Extreme Temperatures and Field-Level Crop Quality, Yield, and Revenue

Sarah C. Smith & Timothy K.M. Beatty*

February 1, 2023

Abstract

Quality is central to agriculture because of its role in price determination and contractual arrangements. Yet prior work on the effect of weather and climate change on agriculture mostly focuses on the yields of rainfed, staple grains. In this article, we quantify the effect of temperature exposure on the revenue of profit-maximizing specialty crop growers through two pathways: quality and yield. We use proprietary data on quality, yield, price, and grower practices for thousands of fields contracted with a large tomato processor operating in California. In contrast to earlier work on irrigated crops, we find extreme temperatures negatively affect both yield and quality leading to reduced grower revenue in a setting where irrigation is the norm. While the yield effect dominates, failing to account for quality significantly underestimates the true effect of temperature exposure on revenue by up to 20%. We find that ignoring the common practice of selectively screening out low-quality products introduces bias in estimates of the effect of weather and climate change on yield and quality. A back-of-the-envelope counterfactual analysis finds that warming trends over the last three decades reduced grower revenue by an average of 2.4% per year.

Keywords: quality, agriculture, irrigation, climate change.

JEL codes: Q12, Q54, C23

*Department of Agricultural and Resource Economics, University of California, Davis.

1 Introduction

The value of every agricultural product depends on its quality. Grain, meat, and milk are graded according to USDA quality standards, and fresh produce is sorted by size, color, and defects. Products grown under contract often face quality incentives and low quality can violate contractual obligations or make products unmarketable. But work quantifying the effects of extreme weather and climate change on agriculture focuses almost exclusively on yield (see Carter et al. (2018) or McCarl & Hertel (2018) for summaries). Ignoring quality may bias estimates of the impact of weather and climate change on agricultural productivity and farm income.

We ask three related questions: Has historical weather impacted the incomes of specialty crop producers through both yield and quality? What are the relative magnitudes of the yield and quality effects? To what extent does failing to account for quality bias estimates of weather's impact on farm revenue? We answer these questions using 12,000 field-level observations of processing tomato yield, quality, and grower practices from across California between 2011 and 2020. These data are generated by a large tomato processor for the purposes of contracting and payment and capture the profit-maximizing behavior of hundreds of farmers. We use standard gridded weather data from PRISM and the well-established panel specification (Deschênes & Greenstone (2007); Schlenker & Roberts (2009)) to facilitate comparison to earlier work. Finally, we implement a restricted cubic spline to flexibly estimate the effect of temperature exposure on yield and quality.

Despite being irrigated, we find that extreme weather conditions affect the revenue of growers. Yield responds negatively to exposure to hot temperatures and, to a lesser extent, cool temperatures. An additional 24 hours of exposure above 30°C causes yields to decrease by up to 1.8% relative to 24 hours of average temperatures. Further, quality declines with exposure to hot temperatures that causes growers to receive a lower price. Taken as a whole, we find that, relative to 24 hours of average temperatures, exposure to temperatures in excess of 30°C decreases revenue up to 2.3%. Exposure to cool temperatures below 10°C causes a statistically significant, but smaller, decrease in revenue. Failing to account for quality effects would bias downward the effect of exposure on revenue by up to 20%.

This paper is among the first to document the effect of weather and climate change on agricultural product quality. Kawasaki & Uchida (2016), Dalhaus et al. (2020), and Ramsey et al. (2020) all find negative, economically important, effects of weather on grower income with quality being a key pathway. Our setting and novel data offer advantages relative to earlier work.

First, quality is precisely measured, rather than inferred, and linked to price using a schedule of bonuses and deductions established prior to planting. The Californian processing tomato industry has mandatory quality testing by an independent third party, so neither grower nor processor can accidentally or intentionally misstate quality. Growers are paid a price per ton that depends on observed quality, which introduces significant variation in price of plus or minus 20%. For each field-year observation, we calculate an economically meaningful quality index with weights equal to each quality attribute's effect on price established in observable pre-planting contract terms common across all growers.

Second, in settings where quality is important, researchers typically observe the subset of production that exceeds a quality threshold and is selected to be graded or harvested. This selection biases unconditional measures of both quality and yield. As detailed below, selection in the Californian processing tomato industry is rare and we observe the little selection that occurs. We simulate the sample selection problem common in other settings and find that selection bias affects estimates of the effect of weather on quality, yield, and revenue.

Taking a step back, prior work has mostly focused on staple crop yields (Schlenker & Roberts (2009); Lobell et al. (2011b); Tack et al. (2015); Gammans et al. (2017); Shew et al. (2020), Schmitt et al. (2022)). In contrast, specialty crops are understudied despite making up 40% of the total value of U.S. crops (USDA NASS, 2017). By focusing on an irrigated specialty crop, we extend a literature that has largely focused on rain-fed staple crops. Specialty crops have distinct production functions and likely respond differently to weather shocks than rain-fed field crops. Prior work finds that irrigated water application essentially eliminates the negative effect of extreme heat on wheat yields (Tack et al., 2017), corn yields (Shaw et al. (2014); Carter et al. (2016)), and agricultural total factor productivity (Ortiz-Bobea et al., 2018). But in a setting where irrigation has long been the rule rather than the exception, we find both yield and quality are affected by exposure to hot temperatures, leading to lower grower revenue.

Finally, we explore how historical climate trends have already contributed to declining grower revenue through effects on quality and yield. California's climate has become hotter and drier, with rain becoming less frequent and more variable from year-to-year (Zhang et al., 2021). A back-of-the-envelope counterfactual analysis compares outcomes using actual temperatures with predicted outcomes had temperatures remained at cooler, historical levels. We find that yield, quality, and revenue have all been negatively affected by warming trends. On average, annual revenue was 2.4% lower in 2011-2020 than it would have been had temperatures stayed at 1991-2010 levels. The effect worsens over the decade – in the last year in our sample (2020) we estimate revenue would have been 5.3% higher had temperature exposure remained at 1991-2010 levels. This is non-negligible loss for farmers who already operate on tight margins.

2 The Setting

Tomatoes are the third most produced fruit or vegetable globally (behind potatoes and cassava) (FAO, 2022) and the second most consumed in the United States (USDA ERS, 2020). Tomatoes can be either consumed fresh or processed into paste, ketchup, or a canned product. They contain nutrients like vitamin E, potassium and lycopene that are important to human health but often under consumed (Wu et al., 2022).

Tomatoes destined for processing (henceforth processing tomatoes) are specific varieties, distinct from fresh tomatoes, bred and grown to enhance qualities desirable for processing into paste or for canning. California's $1 billion processing tomatoes industry produces more than 90% of U.S. processing tomato production (California Department of Food and Agriculture, 2019). In California, processing tomatoes are planted between February and June to facilitate continuous harvesting between July and October. They are mostly grown in outdoor fields in the San Joaquin and Sacramento Valleys (California Department of Food and Agriculture, 2019). Processing tomatoes are a warm-season crop – during the Californian growing season, maximum temperatures average around 30°C and precipitation is scant. Growers irrigate to ensure crops receive enough water and can tolerate high temperatures during the height of summer (Hartz et al., 2008).

Our data are from a large tomato processor that purchases processing tomatoes under contract from growers in the San Joaquin Valley, Sacramento Valley, and Central Coast regions of California. Processors contract with growers because input quality is crucial to produce consistent and high-quality output. They incentivize growers by paying a price that depends on the quality of tomatoes delivered. Contracts are negotiated between individual processors and the California Tomato Growers Association (CTGA) on behalf of all growers.

Negotiations establish each processor's season's base price, quality adjustments, and bonuses, which processors then offer to growers on a take it or leave it basis.

Table 1 summarizes the eight quality attributes we observe and their effect on price. The processor deducts a percentage of the base price for the presence of defects (mold, green tomatoes, worms, material other than tomatoes (MOT) and limited use (LU) tomatoes). The processor has a incentive program where growers receive a bonus (or penalty) if the brix (soluble solids or sugar content) of delivered tomatoes is more (or less) than the average for the same variety in the same county in which they were grown. Quality adjustments are proportional to the quality achieved by growers as measured by Processing Tomato Advisory Board (PTAB). Finally, the processor values staggered harvesting and delivery to minimize bottlenecks at processing facilities. Producers receive a bonus for delivering tomatoes early or late in the season. Quality incentives are economically important and introduce price variation of plus or minus 20% relative to average prices.

Prior work uses observational data to analyze effects of weather and climate on different agricultural products, including rice (Kawasaki & Uchida, 2016), apples (Dalhaus et al., 2020), peanuts (Ramsey et al., 2020), and wine (see Ashenfelter & Storchmann (2016)). However, studying quality with observational data presents identification challenges. Data are typically only available for the subset of production that growers choose to market because it exceeds an implicit or explicit quality threshold. This has two consequences. First, observations of quality are biased measures of quality at harvest. Kawasaki & Uchida (2016) find low quality rice is less likely to undergo the costly process of being graded and more likely to be withheld for self-consumption or sale on informal markets (such as to local households or for animal feed). Quality measures for hand-harvested crops, such as berries, stone fruit, apples, and leafy greens, may also be biased. Harvesting guidelines (e.g. instructions to only pick high-quality produce), adequate access to labor, and how labor is paid can all affect the observed quality of hand-harvested crops (e.g. Hill & Beatty (2020)).

Second, yield estimates will be biased if weather affects quality and quantities are measured after farmers screen out low-quality products. The result is quality effects being attributed to yield and biased estimates of the effect of weather on both yield and quality. For example, wet conditions around harvest can result in quality problems in grain that may cause growers to withhold low quality grain from sale. If this withheld grain is not counted towards production totals, an outside observer will incorrectly ascribe the effects of wet con-

Table 1: Summary of quality measures

Quality attribute	Explanation	Effect on price
Brix	Brix is a measure of soluble solids or sugar content	Bonus (penalty) if brix is more (less) than the average for the same variety in the same county
Limited use (LU) percent	Tomatoes that are soft, split or squashed and have limited processing use	Deduction from base price in proportion to percentage
Material other than tomatoes (MOT) percent	Mainly dirt and sometimes vines	Deduction from base price in proportion to percentage
Green percent	Tomatoes that are unripe	Deduction from base price in proportion to percentage
Mold percent		Deduction from base price in proportion to percentage
Worm percent		Deduction from base price in proportion to percentage
Color score	Red-ripe fruit are given a color score where low scores indicate a better color	No effect
pH	Higher pH usually indicates the fruit is more ripe	No effect

ditions around harvest to yield instead of quality. The overall effect on revenue or profit estimates is ambiguous since bias in yield is at least partially offset by bias in quality. This concern is not limited to papers focusing on quality but also applies to *all* studies estimating the effect of weather on yield or revenue that don't address selection.

There is limited opportunity to selectively harvest and grade processing tomatoes. Unlike many specialty crops, processing tomatoes are mechanically harvested[1]. The processor closely manages harvesting logistics because operating processing plants near full capacity is key to profitability. Almost all Californian processing tomato production is grown under a contract between a grower and processor (USDA NASS, 2021b). Contracts are written for specific fields and the processor typically harvests and transports tomatoes to the processing plant. As a consequence, growers have little opportunity to strategically sort prior to sale.

Further, each and every truckload of processing tomatoes in California undergoes mandatory grading, including quality measurement, at an independent state inspection station prior to delivery. Administered by PTAB, the California Processing Tomato Inspection Program was established in 1987 to create and uphold quality standards for Californian processing tomatoes. Quality observations are shared with processors and growers and are used for price determination.

Finally, the processor engages in a small degree of sorting in which foreign matter and some very low quality tomatoes are deducted from the quantity used for payment. However, we are able to fully observe this sorting. The processor records total tons prior to sorting as well as paid tons after sorting. In sum, the institutional setting means quality and yield observations accurately reflect conditions in the field and do not suffer from selection problems common in most settings.

The processing tomato agronomic literature (Hartz et al., 2008) finds that maximum temperatures between 25°C and 35°C are ideal for vegetative growth, plant development, and fruit set, so long as plants have sufficient access to water. Hot temperatures without adequate moisture will cause tomato plants to become stressed, affecting yield and quality. Minimum temperatures below 10°C slow development and also affect quality. The extent of damage caused by extreme temperature depends on its timing in the phenological cycle. Cool temperatures at the beginning of the growing season are believed to benefit brix, while

[1] See Just & Chern (1980) for details on the introduction and widespread adoption of mechanical harvesters in the Californian processing tomato industry during the 1960s.

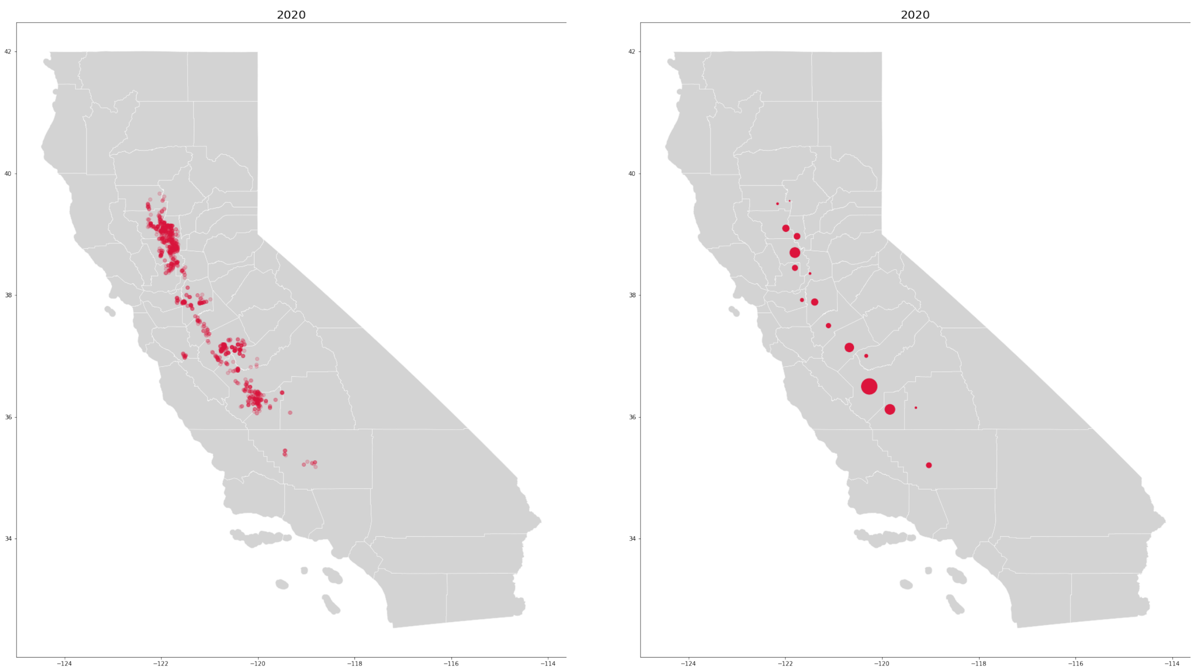
hot temperatures at the beginning and end of the growing season worsen yield and LU share respectively (Personal correspondence with the processor, 2020).

Lobell et al. (2007) find that maximum temperatures in April and June explain 58% of yield variability in Californian processing tomatoes between 1980 and 2003. Hot temperatures benefit seedling growth during April. However, yields decrease when exposed to maximum temperatures above 32°C in June. Marklein et al. (2020) estimate that 34–87% of land on which tomatoes are grown in California will no longer be suitable by mid-century due to increasing summer temperatures. This may be an upper bound as they assume hot temperatures translate directly to heat stress. Tomato plants can withstand considerable heat so long as they have access to water (Hartz et al., 2008). Cammarano et al. (2022) project a decrease in global processing tomato production by 2050, driven by temperatures rising above the optimal threshold (28°C) in Californian and Italian growing regions.

When using proprietary data, there is a tradeoff between internal and external validity. In this case, we obtain a level of detail not available in public data. Data are not from surveys but rather from administrative records of every field contracting with the processor between 2011 and 2020. Our data is at the field-year level and contains a range of information about fields so that we can observe and control for field-specific factors. Detail enhances internal validity but may limit external validity.

One concern is that the hundreds of growers in this proprietary dataset are not representative of the broader processing tomato industry. Several factors argue against this concern. The fields in our sample are geographically dispersed across 18 counties in California and closely match patterns of production locations in California (Figure 1). Field-level yields averaged to the county level are 10% higher than county yields reported by NASS. However, the series are highly correlated. That said, it is always possible that there is selection on unobservables that may affect the external validity of our results.

Another concern is that our results are only relevant for processing tomatoes. However, almost all agricultural producers are paid a price that depends on quality and face similar incentives to processing tomato growers, albeit with different contract incentives. As a result, findings from this case study offer insight into the effects of weather on yield and quality in other agricultural settings.



(a) Location of fields contracted with processor in 2020

(b) Californian production in 2020, by county (USDA NASS, 2021a)

Figure 1: Californian processing tomato production

3 Data

3.1 Field-level dataset

As described above, we use data on all tomatoes grown and sold under contract to a large tomato processor in California, between 2011 and 2020. Observations are at the field-year level ($n = 11,926$) and include information on field acreage, variety, total tons, paid tons, yield, quality attributes, and the latitude and longitude of the field centroid. Observations of quality attributes and tonnage are from PTAB mandatory testing that occurs prior to delivery. The data also includes information about the growing practices, including planting and harvesting dates, irrigation technology, and the crop previously planted on a field. Field-level observations are linked to an unbalanced panel of 438 growers and 247 grower groups. The grower group identifier links growers within the same network or organization[2].

3.2 Contract terms

We also observe pricing terms negotiated between the processor and the California Tomato Growers Association for each year. All growers are offered the same contract in a year and do not negotiate individual pricing terms with the processor. Every year, a base price is established at the start of the growing season that reflects current and expected market conditions. Contracts also establish bonuses and deductions. As summarized in Table 1, there are eight measures of quality of which six are linked to price bonuses or deductions. In addition, growers receive an early (late) season bonus if they deliver tomatoes at the beginning (end) of the growing season.

3.3 Outcomes

Yield for field i in year t is defined as total tons divided by acres:

$$\text{yield}_{it} = \frac{\text{total tons}_{it}}{\text{acres}_{it}} \tag{1}$$

Next, we calculate price for each field-year observation using observed quality and the schedule of quality bonuses and deductions established at the beginning of the growing

[2]An example of a grower group is four children dividing a family farm. Each child would have a distinct grower ids and share a common grower group id.

season. Processing tomatoes are used to produce a storable commodity and therefore base prices are temporally correlated. For example, a negative production shock in year $t-1$ will induce processors to draw down stocks and offer a higher base price in year t. This could introduce endogeneity into our panel model. Instead, we isolate variation in price driven by quality by applying the estimated adjustments to the 10-year average base price. This removes common price movements driven by potentially endogenous market shocks while preserving common and individual quality shocks. This "quality-adjusted price" is effectively a quality index with weights equal to each quality attribute's effect on price.

$$\text{price}_{it}^{\text{quality adjust}} = \overline{\text{base price}} \times (1 - \text{deducts}_{it}) + \text{bonus}_{it} \tag{2}$$

where $\overline{\text{base price}}$ is the average base price over 10 year sample and deducts_{it} and bonus_{it} are adjustments that depend on observed quality and date of delivery of tomatoes from field i in year t.

Finally, we estimate field-level revenue per acre (henceforth revenue) by multiplying quality-adjusted price by paid tons and dividing by acreage. The processor doesn't count some flawed tonnage towards the quantity on which producers are paid. The revenue estimate uses paid tons is which on average 7% smaller than total tons used in the yield calculation.

$$\text{revenue}_{it} = \frac{\text{paid tons}_{it}}{\text{acres}_{it}} \times \text{price}_{it}^{\text{quality adjust}} \tag{3}$$

3.4 Weather data

We collect weather data from PRISM (PRISM Climate Group, Oregon State University, 2020), which publishes daily temperature and precipitation data interpolated to 4km grids for the whole time span. We match weather data to each field-level observation by identifying the PRISM grid in which the field centroid falls. In Section 4.1, we explain how we translate daily observations of temperatures into measures of temperature exposure for each field-year observation.

3.5 Control variables

We also gather data on several controls. We source information on tomato varieties from AgSeeds (AgSeeds, 2020), which includes key attributes and use categories for each of the 159 varieties in the processor dataset. Finally, we match each field to its soil type in the National Cooperative Soil Survey (NRCS USDA, 2020).

Table 2: Summary statistics

	units	count	mean	sd	min	max
Area	acres	11926	57.31	43.47	0.30	323.20
Growing days	no.	11926	133.71	9.50	96.00	175.00
Yield	tons per acre	11926	52.49	13.32	6.70	99.86
Brix		11926	5.08	0.49	0.00	7.17
Limited use percent		11926	1.42	1.17	0.00	13.05
Material other than tomatoes percent		11926	1.64	1.23	0.00	13.33
Green percent		11926	3.02	2.33	0.00	24.29
Mold percent		11926	1.66	1.92	0.00	27.53
Worm percent		11926	0.00	0.01	0.00	0.50
Color score		11926	20.98	1.75	0.00	34.43
pH		11926	4.41	0.10	0.00	4.77
Average minimum temperature	° C	11926	13.79	1.08	9.81	17.61
Average maximum temperature	° C	11926	31.19	1.62	24.50	35.39
Total precipitation	mm	11926	24.45	25.80	0.00	198.85
Alluvium	prop.	11926	0.96	0.20	0.00	1.00
Eolian	prop.	11926	0.00	0.04	0.00	1.00
Organic material	prop.	11926	0.03	0.17	0.00	1.00
Lacustrine	prop.	11926	0.00	0.03	0.00	1.00
Residuum	prop.	11926	0.01	0.08	0.00	1.00
Drip irrigation	prop.	11926	0.75	0.43	0.00	1.00
Furrow irrigation	prop.	11926	0.13	0.33	0.00	1.00
Missing irrigation tech.	prop.	11926	0.10	0.31	0.00	1.00
Sprinkler irrigation	prop.	11926	0.02	0.14	0.00	1.00
Extended field storage variety	prop.	11926	0.55	0.50	0.00	1.00
Tomato spotted wilt resistant	prop.	11926	0.45	0.50	0.00	1.00
Fusarium Wilt resistant	prop.	11926	0.15	0.36	0.00	1.00
Powdery Mildew resistant	prop.	11926	0.04	0.19	0.00	1.00
Fusarium Crown Rot resistant	prop.	11926	0.00	0.05	0.00	1.00
Bacterial Spot resistant	prop.	11926	0.00	0.05	0.00	1.00
High solids	prop.	11926	0.06	0.23	0.00	1.00
High yield	prop.	11926	0.06	0.23	0.00	1.00
Early	prop.	11926	0.13	0.33	0.00	1.00
Thin	prop.	11926	0.14	0.35	0.00	1.00
Intermediate	prop.	11926	0.23	0.42	0.00	1.00
Thick	prop.	11926	0.58	0.49	0.00	1.00
Pear-shaped	prop.	11926	0.01	0.08	0.00	1.00

4 Methods

The aim is to estimate the marginal effect of weather on processing tomato yield, quality, and revenue. Since precipitation is scant and growers control the amount of water applied through irrigation, we focus on the marginal effect of temperature exposure and include precipitation as a control.

Since we are interested in the direct and indirect effects of temperature exposure, we take care not to introduce bad controls – variables that are an outcome of the experiment at hand. A key example is irrigation volumes, which is itself a function of temperature and also affects the outcome variable. Including irrigation volumes would bias the coefficient on temperature exposure because some explanatory power of temperature may be incorrectly assigned to irrigation volume.

We follow the standard approach proposed by Schlenker & Roberts (2009) and adopted by Gammans et al. (2017), Shew et al. (2020) among others. Ortiz-Bobea (2021) provides a comprehensive summary. We take an off-the-shelf econometric approach to emphasize that results are driven by differences in focus and setting rather than differences in methodology. We estimate:

$$y_{it} = \int_h f(h)\phi_{it}(h)d(h) + \delta z_{it} + \alpha_{g(i)} + \psi(t) + \epsilon_{it} \tag{4}$$

where y_{it} is a log-transformed outcome variable (yield, quality, and revenue) in field i in year t, $\alpha_{g(i)}$ is a grower fixed effect, and $\psi(t)$ is a quadratic year trend. The first term characterizes the relationship between temperature exposure and the outcome variable, where $f(h)$ is the marginal effect of temperature h and $\phi_{it}(h)$ is the growing-season density at h for field i in year t. This continuous representation is not tractable for estimation, but can be approximated using the restricted cubic spline specification detailed in Section 4.2. Field-year specific control variables z_{it} include variety attributes (extended field storage, various disease resistant traits, high solids, high yield, early, thin, intermediate, thick, and pear-shaped), irrigation technology (drip, sprinkler, furrow), soil type (alluvium, eolian, organic material, lacustrine, and residuum), growing season precipitation, a dummy for planting week, and the difference between actual growing days and estimated growing days specified by the seed manufacturer.

As with any annual crop, growers influence expected weather through their choice of planting date. In our setting, tomatoes planted earlier are exposed to cooler temperatures on average, whereas tomatoes planted later are exposed to hotter temperatures on average. The implication is that weather is endogenous and coefficients on temperature exposure may be biased. Overall, this is less of a concern in our setting because planting date is determined by the processor in consultation with the grower. Grower's potentially endogenous preferences over planting dates is mediated by the processors preferences for staggered planting to facilitate continuous processing at harvest. We include dummies for planting week-of-year to account for this endogeneity in weather. Failing to control for planting date leads to bias in quality results, but yield and revenue results are largely unchanged (see Appendix A).

The error term ϵ_{it} is likely heteroskedastic, spatially correlated, and temporally correlated within similar growers over time. We use heteroskedatic robust standard errors two-way clustered by grower group[3] and county by year. We cluster at the grower group level to account for possible dependence between growers within the same grower group. We do not cluster by year as we only observe 10 years of data. Even in a multiway cluster, too few clusters in any cluster group will result in incorrect statistical inference (Cameron et al., 2011). Instead, we cluster county by year to account for spatial correlation. Results are robust to using spatial heteroscedasticity and autocorrelation consistent errors that allow for spatial correlation between nearby fields and serial correlation in panel data (see Appendix B).

We include grower fixed effects $\alpha_{g(i)}$, where individual growers can be associated with multiple fields – on average, each grower is associated with 27 field-year observations. This controls for time-invariant, grower-specific factors that may be related to outcome or explanatory variables. Our preferred specification uses grower fixed effects as regular crop rotation results in an unbalanced panel of field-year observations. Note that results are robust to using field fixed effects in place of grower fixed effects (see Appendix C). Results are also robust to replacing quadratic year trends with (a) linear year trend, and (b) county-specific quadratic trends[4].

[3]Recall that the grower group identifier links growers within the same network or organization.

[4]We refrain from using year fixed effects. Given our focus on a relatively small geographic region, year fixed effects absorb most of the variation in weather used to identify the effects of interest.

4.1 Estimating temperature exposure

We translate daily observations of minimum and maximum temperatures to a measure of temperature exposure for each field-year observation. For each day of the growing season[5], we estimate how many hours are spent in 1°C temperature intervals by fitting a sinusoidal curve between each day's minimum and maximum temperature. We then sum over days to estimate the number of days spent in each 1°C temperature interval during the entire growing season. The result is x_{it}, a 1-by-J vector of temperature exposures for field i during the growing season in year t, where J is the number of temperature bins. In our setting, we bin temperatures from 5°C to 41°C, so $J = 37$.[6]

$$x_{it} = \begin{pmatrix} x_{it,5} & x_{it,6} & \dots & x_{it,40} \end{pmatrix} \tag{5}$$

where $x_{it,j}$ is the number of days spent between j°C and $(j+1)$°C during the growing season in year t.

This approach has several advantages. First, it addresses the empirical challenge of mixed frequency between regressor and outcome variables. We have many daily observations of minimum and maximum temperatures to match with one annual observation of an outcome variable. Averaging daily temperatures across the growing season would mask differences in exposure to extreme temperatures. The second advantage of this approach is that it preserves the temperature distribution. This allows us to uncover the marginal effect of exposure to different temperatures.

4.2 Restricted cubic spline specification

Next, we choose a functional form to characterize the relationship between outcome variables and temperature exposure. The agronomic literature suggests that mid-range temperatures are ideal for yield and quality outcomes, but these may be damaged by hot or cool temperatures if exposure occurs during key stages of the plant's growth cycle. The implication is that the relationship between temperature exposure and outcome variables is nonlinear.

To capture nonlinearity in the response of the outcome variable to temperature, we estimate a restricted cubic spline model (otherwise known as a natural cubic spline). Restricted

[5] The growing season starts on the day of planting and ends on the last day of harvest

[6] Temperatures range from −1°C to 45°C. We aggregate temperature exposure below 5°C and above 41°C to avoid bins with little exposure.

cubic splines offers several benefits over alternative methods of estimating nonlinear temperature effects that have led to its increase in popularity (Berry et al. (2014); D'Agostino & Schlenker (2016); Ortiz-Bobea et al. (2019); Blanc & Schlenker (2020); Bucheli et al. (2022)). First, it offers smooth parsimonious semiparametric estimation without needing to define critical temperature thresholds. Second, it imposes a restriction that its tails (i.e. before the first knot and after the last knot) are linear. This reduces overfitting in the data-sparse tails of the temperature distribution, an issue with the polynomial and cubic spline functional forms. We estimate a piecewise linear degree day model as a robustness check. Overall, results from the two specifications – piecewise linear degree day model and restricted cubic spline model – are consistent in terms of magnitude and significance (see Appendix D).

We identify $K = 4$ temperatures that split the distribution of temperature exposure by interval into quintiles. This accounts for the fact that relatively less time is spent at extreme temperatures. Unlike other models, knot placement does not strongly influence the cubic spline results because the marginal effect of exposure is allowed to vary smoothly between knots.

Next, we introduce a basis matrix B and a vector of coefficients Γ. B is J-by-P while Γ P-by-1, where P is the number of parameters on temperature to be estimated, and is directly related to the number of knots K. A restricted cubic spline with $K = 4$ knots results in $P = 3$, which is smaller than the number of temperature exposure bins $J = 37$. This is an advantage of the spline model – it reduces the dimensionality while still allowing for flexible semiparametric estimation. The derivation of the basis matrix B that corresponds to the restricted cubic spline is shown in Appendix E.

Under these assumptions, we can write Equation 4 as:

$$y_{it} = x_{it}B\Gamma + \delta z_{it} + \alpha_{g(i)} + \psi(t) + \epsilon_{it} \tag{6}$$

Stacking n observations across fields and years gives Equation 6 in matrix notation:

$$Y = XB\Gamma + \delta Z + \alpha + \psi + \epsilon \tag{7}$$

where X is a n-by-J matrix of temperature exposures, and Y, δZ, α, ψ and ϵ are n-by-1 vectors of outcomes, controls, grower fixed effects, quadratic time trends, and errors respectively.

After estimation, we recover the marginal effect of temperature exposure evaluated at each interval. We pre-multiply the vector of estimated coefficients $\hat{\Gamma}$ by the corresponding B matrix. The resulting J-by-1 vector $\hat{\beta}$ is the marginal effect of spending one additional day at each temperature bin $j = 1, ..., J$.

$$\underset{J\times 1}{\hat{\beta}} = \underset{J\times P}{B} \times \underset{P\times 1}{\hat{\Gamma}} \tag{8}$$

Last, we derive an estimate of the variance-covariance matrix for $\hat{\beta}$:

$$\underset{J\times J}{\widehat{var}(\hat{\beta})} = \underset{J\times P}{B} \times \underset{P\times P}{\widehat{var}(\hat{\Gamma})} \times \underset{P\times J}{B'} \tag{9}$$

5 Results

Figure 2 contains our main results. In each graph, the top graph shows the effect of an additional 24 hours spent at a given temperature interval on the outcome variable relative to 24 hours spent at 26°C. This temperature has the greatest exposure in our sample and represents average temperatures. The 95% confidence intervals account for the possibility of heteroskedasticity, spatial correlation, and temporal correlation in the errors. The gray vertical lines show the position of the knots. The histogram at the bottom of the frame shows the average exposure to each temperature interval during the growing season across all fields in all years.

For yield, we find that the optimal temperature is around 28°C. Exposure to temperatures above 35°C leads to significantly lower yields. An additional 24 hours of exposure to 40°C decreases yield by almost 2% on average relative to 24 hours of average temperatures. Exposure to cold temperatures below 10°C causes a small but significant decline in yield by 0.7% relative to 24 hours at average temperatures.

We isolate variation in price driven by quality by applying quality adjustments to the 10-year average base price. For quality, we find that the optimal temperature is around 20°C. Quality declines with exposure to hot conditions. An additional 24 hours of exposure above 30°C causes quality to drop by up to 0.2% relative to 24 hours spent at average temperatures.

Revenue is maximized with exposure to temperatures around 27°C. An additional 24 hours of exposure to 40°C decreases revenue by 2.3% compared to 24 hours of average temperatures. This is expected since both yield and quality respond negatively to hot conditions. Exposure to cool temperatures below 10°C causes a smaller but significant decrease in revenue by almost 1% relative to 24 hours at average temperatures.

5.1 Decomposition

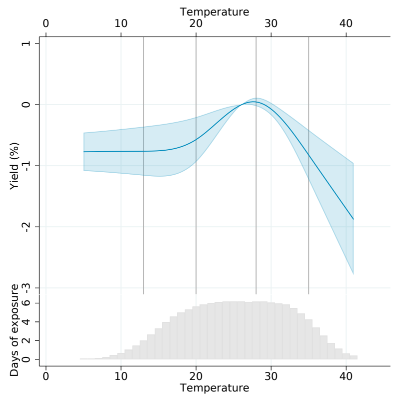
We find that temperature exposure significantly affects both yield and quality. However, the relative importance of each pathway is unclear a priori. Here, we decompose the effect of temperature exposure on revenue per acre (total effect) into the effect on revenue driven by yield (yield effect) and effect on revenue driven by quality (quality effect) as shown in Equation 10. This allows us to answer two questions. What is the relative importance of the yield and quality effects? And, would estimates be biased if quality were omitted?

$$\text{total effect} = \text{yield effect} + \text{quality effect} \tag{10}$$

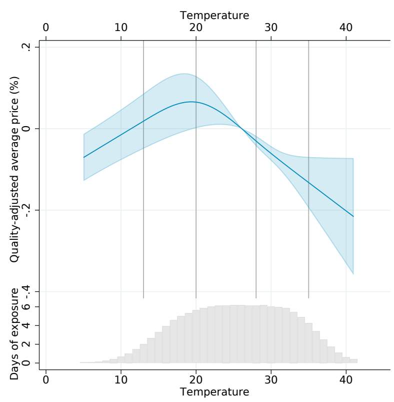
The total effect, which captures both yield and quality pathways, is equal to the effect of temperature exposure on revenue estimated above. Since we use log-transformed variables, the yield effect is similarly equal to the effect of exposure on yield estimated above. The quality effect, however, is slightly larger than the effect of exposure on quality index shown above. Recall revenue is a function of paid tons – the processor won't pay producers for some poor quality tonnage. Therefore, quality can affect revenue via (a) changes in price captured in the quality index, and (b) changes in paid tonnage. By rearranging Equation 10 and estimating the quality effect as the difference between the yield effect and total effect in Equation 11, we capture the effect of temperature on both price and unpaid tons.

$$\begin{aligned} \text{quality effect} &= \text{total effect} - \text{yield effect} \\ &= \hat{\beta}^{\ln(\text{revenue per acre})} - \hat{\beta}^{\ln(\text{yield})} \end{aligned} \tag{11}$$

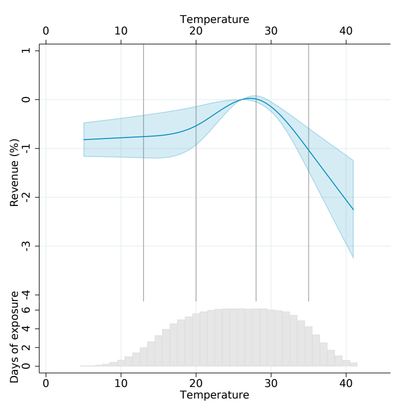
As shown in Figure 3, while the yield effect dominates grower revenue, quality also plays an important role. Without access to data on quality, a researcher can only recover the yield effect: an additional 24 hours of exposure to 40°C decreases revenue by 1.8% compared to 24 hours of average temperatures. This underestimates the effect of exposure on revenue by up to 0.5 percentage points, or 20% of the point estimate at 40°C. Failing to account for quality's effect on revenue biases estimates of temperature on revenue.



(a) Yield



(b) Quality



(c) Revenue per acre

Figure 2: Restricted cubic spline results

For each figure, the graph at the top of the frame shows the effect of an additional 24 hours spent at a given temperature interval on the outcome variable relative to 24 hours spent at 26°C. The histogram at the bottom of the frame shows the average exposure to each temperature interval during the growing season across all fields in all years.

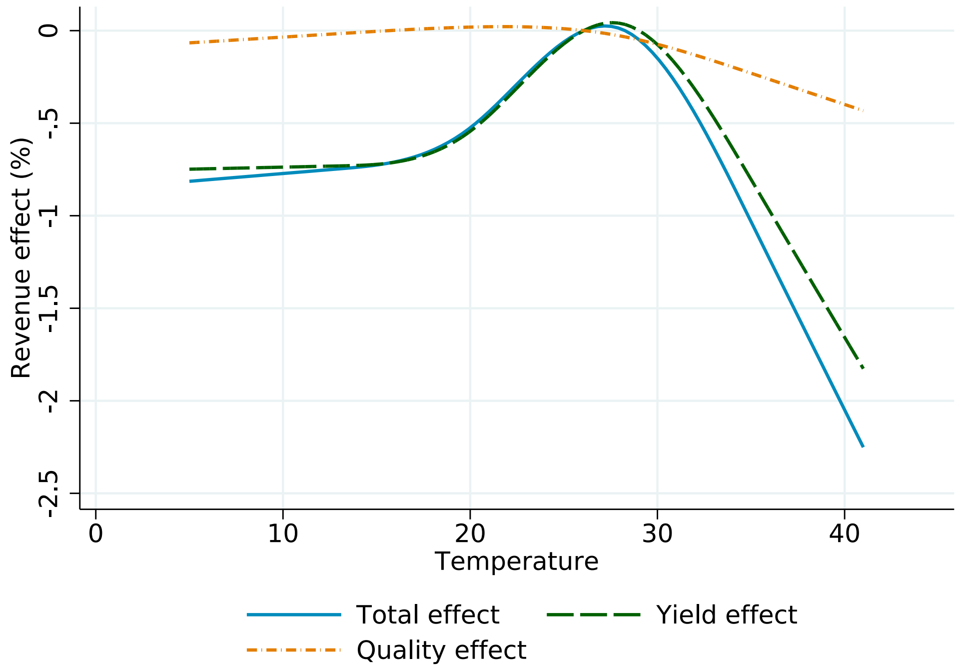


Figure 3: Decomposing the effect of exposure on revenue into the yield effect and quality effect

5.2 Selection

A benefit of our setting is that we observe the small amount of selection that is occurring. Recall from Section 2 that observations of quality and yield reflect conditions in the field because of mandatory grading, as well as contracting and harvesting practices. Moreover, the processor engages in sorting in which some defective tonnage isn't counted towards the quantity on which producers are paid. We are able to observe sorting through differences in total tons (before sorting) and paid tons (after sorting) – paid tons is on average 7% smaller than total tons. This gives us the opportunity to simulate selection and consider the consequences for the resulting estimates.

To simulate selection, we assume that the researcher only observes yield and quality of the paid tons i.e. a portion of total tons is now unobserved. Consistent with the processor sorting defective tonnage, we assume that unobserved tons had defects. Yield under selection is artificially reduced because it is calculated using a smaller tonnage (paid tons) than what is actually harvested from the field (total tons). Quality under selection is artificially improved

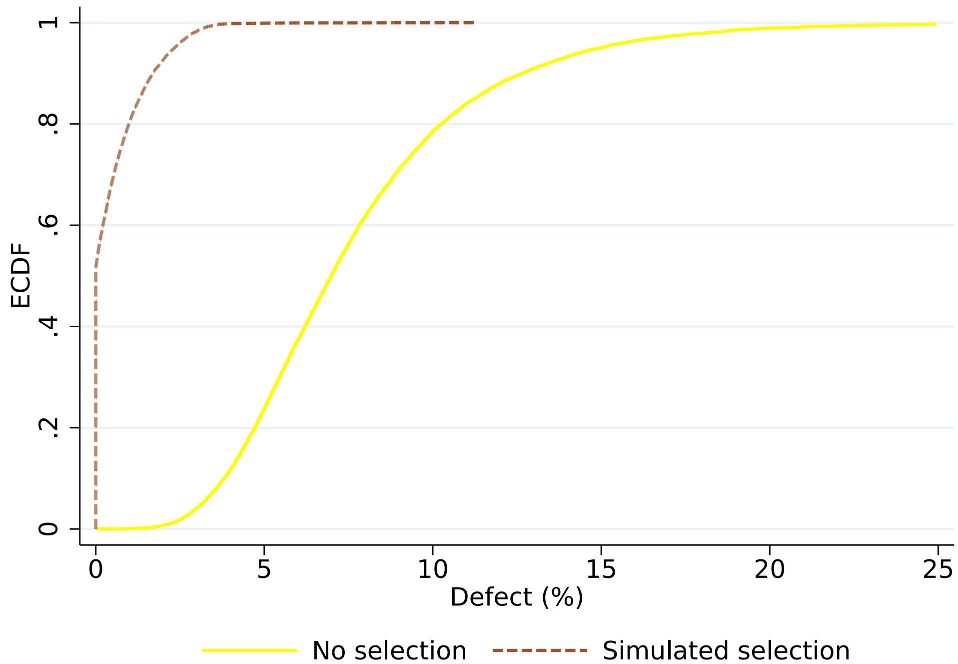


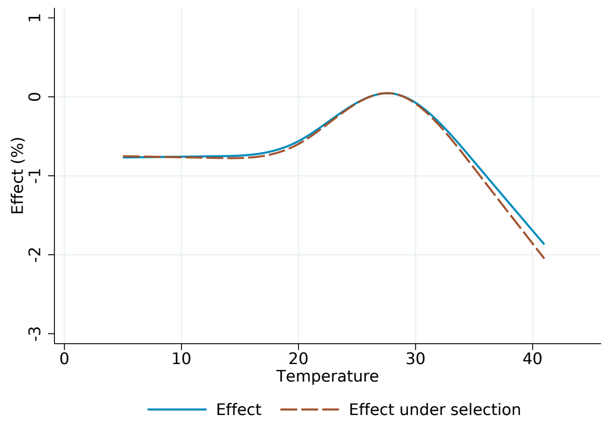
Figure 4: Empirical CDF of defects

because some tonnage with defects is no longer observed. Figure 4 compares the actual distribution of defects to its distribution when we simulate selection.

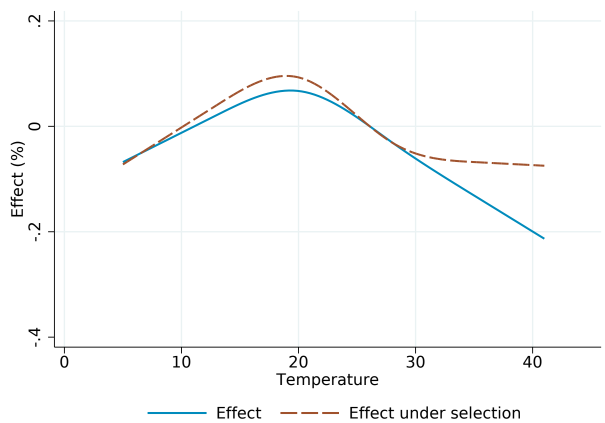
Our first hypothesis is that the effect of exposure to high temperatures on quality will be biased towards zero under selection. This follows from the result that high temperatures negatively affect quality, and the assumption that lower quality products are being withheld. Our second prediction is that the negative effect on quality will be incorrectly assigned to yield, causing the negative yield effect to increase in magnitude.

We compare our preferred specification without selection and an estimate under "simulated selection." Consistent with our hypothesis, the effect of exposure on yield is biased upwards by up to 10%. The effect of high temperatures on quality is indeed biased towards zero under selection by up to 66%. The upward bias in yield partially offsets the downward bias in quality. However, bias still remains and the effect of exposure to high temperatures on revenue is attenuated by up to 6%. Our results suggest that estimates will be biased in settings with selection but that the magnitude will depend on (a) the actual effect of weather on quality, and (b) how much selection is occurring.

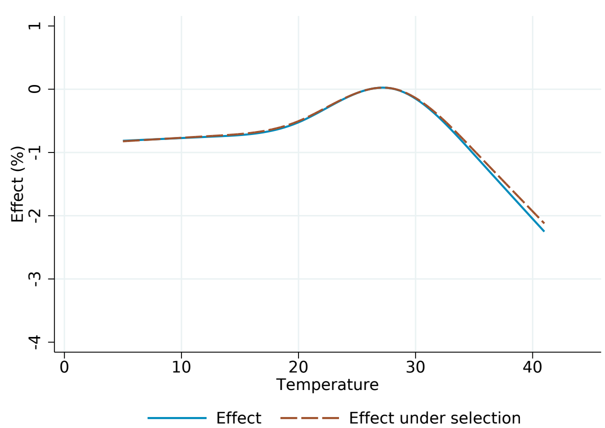
In sum, we simulate selection and find that selection biases estimates of the effect of



(a) Yield



(b) Quality



(c) Revenue per acre

Figure 5: Simulated selection bias

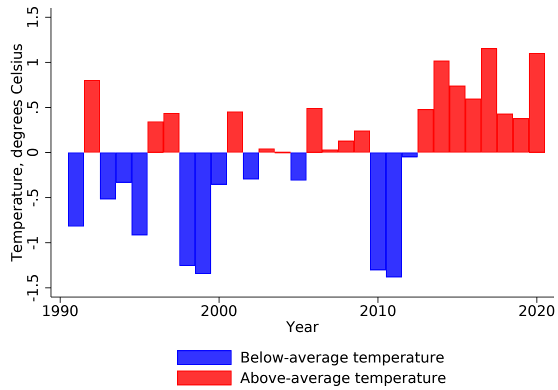
temperature on quality, yield, and revenue. In almost all published work, researchers rely on the assumption that observations of quality and yield accurately reflect conditions at harvest to recover unbiased estimates of temperature and climate change damages. For many agricultural products, data is only available for the subset of production that producers select to be graded. This may bias observations of both quality and yield – e.g. if only high quality product is graded, yield will be underestimated and quality overestimated. Focusing on processing tomatoes means that results are unlikely to be biased by selective harvesting or grading common in other settings where quality is crucial.

6 Discussion and Conclusion

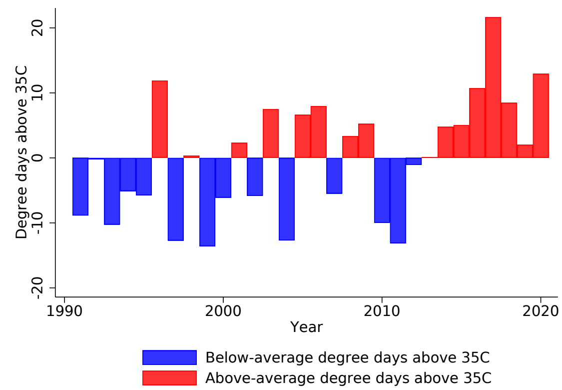
We use novel data from a large tomato processor to study the effect of temperature exposure on yield, quality, and grower revenue. We find exposure to hot temperatures reduces grower revenue through two channels: yield and quality. Quality matters because of its role in contractual arrangements and price determination but is either ignored or mismeasured in earlier work. Failing to account for quality's effect on revenue would bias estimates of temperatures on revenue by up to 20%.

Armed with the knowledge that processing tomatoes are susceptible to extreme temperatures, a natural follow-up question is: Has warming under a changing climate affected processing tomatoes? In the spirit of Lobell et al. (2011a) and Ortiz-Bobea et al. (2021), we investigate the extent to which recent warming trends have affected yield, quality, and revenue[7]. Over the last 30 years, average temperatures in California tomato growing regions have increased by around 0.04°C per year. Californian tomato growing regions are also exposed to an increasing duration of extreme heat, measured by degree days above 35°C (see Appendix D for an explanation of how we calculate degree days above 35°C).

[7]We focus on historical damages of warming trends. To accurately estimate future climate change damages would require careful consideration of water availability that is beyond the scope of the current paper.



(a) Deviation in annual average temperature from its 30-year average



(b) Deviation in annual degree days above 35°C from its 30-year average

Figure 6: Temperature trends

We conduct a back-of-the-envelope counterfactual analysis, comparing outcomes using realized temperatures with predicted outcomes had temperatures remained at the cooler levels observed between 1991 and 2010. This likely understates the true damages from warming trends as it does not account for how producers, processors, seed manufacturers, and other industry participants might have re-optimized absent the warming trend. That said, this exercise offers a rough lower bound of damages due to recent trends.

From our preferred specification, we estimate $\hat{y}_{it}$ in the usual manner:

$$\hat{y}_{it} = x_{it}B\hat{\Gamma} + \hat{\delta}z_{it} + \hat{\alpha}_{g(i)} + \hat{\psi}(t) \tag{12}$$

Next, we predict $\hat{y}_{itp}$, which replaces x_{it}, temperature exposure experienced at field i during year t, with x_{ip}, temperature exposure experienced at field i in each counterfactual year $p = 1991, ..., 2010$. All other controls and fixed effects are kept at their year t levels. This simulates outcomes as if the field was exposed to temperatures from counterfactual year p instead of the actual temperatures experienced in year t.

$$\hat{y}_{itp} = x_{ip}B\hat{\Gamma} + \hat{\delta}z_{it} + \hat{\alpha}_{g(i)} + \hat{\psi}(t) \tag{13}$$

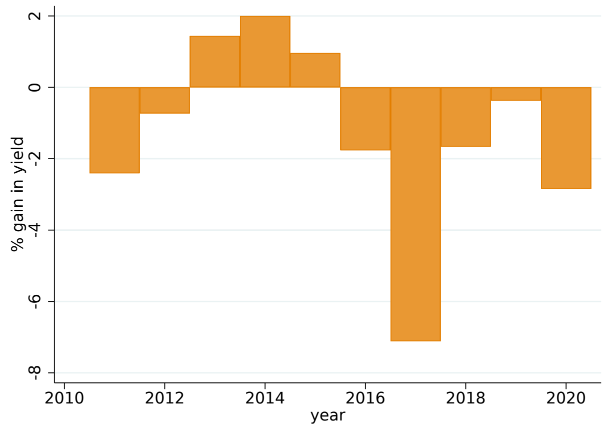
We then estimate annual gain or loss from warming trends. We define Δ_t as the difference between the predicted outcome using actual temperatures and predicted outcomes using counterfactual temperatures. The result is averaged across counterfactual years and fields for each year $t = 2011, ..., 2020$, where n_t is the number of field observations in year t. Using multiple counterfactual years ($P = 20$) reduces the effect of any one counterfactual year on the results. When $\Delta_t < 0$, temperatures experienced in year t caused a loss in the outcome variable relative to temperatures experienced in counterfactual years.

$$\Delta_t = \sum_{p=1}^{P} \sum_{i=1}^{n_t} (\hat{y}_{it} - \hat{y}_{itp}) / P n_t \tag{14}$$

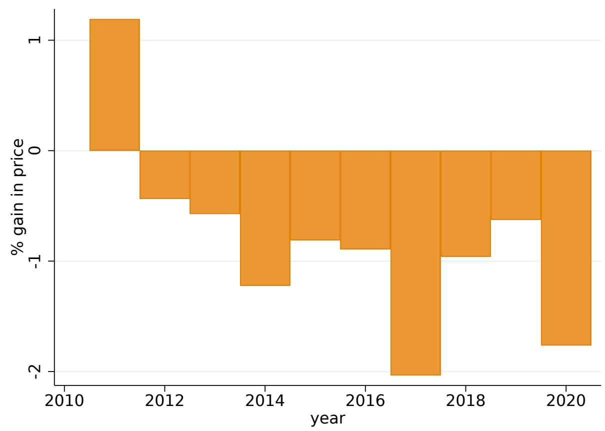
Over our 10 year sample, we find that yields were 1.3% lower on average due to warming temperatures. Put differently, yield would have averaged 1.3% higher had temperature exposure remained at 1991-2010 levels. Quality and revenue averaged 0.8% and 2.4% lower respectively due to warming trends. The effect varies from year to year, but losses are increasing over time. In 2020, revenue would have been 5.3% higher had temperature exposure remained at 1991-2010 levels. Our counterfactual analysis suggests that warming over the last 20 years resulted in economic damages for processing tomato growers.

The clear implication is that the continued warming predicted under climate change is a cause for concern. Growers cannot fully mitigate damages caused by extreme heat, indicating that the processing tomato industry will be susceptible to the worsening droughts and continued warming predicted as the climate changes. This stands in contrast to earlier work that finds irrigation can mitigate the effects of heat on yield. Rather, our results suggest that relying on irrigation to manage the impacts of climate change will be ineffective. These finding reinforce the need for investment in research into and development of heat tolerant varieties.

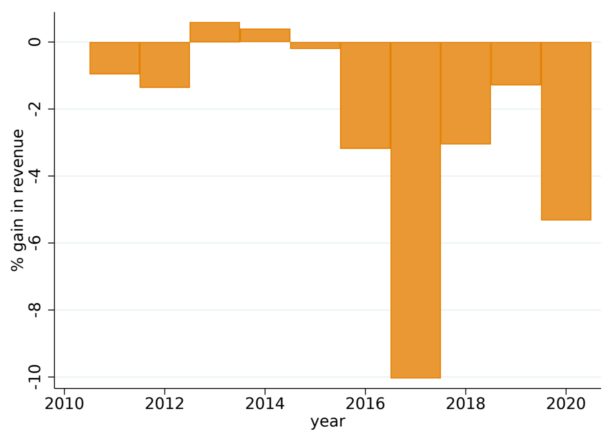
A caveat of our research is its focus on a single processor operating in one specific agricultural industry. While our narrow focus comes at the expense of potentially limiting the external validity of our results, we benefit from detailed and reliable observations of profit-maximizing farmers. This provides new insights into the effect of weather and climate on individual commercial agricultural producers.



(a) Yield



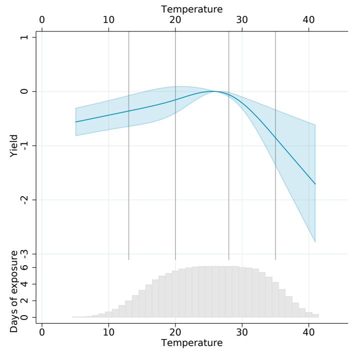
(b) Quality



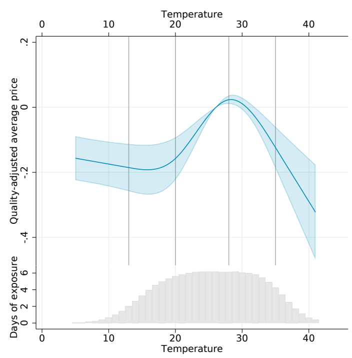
(c) Revenue per acre

Figure 7: Annual gains/losses from warming trends, relative to 1991-2010 weather

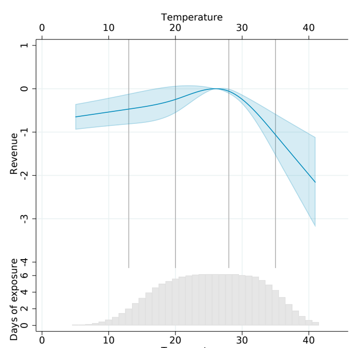
A Endogeneity in weather



(a) Yield



(b) Quality



(c) Revenue per acre

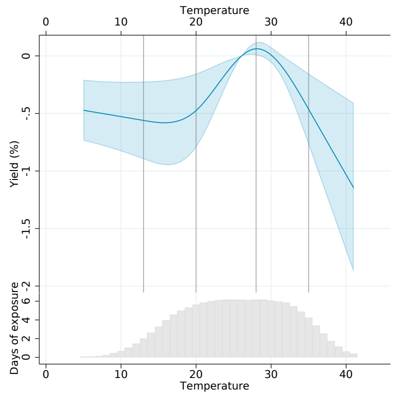
Figure 8: Results without a planting date control

For each figure, the graph at the top of the frame shows the effect of an additional 24 hours spent at a given temperature interval on the outcome variable relative to 24 hours spent at 26°C. The histogram at the bottom of the frame shows the average exposure to each temperature interval during the growing season across all fields in all years.

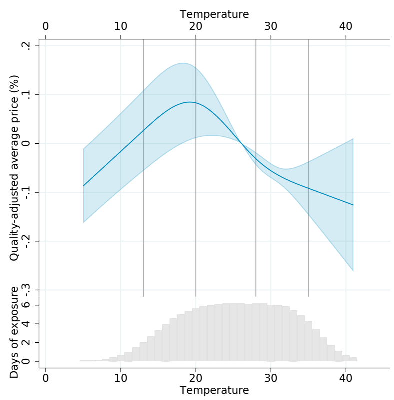
B Spatial Heteroscedasticity and Autocorrelation Consistent Errors

Our preferred specification clusters standard errors by grower group and county by year to account for the possibility of heteroskedasticity, spatial correlation, and temporal correlation in the errors. An alternative method to correct for possible dependence in standard errors is to estimate spatial heteroscedasticity and autocorrelation consistent (HAC) errors that allow for spatial correlation and serial correlation in panel data (Conley, 1999). Using code from Hsiang (2010), we allow for spatial correlation for field observations that are within 200km (124 miles) of each other. The correlation between observations is assumed to decay linearly with distance.

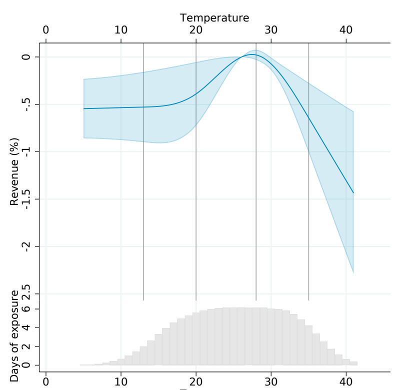
Results are robust to using spatial HAC errors, however its appropriateness on an unbalanced panel remains an open question.



(a) Yield



(b) Quality



(c) Revenue per acre

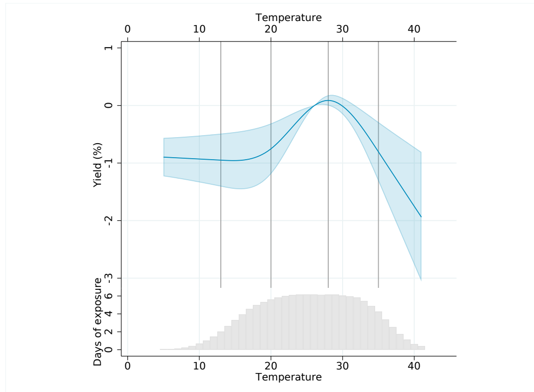
Figure 9: Results using spatial HAC errors

For each figure, the graph at the top of the frame shows the effect of an additional 24 hours spent at a given temperature interval on the outcome variable relative to 24 hours spent at 26°C. The histogram at the bottom of the frame shows the average exposure to each temperature interval during the growing season across all fields in all years.

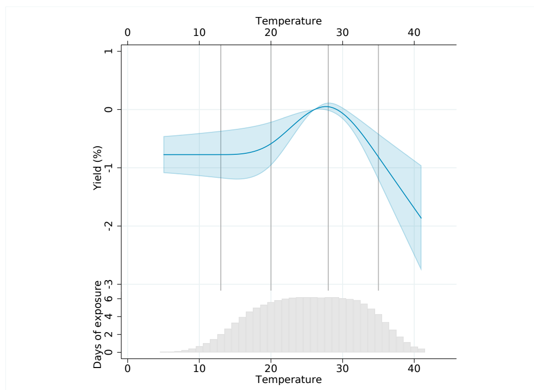
C Robustness checks

Our preferred specification uses grower fixed effects. Grower fixed effects capture time-invariant characteristics about the grower and characteristics of their fields which are both time-invariant and common across fields. One might be concerned that fields differ in ways that are correlated with weather, which would introduce omitted variable bias into our estimation. To alleviate this concern, we estimate Equation 18 using field fixed effects instead of grower fixed effects. Some fields do not appear multiple times in our sample because of crop rotation. We drop around 30% of field-year observations because they do not make a field-level panel. The results from this robustness check are similar to results using grower fixed effects.

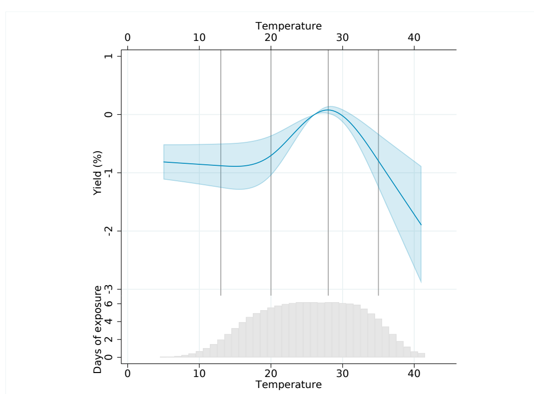
We also replace the quadratic year trend with a (a) linear year trend, and (b) county-specific quadratic year trends. The results are robust to the choice of functional form.



(a) Restricted spline with field fixed effects



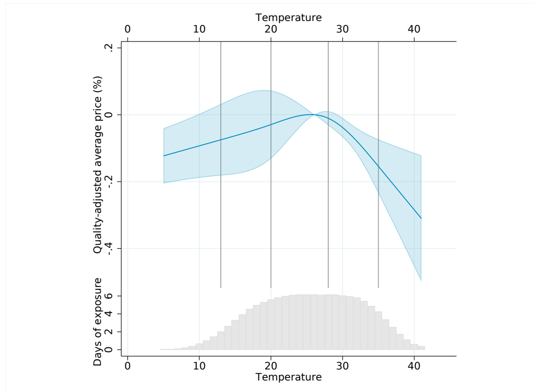
(b) Restricted spline with linear year trend



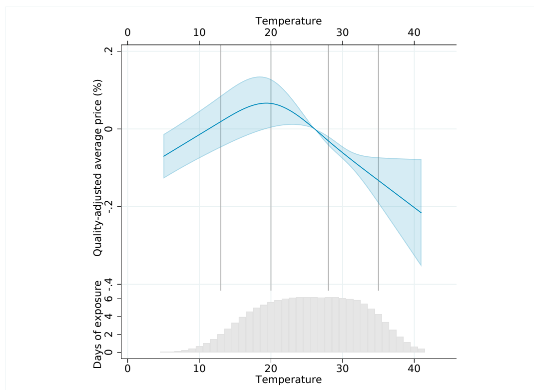
(c) Restricted spline with county-specific quadratic year trends

Figure 10: Yield

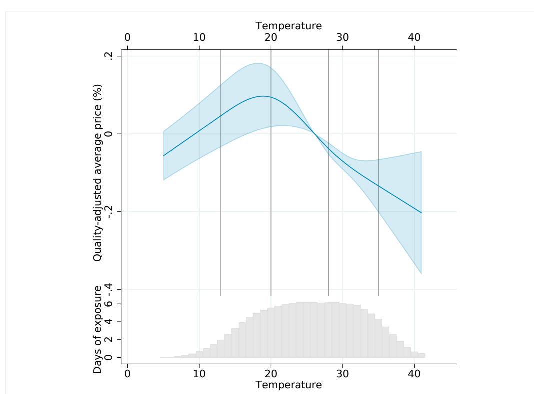
For each figure, the graph at the top of the frame shows the effect of an additional 24 hours spent at a given temperature interval on the outcome variable relative to 24 hours spent at 26°C. The histogram at the bottom of the frame shows the average exposure to each temperature interval during the growing season across all fields in all years.



(a) Restricted spline with field fixed effects



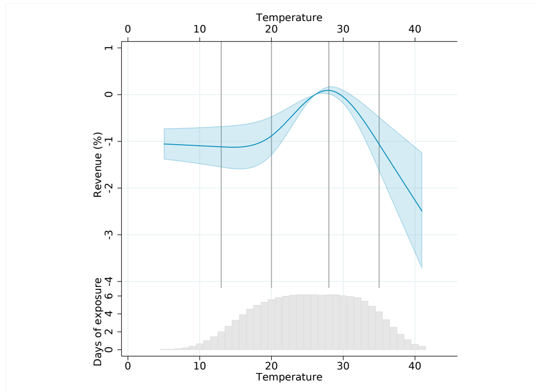
(b) Restricted spline with linear year trend



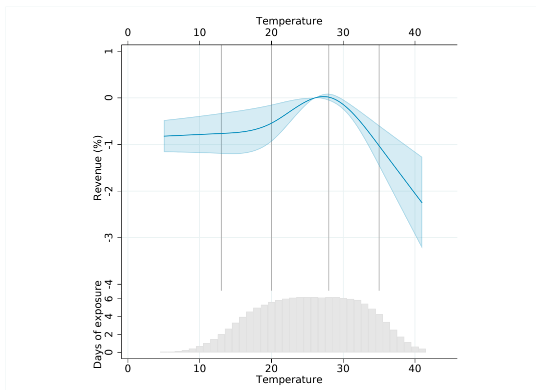
(c) Restricted spline with county-specific quadratic year trends

Figure 11: Quality

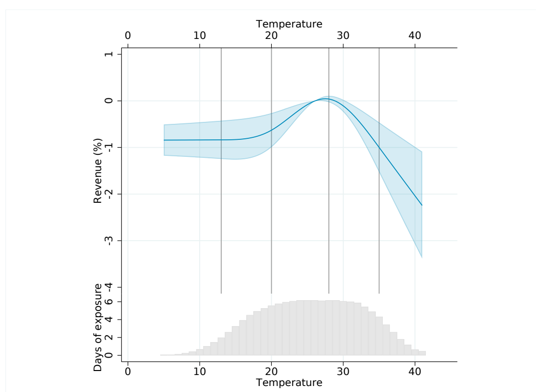
For each figure, the graph at the top of the frame shows the effect of an additional 24 hours spent at a given temperature interval on the outcome variable relative to 24 hours spent at 26°C. The histogram at the bottom of the frame shows the average exposure to each temperature interval during the growing season across all fields in all years.



(a) Restricted spline with field fixed effects



(b) Restricted spline with linear year trend



(c) Restricted spline with county-specific quadratic year trends

Figure 12: Revenue per acre

For each figure, the graph at the top of the frame shows the effect of an additional 24 hours spent at a given temperature interval on the outcome variable relative to 24 hours spent at 26°C. The histogram at the bottom of the frame shows the average exposure to each temperature interval during the growing season across all fields in all years.

D Linear piecewise degree day specification

The piecewise linear degree day model is widely used the agronomic and agricultural economic literature. It imposes more structure than flexible, semiparametric models. It also relies on the econometrician to correctly choose knot locations where the marginal effects change. However, it is less likely to overfit the data and has been shown in some contexts to perform better out-of-sample (Schlenker & Roberts, 2009).

To implement the piecewise linear degree day functional form, we first need to calculate degree days. Degree days are related to, but different from, temperature exposure. Temperature exposure measures how long is spent in a given temperature interval. Degree days measure how long and by how much temperatures exceed the lower bound of a temperature interval while being truncated at an upper bound (Snyder, 1985). When the temperature interval is small (e.g. 1°C), the difference between the two methods is relatively small because the "how much" dimension is unimportant relative to the "how long" dimension. When the temperature interval is large, as is the case in a piecewise linear model, the difference between the two methods will be large. For example, if we use temperature exposure, we assume that the damage of one day of exposure at 35°C is equal to the damage of one day at 40°C. If we use degree days, we assume the damage of five days at 35°C is equal to the damage of one day at 40°C. The underlying assumption of degree days is that the effect of temperature exposure increases linearly with temperature between lower and upper bound.

Degree days can be computed from the temperature exposure vector x_{it}. The expression for calculating degree days between a lower bound of $\underline{h}$ and upper bound of $\overline{h}$ is:

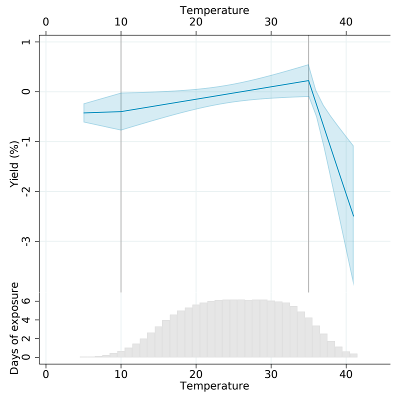
$$DD_{it,[\underline{h},\overline{h}]} = \sum_{j=\underline{h}}^{\overline{h}-1} x_{it,j} \times (j - \underline{h} + 1) \tag{15}$$

Next, we choose knot locations. In the first set of results, we use knot locations suggested by the agronomic literature. Mid-range temperatures are ideal for yield and quality outcomes, but these outcomes may be damaged by hot (greater than 35°C) or cool (less than 10°C) temperatures (Hartz et al., 2008). As a result, we choose two knots at $\kappa_1 = 10$°C and $\kappa_2 = 35$°C. We estimate degree days using Equation 15 for each of the three "segments": below 10°C, between 10°C and 35°C, and above 35°C.

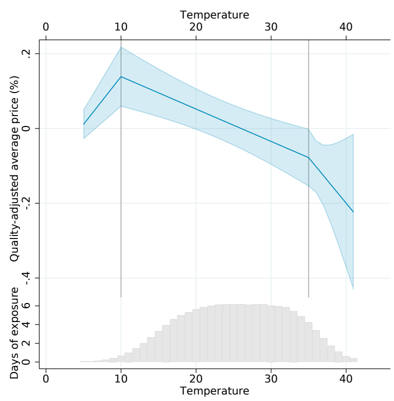
Equation 4 can then be modelled as:

$$y_{it} = \beta_0 + \beta_1 DD_{it,(-\infty,10]} + \beta_2 DD_{it,[10,35]} + \beta_3 DD_{it,[35,\infty)} + \delta z_{it} + \alpha_{g(i)} + \psi(t) + \epsilon_{it} \tag{16}$$

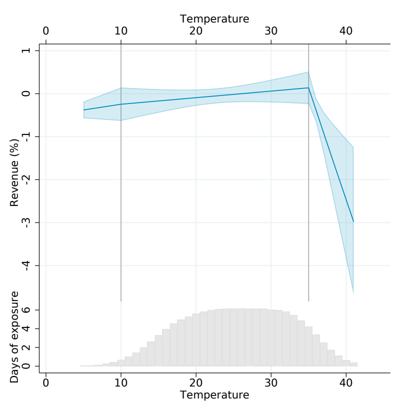
In the second set of results, we use knot locations that correspond to the turning points in the restricted cubic spline estimates.



(a) Yield



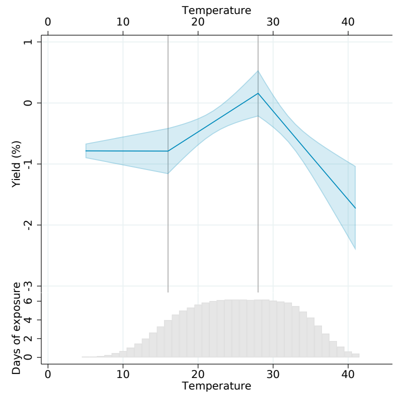
(b) Quality



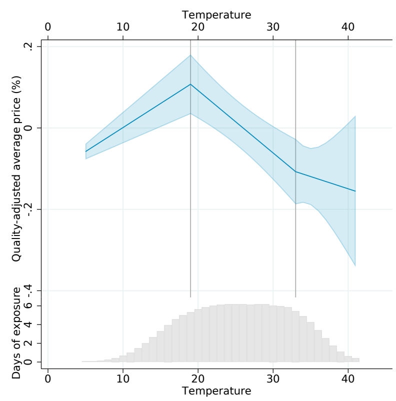
(c) Revenue per acre

Figure 13: Piecewise linear degree day results, agronomic knots

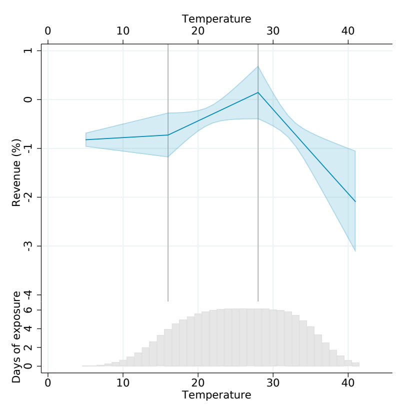
For each figure, the graph at the top of the frame shows the effect of an additional 24 hours spent at a given temperature interval on the outcome variable relative to 24 hours spent at 26°C. The histogram at the bottom of the frame shows the average exposure to each temperature interval during the growing season across all fields in all years.



(a) Yield



(b) Quality



(c) Revenue per acre

Figure 14: Piecewise linear degree day results, spline knots

For each figure, the graph at the top of the frame shows the effect of an additional 24 hours spent at a given temperature interval on the outcome variable relative to 24 hours spent at 26°C. The histogram at the bottom of the frame shows the average exposure to each temperature interval during the growing season across all fields in all years.

E Details on restricted cubic spline specification

The B matrix for a cubic spline combines the B matrix for the cubic polynomial and a $K \times J$ matrix Z. First, define $1 \times J$ temperature vector $W = \begin{pmatrix} 5 & 6 & 7 & ... & 41 \end{pmatrix}$. Z is the resulting matrix after applying the cubic spline basis function to W with K knots at $\kappa_1, ..., \kappa_K$.

$$\begin{aligned} B_{cubic.spline} &= \begin{pmatrix} B_{cubic.polynomial} & Z \end{pmatrix} \\ &= \begin{pmatrix} 5 & 25 & 125 & (5-\kappa_1)^3_+ & (5-\kappa_2)^3_+ & \dots & (5-\kappa_K)^3_+ \\ 6 & 36 & 216 & (6-\kappa_1)^3_+ & (6-\kappa_2)^3_+ & \dots & (6-\kappa_K)^3_+ \\ \vdots & \vdots & \vdots & \vdots & & \ddots & \vdots \\ 41 & 41^2 & 41^3 & (41-\kappa_1)^3_+ & (41-\kappa_2)^3_+ & \dots & (41-\kappa_K)^3_+ \end{pmatrix} \end{aligned} \tag{17}$$

Stone & Koo (1985) use the linearity constraints to develop a restricted cubic spline function. Using this function, a restricted cubic spline with K knots requires the estimation of only $K-1$ parameters on temperature (as opposed to $K+3$ parameters for the cubic spline). Equation 18 is the restricted spline function that we apply to W for $i = 1, 2, ..., K-2$ [8].

$$\begin{aligned} V_1 &= W \\ V_{i+1} &= \frac{(W-\kappa_i)^3_+ - (\kappa_K - \kappa_{K-1})^{-1}\{(W-\kappa_{K-1})^3_+(\kappa_K-\kappa_i) - (W-\kappa_K)^3_+(\kappa_{K-1}-\kappa_i)\}}{(\kappa_K-\kappa_1)^2} \\ &\quad \text{for i=1,2,..., K-2} \end{aligned} \tag{18}$$

Together, W and V_2 to V_{K-1} make up the B matrix for the restricted cubic spline, where $V_{i,j}$ is the j-th element of the V_i vector.

$$\begin{aligned} B_{rest.cubic.spline} &= \begin{pmatrix} W & V_2 & \dots & V_{K-1} \end{pmatrix} \\ &= \begin{pmatrix} 5 & V_{2,0} & \dots & V_{K-1,0} \\ 6 & V_{2,1} & \dots & V_{K-1,1} \\ 7 & V_{2,2} & \dots & V_{K-1,2} \\ \vdots & \vdots & \ddots & \vdots \\ 41 & V_{2,45} & \dots & V_{K-1,45} \end{pmatrix} \end{aligned} \tag{19}$$

[8] For more details, see the Stata manual for the function mkspline

References

AgSeeds (2020). Season Review. Available at http://www.agseeds.com/.

Ashenfelter, O. & Storchmann, K. (2016). Climate change and wine: A review of the economic implications. *Journal of Wine Economics*, 11(1), 105–138.

Berry, S. T., Roberts, M. J., & Schlenker, W. (2014). Corn production shocks in 2012 and beyond: Implications for harvest volatility. In *The Economics of Food Price Volatility* (pp. 59–81). University of Chicago Press.

Blanc, E. & Schlenker, W. (2020). The use of panel models in assessments of climate impacts on agriculture. *Review of Environmental Economics and Policy.*

Bucheli, J., Dalhaus, T., & Finger, R. (2022). Temperature effects on crop yields in heat index insurance. *Food Policy*, 107, 102214.

California Department of Food and Agriculture (2019). 2018-2019 California Agricultural Statistics Review. Available at https://www.cdfa.ca.gov/statistics/PDFs/2018-2019AgReportnass.pdf.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2), 238–249.

Cammarano, D., Jamshidi, S., Hoogenboom, G., Ruane, A. C., Niyogi, D., & Ronga, D. (2022). Processing tomato production is expected to decrease by 2050 due to the projected increase in temperature. *Nature Food*, (pp. 1–8).

Carter, C., Cui, X., Ghanem, D., & Mérel, P. (2018). Identifying the economic impacts of climate change on agriculture. *Annual Review of Resource Economics*, 10, 361–380.

Carter, E. K., Melkonian, J., Riha, S. J., & Shaw, S. B. (2016). Separating heat stress from moisture stress: analyzing yield response to high temperature in irrigated maize. *Environmental Research Letters*, 11(9), 094012.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of econometrics*, 92(1), 1–45.

D'Agostino, A. L. & Schlenker, W. (2016). Recent weather fluctuations and agricultural yields: implications for climate change. *Agricultural economics*, 47(S1), 159–171.

Dalhaus, T., Schlenker, W., Blanke, M. M., Bravin, E., & Finger, R. (2020). The effects of extreme weather on apple quality. *Scientific reports*, 10(1), 1–7.

Deschênes, O. & Greenstone, M. (2007). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *American Economic Review*, 97(1), 354–385.

FAO (2022). Food and Agriculture Organization Corporate Statistical Database (FAOSTAT). Available at `https://www.fao.org/faostat/en/`.

Gammans, M., Mérel, P., & Ortiz-Bobea, A. (2017). Negative impacts of climate change on cereal yields: statistical evidence from France. *Environmental Research Letters*, 12(5), 054007.

Hartz, T., Miyao, G., Mickler, J., Lestrange, M., Stoddard, S., Nuñez, J., & Aegerter, B. (2008). Processing tomato production in California. *UC Vegetable Research & Information Center*.

Hill, A. E. & Beatty, T. K. M. (2020). Haste makes waste: Evidence on speed and quality tradeoffs in the workplace. Working paper.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *Proceedings of the National Academy of sciences*, 107(35), 15367–15372.

Just, R. E. & Chern, W. S. (1980). Tomatoes, technology, and oligopsony. *The Bell Journal of Economics*, (pp. 584–602).

Kawasaki, K. & Uchida, S. (2016). Quality matters more than quantity: asymmetric temperature effects on crop yield and quality grade. *American Journal of Agricultural Economics*, 98(4), 1195–1209.

Lobell, D. B., Cahill, K. N., & Field, C. B. (2007). Historical effects of temperature and precipitation on California crop yields. *Climatic change*, 81(2), 187–203.

Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011a). Climate trends and global crop production since 1980. *Science*, 333(6042), 616–620.

Lobell, D. B., Torney, A., & Field, C. B. (2011b). Climate extremes in California agriculture. *Climatic change*, 109(1), 355–363.

Marklein, A., Elias, E., Nico, P., & Steenwerth, K. (2020). Projected temperature increases may require shifts in the growing season of cool-season crops and the growing locations of warm-season crops. *Science of The Total Environment*, 746, 140918.

McCarl, B. A. & Hertel, T. W. (2018). Climate change as an agricultural economics research topic. *Applied Economic Perspectives and Policy*, 40(1), 60–78.

NRCS USDA (2020). Soil Survey Geographic (SSURGO) Database for California. Available at `https://websoilsurvey.nrcs.usda.gov/`.

Ortiz-Bobea, A. (2021). The empirical analysis of climate change impacts and adaptation in agriculture. In *Handbook of agricultural economics*, volume 5 (pp. 3981–4073). Elsevier.

Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., & Lobell, D. B. (2021). Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change*, 11(4), 306–312.

Ortiz-Bobea, A., Knippenberg, E., & Chambers, R. G. (2018). Growing climatic sensitivity of us agriculture linked to technological change and regional specialization. *Science advances*, 4(12), eaat4343.

Ortiz-Bobea, A., Wang, H., Carrillo, C. M., & Ault, T. R. (2019). Unpacking the climatic drivers of us agricultural yields. *Environmental Research Letters*, 14(6), 064003.

Personal correspondence with the processor (2020). 2020-07-06.

PRISM Climate Group, Oregon State University (2020). PRISM Climate Data. Available at `http://prism.oregonstate.edu`.

Ramsey, A. F., Tack, J. B., & Balota, M. (2020). Double or nothing: Impacts of warming on crop quantity, quality, and revenue. *Journal of Agricultural and Resource Economics*, (pp. 1–22).

Schlenker, W. & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of sciences*, 106(37), 15594–15598.

Schmitt, J., Offermann, F., Söder, M., Frühauf, C., & Finger, R. (2022). Extreme weather events cause significant crop yield losses at the farm level in German agriculture. *Food Policy*, 112, 102359.

Shaw, S. B., Mehta, D., & Riha, S. J. (2014). Using simple data experiments to explore the influence of non-temperature controls on maize yields in the mid-West and Great Plains. *Climatic change*, 122(4), 747–755.

Shew, A. M., Tack, J. B., Nalley, L. L., & Chaminuka, P. (2020). Yield reduction under climate warming varies among wheat cultivars in South Africa. *Nature communications*, 11(1), 1–9.

Snyder, R. L. (1985). Hand calculating degree days. *Agricultural and forest meteorology*, 35(1-4), 353–358.

Stone, C. J. & Koo, C.-Y. (1985). Additive splines in statistics. In *Proceedings of the Statistical Computing Section ASA. Washington, DC* (pp. 45–48).

Tack, J., Barkley, A., & Hendricks, N. (2017). Irrigation offsets wheat yield reductions from warming temperatures. *Environmental Research Letters*, 12(11), 114027.

Tack, J., Barkley, A., & Nalley, L. L. (2015). Effect of warming temperatures on US wheat yields. *Proceedings of the National Academy of Sciences*, 112(22), 6931–6936.

USDA ERS (2020). Ag and Food Statistics: Charting the Essentials, February 2020. Available at `https://www.ers.usda.gov/webdocs/publications/96957/ap-083.pdf?v=1384.6`.

USDA NASS (2017). 2017 Census of Agriculture. Available at www.nass.usda.gov/AgCensus.

USDA NASS (2021a). 2020 California Processing Tomato County Estimates. Available at https://www.nass.usda.gov/Statistics_by_State/California/Publications/County_Estimates/2021/PTOMCounty_0321.pdf.

USDA NASS (2021b). 2021 California Processing Tomato Report. Available at https://www.nass.usda.gov/Statistics_by_State/California/Publications/Specialty_and_Other_Releases/Tomatoes/2021/PTOMInt_0121.pdf.

Wu, X., Yu, L., & Pehrsson, P. R. (2022). Are processed tomato products as nutritious as fresh tomatoes? Scoping review on the effects of industrial processing on nutrients and bioactive compounds in tomatoes. *Advances in Nutrition*, 13(1), 138–151.

Zhang, F., Biederman, J. A., Dannenberg, M. P., Yan, D., Reed, S. C., & Smith, W. K. (2021). Five decades of observed daily precipitation reveal longer and more variable drought events across much of the western United States. *Geophysical Research Letters*, 48(7), e2020GL092293.