

## ONLINE APPENDIX

### Climate Change and Agriculture: Subsistence Farmers' Response to Extreme Heat

Fernando M. Aragón and Francisco Oteiza and Juan Pablo Rud

#### ADDITIONAL FIGURES AND TABLES

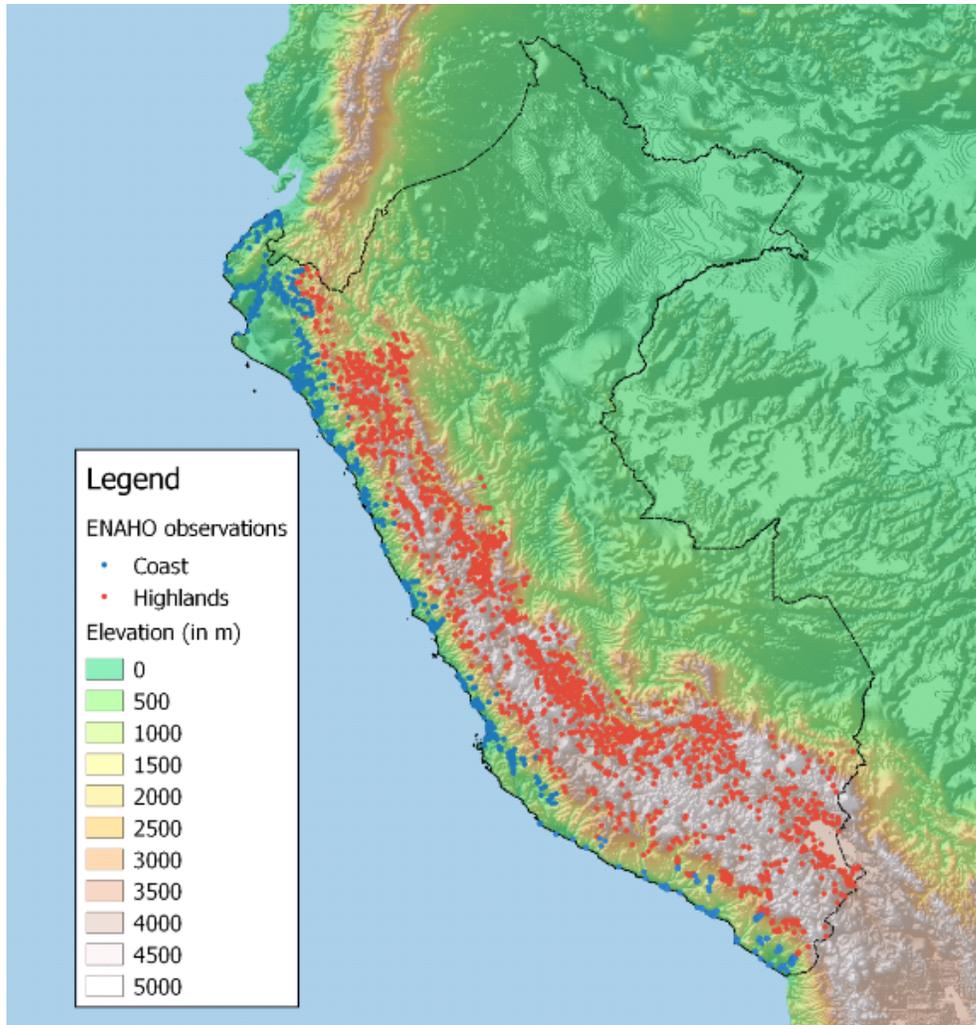


FIGURE A.1. ENAHO OBSERVATIONS 2007-2015

*Notes:* Map depicts Peru's climatic regions and location of the ENAHO clusters used in this study.

TABLE A.1—EFFECT OF HDD ON AREA PLANTED IN A GIVEN MONTH

	Area planted with transitory crops (has) in month $t$					% annual area planted
	(1)	(2)	(3)	(4)	(5)	(6)
HDD $_{t-4}$	-0.0003 (0.000)	-0.0016** (0.001)	-0.0003 (0.000)	-0.0024 (0.002)	-0.0014*** (0.000)	0.0388*** (0.011)
HDD $_{t-3}$	0.0002 (0.000)	-0.0001 (0.001)	0.0004 (0.000)	0.0020 (0.002)	-0.0014** (0.001)	0.0831*** (0.013)
HDD $_{t-2}$	0.0024*** (0.001)	0.0039*** (0.001)	0.0010 (0.001)	0.0033** (0.001)	0.0017*** (0.001)	0.1257*** (0.015)
HDD $_{t-1}$	0.0018*** (0.001)	0.0022** (0.001)	0.0017* (0.001)	0.0024** (0.001)	0.0024*** (0.001)	0.0486*** (0.016)
HDD $_t$	-0.0003 (0.001)	-0.0003 (0.001)	0.0002 (0.001)	-0.0004 (0.001)	0.0017** (0.001)	-0.1069*** (0.015)
HDD $_{t+1}$					-0.0028*** (0.001)	
HDD $_{t+2}$					-0.0011* (0.001)	
HDD $_{t+3}$					-0.0021*** (0.001)	
HDD $_{t+4}$					0.0005 (0.001)	
Specification	Baseline	Only Coast	Only Highlands	Spring planting	Adding leads	Alternative outcome
No. obs.	480,462	98,317	382,145	192,348	438,298	480,280
R-squared	0.023	0.023	0.033	0.012	0.026	0.118
No. farmers	38,485	7,908	30,577	38,467	38,471	38,472

*Notes:* Standard errors clustered at farmer level (in parenthesis). Stars indicate statistical significance: \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Regression uses data from the Peruvian National Agricultural years 2014 to 2017. This dataset has farm-level data of monthly planting over a 12-month period. All specifications include farmer, month-by-strata, and year-by-strata fixed effects. Columns 2 and 3 restrict sample to a climatic region (Coast or Highlands). Column 4 restricts sample to planting done in months of August to December. Column 5 adds leads of HDD, while column 6 uses the share of annual area planted in a given month (i.e. area planted in month  $t$  / total area planted in a year) as outcome variable.

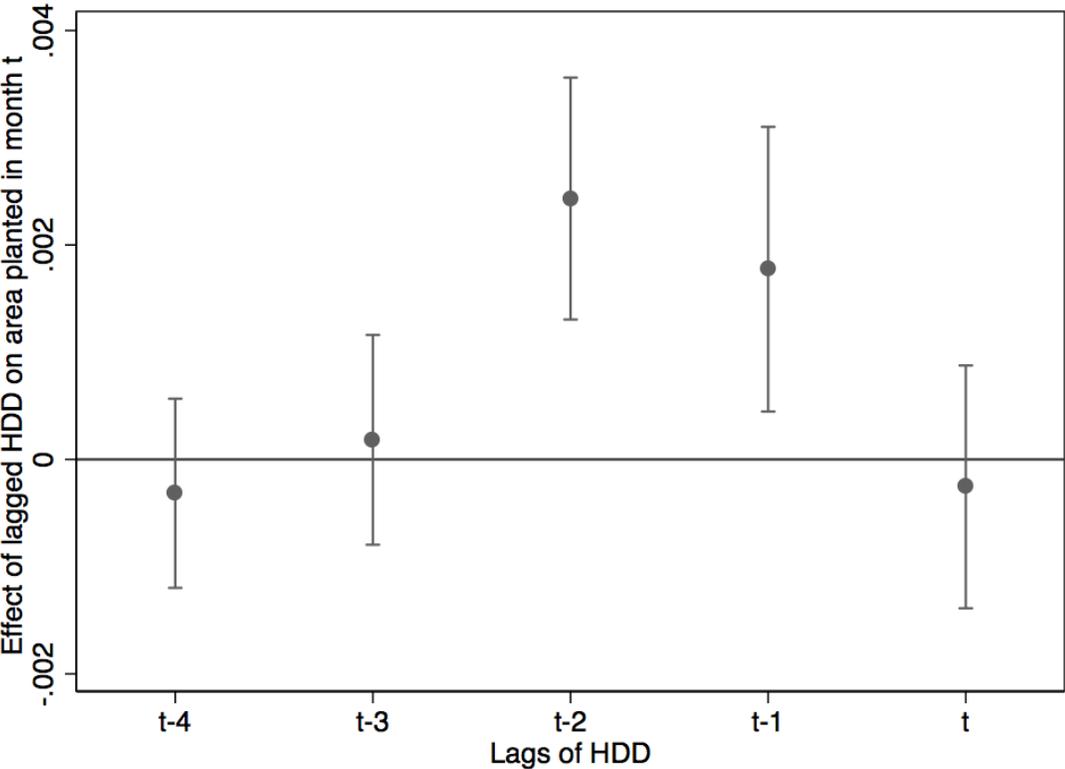


FIGURE A.2. EFFECT OF LAGGED HDD ON AREA PLANTED IN A GIVEN MONTH

Notes: Figure displays results of regressing area planted with transitory (annual) crops in month  $t$  on lagged values of HDD ( $t$  to  $t - 4$ ). Regression uses data from the Peruvian National Agricultural years 2014 to 2017. This dataset has farm-level data of monthly planting over a 12-month period. Regression includes farmer, month-by-strata, and year-by-strata fixed effects. Dots are point estimates and lines indicate 95% confidence intervals. Standard errors clustered at the farmer level. Estimates and additional checks are available in Table A.1.

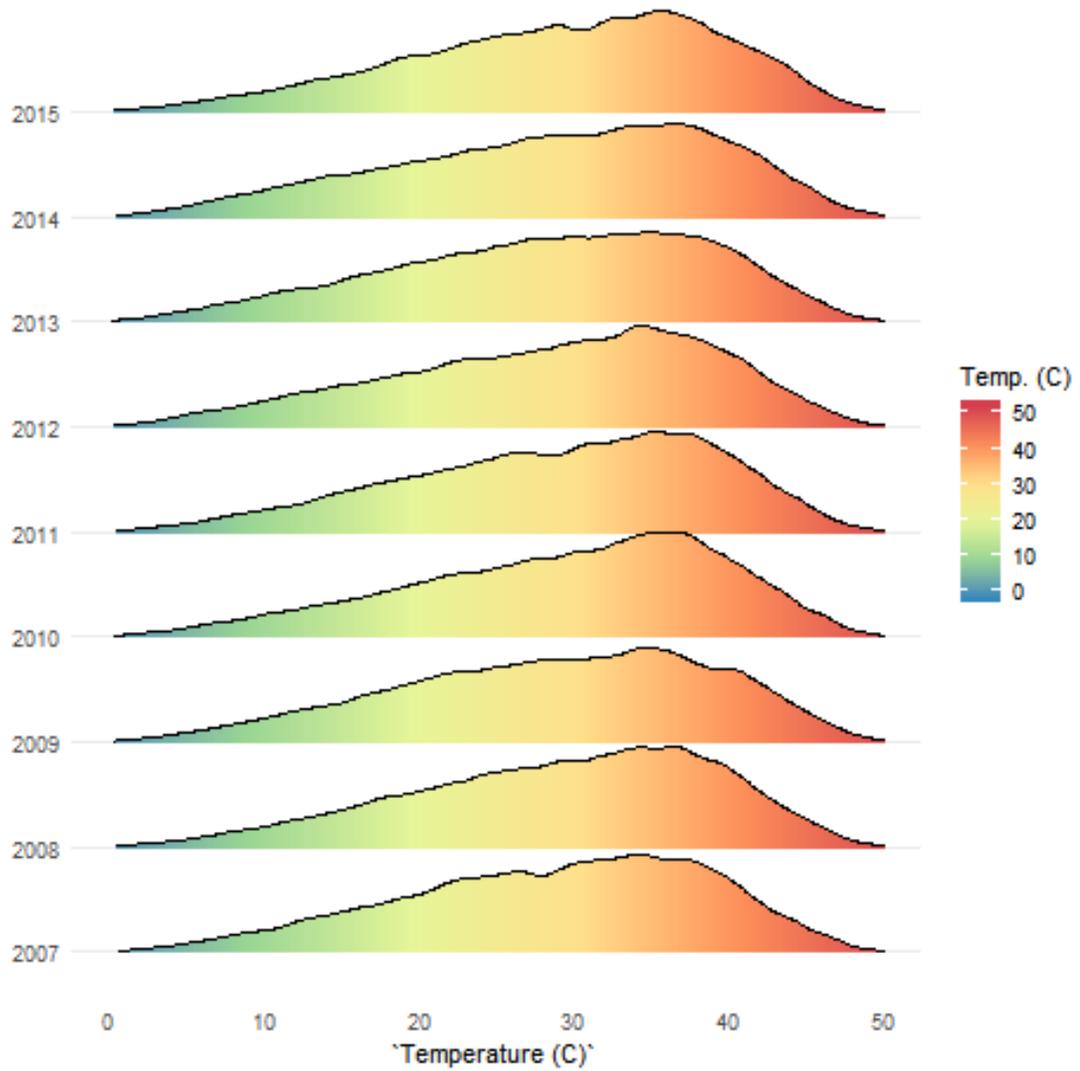


FIGURE A.3. DISTRIBUTION OF DAILY AVERAGE TEMPERATURE BY GROWING SEASON

*Notes:* Figure depicts the share of days spent in each temperature bin by the farmers in our sample, during the 2007-2015 growing seasons (i.e., October to March).

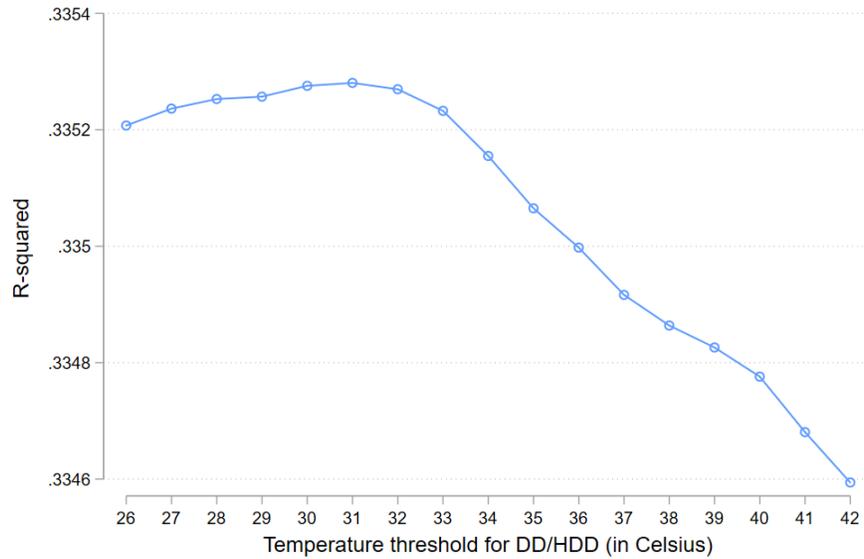


FIGURE A.4. OPTIMAL TEMPERATURE THRESHOLD USING THE ITERATIVE REGRESSION APPROACH

TABLE A.2—TEMPERATURE VARIATION UNDER VARIOUS SETS OF FIXED EFFECTS (IN °C)

	DD			HDD		
	$R^2$ (1a)	$\sigma_e$ (1b)	$ e  > 1^\circ\text{C}$ (1c)	$R^2$ (2a)	$\sigma_e$ (2b)	$ e  > 1^\circ\text{C}$ (2c)
No fixed effects (FE)		4.81	100.0%		1.34	28.6%
District FE	0.90	1.50	37.7%	0.86	0.44	23.2%
District + growing season FE	0.91	1.41	36.5%	0.86	0.43	23.1%
District + growing season-by-region FE	0.92	1.40	36.2%	0.87	0.42	23.1%

*Notes:* This table replicates Table 2 of ?, It summarises regressions of measures of temperature on various sets of fixed effects and shows how much of the variation they absorb. The first three columns use average degree days (DD), and the last three columns use harmful degree days(HDD), using a threshold of 33°C. Columns (a) report the  $R^2$  of the regression; columns (b) report the standard deviation of the residuals (remaining temperature variation) in degrees Celsius during the growing season; and columns (c) report what fraction of the observations have a residual that is larger than 1°C over the growing season.

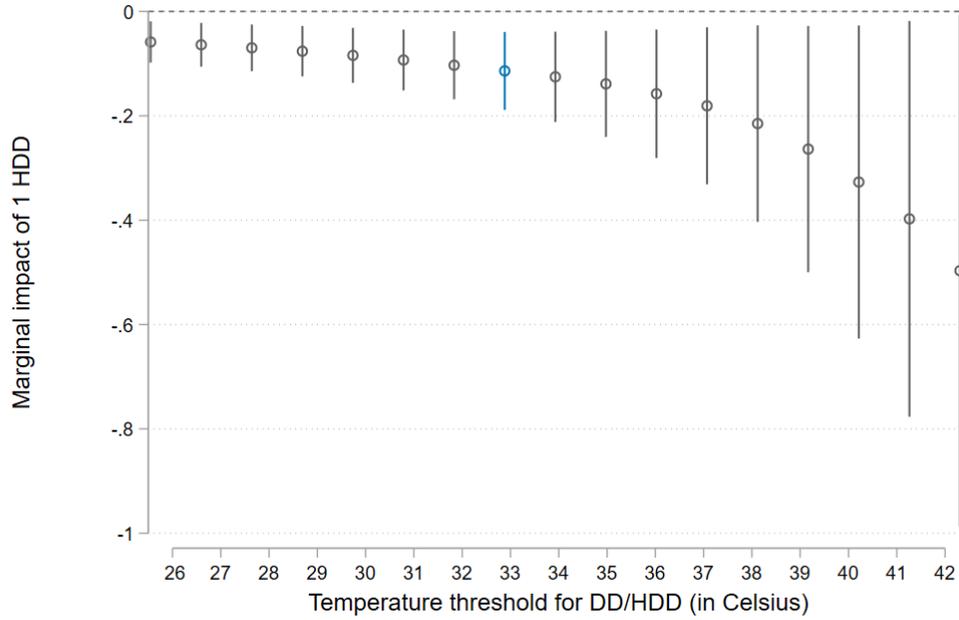


FIGURE A.5. EFFECT OF HDD ON LN(OUTPUT PER HECTARE) USING ALTERNATIVE DD/HDD THRESHOLDS

TABLE A.3—TEMPERATURE VARIATION UNDER VARIOUS SETS OF FIXED EFFECTS (IN °F )

Variable: Measure:	DD			HDD		
	$R^2$ (1a)	$\sigma_e$ (1b)	$ e  > 1^\circ\text{F}$ (1c)	$R^2$ (2a)	$\sigma_e$ (2b)	$ e  > 1^\circ\text{F}$ (2c)
No fixed effects (FE)		8.66	100.00		2.40	39.27
District FE	0.90	2.70	46.30	0.86	0.79	25.73
District + growing season FE	0.91	2.54	45.54	0.86	0.77	25.43
District + growing season-by-region FE	0.92	2.52	44.98	0.87	0.76	25.34

*Notes:* This table replicates Table 2 of ?, It summarises regressions of measures of temperature on various sets of fixed effects and shows how much of the variation they absorb. The first three columns use average degree days (DD), and the last three columns use harmful degree days(HDD), using a threshold of 33°F . Columns (a) report the  $R^2$  of the regression; columns (b) report the standard deviation of the residuals (remaining temperature variation) in Fahrenheit degrees during the growing season; and columns (c) report what fraction of the observations have a residual that is larger than 1°F over the growing season.

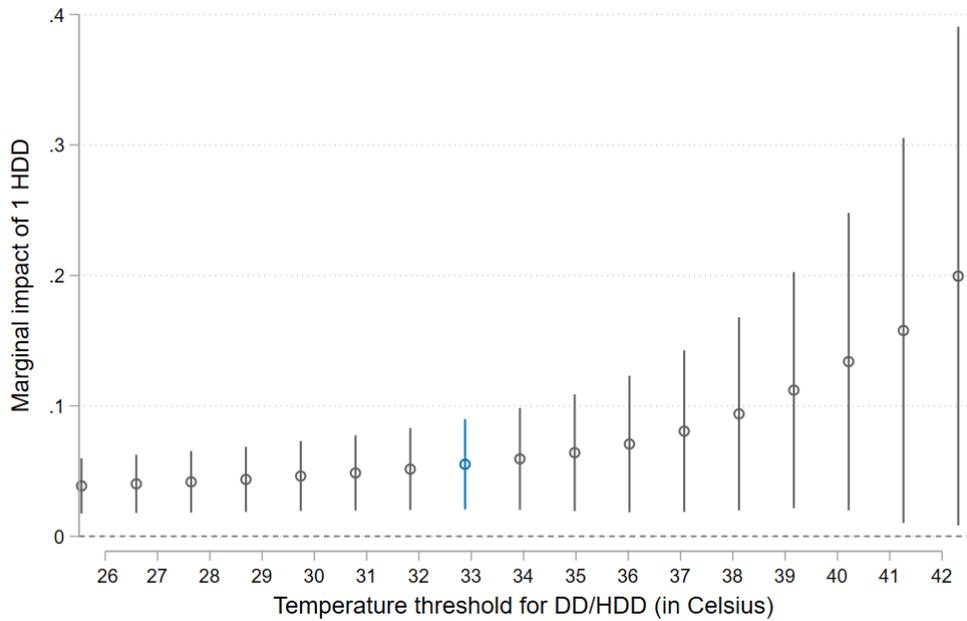


FIGURE A.6. EFFECT OF HDD ON LN(AREA PLANTED) USING ALTERNATIVE DD/HDD THRESHOLDS

TABLE A.4—FIRST STAGE OF 2SLS REGRESSION (COLUMN 3 IN TABLE ??)

Dep. Variable:	ln(area planted)	ln(no. HH members members work in farm)
	(1)	(2)
ln(area owned)	0.165*** (0.005)	0.007*** (0.001)
ln(HH size)	0.195*** (0.013)	0.494*** (0.006)
No. obs.	53,487	53,487
R-squared	0.478	0.481

Notes: Standard errors clustered at district level (in parenthesis). Stars indicate statistical significance: \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Table presents first stage of 2SLS regression presented in column 3 in Table ?. Regression has all included such as district, month of interview, and climatic region-by-growing season fixed effects, and a set of farmer controls.

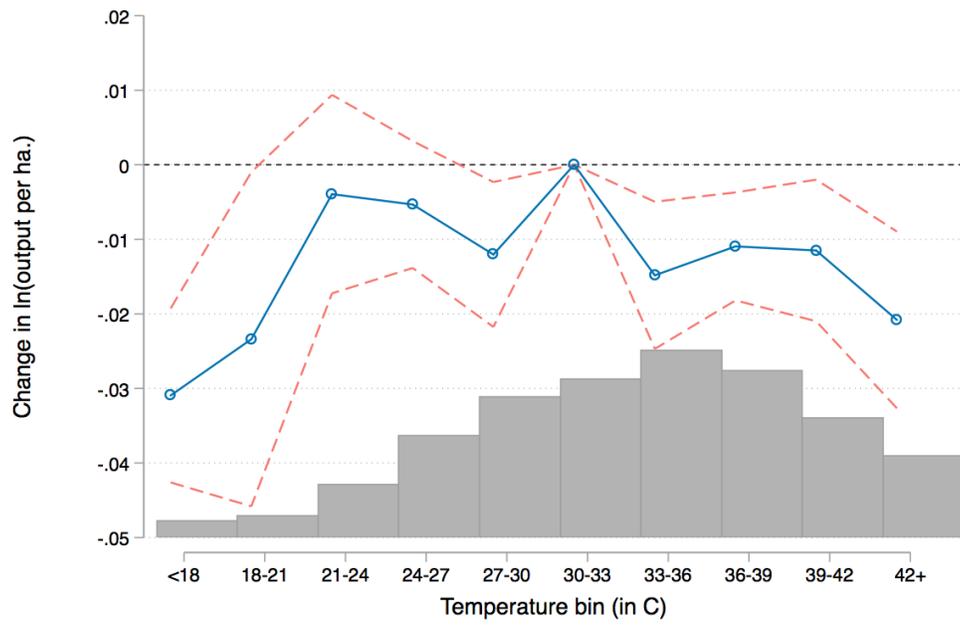


FIGURE A.7. NON-LINEAR RELATIONSHIP BETWEEN TEMPERATURE AND AGRICULTURAL YIELDS BY REGION  
- COAST

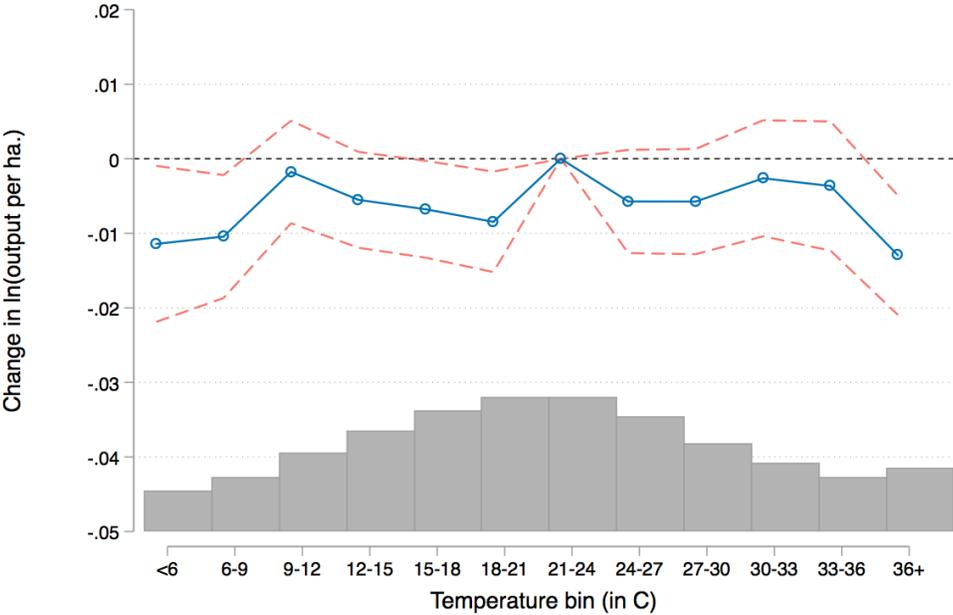


FIGURE A.8. NON-LINEAR RELATIONSHIP BETWEEN TEMPERATURE AND AGRICULTURAL YIELDS BY REGION - HIGHLANDS

TABLE A.5—TEMPERATURE AND AGRICULTURAL PRODUCTIVITY (TFP), ALTERNATIVE SPECIFICATIONS

Dep. Variable:	ln(output)		
	(1)	(2)	(3)
Average DD in growing season	0.014* (0.007)	0.014* (0.007)	0.013* (0.007)
Average HDD in growing season	-0.064* (0.033)	-0.063* (0.033)	-0.062* (0.033)
Inputs controls	Yes	Yes	Yes
Endowment controls	Yes	No	Yes
3rd degree Taylor expansion of inputs	No	Yes	Yes
No. obs.	53,487	53,487	53,487
R-squared	0.550	0.552	0.552

*Notes:* Standard errors clustered at district level (in parenthesis). Stars indicate statistical significance: \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All specifications are estimated using OLS and include district, month of interview, and climatic region-by-growing season fixed effects, and the same farmer controls as regression in column 2 of Table ???. Input controls: log of area planted, number of household members working in agriculture, and amount spent on hired labor. Endowment controls: log of household size and area of land owned. Columns 2 and 3 include a 3rd degree Taylor expansion of two inputs: log of area planted, number of household members working in agriculture.

TABLE A.6—EFFECT OF HDD ON OTHER FARM INPUTS

Dep var:	Fertilizers		Pesticides	
	(1) Extensive	(2) Intensive	(3) Extensive	(4) Intensive
Average DD	-0.003 (0.003)	-0.021 (0.022)	0.001 (0.004)	0.002 (0.018)
Average HDD	0.003 (0.010)	0.002 (0.052)	0.005 (0.008)	0.029 (0.043)
No. obs.	53,619	53,618	53,619	53,618
R-squared	0.272	0.375	0.245	0.354

*Notes:* Extensive margin use is studied using a dummy variable equal to one if the farmer reports to have used fertilizers/pesticides during the last growing season. Intensive margin use is defined as the logarithm of total amounts spent on fertilizers/pesticides. Standard errors clustered at the district level (in parenthesis). Stars indicate statistical significance: \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All specifications include district, month of interview, and climatic region-by-growing season fixed effects, and the same farmer controls as baseline regression in Table ???.

TABLE A.7—EFFECT OF TEMPERATURE ON FARM LABOR INPUTS, BY TYPE OF FARMER

Dep var:	Household Labor			Hired Labor
	(1) HH members in farm	(2) HH hours in farm	(3) Child labor	(4) ln(wage bill)
Average HDD x Owns livestock	0.019 (0.012)	0.032* (0.019)	0.024* (0.012)	-0.095 (0.061)
Average HDD x No livestock	0.014 (0.014)	0.016 (0.024)	0.029* (0.015)	-0.038 (0.055)
No. obs.	26,724	26,726	14,358	53,618
R-squared	0.513	0.361	0.315	0.247

*Notes:* Standard errors clustered at the district level (in parenthesis). Stars indicate statistical significance: \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All specifications include district, month of interview, and climatic region-by-growing season fixed effects, and the same farmer controls as baseline regression in Table ???. Sample restricted to interviews conducted during the growing season (i.e. October to March) in columns 1 and 2, since dependent variable is defined as work conducted over the past week. In column 3, we restrict the sample to households with children between the ages of 6 and 15.

TABLE A.8—EFFECT OF HDD ON LAND PRODUCTIVITY, OUTPUT AND LAND USE - BY BASELINE CLIMATE

	ln(output per ha) (1)	ln(total output) (2)	ln(area planted) (3)	Tubers % output (4)
(A) Average HDD x Hot areas	-0.126*** (0.040)	-0.066 (0.041)	0.039** (0.018)	0.009** (0.004)
(B) Average HDD x Cool areas	-0.228** (0.102)	0.038 (0.068)	0.221** (0.097)	0.050*** (0.019)
Diff. (B)-(A) p-value	0.305	0.130	0.054	0.026
No. obs	53,493	53,619	53,493	53,619
R-squared	0.336	0.348	0.443	0.527

*Notes:* Standard errors clustered at the district level (in parenthesis). Stars indicate statistical significance: \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All specifications include district, month of interview, and climatic region-by-growing season fixed effects, and the same farmer controls as baseline regression in Table ???. Cool areas = clusters with average growing season temperature in period 2007-2015 below the sample median (22.4°C). Hot areas = clusters with average growing season temperature in period 2007-2015 above the sample median.

TABLE A.9—EFFECT OF LAGGED HDD ON LAND PRODUCTIVITY

Dep. Variable:	ln(output/ha)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Average HDD in growing season t	-0.105** (0.042)	-0.106** (0.041)	-0.096** (0.038)	-0.092* (0.047)	-0.111*** (0.041)	-0.102*** (0.039)	-0.071** (0.034)	-0.077** (0.033)	-0.042 (0.039)
Average HDD in growing season t-1	-0.015 (0.025)								0.026 (0.033)
Average HDD in growing season t-2		-0.016 (0.027)							0.028 (0.039)
Average HDD in growing season t-3			-0.031 (0.028)						0.003 (0.039)
Average HDD in growing season t-4				-0.042 (0.031)					-0.043 (0.038)
Average HDD in growing season t-5					-0.007 (0.020)				0.014 (0.026)
Average HDD in growing season t-6						-0.036 (0.024)			-0.020 (0.027)
Average HDD in growing season t-7							-0.049* (0.028)		-0.045 (0.032)
Average HDD in growing season t-8								-0.044 (0.030)	-0.047 (0.040)
No. obs.	53,493	53,493	53,493	53,493	53,493	52,056	46,636	41,465	41,465
R-squared	0.335	0.335	0.335	0.335	0.335	0.332	0.333	0.330	0.330

Notes: Standard errors clustered at the district level (in parenthesis). Stars indicate statistical significance: \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All specifications include district, month of interview, and climatic region-by-growing season fixed effects, and the same farmer controls as baseline regression in Table ??.

TABLE A.10—EFFECT OF TEMPERATURE ON LAND USES, BY PROXIES OF MARKET DEVELOPMENT

	ln(area planted)					
	(1)	(2)	(3)	(4)	(5)	(6)
Average HDD	0.180*** (0.068)	0.075** (0.034)	0.279** (0.138)	0.211* (0.113)	0.150** (0.068)	0.573** (0.280)
Average HDD × $W$ (region level)	-0.297** (0.143)	-0.098 (0.142)	-0.438* (0.255)	-0.315 (0.224)	-0.631 (0.397)	-0.184* (0.098)
$W =$	% output sold	% land with registered title	% farmers hire workers	% use pesticides	% apply to agric. credit	ln(no. branches per 100,000 inhab. 2009)
Mean $W$	0.276	0.158	0.497	0.449	0.082	2.570
No. obs	53,493	53,493	53,493	53,493	53,493	53,493
R-squared	0.443	0.443	0.443	0.443	0.443	0.443

*Notes:* Standard errors clustered at the district level (in parenthesis). Stars indicate statistical significance: \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All specifications include district, month of interview, and climatic region-by-growing season fixed effects, and the same farmer controls as baseline regression in Table ???.  $W$  are proxies of market distortions calculated at region ( $n=24$ ) level. Data for constructing these measure comes from the ENAHO survey, except in columns 5 and 6 which were obtained from the National Agricultural Census 2007 and the Superintendencia de Banca, Seguros y AFP (SBS). Column 6 refers to the number if of branches of banks providing credit to farmers.

TABLE A.11—EFFECT OF TEMPERATURE ON HOUSEHOLD INCOME, CONSUMPTION AND POVERTY RATES

Sample:	ln(inc/capita)			ln(cons/capita)			Poor (Yes=1)		
	(1) All	(2) Coast	(3) Highlands	(4) All	(5) Coast	(6) Highlands	(7) All	(8) Coast	(9) Highlands
Average DD	0.023*** (0.004)	0.010 (0.012)	0.024*** (0.004)	0.021*** (0.004)	0.013 (0.014)	0.021*** (0.004)	-0.014*** (0.003)	-0.009 (0.012)	-0.013*** (0.003)
Average HDD	-0.017 (0.013)	-0.015 (0.013)	-0.008 (0.022)	-0.014 (0.010)	-0.016 (0.010)	0.001 (0.017)	0.003 (0.007)	0.009 (0.008)	-0.008 (0.015)
No. obs.	53,619	7,439	46,180	53,619	7,439	46,180	53,619	7,439	46,180
R-squared	0.380	0.388	0.335	0.452	0.451	0.416	0.264	0.282	0.244

Notes: Standard errors clustered at the district level (in parenthesis). Stars indicate statistical significance: \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All specifications include district, month of interview, and climatic region-by-growing season fixed effects, and the same farmer controls as baseline regression in Table ??.

TABLE A.12—PREDICTED EFFECTS OF TEMPERATURE AND PRECIPITATION ON AGRICULTURE UNDER TWO CLIMATE CHANGE SCENARIOS

	RCP 4.5			RCP 8.5		
	All (1)	Coast (2)	Highlands (3)	All (4)	Coast (5)	Highlands (6)
<i>A. Predicted change of temperature</i>						
$\Delta$ DD	2.001	0.820	2.196	4.385	1.565	4.850
$\Delta$ HDD	0.305	1.204	0.157	0.950	2.910	0.627
$\Delta$ Precipitation	0.910	0.137	1.038	0.122	-0.560	0.235
<i>B. Predicted effect on agriculture</i>						
$\Delta$ Yields (ln Y/T)	-0.058	-0.110	-0.050	0.093	-0.326	0.163
$\Delta$ Output (ln Y)	0.049	-0.066	0.068	0.089	-0.271	0.148
<i>C. Differences on estimate of damages</i>						
$\Delta$ yields - $\Delta$ output	-0.107	-0.044	-0.118	0.004	-0.055	0.015

*Notes:* Table presents predictions of the effect of increased temperatures on agriculture under two climate change scenarios (RCP 4.5 and 8.5). Predictions uses region-specific estimates of the effect of temperature and precipitation on yields and output from columns 1 and 3 in Table ??.