Supplement to "Climate change and U.S. agriculture: Accounting for multidimensional slope heterogeneity in panel data"

(*Quantitative Economics*, Vol. 11, No. 4, November 2020, 1391–1429)

Michael Keane
Department of Economics, University of New South Wales

Timothy Neal
Department of Economics, University of New South Wales

Appendix A: Shows that the key results of the farmer-level model in Section 1 still hold if we aggregate to county-level data.

Appendix B: Shows that the conventional FE-OLS econometric model of crop yield is likely understating the negative impact of high temperatures on yield.

Appendix C: Compares the Computational Performance of MO-OLS with Brute Force OLS.

Appendix D: Presents Monte Carlo Simulations for MO-OLS.

Appendix E: Shows how the sources of identifying information for the slope parameters differ between the MO-OLS and FE-OLS models.

Appendix F: Presents projections of future GDD_{it} and KDD_{it} from the nineteen climate models.

Appendix G: Incorporates projections of technical progress into our projections of future crop yields.

Appendix H: Presents estimation and projections for soybean yield.

Michael Keane: m.keane@unsw.edu.au
Timothy Neal: timothy.neal@unsw.edu.au

 https://doi.org/10.3982/QE1319

Appendix A: Aggregation of the farm production function

Here, we verify that the main predictions of the farm-level model in Section 1 aggregate to the collection of farmers within each county. We observe county-level crop yield:

$$y_{ct} = \frac{\sum_{i=1}^{N} Y_{it}}{\sum_{i=1}^{N} C_{it}} \tag{A1}$$

for county c at time t, where Y_{it} and C_{it} are output and acreage of farm i, respectively. Assume that KDD_{it} is common to farms within counties (as is true for existing weather measures), giving $\text{KDD}_{it} = \text{KDD}_{ct}\ \forall i$ in c. Thus, the county-level sensitivity to KDD_{ct} is an average of the farm-level sensitivities weighted by the size of the farm in terms of land and other inputs:

$$\beta_{2ct} = \frac{\sum_{i=1}^{N} C_{it} I_{it} \beta_{2it}}{\sum_{i=1}^{N} C_{it}} = \frac{\sum_{i=1}^{N} C_{it} I_{it} s \left(\frac{-p C_{it} A_t \mu_i I_{it} s \text{KDD}_{ct}}{\gamma} \right)^{-1/2}}{\sum_{i=1}^{N} C_{it}}. \tag{A2}$$

Since $s < 0$, $C_{it} > 0$, and $I_{it} > 0$ for all i and t, we can verify that the county sensitivity to harsh temperature β_{2ct} is strictly concave in KDD_{ct}, just as β_{2it} is strictly concave in KDD_{it}:

$$\frac{\partial \beta^*_{2ct}}{\partial \text{KDD}_{ct}} = \frac{\sum_{i=1}^{N} \left(\frac{s^2 A_t \mu_i C_{it} I_{it}^2 p}{2\gamma \left(\frac{-sp A_t \mu_i C_{it} I_{it} \text{KDD}_{ct}}{\gamma} \right)^{3/2}} \right)}{\sum_{i=1}^{N} C_{it}} > 0 \quad \forall \text{KDD}_{ct},$$

$$\frac{\partial^2 \beta^*_{2ct}}{\partial \text{KDD}_{ct}^2} = \frac{\sum_{i=1}^{N} \left(\frac{3 s^3 A_t^2 \mu_i^2 C_{it}^2 I_{it}^3 p^2}{4\gamma^2 \left(\frac{-sp A_t \mu_i C_{it} I_{it} \text{KDD}_{ct}}{\gamma} \right)^{5/2}} \right)}{\sum_{i=1}^{N} C_{it}} < 0 \quad \forall \text{KDD}_{ct}$$

and the similarity to a log-linear function is also maintained. Thus, the two main predictions of the theoretical model carry over to the county-level (even if farm-level production functions cannot be aggregated into a "representative producer").

Appendix B: Bias in fixed effects regressions under slope heterogeneity

In this appendix, we demonstrate how adaptation leads to bias in conventional estimates of crop yield equations. As we showed in Section 1, the use of adaptation techniques will generate heterogeneity in the parameter β_2 that captures sensitivity to high

temperatures in equation (5), versions of which have been estimated in several prior papers in the literature. This heterogeneity may arise in both the i and t dimensions. We can capture this heterogeneity by generalizing the conventional specification (5) as follows:

$$y_{it} = c_i + \tau_t + \beta_1 \text{GDD}_{it} + \beta_{2it} \text{KDD}_{it} + \beta_3 \text{PREC}_{it} + \beta_4 \text{PREC}_{it}^2 + \epsilon_{it}, \tag{B1}$$

where, for identification, we impose the additive structure:

$$\beta_{2it} = \beta_2 + \lambda_i + \theta_t. \tag{B2}$$

Note that this specification allows for unit and time fixed effects in the sensitivity of crop yield to temperature, as well as unit and time fixed effects in the intercept.

To see the bias of a conventional fixed effects estimator—with fixed effects only in the intercepts—in a context where adaptation generates slope heterogeneity, we first simplify (B1) by excluding precipitation and stacking the variables:

$$y_{it} = c_i + \tau_t + z_{it}' \boldsymbol{\beta}_{it} + \epsilon_{it}, \tag{B3}$$

where $z_{it}' = (\text{GDD}_{it}, \text{KDD}_{it})$ and $\boldsymbol{\beta}_{it} = (\beta_1, \beta_{2it})' = (\beta_1, \beta_2 + \lambda_{2i} + \gamma_{2t})'$. Consider a two-way within transformation of (B3) to remove the fixed effects in the intercept term:

$$\tilde{y}_{it} = \tilde{z}_{it}' \boldsymbol{\beta} + v_{it}, \tag{B4}$$

where $\tilde{y}_{it} = y_{it} - N^{-1} \sum_{i=1}^{N} y_{it} - T^{-1} \sum_{t=1}^{T} y_{it} + NT^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it}$ is the two-way within transformation of y_{it}, and similarly for $\tilde{z}_{it}$, $\boldsymbol{\beta} = (\beta_1, \beta_2)'$, and v_{it} is defined as

$$v_{it} = \widetilde{z_{it}' \boldsymbol{\lambda}_i} + \widetilde{z_{it}' \boldsymbol{\theta}_t} + \tilde{\epsilon}_{it}, \tag{B5}$$

where $\widetilde{z_{it}' \boldsymbol{\lambda}_i} = z_{it}' \boldsymbol{\lambda}_i - N^{-1} \sum_{i=1}^{N} z_{it}' \boldsymbol{\lambda}_i - T^{-1} \sum_{t=1}^{T} z_{it}' \boldsymbol{\lambda}_i + NT^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} z_{it}' \boldsymbol{\lambda}_i$ is the two-way within transformation of $z_{it}' \boldsymbol{\theta}_i$, similarly for $\widetilde{z_{it}' \boldsymbol{\theta}_t}$, $\boldsymbol{\lambda}_i = (0, \lambda_i)'$, and $\boldsymbol{\theta}_t = (0, \theta_t)'$.

The FE-OLS estimate of $\boldsymbol{\beta}$ will be

$$\hat{\boldsymbol{\beta}} = \left(\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{z}_{it} \tilde{z}_{it}' \right)^{-1} \left(\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{z}_{it} \tilde{y}_{it} \right). \tag{B6}$$

Expanding on $\tilde{y}_{it}$ and simplifying yields

$$\begin{aligned} \hat{\boldsymbol{\beta}} = \boldsymbol{\beta} &+ \boldsymbol{Q}_{zz,NT}^{-1} \left(\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{z}_{it} \widetilde{z_{it}' \boldsymbol{\lambda}_i} \right) \\ &+ \boldsymbol{Q}_{zz,NT}^{-1} \left(\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{z}_{it} \widetilde{z_{it}' \boldsymbol{\theta}_t} \right) + \boldsymbol{Q}_{zz,NT}^{-1} \left(\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{z}_{it} \tilde{\epsilon}_{it} \right), \end{aligned} \tag{B7}$$

where $\boldsymbol{Q}_{zz,NT}^{-1} = (\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{z}_{it} \tilde{z}_{it}')^{-1}$. If z_{it} is not independent to $\boldsymbol{\lambda}_i$ and $\boldsymbol{\theta}_t$, then the two bias terms generated by the slope heterogeneity will not vanish as the sample size increases, and the fixed effects estimator will be inconsistent.

There is every reason to believe that the KDD_{it} component of z_{it} is not independent of $\boldsymbol{\lambda}_i$ and $\boldsymbol{\theta}_t$. The theoretical model presented in Section 1 predicts a positive relationship between the two, which will result in the standard FE-OLS estimator overestimating (i.e., estimates closer to zero in this case) the average sensitivity to harsh temperatures β_2. This is intuitive as farmers who experience hotter temperatures in their region or time period have more incentive to adopt adaptation techniques. Indeed, Butler and Huybers (2013) provided evidence that the estimate of β_2 for an individual county is significantly positively correlated with its relative experience of heat. A key aim of our paper is to more precisely ascertain the nature of the relationship between KDD_{it} and β_{2it}, and to use this information to help predict the impact of climate change and how it may be mitigated by adaptation.

Appendix C: "Brute force" OLS estimation and its computational requirements

In principle, we could estimate the $\boldsymbol{\beta}_{it}$ parameters in equation (12) by what we call "brute force" OLS. That is, run OLS on a model that includes: (i) dummy variables for each i and t, and (ii) a complete set of interaction terms between the regressors and the i and t dummies (to capture unit/time fixed effects in slopes). Specifically, we have

$$y_{it} = \boldsymbol{\beta}' \boldsymbol{x}_{it} + \sum_{j=2}^{N} \left(\boldsymbol{x}'_{it} d_{ij}\right)' \boldsymbol{\gamma}_j + \sum_{j=2}^{T} \left(\boldsymbol{x}'_{it} d_{tj}\right)' \boldsymbol{\delta}_j + u_{it}, \tag{C1}$$

where $d_{ij} = 1$ if $i = j$ and 0 otherwise, and similarly for d_{tj}. Here, $\boldsymbol{\gamma}_j$ and $\boldsymbol{\delta}_j$ are $(K+1) \times 1$ vectors of coefficients to be estimated. After estimating $\boldsymbol{\beta}$, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_N)$, and $\boldsymbol{\delta} = (\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \ldots, \boldsymbol{\delta}_T)$ through OLS, it is possible to form $\hat{\boldsymbol{\beta}}_{it} = \hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\gamma}}_i + \hat{\boldsymbol{\delta}}_t$. These are consistent and asymptotically efficient estimates of the $\boldsymbol{\beta}_{it}$ given the usual OLS assumptions.[1]

To clarify the computational burden of this procedure, write (C1) in matrix form:

$$\underset{NT\times 1}{Y} = \underset{NT\times R}{\boldsymbol{Z}} \underset{R\times 1}{\boldsymbol{\Xi}} + \underset{NT\times 1}{U}, \tag{C2}$$

where $R = (K+1)(N+T-1)$ is the number of regressors, including interactions, while $\underset{R\times 1}{\boldsymbol{\Xi}} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta})'$, $\underset{NT\times R}{\boldsymbol{Z}} = (\boldsymbol{X}, \boldsymbol{X} \circ \boldsymbol{D}^{i=1}, \ldots, \boldsymbol{X} \circ \boldsymbol{D}^{i=N}, \boldsymbol{X} \circ \boldsymbol{D}^{t=1}, \ldots, \boldsymbol{X} \circ \boldsymbol{D}^{t=T})$, $\underset{NT\times(K+1)}{\boldsymbol{X}}$ is the stacked set of regressors, $\circ$ is the Hadarmard product, $\underset{NT\times(K+1)}{\boldsymbol{D}^{i=j}}$ is a set of stacked dummy variables for each regressor that is 1 when $i = j$ and 0 otherwise, and similarly $\underset{NT\times(K+1)}{\boldsymbol{D}^{t=j}}$ is a set of stacked dummy variables for each regressor that is 1 when $t = j$ and 0 otherwise. The OLS estimator is $\hat{\boldsymbol{\Xi}} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'Y$.

Note that $\boldsymbol{Z}'\boldsymbol{Z}$ in equation (C2) is a matrix of rank $R = (K+1)(N+T-1)$. To calculate the OLS estimator, a computer must be able to hold $(Z'Z)$ in memory, and then calculate

[1] See Hsiao (1975, Section 7) for more details. Note that identification of the separate components ($\boldsymbol{\beta}$, $\boldsymbol{\gamma}_i$, $\boldsymbol{\delta}_t$) is achieved here via the location normalization $\boldsymbol{\gamma}_1 = 0$, $\boldsymbol{\delta}_1 = 0$. Other normalizations are of course possible.

TABLE C1. Memory requirements to hold $\boldsymbol{Z}$.

N	T	K	$\boldsymbol{Z}$ in GB
500	200	5	3.36
1000	200	8	17.29
3000	300	15	380.28

$(\boldsymbol{Z}'\boldsymbol{Z})^{-1}$ or otherwise solve the linear system $(Z'Z)\hat{\boldsymbol{\Xi}} = Z'Y$. The memory and/or time requirements for these computations render a "brute force" OLS approach infeasible in large panels. For example, in our case we have $R = (4+1)(2209+65-1) = 11{,}365$ regressors, making the "brute force" OLS approach quite impractical.

Table C1 gives examples of space required to hold $\boldsymbol{Z}$ in memory (in gigabytes), while Figure C1 reports computation times for MO-OLS vs. "brute force" OLS in Stata. The results in Figure C1 are based on simulated data (Appendix D). We fix $K = 4$ and $T = 65$ as in our corn yield data, and vary N. Figure C1 shows how computation time increases rapidly with N for "brute force" OLS, while increasing slowly with N for MO-OLS. Beyond $N = 950$ it is no longer possible to run the regression in Stata due to memory limitations.[2]

Extrapolating from Figure C1, we estimate that, for a panel of the size used in our corn yield application ($N = 2209$), the "brute force" OLS approach, if it were feasible in terms of memory, would require 76 minutes. In contrast, MO-OLS requires roughly 90 seconds.

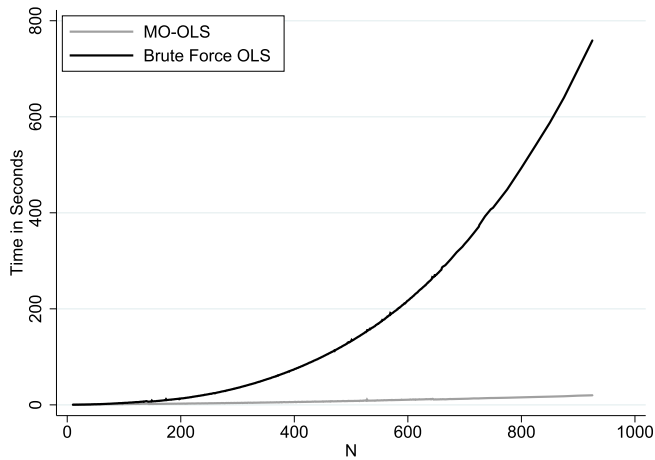

FIGURE C1. Computation time of MO-OLS and brute force OLS. *Note*: The data generating process for this example simulates (5) with $K = 4$ and $T = 65$.

[2]At $N = 950$ "brute force" OLS requires roughly 13 minutes while MO-OLS requires roughly 30 seconds.

Appendix D: Monte Carlo simulations

Here, we examine the finite sample performance of the MO-OLS estimator proposed in Section 2.3. We conduct Monte Carlo simulations based on environments characterized by multidimensional slope heterogeneity, with fixed effects in both intercepts and slopes. We study MO-OLS estimates of both the observation-level coefficients $\boldsymbol{\beta}_{it}$ and the mean coefficient vector $\bar{\boldsymbol{\beta}}$. For the mean, we also report results using three traditional panel data estimators: one-way and two-way fixed effects and mean group OLS ("MG-OLS").

We consider two alternative data generating processes. In each case, the dependent variable is generated by

$$y_{it} = c_{it} + \rho_{it} y_{it-1} + \beta_{it} x_{it} + \epsilon_{it}, \tag{D1}$$

where $i = 1, 2, \ldots, N$ and $t = -10, \ldots, 0, 1, \ldots, T$. We set $y_{i,-11} = 0$ and discard the first 10 observations for each unit i prior to estimation. The fixed effects in the intercept are generated by $c_{it} = 1 + f_i + f_t$ where $f_i \sim f_t \sim N(0, 0.353)$, and the idiosyncratic errors are generated as $\epsilon_{it} \sim N(0,1)$. In both scenarios, we generate the heterogeneous coefficient on the covariate x as $\beta_{it} = \beta + \lambda_i + \theta_t$ where $\beta = 1$ and $\lambda_i \sim N(0, 0.353)$, and we generate the heterogeneous autoregressive coefficient as $\rho_{it} = \rho + \delta_i + \phi_t$, where $\rho = 0.5$ and $\delta_i \sim N(0, 0.104)$. We will consider two scenarios with different specifications of the time effects θ_t and ϕ_t, as we discuss in detail below.

The regressor x is generated by

$$x_{it} = 0.5 x_{it-1} + \alpha_1 c_{it} + \alpha_2(\rho_{it} + \beta_{it}) + e_{it}. \tag{D2}$$

The parameter α_1 governs the correlation between the intercepts c_{it} and the regressor, while α_2 governs the correlation between the slope coefficients ρ_{it} and β_{it} and the regressor. In both scenarios, we set $\alpha_1 = \alpha_2 = 1$ and $e_{it} \sim N(0,1)$.

Finally, consider the time effects. These are of central importance, as a key advantage of MO-OLS is that it can accommodate time effects while MG-OLS in general cannot. We consider two scenarios: In the first, we set $\theta_t \sim N(0, 0.353)$ and $\phi_t \sim N(0, 0.104)$, so aggregate time effects are stochastic and correlated with the regressors (as $\alpha_2 \neq 0$). In the second scenario, the aggregate time effects are constant except for a jump half-way through the sample period:

$$\phi_t = \begin{cases} -0.104 & \text{if } t < T/2, \\ 0.104 & \text{if } t \geq T/2, \end{cases} \qquad \theta_t = \begin{cases} -0.353 & \text{if } t < T/2, \\ 0.353 & \text{if } t \geq T/2. \end{cases}$$

In our Monte Carlo analysis, we consider panels that are small relative to the county-level yield dataset we analyze in the main text. In fact, they are small enough that the 'brute force' OLS approach is feasible. But we do not report the OLS results as they are equivalent to the MO-OLS results. Recall that MO-OLS generates approximately identical results to "brute force" OLS if the number of terms L in the Cauchy sequence (20) is sufficiently large. We specify that the summation terminates when a tolerance of 1.0E-6 is achieved for the maximum change in $\hat{\boldsymbol{\beta}}_{it}$ from l to $l+1$, and we report for each experiment the mean and standard deviation of the L required for convergence.

TABLE D1. Simulation results—Scenario 1.

	Mean					Std. Dev.				
$(N = 50, T)$	30	50	70	100	200	30	50	70	100	200
				Results for $\bar{\boldsymbol{\rho}}$						
True values	0.500	0.500	0.500	0.500	0.501	0.025	0.022	0.020	0.018	0.016
Pooled OLS										
One-way FE	0.412	0.450	0.462	0.477	0.491	0.146	0.116	0.104	0.085	0.067
Two-way FE	0.597	0.629	0.639	0.653	0.664	0.098	0.085	0.078	0.068	0.063
MG-OLS	0.356	0.393	0.407	0.420	0.434	0.138	0.107	0.094	0.075	0.055
MO-OLS	0.474	0.485	0.490	0.493	0.497	0.030	0.024	0.022	0.019	0.017
				Results for $\bar{\boldsymbol{\beta}}$						
True values	0.998	1.003	1.000	0.998	1.002	0.082	0.071	0.066	0.059	0.056
Pooled OLS										
One-way FE	2.381	2.370	2.343	2.325	2.327	0.465	0.361	0.304	0.250	0.194
Two-way FE	0.949	0.932	0.920	0.906	0.900	0.109	0.095	0.083	0.073	0.066
MG-OLS	2.541	2.492	2.452	2.421	2.410	0.493	0.372	0.312	0.255	0.191
MO-OLS	1.010	1.010	1.007	1.002	1.004	0.087	0.075	0.069	0.061	0.057
Iterations (L)	356	335	325	320	313	82	62	56	47	39

Note: 1000 Monte Carlo Simulations with $N = 50$ and varied T.

For each scenario, we report results from 1000 Monte Carlo replications. In each case, the average β_{it} over all Monte Carlo repetitions ($\bar{\beta}$) is approximately 1.0, while the average ρ_{it} ($\bar{\rho}$) is approximately 0.5 (with slight deviations due to sampling variation).

Consider first the results for the first scenario, the DGP with stochastic time effects in slopes. Table D1 reports results for the estimates of the mean slope parameters $\bar{\beta}$ and $\bar{\rho}$. We set the number of cross-sectional units N to 50 in all experiments. We vary T from 30 to 200, as it is primarily T that affects performance due to the $O(T^{-1})$ Hurwicz bias. The left panel reports the mean of the estimates (across the Monte Carlo replications) obtained using all four estimators under consideration, while the right panel lists the empirical standard deviations of the estimates.[3]

The results in Table D1 show that ignoring the time heterogeneity in the slope coefficients can have dramatic implications for statistical inference. Both the one-way and two-way FE estimators and MG-OLS exhibit serious biases for the mean slope and autoregressive coefficients. These biases are not eliminated by increasing T from 30 to 200. In contrast, the bias in the MO-OLS estimates of the mean slope parameter $\bar{\beta}$ is negligible in all cases. The MO-OLS estimates of the mean autoregressive parameter $\bar{\rho}$ are slightly downward biased when $T = 30$, but the bias becomes negligible as T increases (removing the Hurwicz bias). The empirical standard deviations in the right side of the panel show that the MO-OLS estimates of the mean coefficients are also much more efficient.

[3]The true mean slope coefficients vary between simulated samples (because they depend on random variables). So the right panel also reports the empirical standard deviation of the true mean coefficients.

TABLE D2. Simulation results for $\hat{\boldsymbol{\theta}}_{it}$—Scenario 1.

(N, T)	30	50	100	200	350
		Results for $\hat{\boldsymbol{\rho}}_{it}$			
Mean Bias	−0.025	−0.014	−0.007	−0.004	−0.002
Abs. Bias					
Mean	0.038	0.025	0.015	0.010	0.007
Std. Dev.	0.027	0.018	0.011	0.007	0.005
Corr. with $\boldsymbol{\rho}_{it}$	0.970	0.985	0.994	0.997	0.998
Corr. with y_{it-1}					
True	0.129	0.133	0.132	0.135	0.135
Estimated	0.135	0.138	0.135	0.136	0.135
		Results for $\hat{\boldsymbol{\beta}}_{it}$			
Mean bias	0.013	0.007	0.004	0.002	0.001
Abs. bias					
Mean	0.188	0.136	0.091	0.063	0.047
Std. Dev.	0.142	0.103	0.069	0.048	0.036
Corr. with $\boldsymbol{\beta}_{it}$	0.903	0.945	0.975	0.988	0.991
Corr. with x_{it}					
True	0.446	0.446	0.452	0.453	0.453
Estimated	0.403	0.423	0.441	0.448	0.450

Note: 1000 Monte Carlo Simulations with varied N and T.

Next, we evaluate the MO-OLS estimates of the observation level slope coefficients β_{it} and ρ_{it}, estimated using (20). Here, we let (N, T) grow together, as this is required for consistency of the $\hat{\boldsymbol{\beta}}_{it}$. We report the results in Table D2, which shows the mean bias, the mean and standard deviation of the absolute bias, the correlation of the estimates with the true values, and the true and estimated correlation of the estimates with the covariates.

The bias of each (i, t)-level estimate is defined as $B_{it} = S^{-1}\sum_{s=1}^{S}(\hat{\beta}_{its} - \beta_{its})$ where $S = 1000$, and similarly for ρ_{it}. The mean bias averages this over all observations (i, t). As we see in Table D2, the mean bias is small even when $(N, T) = 30$, and quickly vanishes as (N, T) increase. Of course, this is just another way of expressing our previous result from Table D1 that the bias in estimates of the mean coefficients $\bar{\beta}$ and $\bar{\rho}$ is small. When $(N, T) = 30$ the mean absolute bias is 0.038 (or 7.6%) for the ρ_{it} estimates, and 0.188 (or 18.8%) for the β_{it}. Both the mean absolute bias and its standard deviation decline rapidly as (N, T) increase.

Impressively, we see that $\text{corr}(\hat{\rho}_{it}, \rho_{it}) = 0.97$ even when $(N, T) = 30$. And we have that $\text{corr}(\hat{\beta}_{it}, \beta_{it}) = 0.90$ when $(N, T) = 30$, and this increases to 0.975 when $(N, T) = 100$. Finally, the estimates capture the correlations between the slope coefficients and the covariates (i.e., $\text{corr}(\beta_{it}, x_{it})$ and $\text{corr}(\rho_{it}, y_{i,t-1})$) very accurately, even when $(N, T) = 30$.

Now we turn to the results for the second scenario, the DGP where the aggregate time effects that enter the slope coefficients are constant except for a jump half-way through the sample. These results are presented in Tables D3 and D4. The one-way and two-way

TABLE D3. Simulation results—Scenario 2.

	Mean					Std. Dev.				
$(N = 50, T)$	30	50	70	100	200	30	50	70	100	200
				Results for $\bar{\boldsymbol{\rho}}$						
True Values	0.499	0.500	0.501	0.500	0.499	0.016	0.016	0.015	0.015	0.015
Pooled OLS										
One-way FE	0.950	0.956	0.961	0.963	0.967	0.017	0.012	0.012	0.011	0.009
Two-way FE	0.991	0.983	0.982	0.981	0.981	0.020	0.015	0.014	0.013	0.011
MG-OLS	0.920	0.930	0.936	0.941	0.947	0.016	0.010	0.008	0.006	0.004
MO-OLS	0.461	0.478	0.484	0.488	0.493	0.023	0.020	0.018	0.018	0.016
				Results for $\bar{\boldsymbol{\beta}}$						
True values	0.999	0.999	0.998	0.998	1.003	0.051	0.051	0.051	0.049	0.049
Pooled OLS										
One-way FE	1.310	1.098	0.992	0.919	0.829	0.213	0.150	0.128	0.107	0.080
Two-way FE	0.784	0.755	0.740	0.727	0.718	0.069	0.055	0.050	0.042	0.039
MG-OLS	1.571	1.309	1.186	1.094	0.978	0.221	0.156	0.117	0.089	0.058
MO-OLS	1.010	1.008	1.006	1.003	1.006	0.059	0.054	0.054	0.050	0.050
Iterations (L)	820	839	881	905	932	232	214	222	210	214

Note: 1000 Monte Carlo Simulations with $N = 50$ and varied T.

TABLE D4. Simulation results for $\hat{\boldsymbol{\theta}}_{it}$—Scenario 2.

(N, T)	30	50	100	200	350
		Results for $\hat{\boldsymbol{\rho}}_{it}$			
Mean bias	−0.039	−0.022	−0.011	−0.006	−0.003
Abs. bias					
Mean	0.044	0.026	0.014	0.008	0.006
Std. Dev.	0.028	0.019	0.011	0.007	0.005
Corr. with $\boldsymbol{\rho}_{it}$	0.980	0.991	0.996	0.998	0.999
Corr. with y_{it-1}					
True	0.420	0.428	0.430	0.432	0.434
Estimated	0.421	0.430	0.431	0.432	0.434
		Results for $\hat{\boldsymbol{\beta}}_{it}$			
Mean bias	0.013	0.009	0.006	0.003	0.002
Abs. bias					
Mean	0.194	0.139	0.094	0.065	0.048
Std. Dev.	0.147	0.105	0.071	0.049	0.037
Corr. with $\boldsymbol{\beta}_{it}$	0.899	0.944	0.973	0.987	0.993
Corr. with x_{it}					
True	0.606	0.616	0.619	0.624	0.625
Estimated	0.542	0.580	0.602	0.616	0.620

Note: 1000 Monte Carlo Simulations with varied N and T.

fixed effects and MG-OLS estimators all exhibit severe biases in this case. The estimates of $\bar{\rho}$ are in all cases biased upward very severely, implying the process is close to a unit root when in fact $\bar{\rho}$ is approximately one-half. All three estimators also generate severely biased estimates of $\bar{\beta}$. In contrast, the MO-OLS estimates of $\bar{\rho}$ are just slightly downward biased when $T = 30$, as we would expect given the Hurwicz bias, but this bias vanishes quickly as T increases. And there is no evidence of bias in the MO-OLS estimates of $\bar{\beta}$, even when $T = 30$. Turning to an evaluation of the observation (i, t)-level estimates in Table D4, we see the results are very similar to those in Table D2. For example, when $(N, T) = 50$ we have that $\text{corr}(\hat{\rho}_{it}, \rho_{it}) = 0.99$ and $\text{corr}(\hat{\beta}_{it}, \beta_{it}) = 0.94$.

In summary, the Monte Carlo results show that, in panel data with both spatial and time fixed effects in both intercept and slopes, the MO-OLS estimator does a good job of uncovering both the mean and observation-level coefficients, even in relatively small samples. It achieves this even when the slope heterogeneity is correlated with the regressors, and whether time effects are randomly or deterministically generated. We also find that conventional fixed effects and MG-OLS estimators are severely biased and inconsistent in the same contexts.

Appendix E: Understanding identification in MO-OLS relative to FE-OLS

In panel data models with unit and time fixed effects in the intercept (i.e., two-way fixed effects), it is simple to demonstrate that the slope parameters are identified from idiosyncratic variation in the regressors associated with particular $\{i, t\}$ pairs. In our application, this implies that the model identifies a relationship between yield and temperature from idiosyncratic variation in local weather—variation that is distinct from county-specific climate variation or longer term temperature trends.[4] This leads to the common interpretation that two-way fixed effects identifies "short-run" relationships.

Here, we demonstrate that this result does not carry over to panel data models containing two dimensions of fixed effects in both the intercept and slope parameters. Instead, we show that MO-OLS is capable of identifying slope heterogeneity that is driven by long-term climatic differences between counties, or from time-specific shocks that are common across all counties (e.g., a heatwave). This point is important, as the results in our paper rely on the fact that we are identifying long-term adaptations by the farmer that are driven by spatial variation in climate and/or temperature trends over time, and not simply responses to idiosyncratic local weather shocks.

Consider the following panel data model:

$$y_{it} = c_i + \tau_t + \beta x_{it} + e_{it}, \tag{E1}$$

where c_i and τ_t are fixed effects in the intercept. This model can be estimated (in principle) by running an OLS regression that includes dummies for each unit i and time period t. Simple algebra shows this is equivalent to estimating the following regression by OLS:

$$\tilde{y}_{it} = \beta \tilde{x}_{it} + \tilde{e}_{it}, \tag{E2}$$

[4] For instance, if a hot county experiences hotter or cooler temperature than usual in a given year that is also above or below the temperature shock experienced by the entire country in that given year.

where $\tilde{y}_{it} = y_{it} - N^{-1}\sum_{i=1}^{N} y_{it} - T^{-1}\sum_{t=1}^{T} y_{it} + NT^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}$ is the two-way within transformation, and likewise for $\tilde{x}_{it}$ and $\tilde{e}_{it}$. It is this equivalence that leads to the result that β in the two-way fixed effects model is identified from shocks to x_{it} that are distinct from county-specific or time-specific variation in the regressors.

In contrast, consider a model with multidimensional slope heterogeneity:

$$y_{it} = c_i + \tau_t + \beta_{it} x_{it} + e_{it}, \tag{E3}$$

where $\beta_{it} = \beta + \lambda_i + \theta_t$. In this case, it can be shown that $\tilde{y}_{it}$ becomes

$$\begin{aligned}\tilde{y}_{it} = \beta\tilde{x}_{it} + \tilde{e}_{it} + \lambda_i x_{it} + \theta_t x_{it} - N^{-1}\left(\sum_{i=1}^{N}\lambda_i x_{it} + \theta_t\sum_{i=1}^{N} x_{it}\right) \\ - T^{-1}\left(\sum_{t=1}^{T}\theta_t x_{it} + \lambda_i\sum_{t=1}^{T} x_{it}\right) + NT^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_i x_{it} + \sum_{i=1}^{N}\sum_{t=1}^{T}\theta_t x_{it}\right).\end{aligned}$$

Correlations between the slope heterogeneity and the regressors prevent this from being simplified to a regression on transformed data $\tilde{y}_{it} = \beta_{it}\tilde{x}_{it} + \tilde{e}_{it}$ where the $\tilde{e}_{it}$ satisfy OLS assumptions. Thus, the β_{it} are not identified from idiosyncratic variation in the regressors.

In fact, as Figures 2 and 4 in the main text clearly show, in our application the MO-OLS estimates capture variation in the KDD coefficients across counties that is systematically related to their long-run differences in KDD levels, as well as trends in the mean KDD coefficient over time. To gain intuition for how coefficients are identified in MO-OLS, it is useful to consider the case where x_{it} follows a permanent/transitory structure, as in

$$x_{it} = \delta_i + \eta_t + \xi_{it}, \tag{E4}$$

where δ_i, η_t, and ξ_{it} are mutually independent, and ξ_{it} is independent over both i and t. To simplify exposition, consider the case where $E(c_i) = E(\tau_t) = \beta = E(\lambda_i) = E(\theta_t) = E(\delta_i) = E(\eta_t) = E(\xi_{it}) = 0$. Then the conditional expectations of y_{it} become

$$\begin{aligned}E[y_{it}|c_i, \delta_i] \equiv E(\bar{y}_i) = c_i + \lambda_i\delta_i, \\ E[y_{it}|\tau_t, \eta_t] \equiv E(\bar{y}_t) = \tau_t + \theta_t\eta_t,\end{aligned}$$

where $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$ and $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$. Thus, the unit-specific slope heterogeneity parameter λ_i is the effect of the unit-specific component of the regressor δ_i on $E(\bar{y}_i)$, while the time-specific slope heterogeneity parameter θ_t is the effect of the time-specific shock to the regressor η_t on $E(\bar{y}_t)$. In other words, the unit-specific slope parameter λ_i is identified from the cross-unit variation in the regressor (the δ_i), while the time-specific parameter θ_i is identified from the common over-time variation in the regressor (the η_t).

In our application, the assumption that x_{it} (i.e., weather variables) follow a permanent transitory structure is not unrealistic. Table E1 presents the results of a series of

TABLE E1. Analyzing the dependence of KDD_{it} and GDD_{it}.

	KDD_{it}			GDD_{it}		
Regression Results	(1)	(2)	(3)	(1)	(2)	(3)
Dep. var.: KDD_{it} *or* GDD_{it}						
Lagged dep. var.			0.106			0.164
			(0.021)			(0.027)
R^2	0.748	0.851	0.852	0.944	0.980	0.981
County dummies	Yes	Yes	Yes	Yes	Yes	Yes
Time dummies	No	Yes	Yes	No	Yes	Yes
Dep. var.: Time dummies						
Lagged time dummy		0.183	0.089		0.067	−0.045
		(0.126)	(0.126)		(0.127)	(0.129)
Time trend		−0.138	−0.096		0.404	0.518
		(0.119)	(0.120)		(0.629)	(0.642)
R^2		0.063	0.022		0.012	0.012

Note: This table presents OLS parameter estimates for regressions of KDD_{it} and GDD_{it} on their lag and a series of county and time dummies, as well as regressions of those estimated time dummies on their lag and a linear time trend. Standard errors are in parentheses, and are clustered at the state level in the panel data regressions.

regressions of KDD_{it} and GDD_{it} on county and time effects and lagged dependent variables. The table presents regression results from three different models: model (1) includes only county dummies, model (2) includes county and time dummies, and finally model (3) includes both sets of dummy variables and the lagged temperature variable.

The results show that the vast majority of variation in the temperature variables can be explained by unit-specific and time-specific components. The R^2 of the regressions with county and time dummies is 0.85 for KDD_{it} and 0.98 for GDD_{it}. Moreover, the results for model (3) reveal that there is very little serial dependence in local weather shocks (i.e., ξ_{it} in (E4)). The estimated AR coefficients are 0.11 for KDD_{it} and 0.16 for GDD_{it}, and their inclusion into the model adds almost nothing to the R^2.

We also find little evidence for serial dependence in the time-specific aggregate temperature shocks. The bottom half of Table E1 presents regressions of the time dummies estimated in models (2) and (3) on their lag and a linear time trend. The estimated coefficients from this exercise are statistically insignificant. These results support the decomposition of our temperature variables into a permanent transitory structure, and also suggest that farmers would find it very difficult to project either local weather shocks or time-specific weather shocks when planning their sowing or adaptation behavior.

Given the permanent/transitory structure is reasonably accurate, we expect that in our application the county-specific component of the KDD coefficient will pick up the extent to which sensitivity to high temperature is lower in counties with higher mean KDD levels, while the time-specific component will pick up how sensitivity to high temperature shocks (common across counties) has varied over time. This is exactly what we see in Figures 2 and 4 in the main text.

Appendix F: Projections of GDD_{it} and KDD_{it}

This appendix describes the climate models' projections of annual KDD and GDD in the corn-growing regions of the U.S., as well as their variability across climate models and RCP scenarios. Figure F1 presents the average (over counties) projection of KDD across the ensemble of models for the three RCPs, where the solid line is the average projection and the shaded areas are the 80% (1.28 standard deviation) prediction intervals.

Over the historical period (1950–2015) mean KDD has decreased slightly, due to mild recent weather. The mean annual KDD level across all years/counties in the historical data is 41. In the optimistic RCP26 scenario (green) the ensemble average prediction is that KDD will increase slowly until 2050, when its projected mean (across models) reaches 69. It then plateaus, and is 70 in 2100, or 75% above the historical mean.

Under the RCP45 scenario (orange), mean KDD continues to grow until about 2080, and only then does it plateau. The projected mean of KDD (across models) reaches 82 in 2050, which is double the historical mean, and increases to 115 in 2100, which is almost triple the historical mean. But there is great disagreement between climate models, with some predicting more modest changes.

Finally, in the "business as usual" RCP85 scenario (purple) mean KDD increases to 107 by 2050 and an extraordinary 306 in 2100 (almost eight times the historical mean). Despite wide prediction intervals, every model predicts sharp increases relative to historical experience.

Figure F2 presents the distribution of KDD across counties. The recent historical period of 2010–2015 is plotted, as well as projections for the distribution of KDD under the

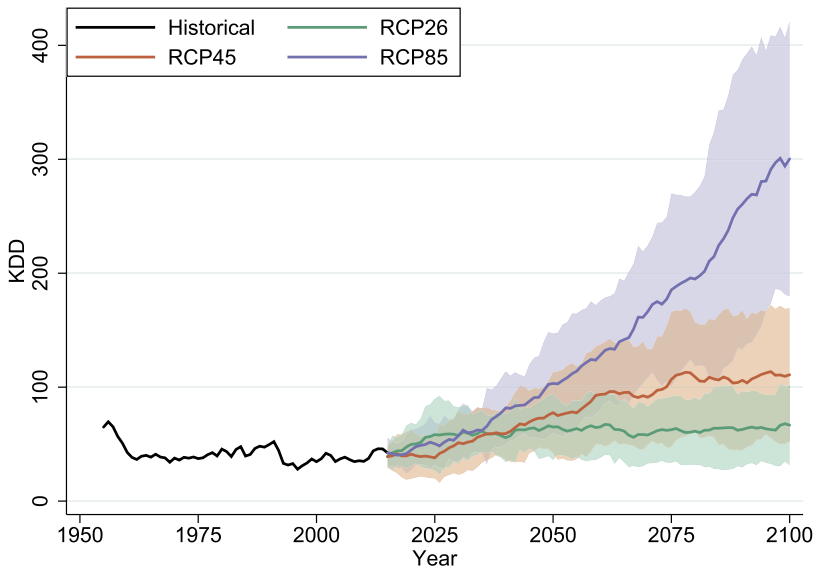

Figure F1. Projected KDD_{it} by representative concentration pathway. *Note*: This graph presents projections of KDD (averaged over counties and weighted by corn acreage) across three RCPs, where the solid line is the average projection across 19 CMIP5 climate models, and the shaded areas are the 80 percent prediction intervals.

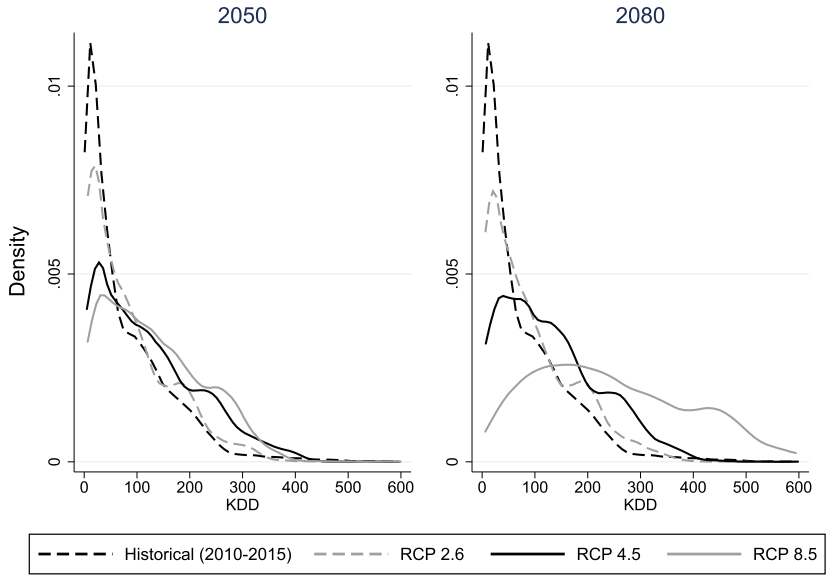

FIGURE F2. Density of KDD_{it} across counties by scenario. *Note*: This graph presents the density of KDD across counties over the 2010–2015 historical period, and compares it with three projections by RCP scenario and averaged across the 19 CMIP5 climate models.

three RCP scenarios for both 2050 (left) and 2080 (right). As we already noted, the mean of KDD in the historical data is 41. Here, we see that most of the mass of the county/year observations is between 0 and 100, but there is considerable mass in the 100 to 250 range, and nonnegligible mass up to about 400. Only 33 historical observations exceed 400.

For 2050, both the RCP 4.5 and RCP 8.5 scenarios have much less mass in the distribution below a KDD of 50, relative to the historical period. Conversely, there is a sharp increase in the incidence of KDD between 50 and 400, after which the density approaches zero. For example, in the historical data only 2% of county/year observations have KDD over 200. But the average frequencies (across models) of KDD $>$ 200 in 2050 are 10%, 21%, and 23% under the RCP 2.6, 4.5, and 8.5 scenarios, respectively. However, while the mass of the distribution shifts substantially to the right, the vast majority of observations remain within the support of the historical distribution (i.e., below 400). This gives us some confidence in using our models to forecast out to 2050 under all three RCP scenarios.

The densities at 2080 reveal a much larger range of outcomes across the RCP scenarios. The distributions of outcomes under RCP 2.6 is almost identical in 2080 and 2050, because under this scenario warming ceases after 2050. But under RCP 4.5, the distribution of KDD flattens considerably between 2050 and 2080, with much more mass in the 100 to 400 range. Nevertheless, the density still goes to zero quickly above 400. Thus, the majority of observations remain within the range of the historical data.

Under RCP 8.5, however, the change in the distribution is far more dramatic. The mode increases to about 150, and there is now considerable mass above 400.[5] Thus, forecasting climate change impacts as far out as 2080 under the RCP 8.5 scenario requires extrapolation beyond the historical data for many of the county/year observations.

Recall that the log-linear heat sensitivity curve that we estimate, reported in Figure 4, implies substantial scope for adaptation as KDD increases from 0 to 100, but the curve becomes much flatter at higher levels of KDD. It is very flat indeed by the time KDD reaches 400, and, as we have very little historical data above that level, our projections assume KDD sensitivity is unchanging at higher KDD levels.

Figure F3 presents projections for GDD. Mean annual GDD is about 3450 in the historical period. It is predicted to plateau at about 3600 in the RCP26 scenario, or more in the RCP45 and RCP85 scenarios. Thus, the climate models predict that KDD and GDD will *both* increase in the absence of emissions reductions. As a result, the increase of KDD in cooler corn-growing regions is mitigated by the increase in GDD. The effect of climate change on yields that we project in these regions will be the net balance between the negative effect of higher KDD and the positive effect of higher GDD.

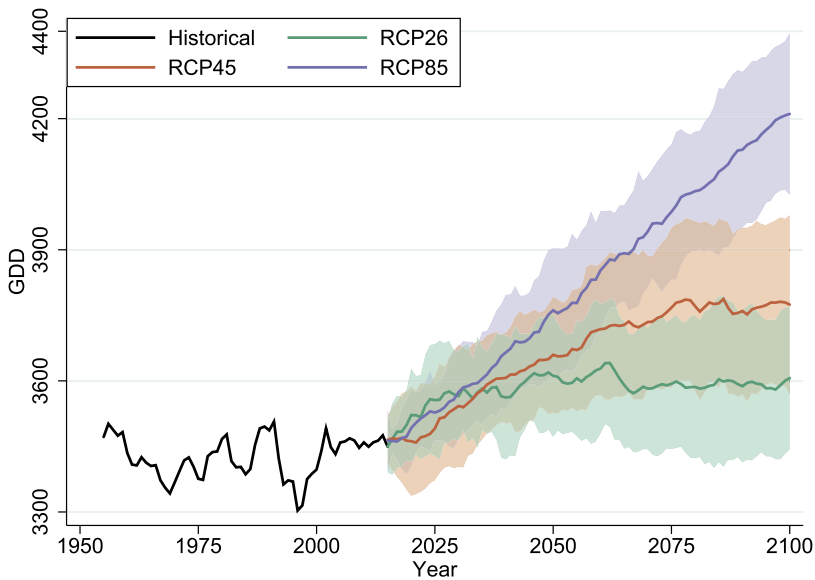

FIGURE F3. Projected GDD_{it} by representative concentration pathway. *Note*: This graph presents projections of GDD (averaged over counties and weighted by corn acreage) across three RCPs, where the solid line is the average projection across 19 CMIP5 climate models, and the shaded areas are the 80% prediction intervals.

[5]Under the RCP 2.6 and 4.5 scenarios, the unweighted mean levels of KDD in 2080 are 91 and 129, and the average frequencies (across models) of KDD $>$ 200 in 2080 are 12% and 22%, respectively. But under the RCP 8.5 scenario, the unweighted mean level of KDD in 2080 is 246, and the average frequency (across models) of KDD $>$ 200 is a staggering 57%.

Appendix G: Projecting corn yield with technical change

Here, we report projections of corn yield that incorporate projected technical progress. In the main text, we reported predicted changes in yield that account for adaptation to high temperatures, but that hold other aspects of technology fixed. We have a firm historical basis for predicting adaptation, based on comparing historical production function parameters for hot versus cold counties and time periods. But projecting more general forms of technical changed into the future is a more speculative exercise. (Indeed, the published research that projects effects of future climate change on yields has generally avoided projecting technical progress, instead making yield projections holding technology fixed).

The econometric models in the main text do provide estimates of technological progress over the 1950–2015 period. In the FE-OLS models, neutral technical progress is captured by the time fixed effect, while in the MO-OLS model technology is more complex, as it may alter the time fixed effect or the time effects in the slope coefficients for the temperature and precipitation inputs. Accordingly, we can attempt to extrapolate trends in these time effects into the future to project future technical progress.

Of course, our sample period of 1950 to 2015 saw dramatic improvements in agricultural technology from the spread of machines including tractors, cutters, harvesters, planters and trucks, commercial fertilizer, insect and weed-resistant hybrid seeds, computers and satellite technology, and genetic modification. It may be too optimistic to assume that advances in agricultural technology can continue at such an impressive pace through 2100, but that is in effect what we do here by extrapolating these historical trends.

Here, we focus on the MO-OLS model, and project both the fixed effect and the coefficient on GDD_{it} into the future using a VAR(1) system of two equations. We assume that time effects in the KDD_{it} coefficient are already captured by our adaptation process, and we ignore time effects in the PREC_{it} coefficients because we find no significant time trends on these coefficients in the historical data.

An important issue in specifying the VAR is nature of the time trends. As equation (8) is for log yield, we rule out using trends in t or t^2 as this permits exponential yield growth in levels. Accordingly, we consider two specifications for the trend terms: The first is to include just $\log(t)$, which we call the "pessimistic" scenario of technical progress. The second is to include both $\log(t)$ and $\sqrt{\log(t)}$, which we call the "optimistic" scenario. This is because a negative coefficient on $\sqrt{\log(t)}$ allows diminishing returns to technology to set in more slowly than when using $\log(t)$ alone. It is our intent to present both these scenarios of future technology, without commenting on their likelihood. To conserve on space, we do not present the estimates of the VAR models.[6] We note, however, that the 'optimistic' model had the best in-sample fit of several models we examined. Of course, this does not necessarily make it preferable for out-of-sample projections.

Table G1 presents the results of our projection exercise broken down by technology scenario, adaptation scenario, and RCP scenario. Results are reported in bushels per

[6] Note that we do not include any structural breaks in this VAR model, in contrast to our analysis of the KDD_{it} coefficient in Table 3, as we found no consistent evidence for their existence.

TABLE G1. The effects of climate change on corn yield (MO-OLS).

	Pessimistic Tech Growth		Optimistic Tech Growth	
Year	No Adapt.	Adapt.	No Adapt.	Adapt.
RCP85				
2030	159 (142, 177)	158 (145, 171)	164 (146, 181)	162 (149, 175)
2050	148 (114, 182)	162 (142, 181)	162 (128, 196)	178 (159, 197)
2080	110 (62, 158)	160 (134, 185)	139 (84, 195)	206 (179, 232)
2100	73 (25, 121)	149 (121, 176)	103 (42, 164)	217 (189, 245)
RCP45				
2030	163 (142, 185)	159 (146, 173)	168 (147, 189)	164 (150, 178)
2050	165 (136, 195)	170 (152, 189)	179 (149, 209)	185 (165, 204)
2080	165 (125, 206)	184 (162, 206)	196 (156, 235)	220 (201, 239)
2100	180 (128, 233)	201 (172, 229)	222 (169, 276)	251 (224, 278)
RCP26				
2030	158 (140, 176)	156 (145, 168)	163 (145, 181)	161 (150, 172)
2050	176 (154, 197)	176 (164, 188)	188 (168, 208)	189 (179, 200)
2080	207 (181, 233)	206 (189, 222)	233 (208, 258)	233 (215, 251)
2100	221 (183, 259)	222 (198, 247)	259 (221, 298)	262 (235, 289)

Note: Results are expressed in terms of actual crop yield in bushels per acre. Each number represents a model average over nineteen climate models, while the numbers in brackets represents the 80% prediction interval of that ensemble average.

acre. We present mean projections across 19 climate models, as well as an 80% prediction interval, at four points in time. Figures G1 to G2 present some key results visually.

Turning to our projections, consider first the worst case scenario of pessimistic technology growth and no adaptation. As we see in Table G1, in the RCP85 scenario the mean projection is a catastrophic drop in corn yield from roughly 160 bushels per acre today to 73 in 2100. With adaptation yield stagnates and is slightly lower in 2100 than today.

It is important to understand that a stagnation of yield would itself be a catastrophic outcome. As we see in Figure G1 historical corn yield tripled from roughly 50 in the 1950s to roughly 150 in 2015, so rapid growth was the norm for the past 65 years. Furthermore, Figure G1 plots historical world population along with the median projection from the United Nations Population Division 2017 Revision. Note that U.S. corn yields closely track world population growth from 1950 to 2015, as population also roughly tripled from 2.5 to 7.3 billion. The U.S. provides more than a third of the world's corn exports, so a stagnation of U.S. yields would have a devastating effect on world supply (particularly as it will coincide with yield dropping globally from the effects of climate change).

Next consider the RCP26 scenario of ambitious emissions reductions, while continuing to maintain the pessimistic technology scenario. As we see in Table G1, the mean projected yield in 2100 is 221 or 222, depending on whether we include adaptation.[7] But as we see in Figure G1, yield growth in this scenario is not nearly as rapid as population growth.

[7] Adaptation makes little difference for the mean projection in this case, because temperature increases are moderate. But it does reduce variance.

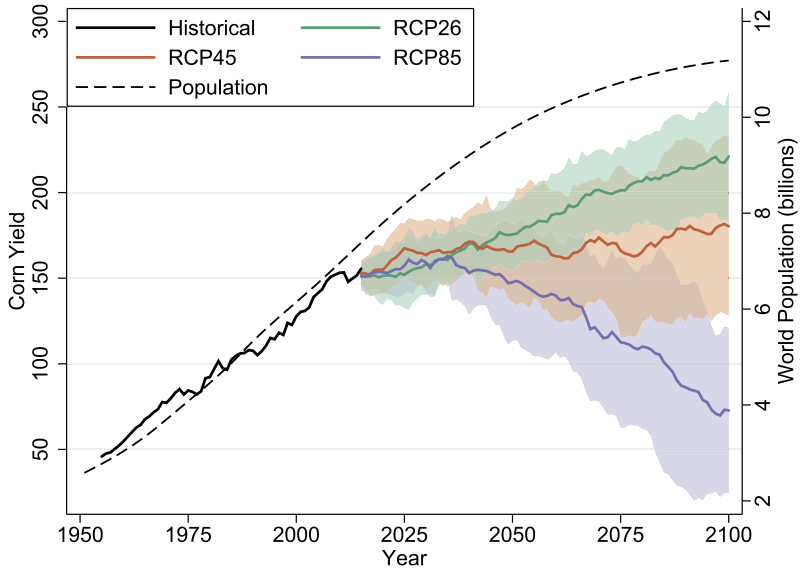

FIGURE G1. Effect of climate change on corn yield with pessimistic tech growth and no adaptation. *Note*: This graph presents projections of corn yield by RCP emission scenario using the MO-OLS model where future adaptation is not modelled and technology is projected using a VAR(1) with a $\log(t)$ trend term. The solid lines are the average projection across nineteen CMIP5 climate models, and the shaded areas are the 80% prediction intervals.

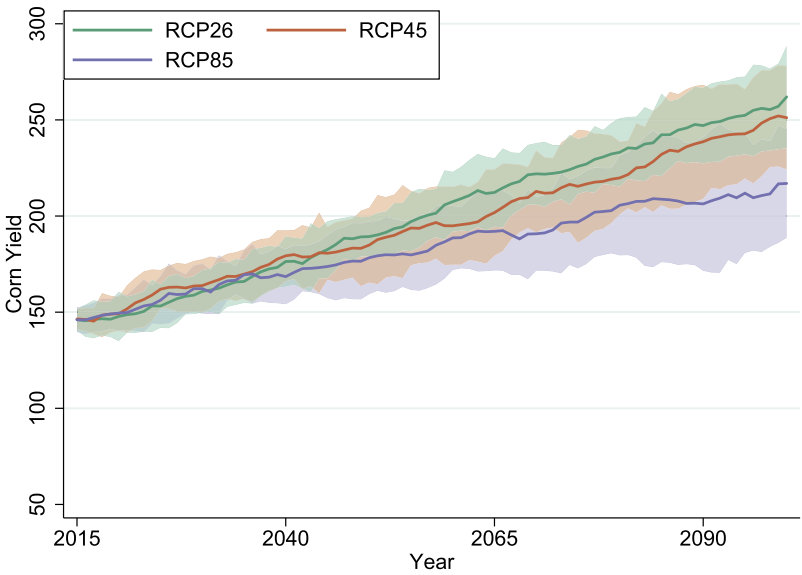

FIGURE G2. The effect of climate change on corn yield with optimistic tech growth and adaptation. *Note*: This graph presents projections of corn yield by RCP emission scenario using the MO-OLS model where future adaptation is not modelled and technology is projected using a VAR(1) with $\log(t)$ and $\sqrt{\log(t)}$ as trend terms. The solid lines are the average projection across nineteen CMIP5 climate models, and the shaded areas are the 80% prediction intervals.

If we turn to the "optimistic" technology assumption, we see in Table G2 that yield projections increase rather substantially under all scenarios. In fact, the mean prediction under RCP85 of a 217 yield in 2100 is similar to the prediction under pessimistic technology and RCP26. Thus, being "optimistic" about technology while assuming "business as usual" emissions leads to similar predictions as being "pessimistic" about technology while assuming "ambitious" emissions reductions. Importantly, however, yield in either case falls well short of keeping pace with population growth.

Finally, consider the best case scenario of optimistic technology growth and ambitious emissions reductions (RCP26). As we see in Figure G2, the mean projected yield in 2100 is 259 or 262, depending on whether we include adaptation. This is roughly what is necessary for yields to increase at a rate that is commensurate with population growth. Even here however, the 80% prediction interval extends down to 235, which is well short of keeping pace with population growth. Thus, even in a best case scenario, climate change creates an environment of considerable *risk* with respect to U.S. agricultural yields.

Appendix H: Model and projection results for soybeans

The case of soybeans offers an interesting comparison to corn. This is the second largest U.S. crop, and both its geographic distribution and growing characteristics are different. Here, we apply the same econometric and projection methods as in the main text, except with soybean yield as the dependent variable.

Table H1 presents the FE-OLS results, in the same format as Table 1 for corn. While the KDD coefficients are again negative and highly significant, we find that soybeans are less sensitive to high temperatures than corn. The model in column (3) includes the nonlinear term to capture adaptation. The coefficient on this nonlinear term is positive and statistically significant. This indicates, as it did with corn, that there is heterogeneity in the marginal effect of KDD_{it} and it has a positive correlation with the level of KDD_{it}.

Table H2 presents the MO-OLS estimates. The mean KDD coefficeint is -0.0055, which is about 25% larger in absolute magnitude that the conventional FE-OLS estimate in Table H1 column (2). So we again find that FE-OLS understates the effect of high temperature, due to the downward bias in the presence of parameter heterogeneity (see Appendix B). But the bias is much smaller than we found for corn. This is because the standard deviation of the KDD coefficient (0.0037) is smaller here. Figure H1 illustrates this, as the 90/10 percentile range of the KDD coefficients cover a narrower range than they did for corn. This is evidence for comparatively little historical adaptation between counties. This finding is intuitive given the greater success in developing corn hybrids that are more resistant to heat.

The time trend of aggregate adaptation in Figure H1 is somewhat difficult to interpret due to the significant degree of random fluctuation in the median KDD coefficient between years. Fitting a linear trend to the median coefficient over time resulted in an extremely small positive trend, and an unreported structural break test could not reject the null hypothesis of no structural trend break.

TABLE H1. FE-OLS estimates of the impacts of temperature on soybean yields.

Specification	(1)	(2)	(3)
GDD	0.0005	0.0005	0.0006
	(0.0001)	(0.0000)	(0.0001)
KDD	−0.0046	−0.0044	−0.0106
	(0.0003)	(0.0002)	(0.0019)
ln(KDD) ∗ KDD − KDD			0.0012
			(0.0004)
Precipitation	0.0013	0.0013	0.0013
	(0.0002)	(0.0002)	(0.0002)
Precipitation^2 (×1000)	−0.0009	−0.0008	−0.0009
	(0.0001)	(0.0001)	(0.0001)
Constant	0.7754	0.9413	0.5169
	(0.2925)	(0.2265)	(0.3227)
Fixed Effects	Cty, Yr	Cty, State-Yr	Cty, Yr
Obs.	88,101	88,101	88,101
R^2	0.77	0.73	0.77

Note: Results exclude counties west of the 100th Meridian. The sample period is 1950–2015, and $N = 1684$. Models (1)–(3) differ by type of fixed effects and whether the adaptation variable is included. Standard errors are reported in parentheses, and are clustered at the state level.

Figure H2 plots the average values of $\hat{\beta}_{2it}$ and KDD_{it} for each county. The correlation between $\hat{\beta}_{2it}$ and KDD_{it} is 0.25, which is weaker than for corn, yet still positive. We fit a log-linear curve to these points to obtain the function we use to model adaptation.

TABLE H2. Mean-observation estimates of the impacts of temperature on U.S. soy yields.

	Mean	Weighted Mean	Median	Standard Deviation	10th Percentile	90th Percentile
GDD	0.0004	0.0004	0.0004	0.0005	−0.0001	0.0010
	(0.0000)					
KDD	−0.0055	−0.0048	−0.0049	0.0037	−0.0096	−0.0014
	(0.0002)					
Precipitation	0.0018	0.0017	0.0015	0.0025	−0.0009	0.0048
	(0.0001)					
Precipitation^2 (×1000)	−0.0014	−1.4e-06	1.2e-06	2.1e-06	−3.8e-06	6.2e-07
	(0.0001)					
Constant	1.5510	1.6671	1.6375	1.9314	−0.8738	3.6623
	(0.1319)					
Obs.	88,101					
R^2	0.84					

Note: Results exclude counties west of the 100th Meridian. The sample period is 1950–2015. Standard errors are reported in parentheses.

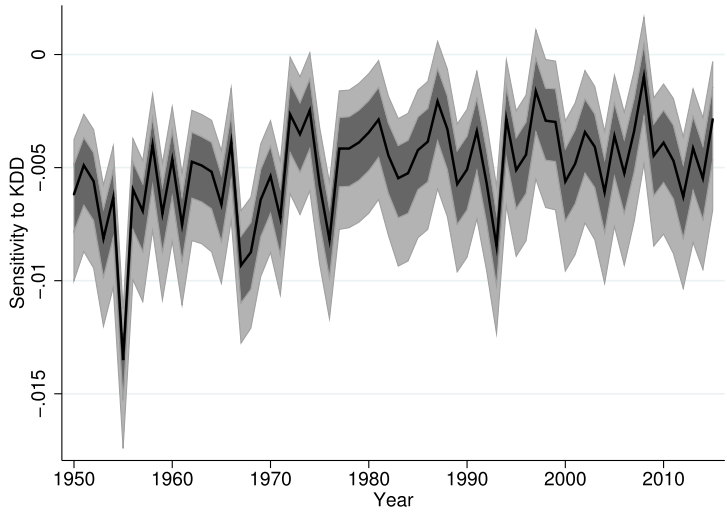

FIGURE H1. Distribution of KDD Slope Coefficients across Time and Counties for U.S. Soy. *Note*: The black line plots the median KDD coefficient from the MO-OLS model in Table H2. The dark (light) grey areas represent the 25th to 75th, and 10th to 90th percentiles, respectively.

Figure H3 compares the log-linear relationships that are obtained using the FE-OLS and MO-OLS models. As with corn, the relationship derived from MO-OLS sits comfortably within the 95% confidence interval for the FE-OLS relationship, indicating that they

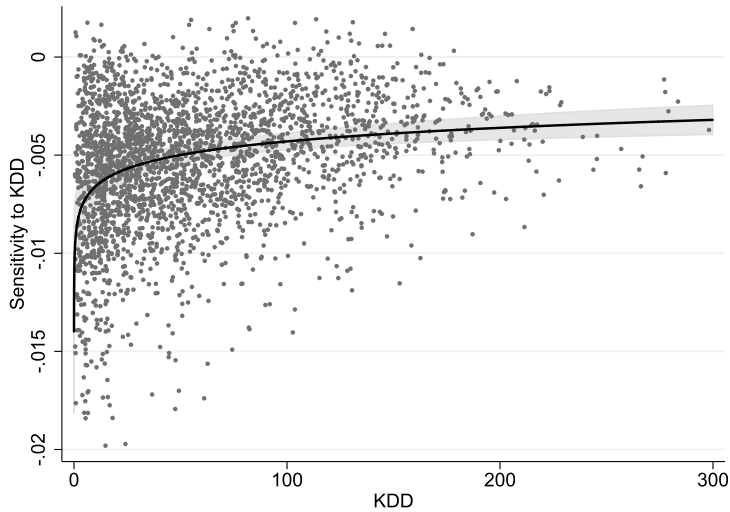

FIGURE H2. Relationship between $\hat{\beta}_{2it}$ and KDD_{it} for U.S. Soybeans. *Note*: This graph is a scatter plot of a random 3% subsample of the coefficients on KDD from the MO-OLS model (see Table H2) against KDD_{it} itself. The fitted line was obtained from the regression $\hat{\beta}_{2it} = \alpha_1 + \alpha_2 \ln(\text{KDD}_{it})$. The estimates are $\alpha_1 = -0.0089$ and $\alpha_2 = 0.0010$ and the 95% confidence interval is shaded.

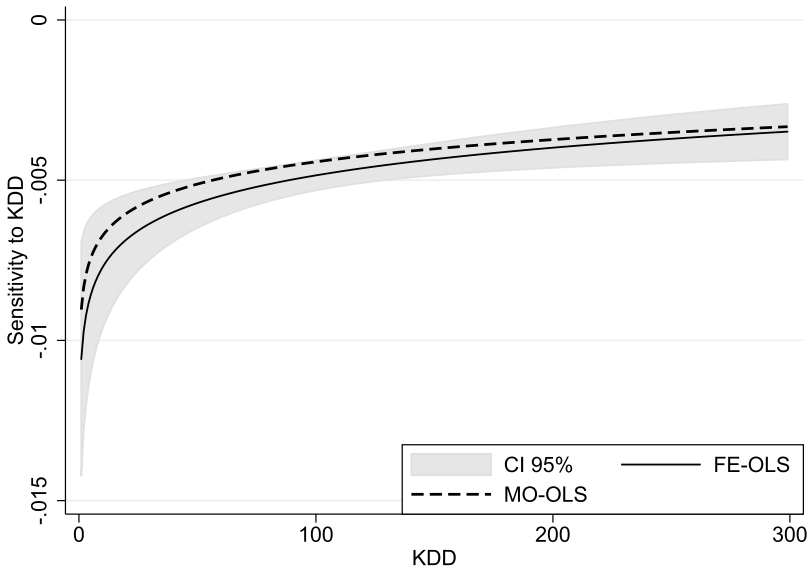

FIGURE H3. Comparison of Log-Linear relationships derived under MO-OLS and FE-OLS for Soybeans. *Note*: This graph compares the fitted log-linear relationships between $\hat{\beta}_{2it}$ and KDD_{it} obtained from (i) the county/time specific coefficients estimated with MO-OLS (see Figure H2) versus (ii) the FE-OLS regression results with adaptation, presented in Table H1, column (3).

are not significantly different from each other. But in contrast to what we find for corn, the MO-OLS relationship here implies less sensitivity to KDD_{it} at every level of KDD_{it}. We can use these results to project soybean yield into the future under climate change.

Turning to the projections, Table H3 presents a summary of the main results for soybean yield with and without adaptation. The format is identical to Table 4 in the text. Without emissions reductions or adaptation, both econometric models predict large reductions in yield. In 2100, according to the MO-OLS model, the ensemble average is a 51% reduction in yield (with an 80% prediction interval of 34% to 69%). The FE-OLS results are very similar. The results under the RCP45 and RCP26 scenarios indicate that emissions reductions are very effective at reducing the impact of climate change on soybean yields, particularly in RCP26 where the prediction interval covers the possibility of no damage over the entire projection horizon.

Figure H4 shows the trajectory of soybean yield under the three RCP scenarios, based on the MO-OLS estimates with no further adaptation. The RCP85 scenario diverges from RCP45 and RCP26 around 2035, leading to a decline in soybean yield of roughly 50 percent (ensemble average) by 2100. The prediction intervals for both the RCP85 and RCP45 scenarios are quite large, especially in the latter half of the century, indicating that the climate models offer varied predictions of growing conditions under these two emission pathways. Meanwhile, as we see in Table H3, the two econometric models give very similar projections across all three RCP scenarios with no adaptation.

We now turn to the projections that incorporate predicted future adaptation. The results are reported in Table H3, right columns, and Figure H5. Similar to the case of

TABLE H3. Effects of climate change on U.S. soybean yield (pct change).

Year	Conventional FE-OLS	MO-OLS w/o future adaptation	FE-OLS with adaptation	MO-OLS with future adaptation
RCP85				
2030	−04 (−11, 04)	−04 (−11, 04)	−04 (−12, 05)	−03 (−10, 04)
2050	−15 (−28, −02)	−15 (−27, −03)	−15 (−28, 00)	−12 (−22, −03)
2080	−36 (−54, −19)	−35 (−51, −20)	−33 (−49, −18)	−28 (−40, −15)
2100	−54 (−74, −35)	−51 (−69, −34)	−48 (−65, −32)	−41 (−55, −26)
RCP45				
2030	−02 (−12, 08)	−02 (−10, 07)	−02 (−13, 09)	−02 (−09, 06)
2050	−09 (−21, 03)	−09 (−20, 02)	−09 (−21, 03)	−08 (−17, 01)
2080	−17 (−29, −05)	−17 (−28, −05)	−17 (−28, −05)	−14 (−23, −05)
2100	−17 (−31, −03)	−17 (−30, −03)	−16 (−30, 00)	−13 (−24, −03)
RCP26				
2030	−05 (−14, 03)	−04 (−12, 03)	−06 (−14, 03)	−04 (−10, 02)
2050	−07 (−15, 01)	−06 (−14, 02)	−07 (−15, 01)	−05 (−12, 01)
2080	−05 (−14, 03)	−05 (−13, 03)	−06 (−15, 03)	−05 (−12, 02)
2100	−07 (−17, 03)	−07 (−17, 04)	−08 (−18, 03)	−06 (−14, 03)

Note: Results are expressed in terms of percentage change from the 2006–2015 historical weighted average crop yield. Each number represents the ensemble average over nineteen climate models, while the numbers in brackets are the the 80% (1.28 standard deviation) prediction interval.

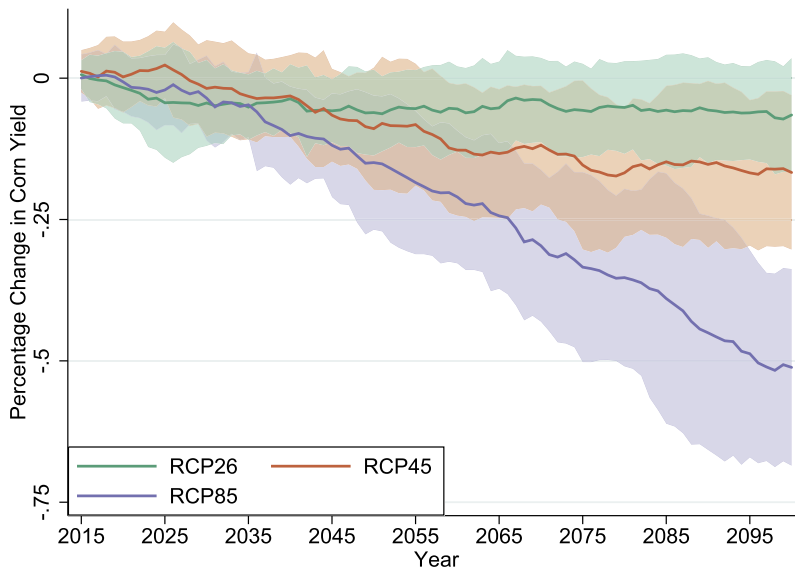

FIGURE H4. The effect of climate change on soy yield by RCP (MO-OLS with no future adaptation). *Note*: This graph presents projections of the percentage change in soybean yield (relative to the 2006–2015 historical average) for three RCPs, where the solid line is the average projection across nineteen CMIP5 climate models, and the shaded areas are the 80% prediction intervals.

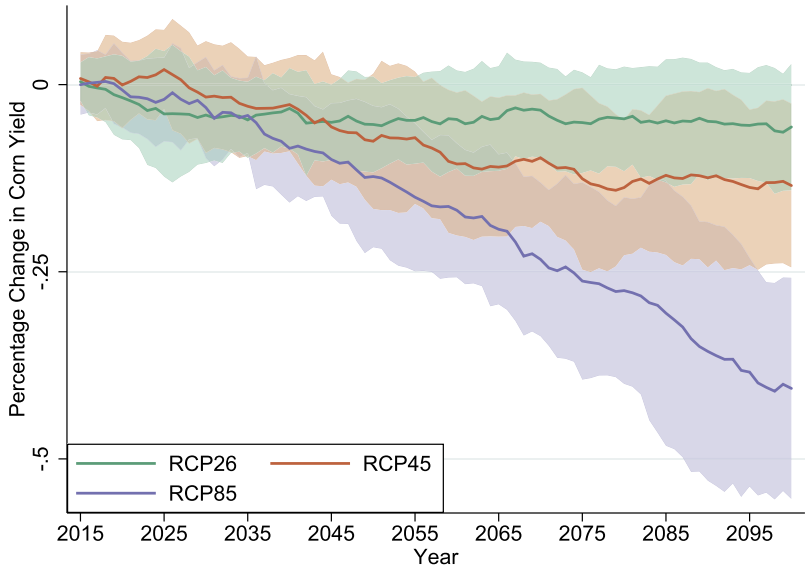

FIGURE H5. The effect of climate change on soybean yield by RCP (MO-OLS with future adaptation). *Note*: This graph presents projections of the percentage change in soybean yield (relative to the 2006–2015 historical average) for three RCPs, where the solid line is the average projection across nineteen CMIP5 climate models, and the shaded areas are the 80% prediction intervals.

corn yield, the MO-OLS model gives a more optimistic projection than FE-OLS of the effectiveness of adaptation at averting future yield damage. Nevertheless, both the FE-OLS and MO-OLS estimates imply that adaptation will do much less to mitigate yield losses for soybeans than for corn. This is intuitive as we find much less heterogeneity in the slope coefficient on KDD for soybeans than we found for corn. Consider first the RCP85 scenario. The MO-OLS model predicts that adaptation will cause the decline in corn yield in 2100 to drop from 51% to 41%. Recall that for corn the yield losses were 70% without adaptation and 36% with adaptation. Thus, in the soybean case, we have both a smaller baseline drop in yield, but also a smaller benefit from adaptation. The FE-OLS model predicts that adaptation will reduce the loss in yield only slightly to 48% in 2100. Of course, the benefits of adaptation become even smaller under the RCP45 and RCP26 scenarios.

Lastly, Table H4 examines the effectiveness of adaptation and alternative emissions reduction scenarios as ways to avert damage from climate change over the whole projec-

TABLE H4. Proportion of climate change damage averted for soybeans (pct).

Estimator	RCP85 + Adapt.	RCP45	RCP45 + Adapt.	RCP26
MO-OLS	17 (13, 22)	60 (36, 85)	64 (42, 86)	78 (53, 103)
FE-OLS	06 (−1, 14)		58 (34, 82)	

Note: Figures are the % reduction in damage relative to the RCP85 scenario with no adaptation using the MO-OLS model.

tion horizon. To construct the percent of damage averted we use the same methodology that we used to construct Table 5 in the main text. The results suggest that adaptation can avert only a small proportion of the total damage, between 13% and 22% according to MO-OLS and −1% to 14% according to FE-OLS. In contrast, shifting from the RCP85 to the RCP45 emissions path offers significantly more damage mitigation, with between 36 to 85% of the damage averted. The RCP26 path averts more damage still, with between half to all of the total damage being averted. It is clear, relative to the results for corn, that the potential for adaptation is far lower for soybeans, while emissions reductions are more effective. Thus, emissions reductions are clearly essential to prevent significant harm to soybean yields.

Reference

Butler, E. and P. Huybers (2013), "Adaptation of US maize to temperature variations." *Nature Climate Change*, 3, 68–72. [4]

Hsiao, C. (1975), "Some estimation methods for a random coefficient model." *Econometrica*, 43, 305–325. [4]

Co-editor Christopher Taber handled this manuscript.

Manuscript received 27 March, 2019; final version accepted 18 April, 2020; available online 15 May, 2020.