rp. 1.2]	*0.060*		
Rice - Humid (N=4416)	0.092***	**-1.873**	0.033***		
Δ y for 0.1 ↑ in SPEI	*0.964*	[p. 1.5; rp. 67.5]	*0.336*		

Table S.6. Summary of coefficients for rice sub-samples

Note: *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. For the irrigation sub-sample, the used cut-off was 70.1% of irrigated rice

Kharif Thresholds					
Sample	β	T1	β	T2	β
Millet - Full sample (N=6256)	0.254***	**-1.724**	0.155***	**0.689**	0.005
Δ y for 0.1 ↑ in SPEI	*2.892*	[p. 4.2; rp. 23.7]	*1.677*	[p. 74.3; rp. 1.3]	*0.050*
Millet - Low irrigation (N=3128)	0.269***	**-1.730**	0.160***	**0.723**	0.003
Δ y for 0.1 ↑ in SPEI	*3.087*	[p. 4.1; rp. 24.2]	*1.735*	[p. 75.3; rp. 1.3]	*0.030*
Millet - High irrigation (N=3128)	0.221***	**-1.443**	0.135***	**0.692**	0.006
Δ y for 0.1 ↑ in SPEI	*2.473*	[p. 8.5; rp. 11.8]	*1.445*	[p. 74.4; rp. 1.3]	*0.060*
Millet - Arid (N=4646)	0.365***	**-1.730**	0.183***	**0.687**	0.002
Δ y for 0.1 ↑ in SPEI	*4.405*	[p. 4.2; rp. 23.3]	*2.008*	[p. 70.7; rp. 1.4]	*0.020*
Millet - Humid (N=1610)	0.073***	**0.725**	0.001		
Δ y for 0.1 ↑ in SPEI	*0.757*	[p. 73.7; rp. 1.4]	*0.010*		

Table S.7. Summary of coefficients for millet sub-samples

Note: *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. For the irrigation sub-sample, the used cut-off was 2.4% of irrigated millet

Kharif Thresholds					
Sample	β	T1	β	T2	β
Maize - Full sample (N=6486)	0.156***	**-1.746**	0.065***	**-0.358**	-0.080***
Δ y for 0.1 ↑ in SPEI	*1.688*	[p. 3.9; rp. 25.6]	*0.672*	[p. 43.2; rp. 2.3]	*-0.769*
Maize - Low irrigation (N=2622)	0.121***	**-1.743**	0.032*	**-0.359**	-0.097***
Δ y for 0.1 ↑ in SPEI	*1.286*	[p. 3.9; rp. 25.6]	*0.325*	[p. 43.2; rp. 2.3]	*-0.924*
Maize - High irrigation (N=2668)	0.238***	**-1.972**	0.094***	**-0.915**	-0.062***
Δ y for 0.1 ↑ in SPEI	*2.687*	[p. 1.8; rp. 54.7]	*0.986*	[p. 24.1; rp. 4.1]	*0.640*
Maize - Arid (N=3128)	0.314***	**-1.764**	0.126***	**0.553**	-0.099***
Δ y for 0.1 ↑ in SPEI	*3.689*	[p. 4.3; rp. 21.6]	*1.343*	[p. 70.1; rp. 1.4]	*-0.943*
Maize - Humid (N=3358)	0.022**	**-0.357**	-0.204***	**0.825**	-0.051***
Δ y for 0.1 ↑ in SPEI	*0.222*	[p. 41.8; rp. 2.4]	*-1.845*	[p. 79.4; rp. 1.3]	*-0.497*

Table S.8. Summary of coefficients for maize sub-samples

Note: *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. For the irrigation sub-sample, the used cut-off was 6.4% of irrigated wheat

Sample	β	T1	β	T2	β
Kharif Thresholds					
Sorghum - Full sample (N=5244)	0.228***	**-1.702**	0.153***	**-0.205**	-0.068***
Δ y for 0.1 ↑ in SPEI	*2.561*	[p. 4.25; rp. 23.6]	*1.653*	[p. 46.4; rp. 2.2]	*-0.657*
Sorghum - Low irrigation (N=1932)	0.191***	**-1.021**	-0.087***		
Δ y for 0.1 ↑ in SPEI	*2.105*	[p. 19.3; rp. 5.2]	*-0.833*		
Sorghum - High irrigation (N=1886)	0.170***	**-1.518**	0.002		
Δ y for 0.1 ↑ in SPEI	*1.853*	[p. 6.7; rp. 14.8]	*0.020*		
Sorghum - Arid (N=4048)	0.308***	**-1.775**	0.178***	**-0.205**	-0.066***
Δ y for 0.1 ↑ in SPEI	*3.607*	[p. 1.7; rp. 51.1]	*1.948*	[p. 41.7; rp. 2.4]	*-0.639*
Sorghum - Humid (N=1196)	0.070***	**-0.739**	-0.068***		
Δ y for 0.1 ↑ in SPEI	*0.725*	[p. 27.2; rp. 3.7]	*-0.657*		
Rabi - Thresholds					
Sorghum - Full sample (N=5244)					
Δ y for 0.1 ↑ in SPEI					
Sorghum - Low irrigation (N=1932)					
Δ y for 0.1 ↑ in SPEI					
Sorghum - High irrigation (N=1886)	0.013	0.580	-0.105***	**1.108129**	0.019
Δ y for 0.1 ↑ in SPEI	*0.131*	[p. 72.7; rp. 1.4]	*-0.997*	[p. 86.6; rp. 1.2]	*0.192*
Sorghum - Arid (N=4048)					
Δ y for 0.1 ↑ in SPEI					
Sorghum - Humid (N=1196)					
Δ y for 0.1 ↑ in SPEI					

Table S.9. Summary of coefficients for sorghum sub-samples

Note: *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. For the irrigation sub-sample, the used cut-off was 1% of irrigated sorghum

Sample	β	T1	β	T2	β
Kharif Thresholds					
Wheat - Full sample (N=8326)					
Δ y for 0.1 $\uparrow$ in SPEI					
Wheat - Low irrigation (N=3036)					
Δ y for 0.1 $\uparrow$ in SPEI					
Wheat - High irrigation (N=2990)					
Δ y for 0.1 $\uparrow$ in SPEI					
Wheat - Arid (N=5152)					
Δ y for 0.1 $\uparrow$ in SPEI					
Wheat - Humid (N=3174)					
Δ y for 0.1 $\uparrow$ in SPEI					
Rabi - Thresholds					
Wheat - Full sample (N=8326)					
Δ y for 0.1 $\uparrow$ in SPEI					
Wheat - Low irrigation (N=3036)	0.024***	**1.550**	-0.034**		
Δ y for 0.1 $\uparrow$ in SPEI	*0.243*	[p. 93.9; rp. 1.1]	*-0.334*		
Wheat - High irrigation (N=2990)					
Δ y for 0.1 $\uparrow$ in SPEI					
Wheat - Arid (N=5152)					
Δ y for 0.1 $\uparrow$ in SPEI					
Wheat - Humid (N=3174)	0.009*	**1.635**	-0.042***		
Δ y for 0.1 $\uparrow$ in SPEI	*0.090*	[p. 95.0; rp. 1.1]	*-0.411*		

Table S.10. Summary of coefficients for wheat sub-samples

Note: *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. For the irrigation sub-sample, the used cut-off was 86.06% of irrigated wheat

	1% (main)	0.50%	1.50%	2.50%	1% (main)	0.50%	1.50%	2.50%
	Kharif - Rice				Rabi - Rice			
	Threshold test p-value							
Single	0.000	0.000	0.000	0.000	0.010	0.030	0.020	0.000
Double	0.000	0.000	0.000	0.000	0.497	0.320	0.510	0.400
Triple	0.938	0.900	0.990	0.860	0.318	0.500	0.280	0.350
	Threshold location and Confidence intervals							
γ_1	-1.348	-1.349	-1.346	-1.303	0.890	0.880	0.876	0.887
γ_2	0.339	0.343	0.343	0.341				
	Kharif - Wheat				Rabi - Wheat			
	Threshold test p-value							
Single	0.130	0.170	0.150	0.130	0.184	0.250	0.160	0.150
Double	0.110	0.140	0.360	0.270	0.352	0.060	0.570	0.570
Triple	0.895	0.750	0.850	0.960	0.715	0.730	0.900	0.910
	Threshold location and Confidence intervals							
γ_1								
γ_2								
	Kharif - Barley				Rabi - Barley			
	Threshold test p-value							
Single	0.722	0.240	0.690	0.680	0.002	0.000	0.010	0.000
Double	0.044	0.280	0.030	0.780	0.042	0.030	0.010	0.010
Triple	0.989	0.960	1.000	0.380	0.852	0.770	0.930	0.850
	Threshold location and Confidence intervals							
γ_1					-0.674	-0.681	-0.688	-0.681
γ_2					0.600	0.656	0.601	0.600
	Kharif - Sorghum				Rabi - Sorghum			
	Threshold test p-value							
Single	0.000	0.000	0.000	0.000	0.833	0.870	0.800	0.830
Double	0.002	0.010	0.000	0.000	0.110	0.120	0.090	0.090
Triple	0.474	0.280	0.420	0.380	0.859	0.810	0.890	0.820
	Threshold location and Confidence intervals							
γ_1	-1.702	-1.689	-1.693	-1.695				
γ_2	-0.205	-0.208	-0.203	-0.289				
	Kharif - Maize				Kharif - Millet			
	Threshold test p-value							
Single	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Double	0.037	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Triple	0.612	0.540	0.580	0.490	0.745	0.790	0.710	0.770
	Threshold location and Confidence intervals							
γ_1	-0.358	-0.362	-0.362	-0.362	-1.724	-1.730	-1.738	-1.738
γ_2	-1.746	-1.730	-1.733	-1.734	0.689	0.687	0.689	0.687

Table S.11. Sensitivity to trimming cut-off point

	Main	Cluster year	Cluster State year	Conley 50km	Conley 100km	Conley 250km	Conley 500km	Conley 1000km
				Kharif - Rice (T1: -1.348; T2: 0.339)				
β<T1	0.185***	0.185***	0.185***	0.185***	0.185***	0.185***	0.185***	0.185***
T1<β<T2	0.125***	0.125***	0.125***	0.125***	0.125***	0.125***	0.125***	0.125***
β>T2	0.042***	0.042**	0.042***	0.042***	0.042***	0.042***	0.042***	0.042***
				Kharif - Maize (T1: -1.746; T2: -0.358)				
β<T1	0.156***	0.156***	0.156***	0.156***	0.156***	0.156***	0.156***	0.156***
T1<β<T2	0.065***	0.065**	0.065**	0.065***	0.065***	0.065***	0.065**	0.065**
β>T2	-0.080***	-0.080**	-0.080**	-0.080***	-0.080***	-0.080***	-0.080***	-0.080***
				Kharif - Millet (T1: -1.724; T2: 0.689)				
β<T1	0.254***	0.254***	0.254***	0.254***	0.254***	0.254***	0.254***	0.254***
T1<β<T2	0.155***	0.155***	0.155***	0.155***	0.155***	0.155***	0.155***	0.155***
β>T2	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
				Kharif - Sorghum (T1: -1.702; T2: -0.205)				
β<T1	0.228***	0.228***	0.228***	0.228***	0.228***	0.228***	0.228***	0.228***
T1<β<T2	0.153***	0.153***	0.153**	0.153***	0.153***	0.153***	0.153***	0.153***
β>T2	-0.068***	-0.068***	-0.068**	-0.068***	-0.068***	-0.068***	-0.068***	-0.068***
				Kharif - Wheat (T1: NA; T2: NA)				
β<T1	0.051***	0.051***	0.051***	0.051***	0.051***	0.051***	0.051***	0.051***
				Kharif - Barley (T1: NA; T2: NA)				
β<T1	0.042***	0.042***	0.042**	0.042***	0.042***	0.042***	0.042***	0.042***
				Rabi - Rice (T1: 0.890; T2: NA)				
β<T1	0.053***	0.053***	0.053**	0.053***	0.053***	0.053***	0.053***	0.053***
T1<β<T2	0.019***	0.019	0.019	0.019***	0.019**	0.019	0.019	0.019
				Rabi - Sorghum (T1: NA; T2: NA)				
β<T1	0.031***	0.031	0.031	0.031***	0.031***	0.031**	0.031*	0.031*
				Rabi - Wheat (T1: NA; T2: NA)				
β<T1	0.015***	0.015*	0.015	0.015***	0.015***	0.015***	0.015**	0.015**
				Rabi - Barley (T1: -0.674; T2: -0.600)				
β<T1	0.041***	0.041**	0.041***	0.041***	0.041***	0.041***	0.041***	0.041***
T1<β<T2	0.095***	0.097***	0.097***	0.097***	0.097***	0.097***	0.097***	0.097***
β>T2	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004

Table S.12. Coefficients – Robustness to SEs

Note: The column labelled main refers to our preferred specification with clustered standard errors at the district level. The columns cluster year and cluster state-year refer to the robustness checks where the standard errors were clustered at the year and two-way state-year clustering. All other columns refer to the coefficients when the regressions are estimated with Conley (1999) standard errors, allowing for correlation at different spatial scales.

	Rice			Barley			Maize			Millet			Sorghum			Wheat		
	Number of SPEI lags																	
	4 (main)	5	6	4 (main)	5	6	4 (main)	5	6	4 (main)	5	6	4 (main)	5	6	4 (main)	5	6
	Threshold test p-value																	
Single	0	0	0	0.722	0.18	0.57	0.000	0	0	0	0	0	0	0	0	0.13	0	0.02
Double	**0**	0	0	0.044	0	0	**0.037**	0	0	**0**	0	0	**0.002**	0	0	0.11	0.25	0.16
Triple	0.938	0.91	0.64	0.989	0.99	0.97	0.612	0.8	0.61	0.745	0.97	0.42	0.474	0.64	0.61	0.895	0.64	0.7
	Threshold location																	
γ_1	-1.348	-1.347	-1.150	NA	NA	NA	-0.358	-0.246	-0.259	-1.724	-1.714	-1.715	-1.702	-1.990	-1.141	NA	-1.348	0.392
γ_2	0.339	0.107	-0.053	NA	NA	NA	-1.746	-1.734	-1.731	0.689	0.696	0.587	-0.205	-0.071	0.737	NA	NA	NA
	Estimated coefficients																	
β<T1	0.185***	0.215***	0.206***				0.156***	0.153***	0.153***	0.254***	0.256***	0.262***	0.228***	0.282***	0.206***		0.073***	0.067***
	(0.016)	(0.016)	(0.016)				(0.025)	(0.025)	(0.024)	(0.031)	(0.030)	(0.032)	(0.025)	(0.039)	(0.022)		(0.007)	(0.006)
T1<β<T2	0.125***	0.159***	0.154***				0.065***	0.060***	0.056***	0.155***	0.161***	0.176***	0.153***	0.177***	0.075***		0.045***	0.034***
	(0.013)	(0.016)	(0.018)				(0.013)	(0.014)	(0.015)	(0.018)	(0.019)	(0.018)	(0.021)	(0.020)	(0.018)		(0.004)	(0.006)
β>T2	0.042***	0.031***	0.028***				-0.080***	-0.083***	-0.087***	0.005	0.008	0.005	-0.068***	-0.080***	-0.045***			
	(0.005)	(0.006)	(0.007)				(0.010)	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)	(0.013)	(0.015)	(0.013)			

Table S.13. Different lag specifications – *kharif* season

Note: *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. For each subsample, the first column ("4 (main)") represents the SPEI index with a 4-month lag at September. Columns "5" and "6" denote the SPEI index with a 5- and 6-month lag, covering the periods between June-October and June-November, respectively.

Kharif Thresholds					
Sample	β	T1	β	T2	β
Maize (N=2208)	0.125***	**-1.752**	0.017*	**1.431**	-0.068***
Δ y for 0.01 ↑ in SPEI	*1.331*		*0.171*		*-0.657*
Millet (N=2162)	0.286***	**-2.053**	0.132***	**0.444**	-0.001
Δ y for 0.01 ↑ in SPEI	*3.311*		*1.411*		*-0.010*
Rice (N=4324)	0.300***	**-2.004**	0.139***	**-1.305**	0.066***
Δ y for 0.01 ↑ in SPEI	*3.499*		*1.491*		*0.682*
Sorghum (N=2300)	0.300***	**-1.895**	0.131***	**-0.255**	-0.073***
Δ y for 0.01 ↑ in SPEI	*2.561*		*1.653*		*-0.657*
Wheat (N=3404)	0.100***	**-1.524**	0.053***		
Rabi - Thresholds					
Sample	β	T1	β	T2	β
Rice	0.068***	**-1.586**	0.183***	**-1.445**	0.019***
Δ y for 0.01 ↑ in SPEI	*0.704*		*2.008*		*0.192*
Sorghum	0.073***	**-0.822**	0.013		
Δ y for 0.01 ↑ in SPEI	*0.757*		*0.131*		
Wheat	0.027***	**-1.537**	-0.026*		
Δ y for 0.01 ↑ in SPEI	*0.274*		*-0.257*		

Table S.14. Summary of coefficients including controls

Note: *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. Numbers in bold denote the estimated threshold values of the SPEI. T1 and T2 denote threshold 1 and 2, respectively. Numbers in italics represent the predicted effect of a 0.1 increase in the SPEI value.