The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather: Comment

Online Appendix

Anthony C. Fisher, W. Michael Hanemann, Michael J. Roberts, and Wolfram Schlenker

Contents

Fisher: University of California at Berkeley, Department of Agricultural and Resource Economics, 207 Giannini Hall, Berkeley, CA 94720 (e-mail: acfisher@berkeley.edu). Hanemann: University of California at Berkeley, Department of Agricultural and Resource Economics, 207 Giannini Hall, Berkeley, CA 94720 (e-mail: hanemann@berkeley.edu). Roberts: North Carolina State University, Department of Agricultural and Resource Economics, Box 8109, Raleigh, NC 27695 (e-mail: michael_roberts@ncsu.edu). Schlenker: Columbia University and NBER, Department of Economics and School of International and Public Affairs, 420 W 118th St, New York, NY 10027 (e-mail: wolfram.schlenker@columbia.edu). We would like to thank two anonymous referees, Bernard Salanie, and Kerry Smith for helpful comments. Financial support from Department of Energy Grant number DE-FG02-08ER64640 is gratefully acknowledged.

In this appendix we document our replication of DG in more detail and report additional analysis not presented in the main article.

A1 Coding Issues

DG model yields y using a quadratic function of degree days d, i.e., $y = \beta_0 + \beta_1 d^2 + \beta_2 d$, so the effect of a change in degree days from d_0 to $d_1 = d_0 + \Delta$ affects yields by

$$\begin{aligned} y_1 - y_0 &= \beta_0 + \beta_1 d_1^2 + \beta_2 d_1 - \beta_0 - \beta_1 d_0^2 - \beta_2 d_0 \\ &= \beta_1 (d_0 + \Delta)^2 + \beta_2 (d_0 + \Delta) - \beta_1 d_0^2 - \beta_2 d_0 \\ &= \beta_1 (2\Delta d_0 + \Delta^2) + \beta_2 \Delta \end{aligned}$$

In their STATA code, DG use the variable dd89_7000 to derive the change Δ but use the variable dd89 to derive the average number of baseline degree days d_0. However, because those two variables are inconsistent (dd89 appears to contain many errors, as described in the main paper), this calculation introduces additional noise in their predictions.

DG include 240 observations where corn yield equals zero and one observation where soybean yield equals zero. This is inconsistent with data from USDA's National Agricultural Statistics Service that show positive production for these observations. We speculate these errors stem from an incomplete merge. In our replication we drop observations where yields equal zero in DG's data.

DG weight observations by acreage, which varies over time and across counties. STATA's xtreg command does not allow weights to vary within groups in a panel so we use the average area of a county across all four years in the panel. Because the planting area is endogenous to year-to-year price fluctuations, weighting by the long-term average seems preferable.

In our replication of DG using their data we obtain slightly different damage estimates. DG derive the area-weighted sum of the changes in the weather variables. Summary statistics are provided using STATA's command `summ` and the mean value is then multiplied with the coefficient estimates of the corresponding weather variable. The log-files posted on the AER website reveal that sometimes the authors multiply the coefficients by numbers that differ from what was obtained in the `summ` command.

A2 Sensitivity Checks

Tables A1 , A2 , and A3 extend the results from Table 1 in the main paper for corn, soybeans, and profits, respectively. Columns labeled (1) in each table use year fixed effects, while columns labeled (2) use state-by-year fixed effects. Similar to Table 1 in the main article, columns (a) replicate the results in DG using their data and specification, while columns (b) replicate their specification using our reconstructed data set. Here we add several additional columns to show results from further sensitivity checks.

Columns (1c) and (2c) of each table include one additional variable, the square root of degree days above 34°C, to account for extreme temperatures. DG include such a variable to measure the potentially harmful effect of extreme heat on profits in Table 6 of their paper, but not in their yield regression. Columns (1c) and (2c) also use a slightly different calculation for degree days that accounts for within-day temperature variation.[1] Adding this additional variable increases the fraction of the variance explained through weather variables. Predicted climate change impacts are comparable under the Hadley II model, but more negative under the Haldey III model, which predicts larger temperature increases. It seems intuitive that nonlinear temperature effects are especially important for larger changes. We include the additional weather variable in subsequent regressions (columns (d) and (e)).

Columns (1d) and (2d) include counties for which data is available but missing in DG's analysis. We follow DG's definition and call a county irrigated if more than 10 percent of the farmland is irrigated. The results are reasonably robust to this change.

In columns (1e) and (2e) we limit the comparison to counties east of the 100 degree meridian, which are predominantly non-irrigated. Non-irrigated agriculture is quite unlike agriculture west of the 100^{th} meridian that predominantly uses heavily subsidized irrigation water. DG account for this difference by including separate weather coefficients for irrigated and non-irrigated counties. The problem with DG's approach is that water rights can (and do in California's agricultural regions) vary considerably on a sub-county level. More fundamentally, in irrigated areas the water input comes from groundwater or from precipitation falling elsewhere, so local precipitation is not a valid measure of water supply, and predicted climatic changes in precipitation do not measure predicted changes in access to irrigation water (see Schlenker, Hanemann & Fisher (2007)).When we limit the sample to predomi-

[1]Construction of our weather data from raw sources is detailed in Schlenker & Roberts (2009). These data use the distribution of temperatures within a day between minimum and maximum instead of just the average for the day. This variance can make a difference due to the way growing degree days are defined, especially since maximum temperatures often exceed 34°C while average temperatures do not.

nantly non-irrigated areas, predicted impacts increase only slightly in the yield regressions. This is not surprising as little corn and soybeans are grown in the western United States and the regression uses area weights following DG. Predicted impacts vary more in the profit regression, but standard errors are also much larger.

A3 Storage and Lagged Weather Variables

DG argue in their reply that including lagged weather variables solves the storage problem described in the main article. We disagree. Consider, for example, the theoretical model of competitive storage by Scheinkman & Schechtman (1983). The model predicts a strong contemporaneous correlation between storage and weather shocks. Lagged weather variables cannot mitigate the problem of storage because lagged weather variables do not break the contemporaneous correlation between weather in period t and storage in period t, which enters the error term. Indeed, if weather is random as argued by DG, lagged weather should bear *no* relation to current weather, so past weather cannot control for endogeneity created by current weather. Besides, the theory indicates current weather has a much stronger influence on storage decisions than does past weather.

Using the notation from DG's original paper on page 367:

$$y_{ct} = \alpha_c + \gamma_{st} + X_{ct}\beta + \sum_i \theta_i f_i(W_{ict}) + u_{ct}$$

where y_{ct} is the economic profit in county c in year t, α_c is a county fixed effect, γ_{st} is a year fixed effect or state-by-year fixed effect, X_{ct} are time-varying county-specific control variables like soil quality, W_{ict} is weather variable i in county c in year t. Finally, u_{ct} is the error term.

The economic profit in period t is the value of production minus expenditures, i.e., $y_{ct} = p_{ct}q^p_{ct} - e_{ct}$, where q^p_{ct} is the amount produced and e_{ct} are expenditures. However, the Census of Agriculture only reports sales s_{ct} in a given period, which does not account for the amount stored q^n_{ct}, i.e., $s_{ct} = p_{ct}(q^p_{ct} - q^n_{ct})$ or $p_{ct}q^p_{ct} = s_{ct} + p_{ct}q^n_{ct}$. Substituting this expression into the previous equation above we obtain what DG estimate:

$$s_{ct} - e_{ct} = \alpha_c + \gamma_{st} + X_{ct}\beta + \sum_i \theta_i f_i(W_{ict}) + \underbrace{u_{ct} - p_{ct}q^n_{ct}}_{v_{ct}}$$

The value of storage $p_{ct}q^n_{ct}$ enters the combined error term v_{ct}. As long as there is a contemporaneous correlation between the weather shock W_{ict} and storage q_{ct}^{n}, this violates the identifying assumption of DG that $\mathbb{E}[f_i(W_{ict})v_{ct}|\alpha_c + \gamma_{st} + X_{ct}\beta] = 0$.

Including lagged weather variable does not break this contemporaneous correlation of W_{ict} and v_{ct}. Independent of the past, bad weather shocks increase price and good weather shocks decrease price, and this contemporaneous price movement induces storage behavior. One might wonder whether temporal fixed effects (year or state-by-year fixed effects) account for such storage behavior. Table 3 in our manuscript shows that this is not the case.

Here we present direct evidence that shows contemporaneous weather shocks influence decisions to store corn and soybeans, the nation's two largest crops, even after controlling for prices using year fixed effects and lagged weather shocks. Because no county-level inventory data exist, we use state-level data from the National Agricultural Statistics Service on stock levels of corn and soybeans by state. We regress state-level log inventory levels on damaging extreme heat (degree days above 29°C for corn and degree days above 30°C for soybeans) for the years 1950-2005.[2] In all regressions we include year fixed effects which will capture variation in price of a commodity. Results are reported in Table A4 .

In the main paper, we cite evidence that local price variation exists due to transportation costs or convenience yields. Local price variation will not be captured by yearly fixed effects, but correlated with weather. We observe evidence of this in our regression results: state-level variation in harmful heat is a significant predictor of state-level changes in inventories both for corn (column 1a) and for soybeans (column 2a). More extreme heat reduces yields and decreases inventories as farmers supplement reductions in output by depleting inventories, even after accounting for national prices using year fixed effects. Moreover, if we include two lags of the weather variable (reported in columns b of Table A4), the contemporaneous correlations between extreme heat and storage in the first row remain significant, and even increases for the case of corn.

A4 Robustness of Hedonic Model

In the first half of their paper, DG argue that results from the cross-sectional hedonic approach are not robust. They contend that cross-sectional studies rely on climate variations that are too closely associated with unobserved factors relating to location, and thus likely to be biased. DG's standard of robustness is the consistency of parameter estimates over

[2]These weather variables have been shown to be the single best predictor of corn and soybean yields in the US (Schlenker & Roberts 2009). In this analysis weather variables are cropland-area weighted averages of our county-level weather measures.

different subsets of the data. However, when DG assess the robustness of the hedonic model, the weather variables they use, following Mendelsohn, Nordhaus & Shaw (1994), are average monthly temperature and precipitation for January, April, July and October; when they demonstrate the robustness of their fixed-effects model, the weather variables they use are summed precipitation and degree days over the growing season, which have been shown to be superior on both economic and econometric grounds in a hedonic analysis (Schlenker, Hanemann & Fisher 2006).[3]

The difference in the representation of weather confounds DG's comparison of robustness. For these reasons, we believe it is appropriate to repeat DG's tests of robustness, shown in their Figure 3, using the degree days representation of temperature and applying the tests to a hedonic model. This analysis is summarized in our Figure A1 . The figure shows predicted impacts from climate change using the hedonic model specified in Schlenker, Hanemann & Fisher (2006), estimated using various sets of control variables, in various census years. We replicate this analysis for farms east of the 100th meridian, an approximate boundary between rainfed and irrigated agriculture in the U.S., comprising 80% of county observations. Farms west of the 100th meridian rely on heavily subsidized irrigation water that is capitalized into farmland values. Some areas east of the 100th meridian also rely on irrigation. For example, 79 percent of the corn acreage in Arkansas was irrigated in 2007. The main difference is that access to water, where it exists, is much less heavily subsidized. Cline (1996) emphasizes that a hedonic approach in which observations are pooled assumes that farmers can obtain irrigation at existing marginal cost, which seems unrealistic. The bias induced by pooling observations from areas with subsidized irrigation with areas not primarily dependent on (subsidized) irrigation is demonstrated in Schlenker, Hanemann & Fisher (2005).[4]

Panel A in Figure A1 includes only climate variables with no other controls. If the climate variables are correlated with other variables, such as soil quality, the coefficient on the climate variables will be biased. Accordingly, panel B includes controls for both soil and socio-economic variables to avoid potential omitted variable problems. Finally, panel C additionally includes state fixed effects. Columns marked with a [0] are unweighted regressions, while columns marked with a [1] are regressions weighted by the square root of acres of farmland.

[3]DG write that they are replicating Schlenker, Hanemann & Fisher (2005), but this is incorrect. They critique Schlenker, Hanemann & Fisher's (2005) *replication* of Mendelsohn, Nordhaus & Shaw (1994).

[4]DG assert that Schlenker, Hanemann & Fisher (2005) force the coefficient on irrigated and nonirrigated areas to be the same. The assertion is incorrect - the main point of that paper is that a *different* model is needed to deal with irrigated agriculture since the water supply there does not depend on local precipitation.

All estimates indicate strong negative impacts from climate change, even when all independent variables except for the climatic variables are dropped (panel A). Based on the tests suggested by DG, a hedonic model using degree days and precipitation summed over the growing season and applied to counties not primarily dependent on irrigation is robust. The model also passes a wide array of additional robustness tests discussed in Schlenker, Hanemann & Fisher (2006).

References

Cline, William R. 1996. “The Impact of Global Warming of Agriculture: Comment.” *American Economic Review*, 86(5): 1309–1311.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw. 1994. “The Impact of Global Warming on Agriculture: A Ricardian Analysis.” *American Economic Review*, 84(4): 753–771.

Scheinkman, Josè A., and Jack Schechtman. 1983. “A Simple Competitive Model with Production and Storage.” *Review of Economic Studies*, 50(3): 427–441.

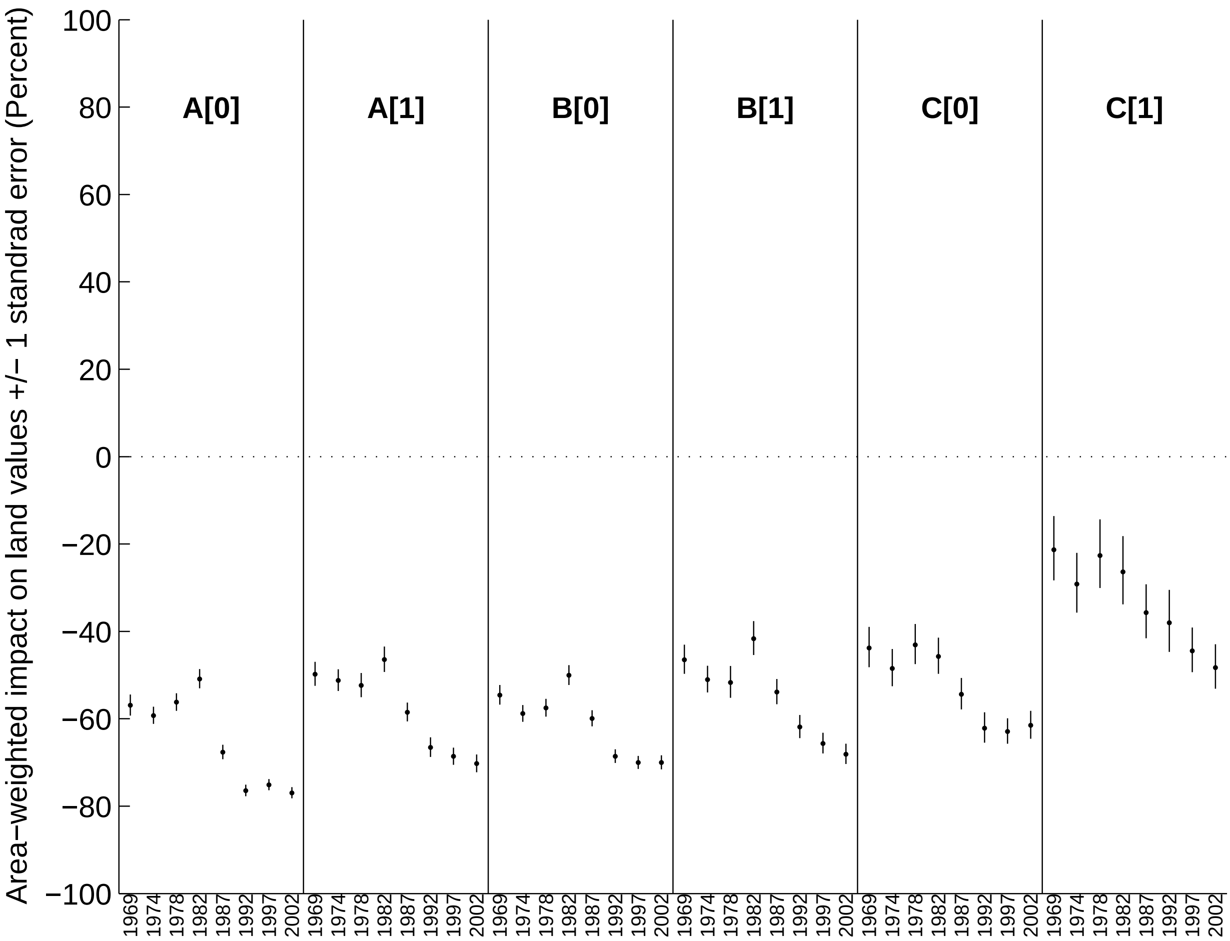
Schlenker, Wolfram, and Michael J. Roberts. 2009. “Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields under Climate Change.” *Proceedings of the National Academy of Sciences*, 106(37): 15594–15598.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher. 2005. “Will U.S. Agriculture Really Benefit from Global Warming? Accounting for Irrigation in the Hedonic Approach.” *American Economic Review*, 95(1): 395–406.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher. 2006. “The Impact of Global Warming on U.S. Agriculture: An Econometric Analysis of Optimal Growing Conditions.” *Review of Economics and Statistics*, 88(1): 113–125.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher. 2007. “Water Availability, Degree Days, and the Potential Impact of Climate Change on Irrigated Agriculture in California.” *Climatic Change*, 81(1): 19–38.

Figure A1 : Robustness of Hedonic Model using Degree Days



Notes: This is a replication of Figure 3 in Deschenes and Greenstone using the degree days from Schlenker, Hanemann & Fisher (2006) instead of average monthly temperatures. Panel A only includes climate variables with no other controls. Panel B includes controls for both soil and socio-economic variables to avoid potential omitted variable problems. Panel C additionally includes state fixed effects. Columns marked with a [0] are unweighted regressions, while columns marked with a [1] are regressions weighted by the square root of acres of farmland. The x-axis lists the year in which the cross-section is estimated.

Table A1 : Comparison of Various Data Sources in Corn Regressions

	(1a)	(1b)	(1c)	(1d)	(1e)	(2a)	(2b)	(2c)	(2d)	(2e)
Regression diagnostics										
Variance explained by weather	11.6%	19.6%	24.1%	24.8%	24.8%	7.0%	12.9%	16.9%	17.9%	20.4%
Climate change impact (Percent)										
Hadley II-IS92a scenario	-0.80	-10.61	-12.74	-14.66	-16.90	0.23	-20.36	-18.58	-16.71	-13.74
(s.e.)	(1.24)	(1.45)	(1.43)	(1.18)	(1.29)	(1.11)	(3.04)	(2.84)	(2.33)	(2.34)
[s.e. clustered by state]	[2.08]	[4.18]	[3.56]	[4.98]	[5.99]	[2.04]	[5.40]	[3.70]	[3.58]	[3.35]
Hadley III-B2 scenario		-42.01	-61.26	-67.36	-75.97		-60.58	-72.30	-71.95	-74.48
(s.e.)		(3.23)	(3.51)	(3.13)	(3.60)		(6.55)	(6.51)	(5.57)	(5.74)
[s.e. clustered by state]		[11.14]	[9.72]	[12.86]	[14.91]		[12.16]	[12.16]	[11.68]	[9.71]
Observations	6623	6623	6623	8562	7538	6623	6623	6623	8562	7538
Soil controls	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
State-by-Year FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Notes: This table summarizes and compares alternative regression models. Columns (a) replicate the results in DG using their code and data (a quadratic in degree days 8-32°C and precipitation); columns (b) are the same models as (a) estimated with our reconstructed data; columns (c) account for within-day temperature variation and extremely warm temperatures (square root of degree days above 34°C as an additional variable); columns (d) include observations that are missing in DG's data; columns (e) use only counties east of the 100 degree meridian that are treated as dryland. The variance explained by weather is 1 minus the ratio of the residual variance of the full specification over the residual variance of the model excluding weather. Standard errors in round brackets cluster by fips code following DG, while standard errors in brackets cluster by state.

Table A2 : Comparison of Various Data Sources in Soybeans Regressions

	(1a)	**(1b)**	**(1c)**	**(1d)**	**(1e)**	**(2a)**	**(2b)**	**(2c)**	**(2d)**	**(2e)**
Regression diagnostics										
Variance explained by weather	14.4%	30.6%	37.0%	35.9%	35.4%	12.0%	19.4%	22.1%	20.8%	20.9%
Climate change impact (Percent)										
Hadley II-IS92a scenario	-2.73	-15.63	-16.74	-14.77	-14.85	1.46	-18.36	-14.42	-13.25	-13.77
(s.e.)	(1.38)	(1.60)	(1.55)	(1.20)	(1.21)	(0.94)	(2.58)	(2.50)	(2.01)	(2.01)
[s.e. clustered by state]	[2.08]	[4.93]	[4.45]	[4.10]	[4.16]	[1.15]	[3.85]	[3.21]	[3.40]	[3.36]
Hadley III-B2 scenario		-51.59	-70.72	-64.25	-65.01		-55.63	-60.30	-56.14	-58.28
(s.e.)		(3.65)	(3.65)	(2.91)	(2.93)		(5.72)	(5.65)	(4.63)	(4.67)
[s.e. clustered by state]		[11.80]	[9.88]	[8.99]	[9.35]		[9.32]	[8.71]	[8.44]	[8.25]
Observations	5140	5140	5140	6742	6504	5140	5140	5140	6742	6504
Soil controls	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
State-by-Year FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Notes: This table summarizes and compares alternative regression models. Columns (a) replicate the results in DG using their code and data (a quadratic in degree days 8-32°C and precipitation); columns (b) are the same models as (a) estimated with our reconstructed data; columns (c) account for within-day temperature variation and extremely warm temperatures (square root of degree days above 34°C as an additional variable); columns (d) include observations that are missing in DG's data; columns (e) use only counties east of the 100 degree meridian that are treated as dryland. The variance explained by weather is 1 minus the ratio of the residual variance of the full specification over the residual variance of the model excluding weather. Standard errors in round brackets cluster by fips code following DG, while standard errors in brackets cluster by state.

Table A3 : Comparison of Various Data Sources in Profit Regressions

	(1a)	**(1b)**	**(1c)**	**(1d)**	**(1e)**	**(2a)**	**(2b)**	**(2c)**	**(2d)**	**(2e)**
Regression diagnostics										
Variance explained by weather	0.4%	1.5%	1.5%	1.1%	1.2%	0.4%	0.6%	0.8%	0.4%	0.3%
Climate change impact (Percent)										
Hadley II-IS92a scenario	-6.63	-36.50	-31.38	-25.44	-28.94	3.75	1.21	4.45	5.60	1.05
(s.e.)	(3.03)	(5.41)	(5.28)	(4.11)	(4.90)	(2.82)	(12.88)	(12.93)	(10.96)	(6.59)
[s.e. clustered by state]	[4.98]	[10.34]	[10.82]	[9.76]	[14.88]	[3.98]	[15.18]	[17.06]	[14.67]	[13.55]
Hadley III-B2 scenario		-55.99	-44.29	-33.89	-69.40		-3.28	6.40	6.75	-7.63
(s.e.)		(8.93)	(9.52)	(7.55)	(9.01)		(20.61)	(20.99)	(17.82)	(12.25)
[s.e. clustered by state]		[16.58]	[19.80]	[18.60]	[26.08]		[25.12]	[28.80]	[25.37]	[26.47]
Observations	9024	9024	9024	11949	9653	9024	9024	9024	11949	9653
Soil controls	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
State-by-Year FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Notes: This table summarizes and compares alternative regression models. Columns (a) replicate the results in DG using their code and data (a quadratic in degree days 8-32°C and precipitation); columns (b) are the same models as (a) estimated with our reconstructed data; columns (c) account for within-day temperature variation and extremely warm temperatures (square root of degree days above 34°C as an additional variable); columns (d) include observations that are missing in DG's data; columns (e) use only counties east of the 100 degree meridian that are treated as dryland. The variance explained by weather is 1 minus the ratio of the residual variance of the full specification over the residual variance of the model excluding weather. Standard errors in round brackets cluster by fips code following DG, while standard errors in brackets cluster by state.

Table A4 : Regressing Storage on Contemporaneous and Lagged Weather

	Corn		Soybeans	
	(1a)	**(1b)**	**(2a)**	**(2b)**
Extreme Heat	-1.57e-3**	-2.57e-3***	-4.60e-3***	-4.18e-3***
(s.e.)	(7.65e-4)	(7.73e-4)	(9.41e-4)	(9.05e-4)
Extreme Heat (Lag 1)		4.02e-4		-2.64e-3***
(s.e.)		(7.85e-4)		(9.18e-4)
Extreme Heat (Lag 2)		7.36e-4		-2.71e-3***
(s.e.)		(7.50e-4)		(9.39e-4)
Observations	1535	1453	1006	971
Year FE	Yes	Yes	Yes	Yes

Notes: This table regresses state-level log inventory levels on extreme heat (degree days 29°C for corn and degree days 30°C for soybeans). Stars indicate significance levels: *** 1%; ** 5%; and * 10%.