



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Climate econometrics: Can the panel approach account for long-run adaptation?

Pierre Mérel[†] and Matthew Gammans[†]

[†] Agricultural and Resource Economics, University of California, Davis, 1 Shields Ave, Davis, CA 95616. Email: merel@primal.ucdavis.edu

Selected Paper prepared for presentation at the 2018 Agricultural & Applied Economics Association Annual Meeting, Washington, D.C., August 5-August 7

Copyright 2018 by Pierre Mérel and Matthew Gammans. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means provided this copyright notice appears on all such copies.

Abstract

The panel data approach with fixed effects has emerged as the preferred method to uncover the effects of climate change on economically relevant outcomes using historical weather data. While the panel method has been criticized for its purported inability to account for long-run adaptation, it has been argued that including nonlinearities in explanatory weather variables makes cross-sectional variation in climate enter coefficient identification, suggesting that the estimates obtained from a nonlinear, fixed-effects panel model at least partially reflect long-run adaptation. We formalize this argument in the context of the popular quadratic specification and show that (i) skewness in the historical weather data conditional on location is an essential driver of the bias in the panel estimates relative to the underlying long-run values, and can result in bias in either direction, (ii) in the absence of such skewness, the panel estimates are a convex combination of the short-run and long-run coefficients, and (iii) the panel estimates reflect the long-run values whenever the cross-sectional variation in climate “dominates” the location-specific weather fluctuations, in a sense that we make explicit. We use our framework to revisit impact estimates from nonlinear panel approaches published in the last decade. We find that for large countrywide or global panels, estimates of the effect of temperature primarily represent the long-run response, due to the large cross-sectional variation within these panels. In contrast, our calculations suggest that estimates of the effect of precipitation on outcomes reflect a more even combination of long- and short-run responses.

[†] Agricultural and Resource Economics, University of California, Davis, 1 Shields Ave, Davis, CA 95616. Email: merel@primal.ucdavis.edu. The authors thank Ariel Ortiz-Bobea for providing data on French weather at the department level.

JEL codes: C23, Q16, Q54

Keywords: climate change, adaptation, long run, crop yield

1 Introduction

The panel data approach with fixed effects has emerged as the preferred method to uncover the effects of climate change on economically relevant outcomes using historical weather data, mainly due to its ability to control for time-invariant omitted

variables that may confound the effect of climate in pure cross-sectional studies (Burke et al., 2015; Blanc and Schlenker, 2017).

Despite its growing popularity, the panel method has been criticized for its purported inability to account for long-run adaptation to climate due to its reliance on weather fluctuations rather than climate differences (Auffhammer et al., 2013; Burke and Emerick, 2016; Mendelsohn and Massetti, 2017). Deschênes and Greenstone (2007), who first introduced the panel approach to climate change impact assessment, write that “it is impossible to estimate the effect of the long-run climate averages in a model with county fixed effects, because there is no temporal variation in [climate variables].” To the extent that warming causes negative effects on the outcome of interest that can be mitigated through climate adaptation, the bias on the panel estimate of the effect of warming (relative to the underlying long-run value) would be away from zero (Deschênes and Greenstone, 2011).¹

Other studies have argued that including nonlinearities in explanatory weather variables makes cross-sectional variation re-enter coefficient identification, suggesting that the estimates obtained from a nonlinear, fixed effects panel model at least partially reflect long-run adaptation (McIntosh and Schlenker, 2006; Lobell et al., 2011; Burke et al., 2015; Schlenker, 2017; Blanc and Schlenker, 2017). Yet the extent to which such estimates should be thought of as inclusive of such adaptation remains unclear.² Whether and how much damage estimates obtained from nonlinear panel data reflect the underlying long-run adaptation potential is critical to their relevance for climate policy. One legitimate fear is that overly pessimistic short-run estimates in a context where significant adaptation potential exists might steer policy makers into making suboptimal policy choices or misdirecting public funding aimed at addressing the impacts of climate change. For instance, unduly pessimistic economic impact estimates could encourage policy makers to engage in costly mitigation efforts with uncertain and distant payoffs.

In this paper, we address the long-run nature of nonlinear panel estimates for a commonly used quadratic specification in weather variables. Although our formal results are derived for this particular specification, the insights they provide are more general. First, we show that in addition to the actual extent of long-run adaptation

¹This is true even in the case where the outcome variable is the value of an optimization problem, in which case the short- and long-run responses are identical to the first-, but not necessarily the second-order (Hsiang, 2016).

²For instance, Burke et al. (2015) write that “using both [...] sources of variation implicitly allows for more historical adaptation to longer-run climate, although the short-run changes in temperature that affect output remain unanticipated.”

undertaken by agents, skewness in the historical weather data conditional on location is an essential driver of the bias in the estimates obtained from the panel model relative to the underlying long-run values. This skewness can actually cause bias in *either* direction.

We then show that in the absence of skewness, the panel coefficient estimates of the quadratic relationship can be written as a convex combination of the underlying short-run and long-run coefficients. The decomposition reveals that the panel estimates reflect long-run values whenever the cross-sectional variation in climate “dominates” the location-specific weather fluctuations, in a sense we make analytically explicit. Said differently, panel estimates of the weather-outcome quadratic relationship can be thought of as a weighted average of short- and long-run responses, with the weight on the long-run parameters increasing with the share of the overall weather variation attributable to cross-sectional differences. In large countries like the US where variation in climate across space dominates location-specific weather fluctuations, existing panel estimates should thus be considered as already reflecting a significant share of the historical climate adaptation. For instance, calculations based on our derivations for quadratic models indicate that panel coefficient estimates obtained from county-level weather data across the years 1950–2015 are heavily weighted towards long-run parameter values, namely 98% for average spring-summer temperature and 67% for precipitation.

2 A simple model of long-run adaptation to climate

Our model of long-run adaptation to climate links an outcome variable y (e.g., the logarithm of farm profits) to weather x and climate μ . There are I locations (e.g., counties) indexed by i and T periods (e.g., years) indexed by t from which observations are drawn. We assume the following regarding the data-generating process (DGP).

First, weather is defined as a random variable centered around climate, that is,

Assumption 1

$$\mathbb{E}[x_{it}|\mu_i] = \mu_i.$$

As in McIntosh and Schlenker (2006), we assume that the outcome depends quadratically on both climate and weather.³ This choice of a quadratic functional form

³This specification is also used in Ihlanfeldt and Willardsen (2017).

in weather allows to capture non-monotonicities and non-linearities in the weather-outcome relationship.⁴

Assumption 2 *The DGP takes the following form:*

$$y_{it} = \alpha_i + \beta_1 x_{it} + \beta_2 x_{it}^2 + \beta_3 (x_{it} - \mu_i)^2 + \epsilon_{it}. \quad (1)$$

Finally, we make the usual strong exogeneity assumption that allows consistent estimation of $\beta = (\beta_1, \beta_2, \beta_3)$ using the fixed-effects estimator for the correctly specified model.

In model (1), climate is fixed over time, that is, the weather realizations x_{it} are realizations of a random variable with time-invariant mean μ_i . Climate μ_i enters through the penalty term $\beta_3 (x_{it} - \mu_i)^2$, where it is expected that $\beta_3 \leq 0$. The interpretation is that economic agents respond both to weather shocks and climate signals, but that, consistent with the classic definition of short-run and long-run production functions, the set of adaptation channels is larger in the long run than in the short run. For instance, as pointed out by Auffhammer and Schlenker (2014) in the context of agriculture, a one-year drought might not warrant the construction of an irrigation channel, yet it may be worth doing so in an arid climate. Similarly, an exceptional heat wave might not warrant installation of an air conditioning system, but it may trigger the purchase of a powerful ceiling fan. These examples parallel the idea that a firm's capital may be fixed in the short run yet variable in the long run. Here, the underlying assumption is that long-run inputs respond to climate while short-run inputs respond to both weather and climate.

Besides the presence of long-run investments, an element that may underlie different short-run and long-run responses is that it may take time and experience to successfully adapt to a particular situation. The basic idea behind the specification with quadratic penalty term is that conditional on *weather*, locations that are used to (have experienced) this weather because their *climate* is closer to it will fare better, other things being equal, than locations for which that particular weather realization happens to be an outlier—because they have had more opportunities to adapt to it. For instance, it is a generally accepted view that locations in hotter climates have adapted to heat better than locations in cooler climates. Similarly, countries exposed to seawater penetration, like the Netherlands, are likely better adapted to sea-level rise than land-locked countries. Note that this specification is admittedly restrictive in the sense

⁴For an example of a different nonlinear model structure that still allows for an interaction between weather and climate, see Schlenker et al. (2013).

that it is the absolute distance between climate and current weather that determines the extent of the penalty, but not the sign of the difference. That is, if a location in a hot climate and another one in a cold climate are both exposed to the mean of their climates, the short-run penalty will be the same for both locations.

[Figure 1 about here.]

The specification in Equation (1) also has a structural behavioral interpretation. It is a special case of the behavioral framework proposed by Schlenker (2017) in the context of crop yields, whereby the coefficient β_3 is allowed to depend on an endogenous index γ_i , interpreted as a crop variety chosen by farmers based on the mean weather μ_i and its variance σ_i^2 , and the index γ_i replaces μ_i in the penalty term. In this setting, the penalty might not be minimized where climate equals the current weather occurrence, as expected-yield-maximizing farmers may choose varieties that are not the best-performing ones under their mean weather ($\gamma_i \neq \mu_i$) if the chosen varieties are also less sensitive to weather fluctuations. If β_3 does not depend on γ_i however, the richer model reduces to specification (1) ($\gamma_i = \mu_i$) as farmers would get no benefit from choosing a variety less suitable for their mean weather. Schlenker (2017) shows that the model flexibility afforded by letting β_3 vary with γ_i does not translate into meaningful differences in long-run coefficient estimates for US corn yields.

The simple specification in equation (1) leads to clearly defined short-run and long-run responses to weather/climate. In the long run, weather shocks are perceived as climate shocks and the whole suite of adaptations is taken by economic agents, resulting in a zero penalty term:

$$\mathbb{E}[y_i^{\text{LR}} | \alpha_i, \mu_i, x_i] = \alpha_i + \beta_1 x_i + \beta_2 x_i^2. \quad (2)$$

In the short run, agents have adapted to their idiosyncratic climate μ_i and therefore

$$\mathbb{E}[y_i^{\text{SR}} | \alpha_i, \mu_i, x_i] = \alpha_i + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - \mu_i)^2. \quad (3)$$

Figure 1 depicts the resulting responses to weather/climate for a collection of locations with different climates in the instance where $\beta_2 < 0$ (dome-shaped long-run response). Response curves are scaled vertically so that α_i is the same across locations. If not, individual short-run and long-run curves are obtained from the depicted ones by vertical translations, with the property that for given location there is tangency between the long-run and short-run responses precisely at the weather realization $x_i = \mu_i$.

With this simple framework in mind, we now explore the consequences of using a “naive” quadratic panel approach that ignores the penalty term for the identification of long-run impacts. This type of model remains a staple of climate change impact analysis (Deschênes and Greenstone, 2007; Lobell et al., 2011; Lobell and Field, 2011; Gourджи et al., 2013; Annan and Schlenker, 2015; Burke et al., 2015; Kawasaki and Uchida, 2016; Cooper et al., 2017; Hsiang et al., 2017),⁵ as it is a convenient way of allowing for non-monotonicities and non-linearities in the weather-outcome relationship. Such non-monotonicities in weather or climate are an essential feature of many real-world phenomena, as has been argued extensively in recent literature (Burke et al., 2015; Carleton and Hsiang, 2016). Quadratic specifications have also remained popular in the cross-sectional Ricardian literature (Schlenker et al., 2006; Fezzi and Bateman, 2015) as well as in pure time-series approaches (Lobell et al., 2007).

We take the long-run coefficients $\beta^{\text{LR}} = (\beta_1, \beta_2)$ as the benchmark against which to evaluate coefficient estimates because they reflect the long-run relationship that is most critical in evaluating the net effects of climate change on economic outcomes (Mendelsohn et al., 1994; Burke and Emerick, 2016; Hsiang, 2016). Importantly, we only explore the impact of ignoring the penalty term and do not investigate the consequences of other types of model misspecification.

To fix ideas, we begin with the simpler case where $\beta_2 = 0$, that is, the long-run impact is linear in x_{it} while the short-run impact is quadratic (and therefore dome-shaped). As will be clear, this simple model is crucial in understanding the extent to which the naive panel estimates may be biased away from the underlying long-run coefficients.

3 Asymptotic bias with a linear long-run effect

In this section we assume that $\beta_2 = 0$, therefore the DGP is

$$\begin{aligned} y_{it} &= \alpha_i + \beta_1 x_{it} + \beta_3 (x_{it} - \mu_i)^2 + \epsilon_{it} \\ &= \alpha_i + \beta_3 \mu_i^2 + (\beta_1 - 2\beta_3 \mu_i) x_{it} + \beta_3 x_{it}^2 + \epsilon_{it} \end{aligned} \quad (4)$$

and the long-run marginal effect of x on y is given by parameter β_1 . The naive model is

$$y_{it} = a_i + b_1 x_{it} + e_{it} \quad (5)$$

⁵Annan and Schlenker (2015) use degree-day regressors to model a crop’s exposure to temperature but still use a quadratic in growing-season precipitation, as did Schlenker and Roberts (2009).

estimated with the fixed-effects (within) estimator.

We adopt the notation of McIntosh and Schlenker (2006). We write the demeaned variables as $\tilde{y}_{it} \equiv y_{it} - \frac{\sum_t y_{it}}{T}$, $\tilde{x}_{it} \equiv x_{it} - \frac{\sum_t x_{it}}{T}$, and $\tilde{x}_{it}^2 \equiv x_{it}^2 - \frac{\sum_t x_{it}^2}{T}$. We consider the asymptotic case where $T \rightarrow \infty$.

We are interested in the asymptotic bias as $T \rightarrow \infty$ of the estimated marginal impact of x_{it} on y_{it} , which in this simple model with linear effect of weather is simply equal to

$$\text{Bias} = \text{plim } \hat{b}_1 - \beta_1,$$

where \hat{b}_1 is the within estimator in model (5).

Define

$$M_{\tilde{x}^2}^i \equiv \text{plim } T^{-1} \sum_t \tilde{x}_{it}^2$$

and

$$M_{\tilde{x}^3}^i \equiv \text{plim } T^{-1} \sum_t \tilde{x}_{it}^3.$$

We show in Appendix A.2 that the resulting asymptotic bias is equal to

$$\text{Bias} = \beta_3 \frac{\sum_i M_{\tilde{x}^3}^i}{\sum_i M_{\tilde{x}^2}^i}. \quad (6)$$

Expression (6) first shows that the size of the bias is proportional to the extent of adaptation being ignored in the estimated model (β_3). This is intuitive, as if there were no long-run adaptation ($\beta_3 = 0$) then the estimated model would be correctly specified. The expression further shows that relying on the panel approach without considering adaptation can result in bias on the marginal effect (relative to the underlying long-run coefficient β_1) in *either* direction. In the case where $\beta_1 < 0$, the panel approach may either under- or over-estimate the negative effects of an increase in, say, average temperature. This result contrasts with the common acceptance that panel models capture short-run effects that overestimate long-run damages due to lack of adaptation. Finally, the expression shows that the bias is entirely driven by the *skewness* in the weather data conditional on location. If the weather data has a systematic positive skew ($\sum_t \tilde{x}_{it}^3 > 0$), and $\beta_1 < 0$, then the panel estimate overestimates the damage from an increase in x_{it} , consistent with the common expectation. But if the weather data has a negative skew, then the panel estimate underestimates this damage.

What is the intuition behind this finding? Clearly, if the omitted variable $(x_{it} - \mu_i)^2$, which asymptotically becomes \tilde{x}_{it}^2 , were uncorrelated with the included regressor x_{it} ,

no bias in b_1 would obtain. This is precisely what happens when the weather data shows no skewness, because then larger values of \ddot{x}_{it} are not systematically associated with larger values of \ddot{x}_{it}^2 .⁶ In particular, if the weather distribution is symmetric, for each positive value of \ddot{x}_{it} there is an equally probable negative one that has the same square. Now assume, for instance, that the weather data shows positive skewness. Assuming the weather distribution is unimodal, this implies that the distribution displays a fat or long tail towards values larger than the mean. Since $\beta_3 \leq 0$, large, positive weather shocks \ddot{x}_{it} will be correlated with large (in magnitude) penalties, and the estimate \hat{b}_1 will thus be biased towards more negative (or less positive) values.

[Figure 2 about here.]

Figure 2 illustrates the possible consequences of the skewness-induced bias on climate change impact predictions when the DGP exhibits a linear long-run response function. Here we have assumed that the weather distributions conditional on location are right-skewed, which results in a negative bias on the marginal effect of weather. Regardless of the point of evaluation, the estimated model will overestimate the negative effects of increases in the weather variable on the outcome.

How much does the skewness-induced bias matter in practice? Because the bias depends on the adaptation potential, as captured by β_3 , it is impossible to answer that question generally. Whether the bias matters also depends on the magnitude of the underlying long-run marginal effect β_1 . Nonetheless, one may look at particular contexts, for instance crop agriculture. Gammans et al. (2017) find that a 1°C uniform warming will result in a 4.1% decrease in wheat yield in France. Assuming that y_{it} denotes the logarithm of wheat yield in Equation (4), this would translate into a value $\beta_1 = -0.041$.⁷ If one believes that a 1°C difference between climate and current weather could cause a 1% yield penalty, then the bias due to skewness would be $-0.01 \times \frac{\sum_i M_{\ddot{x}^3}^i}{\sum_i M_{\ddot{x}^2}^i}$. Using department-level weather data over the period 1950-2016, we calculate that $\frac{\sum_i M_{\ddot{x}^3}^i}{\sum_i M_{\ddot{x}^2}^i} = 0.13$. Therefore, we expect the bias to be -0.0013 for a marginal effect of -0.041, which is negligible.

⁶Formally, the covariance between the included regressor x_{it} , and the omitted variable \ddot{x}_{it}^2 , is equal to $\mathbb{E}[\ddot{x}_{it}^3]$.

⁷Of course, the predicted effect might not be reflecting the underlying long-run parameter, but it provides an order of magnitude.

4 Asymptotic bias with a non-linear long-run effect

In this setting, the naive model is

$$y_{it} = a_i + b_1 x_{it} + b_2 x_{it}^2 + e_{it} \quad (7)$$

while the DGP is still given by equation (1).

4.1 General bias

Here we derive a general expression for the bias of the estimated marginal effect of climate. Because of the nonlinearity in both the DGP (1) and the estimated model (7), this marginal effect now depends on the point of evaluation, which we will denote as μ . In the naive model, the estimated marginal effect is $\hat{b}_1 + 2\hat{b}_2\mu$, whereas the true long-run marginal effect is $\beta_1 + 2\beta_2\mu$. Therefore, the bias evaluated at climate μ is

$$\text{Bias}(\mu) = \text{plim } \hat{b}_1 - \beta_1 + 2\mu \left(\text{plim } \hat{b}_2 - \beta_2 \right). \quad (8)$$

Let us further define

$$M_{(\hat{x}^2 - \bar{x}^2)^2}^i \equiv \text{plim } T^{-2} \sum_{s,t} (\hat{x}_{is}^2 - \hat{x}_{it}^2)^2.$$

We show in Appendix A.3 that the asymptotic bias resulting from the use of the naive model (7) can be written as

$$\text{Bias}(\mu) = -\frac{2\beta_3 N}{D} \quad (9)$$

with

$$\begin{aligned} N = & - \sum_i M_{\hat{x}^3}^i \left[\sum_i (\mu_i - \mu) M_{\hat{x}^3}^i + 2 \sum_i \mu_i (\mu_i - \mu) M_{\hat{x}^2}^i \right] \\ & + \sum_i (\mu_i - \mu) M_{\hat{x}^2}^i \left[2 \sum_i \mu_i M_{\hat{x}^3}^i + \sum_i M_{(\hat{x}^2 - \bar{x}^2)^2}^i \right] \end{aligned} \quad (10)$$

and

$$\begin{aligned}
D = & - \left(\sum_i M_{\bar{x}^3}^i \right)^2 + 4 \sum_i M_{\bar{x}^2}^i \sum_i \mu_i M_{\bar{x}^3}^i - 4 \sum_i M_{\bar{x}^3}^i \sum_i \mu_i M_{\bar{x}^2}^i \\
& + \sum_i M_{\bar{x}^2}^i \sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i + 4 \sum_{i,j} (\mu_i - \mu_j)^2 M_{\bar{x}^2}^i M_{\bar{x}^2}^j
\end{aligned} \tag{11}$$

where the summation $\sum_{i,j}$ in Expression (11) is taken over all un-ordered bundles of indices i and j . Because of the terms involving $M_{\bar{x}^3}^i$, as in the linear case the sign of the bias is generally ambiguous. However, the bias is not purely driven by the skewness of the weather data any more. In order to investigate the other source of bias, in the next section we specialize the analysis to the case $M_{\bar{x}^3}^i = 0$. In that case, we show that a bias still arises that results from the conflation of long-run and short-run responses. Specifically, when $M_{\bar{x}^3}^i = 0$ the panel estimates of β_1 and β_2 can be conveniently written as the *same* convex combination of the long-run and short-run underlying parameters, where the short-run parameters are evaluated at a composite climate μ that we make explicit. As a result, the estimated quadratic curve in the naive model is directly interpretable as a weighted average of long-run and short-run responses, that is, it captures some, but not all, of the long-run response.

4.2 Bias when the weather distribution is not skewed

When $M_{\bar{x}^3}^i = 0 \forall i$, the asymptotic bias as $T \rightarrow \infty$, evaluated at climate μ , simplifies to:

$$\text{Bias}(\mu) = -2\beta_3 \frac{\sum_i (\mu_i - \mu) M_{\bar{x}^2}^i \sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i}{\sum_i M_{\bar{x}^2}^i \sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i + 4 \sum_{i,j} (\mu_i - \mu_j)^2 M_{\bar{x}^2}^i M_{\bar{x}^2}^j}. \tag{12}$$

Furthermore, we show in Appendix A.4 that the fixed-effects estimator $\hat{\mathbf{b}} = \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \end{pmatrix}$ from the naive model (7) converges in probability towards a convex combination of the underlying long-run coefficients $\boldsymbol{\beta}^{\text{LR}} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ and the location-specific short-run coefficients $\boldsymbol{\beta}_i^{\text{SR}} = \begin{pmatrix} \beta_1 - 2\beta_3\mu_i \\ \beta_2 + \beta_3 \end{pmatrix}$:

$$\text{plim } \hat{\mathbf{b}} = (1 - \bar{\theta}) \boldsymbol{\beta}^{\text{LR}} + \bar{\theta} \sum_i \lambda_i \boldsymbol{\beta}_i^{\text{SR}} \tag{13}$$

where

$$\bar{\theta} = \frac{\sum_i M_{\bar{x}^2}^i \sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i}{\sum_i M_{\bar{x}^2}^i \sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i + 4 \sum_{i,j} (\mu_i - \mu_j)^2 M_{\bar{x}^2}^i M_{\bar{x}^2}^j} \quad (14)$$

and

$$\lambda_i = \frac{M_{\bar{x}^2}^i}{\sum_j M_{\bar{x}^2}^j}. \quad (15)$$

[Figure 3 about here.]

Importantly, the contribution of each location's short-run parameters in decomposition (13) can be replaced by the short-run parameters of a "composite" location, with climate equal to the weighted average $\sum_i \lambda_i \mu_i$, that is,

$$\text{plim } \hat{\mathbf{b}} = (1 - \bar{\theta}) \boldsymbol{\beta}^{\text{LR}} + \bar{\theta} \boldsymbol{\beta}_{\sum_i \lambda_i \mu_i}^{\text{SR}} \quad (16)$$

where $\boldsymbol{\beta}_{\sum_i \lambda_i \mu_i}^{\text{SR}} = \begin{pmatrix} \beta_1 - 2\beta_3 \sum_i \lambda_i \mu_i \\ \beta_2 + \beta_3 \end{pmatrix}$.

Expressions (12) and (15) imply that if the *marginal* impact of climate is evaluated at the weighted climate $\bar{\mu} = \sum_i \lambda_i \mu_i$, then the naive estimate has no bias, that is, it reflects the underlying long-run slope. Said differently, at the margin, the estimated relationship is correct when evaluated at the particular climate value $\bar{\mu}$. While previous research has already argued that short-run and long-run responses should be identical at the margin whenever the outcome variable is being optimized (e.g., Hsiang (2016)), our result is both different and more specific. First, we have shown that the estimated relationship is *not* the short-run response, but instead a weighted average of the short-run and long-run responses. Second, our analysis makes explicit at what particular point one should expect tangency between the estimated response and the underlying long-run response: it is a weighted average of the locational climates, where the weight for location i is that location's contribution to the overall time-series variation $\sum_i M_{\bar{x}^2}^i$. Because locations may contribute differently to time-series variation, the tangency will generally not occur at the mean climate.

The quadratic relationship obtained from the naive model departs *globally* from the true, underlying long-run relationship. This departure is illustrated in Figure 3. Because the inferred marginal impact is correct at $\bar{\mu}$, the "true" and "estimated" relationships are tangent at $\bar{\mu}$ (the curves have been vertically scaled so that the value of y is the same at $\bar{\mu}$). However, at any other evaluation point, inference based on first-order effects will be biased. For $\mu < \bar{\mu}$, the slope of the estimated relationship is

less negative, implying *positive* bias (less negative or more positive marginal effect), as illustrated with the evaluation point μ_1 in the figure. In contrast, for $\mu > \bar{\mu}$ there is *negative* bias (more negative or less positive marginal effect). Studies that compute net impacts by aggregating panel-specific impacts are therefore summing positively and negatively biased effects. Depending on the underlying structure of weather fluctuations across panels (as captured by the λ_i parameters), the magnitude of predicted climate changes for each panel, and the weighting scheme used in aggregation, the net impact may be biased in either direction. For instance, if the outcome variable is crop yield, the weather variable is temperature, and planted areas are used as weights in the aggregation, the net bias may be positive if panels with relatively large areas and/or subject to the largest increase in temperature are also those with a cooler climate (a lower value of μ_i).

Importantly, due to the relative positions of the two curves, there is generally bias if one goes beyond first-order effects and uses the globally estimated relationship for counterfactual estimation, even when starting from the climate average $\bar{\mu}$. For instance, moving from $\bar{\mu}$ to the new climate $\mu_2 > \bar{\mu}$, there is a negative bias on the global effect when using the estimated relationship. This gives credence to, while formalizing it, the idea that panel models would tend to overestimate the negative effects of warming due to their (partial) reliance on weather variation.

Expressions (13) and (14) further imply that the estimated response will be close to the underlying long-run response whenever

$$\sum_i M_{\bar{x}^2}^i \sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i \ll \sum_{i,j} (\mu_i - \mu_j)^2 M_{\bar{x}^2}^i M_{\bar{x}^2}^j. \quad (17)$$

This condition has a nice interpretation. First, note that the naive model can only be identified if there is time-series variation conditional on location, that is, $\sum_i M_{\bar{x}^2}^i > 0$.⁸ Also note that $M_{\bar{x}^2}^i = 0 \Rightarrow M_{(\bar{x}^2 - \bar{x}^2)^2}^i = 0$, so that if a location displays no time-series variation in weather, its index can be removed from all summations in condition (17). We can therefore limit ourselves to locations for which $M_{\bar{x}^2}^i > 0$. Condition (17) essentially implies that in order for $\hat{\mathbf{b}}$ to be close to the long-run parameter values, the time-series variation in weather, as captured by the terms $M_{(\bar{x}^2 - \bar{x}^2)^2}^i$, must be small relative to the cross-sectional variation in climate, captured by the terms $(\mu_i - \mu_j)^2$.

Note that if $M_{(\bar{x}^2 - \bar{x}^2)^2}^i = 0 \forall i$, then the identified parameter vector $\hat{\mathbf{b}}$ is consistent

⁸Otherwise, both x_{it} and x_{it}^2 are constant conditional on location in Equation (7), and therefore vector \mathbf{b} cannot be identified due to the inclusion of the fixed effects α_i .

for the vector of long-run parameter values. Given the definition of $M_{(\bar{x}^2 - \bar{x}^2)^2}^i$, this condition is equivalent to saying that, in a given location, weather takes on only two equiprobable values.⁹ In that case, the penalty term $\beta_3 (x_{it} - \mu_i)^2$ is constant conditional on location and is thus collinear to the fixed-effects vector. Therefore, this penalty term is no longer present in the error term e_{it} of the naive fixed-effects model (7) and the bias naturally disappears. Therefore, if the time-series variation in the weather data is purely binary (say either hot or cold weather), then this simple source of variation will allow parameter identification without causing bias.¹⁰ Said differently, it is not the existence of time-series variation *per se* ($\ddot{x}_{it} > 0$) that contaminates the naive estimates (in fact, such variation is essential for parameter identification), but rather the existence of variation in the *absolute departures from climate* in the time series ($\ddot{x}_{is}^2 \neq \ddot{x}_{it}^2$).¹¹

In summary, when estimating a naive model that omits the penalty term capturing adaptation, one should expect to estimate a response that is a weighted average of the underlying long-run response and the locational short-run responses, at least if the weather data is not skewed. Whether the estimated relationship leads to an under- or over-estimate of the impact of a change in climate will depend on the point of evaluation and the size of the change considered. If the initial climate is chosen at the tangency point between the estimated and underlying responses, the estimated relationship will produce impact estimates that are overly pessimistic.

4.3 Decomposition in the general case

Going back to the general case where $M_{\bar{x}^3}^i \neq 0$, Equation (10) implies the existence of a point (climate) where there is no bias on the marginal effect. This climate is given by

$$\tilde{\mu} = \frac{\sum_i \mu_i M_{\bar{x}^2}^i \left[2 \sum_i \mu_i M_{\bar{x}^3}^i + \sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i \right] - \sum_i M_{\bar{x}^3}^i \left[\sum_i \mu_i M_{\bar{x}^3}^i + 2 \sum_i \mu_i^2 M_{\bar{x}^2}^i \right]}{\sum_i M_{\bar{x}^2}^i \left[2 \sum_i \mu_i M_{\bar{x}^3}^i + \sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i \right] - \sum_i M_{\bar{x}^3}^i \left[\sum_i M_{\bar{x}^3}^i + 2 \sum_i \mu_i M_{\bar{x}^2}^i \right]}. \quad (18)$$

In addition, a simple calculation shows that the naive parameter estimate $\hat{\mathbf{b}}$ can be decomposed into

$$\text{plim } \hat{\mathbf{b}} = (1 - \tilde{\theta}) \boldsymbol{\beta}^{\text{LR}} + \tilde{\theta} \boldsymbol{\beta}_{\tilde{\mu}}^{\text{SR}} \quad (19)$$

⁹Because then all deviations from the mean are equal in magnitude, which implies $M_{(\bar{x}^2 - \bar{x}^2)^2}^i = 0$. In all other instances, $M_{(\bar{x}^2 - \bar{x}^2)^2}^i > 0$.

¹⁰Note that this is a consequence of our assumption that the penalty depends only on the absolute distance between weather and climate.

¹¹McIntosh and Schlenker (2006) indicate that the naive model will yield the long-run parameter values only if $\beta_3 = 0$. Here we have essentially shown that there is no bias either if $\beta_3 \neq 0$ but $\sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i = 0$.

for

$$\tilde{\theta} = \frac{\sum_i M_{\bar{x}^2}^i \left[2 \sum_i \mu_i M_{\bar{x}^3}^i + \sum_i M_{(\bar{x}^2 - \bar{x}^2)^2}^i \right] - \sum_i M_{\bar{x}^3}^i \left[\sum_i M_{\bar{x}^3}^i + 2 \sum_i \mu_i M_{\bar{x}^2}^i \right]}{D} \quad (20)$$

where D is given by Equation (11) and $\beta_{\tilde{\mu}}^{\text{SR}} = \begin{pmatrix} \beta_1 - 2\beta_3\tilde{\mu} \\ \beta_2 + \beta_3 \end{pmatrix}$ is the vector of short-run parameters evaluated at climate $\tilde{\mu}$. However, unlike the case where $M_{\bar{x}^3}^i = 0 \forall i$, there is no guarantee that the convexity parameter $\tilde{\theta}$ lies between zero and one, or that the composite climate $\tilde{\mu}$ lies within the range of climates μ_i observed in the sample. In the empirical application below, we show that for three existing weather datasets used in the recent climate change literature, skewness only plays a minor role in the decomposition of the estimated response into its short-run and long-run components.

5 Empirical application

Armed with a better understanding of how climate adaptation is captured by a quadratic specification, we can now revisit existing panel estimates by looking at the nature of the weather variation used. We consider three datasets that have been previously studied in the climate impact assessment literature. First, we consider a 67-year panel of French department-level ($I = 88$) average temperature and cumulative precipitation across the wheat and barley growing season of March through July. This data was previously analyzed in Gammans et al. (2017). Second, we consider a 66-year US county-level ($I = 3,037$) panel of spring-summer average temperature and cumulative precipitation in the contiguous United States. Many studies have used similar data to investigate how various outcomes respond to weather in the US (Schlenker and Roberts, 2009; Burke and Emerick, 2016).¹² Third, we apply our methodology to the global panel of country-level ($I = 167$) annual mean temperature and precipitation used in Burke et al. (2015).¹³ Although not every study has used quadratic specifications (or perhaps not for every weather variable included), we can get a sense of how close to the long-run response the identified relationships could reasonably be expected to lie by computing the parameter $\tilde{\theta}$ in Equation (14).

Results of these calculations are shown in Table 1. Recall that $\tilde{\theta}$ is the weight on the

¹²Our calculations were performed based on the weather data made available by Wolfram Schlenker at <http://www.columbia.edu/~ws2162/dailyData.html>.

¹³Our calculations rely on data made available by Burke et al. at <https://web.stanford.edu/~mburke/climate/data.html>

short-run parameter values and that $\bar{\mu}$ is the composite climate at which the estimated marginal response is equal to the long-run response. Calculations of $\bar{\theta}$ and $\bar{\mu}$ assume that $M_{\tilde{x}^3}^i = 0$, i.e., asymptotically there is no skewness in the weather data. In contrast, calculations of $\tilde{\theta}$ and $\tilde{\mu}$ assume that $M_{\tilde{x}^3}^i = T^{-1} \sum_t \tilde{x}_{it}^3$, i.e., the probability limit of $T^{-1} \sum_t \tilde{x}_{it}^3$ is equal to its sample analog. For the data tested here, we find that allowing for skewness in the weather data does not meaningfully change the calculated values. For calculations of $\bar{\mu}$ and $\tilde{\mu}$, units for temperature are degrees Celsius and units for precipitation are millimeters. For each weather dataset, we construct 1,000 simulated datasets by sampling years with replacement. We sample years rather than individuals in order to preserve the spatial correlation within the data, since weather is highly correlated across space. For each simulated dataset we calculate the value of interest and report the standard deviation of these simulated values in parenthesis. These values represent the uncertainty in our calculated values due to the fact that, for all calculations, we rely on the sample analogs of asymptotic values.

For temperature data, we find that estimates of the effect of growing season temperature mostly reflect the long-run response. In France, estimates of the effect of temperature reflect 86% the long-run response and 14% the weighted short-run responses. For the US and global panels, the weight on the long-run response is 98% and 100% respectively. Estimates of the effect of precipitation on outcomes contain more of the short-run response than those for temperature. For the French, US, and global panels, the weight on the long-run response is 56%, 67%, and 87% respectively. Allowing for skewness in the weather data alters these values by no more than 2%. Across all datasets and both weather variables, $\tilde{\mu}$ falls within the range of sample climates, although in some cases it is relatively far from $\bar{\mu}$, i.e. for US temperature and global precipitation panels.

[Table 1 about here.]

6 Conclusion

This paper has addressed the importance of allowing for climate memory to enter the direct effect of weather on economic outcomes in panel data analysis, thereby allowing for implicit long-run adaptation. Past climate matters to current realizations only if it leads to adaptation by economic agents. Ignoring climate when it matters biases estimates of long-run impacts, whether to the first or second order, except in very specific conditions that we have made explicit in the context of the popular quadratic panel model.

Finally, we should stress that although our analysis was motivated by the measurement of adaptation to climate, it has applications for panel data beyond the climate impact assessment literature. Our framework should be relevant whenever the outcome variable is allowed to depend non-monotonically on the regressor of interest and there is a distinction between within (short-run) and global (long-run) responses.

References

- Annan, F. and Schlenker, W. (2015). Federal Crop Insurance and the Disincentive to Adapt to Extreme Heat. *The American Economic Review*, 105(5):262–266.
- Auffhammer, M., Hsiang, S. M., Schlenker, W., and Sobel, A. (2013). Using Weather Data and Climate Model Output in Economic Analyses of Climate Change. *Review of Environmental Economics and Policy*, 7(2):181–198.
- Auffhammer, M. and Schlenker, W. (2014). Empirical studies on agricultural impacts and adaptation. *Energy Economics*, 46:555–561.
- Blanc, E. and Schlenker, W. (2017). The use of panel models in assessments of climate impacts on agriculture. *Review of Environmental Economics and Policy*, 11(2):258–279.
- Burke, M. and Emerick, K. (2016). Adaptation to Climate Change: Evidence from US Agriculture. *American Economic Journal: Economic Policy*, 8(3):106–140.
- Burke, M., Hsiang, S. M., and Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, 527:235–239.
- Carleton, T. A. and Hsiang, S. M. (2016). Social and economic impacts of climate. *Science*, 353(6304):aad9837.
- Cooper, J., Nam Tran, A., and Wallander, S. (2017). Testing for Specification Bias with a Flexible Fourier Transform Model for Crop Yields. *American Journal of Agricultural Economics*, 99(3):800–817.
- Deschênes, O. and Greenstone, M. (2007). The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather. *The American Economic Review*, 97(1):354–385.
- Deschênes, O. and Greenstone, M. (2011). Climate change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US. *American Economic Journal: Applied Economics*, 3:152–185.

- Fezzi, C. and Bateman, I. (2015). The Impact of Climate Change on Agriculture: Nonlinear Effects and Aggregation Bias in Ricardian Models of Farmland Values. *Journal of the Association of Environmental and Resource Economists*, 2(1):57–92.
- Gammans, M., Mérel, P., and Ortiz-Bobea, A. (2017). Negative impacts of climate change on cereal yields: statistical evidence from France. *Environmental Research Letters*, 12(5):054007.
- Gourdji, S. M., Mathews, K. L., Reynolds, M., Crossa, J., and Lobell, D. B. (2013). An assessment of wheat yield sensitivity and breeding gains in hot environments. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1752):20122190.
- Hsiang, S. (2016). Climate econometrics. *Annual Review of Resource Economics*, 8:43–75.
- Hsiang, S., Kopp, R., Jina, A., Rising, J., Delgado, M., Mohan, S., Rasmussen, D., Muir-Wood, R., Wilson, P., Oppenheimer, M., et al. (2017). Estimating economic damage from climate change in the United States. *Science*, 356(6345):1362–1369.
- Ihlanfeldt, K. and Willardsen, K. (2017). Local public services costs and the geography of development: Evidence from florida counties. *Journal of Regional Science*.
- Kawasaki, K. and 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., and Field, C. B. (2007). Historical effects of temperature and precipitation on California crop yields. *Climatic Change*, 81(2):187–203.
- Lobell, D. B. and Field, C. B. (2011). California perennial crops in a changing climate. *Climatic Change*, 109(1):317–333.
- Lobell, D. B., Schlenker, W., and Costa-Roberts, J. (2011). Climate Trends and Global Crop Production Since 1980. *Science*, 333(6042):616–620.
- McIntosh, C. T. and Schlenker, W. (2006). Identifying Non-linearities in Fixed Effects Models. Working paper, School of International Relations and Pacific Studies, University of California at San Diego.
- Mendelsohn, R., Nordhaus, W. D., and Shaw, D. (1994). The Impact of Global Warming on Agriculture: A Ricardian Analysis. *The American Economic Review*, 84(4):753–771.

- Mendelsohn, R. O. and Massetti, E. (2017). The use of cross-sectional analysis to measure climate impacts on agriculture: theory and evidence. *Review of Environmental Economics and Policy*, 11(2):280–298.
- Schlenker, W. (2017). Inter-annual Weather Variation and Crop Yields. Working Paper, Dept. of Economics and School of International and Public Affairs, Columbia University.
- Schlenker, W., Hanemann, W. M., and Fisher, A. C. (2006). The Impact of Global Warming on U.S. Agriculture: An Econometric Analysis of Optimal Growing Conditions. *The Review of Economics and Statistics*, 88(1):113–125.
- Schlenker, W. and Roberts, M. J. (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, W., Roberts, M. J., and Lobell, D. B. (2013). US maize adaptability. *Nature Climate Change*, 3(8):690.

Appendices

A Derivation of the asymptotic bias

A.1 Useful expressions

We define $\bar{x}_i = \frac{\sum_t \dot{x}_{it}}{T}$. We thus have $\ddot{x}_{it} = x_{it} - \bar{x}_i$, $\sum_t \ddot{x}_{it} = 0$, and $\dot{x}_{it}^2 = (\ddot{x}_{it} + \bar{x}_i)^2 - \frac{\sum_s (\ddot{x}_{is} + \bar{x}_i)^2}{T} = \ddot{x}_{it}^2 + \bar{x}_i^2 + 2\ddot{x}_{it}\bar{x}_i - \frac{\sum_s \ddot{x}_{is}^2 + \bar{x}_i^2 + 2\ddot{x}_{is}\bar{x}_i}{T} = \ddot{x}_{it}^2 + 2\ddot{x}_{it}\bar{x}_i - \frac{\sum_s \ddot{x}_{is}^2}{T}$. We deduce the following:

$$\sum_t \ddot{x}_{it} \dot{x}_{it}^2 = \sum_t \ddot{x}_{it}^3 + 2\bar{x}_i \sum_t \ddot{x}_{it}^2$$

$$\sum_t (\dot{x}_{it}^2)^2 = \sum_t \ddot{x}_{it}^4 + 4\bar{x}_i \sum_t \ddot{x}_{it}^3 + 4\bar{x}_i^2 \sum_t \ddot{x}_{it}^2 - \frac{1}{T} \left(\sum_t \ddot{x}_{it}^2 \right)^2$$

A.2 Linear model

$$\text{Define: } \mathbf{\check{y}} = \begin{pmatrix} \check{y}_{11} \\ \vdots \\ \check{y}_{1T} \\ \vdots \\ \check{y}_{I1} \\ \vdots \\ \check{y}_{IT} \end{pmatrix}, \quad \mathbf{\check{x}} = \begin{pmatrix} \check{x}_{11} \\ \vdots \\ \check{x}_{1T} \\ \vdots \\ \check{x}_{I1} \\ \vdots \\ \check{x}_{IT} \end{pmatrix}, \quad \mathbf{\check{W}} = \begin{pmatrix} \check{x}_{11} & 0 & \dots & 0 & \check{x}_{11}^2 \\ \vdots & 0 & \dots & 0 & \vdots \\ \check{x}_{1T} & 0 & \dots & 0 & \check{x}_{1T}^2 \\ 0 & \check{x}_{21} & \dots & 0 & \check{x}_{21}^2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & \check{x}_{2T} & \dots & 0 & \check{x}_{2T}^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \check{x}_{I1} & \check{x}_{I1}^2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & \check{x}_{IT} & \check{x}_{IT}^2 \end{pmatrix}, \quad \mathbf{\check{\epsilon}} = \begin{pmatrix} \check{\epsilon}_{11} \\ \vdots \\ \check{\epsilon}_{1T} \\ \vdots \\ \check{\epsilon}_{I1} \\ \vdots \\ \check{\epsilon}_{IT} \end{pmatrix},$$

$$\boldsymbol{\beta}^- = \begin{pmatrix} \beta_1 - 2\beta_3\mu_1 \\ \vdots \\ \beta_1 - 2\beta_3\mu_I \\ \beta_3 \end{pmatrix}. \text{ The fixed-effects estimator of } b_1 \text{ in model (5) is}$$

$$\begin{aligned} \hat{b}_1 &= (\mathbf{\check{x}'\check{x}})^{-1} \mathbf{\check{x}'\check{y}} \\ &= (\mathbf{\check{x}'\check{x}})^{-1} \mathbf{\check{x}'(\check{W}\boldsymbol{\beta}^- + \check{\epsilon})} \\ &= \beta_1 + \left(\sum_i \sum_t \check{x}_{it}^2 \right)^{-1} \left(-2\beta_3 \sum_i \mu_i \sum_t \check{x}_{it}^2 + \beta_3 \sum_i \sum_t \check{x}_{it} \check{x}_{it}^2 + \sum_i \sum_t \check{x}_{it} \check{\epsilon}_{it} \right) \\ &= \beta_1 + \left(\sum_i \sum_t \check{x}_{it}^2 \right)^{-1} \left[-2\beta_3 \sum_i \mu_i \sum_t \check{x}_{it}^2 + \beta_3 \sum_i \left(\sum_t \check{x}_{it}^3 + 2\bar{x}_i \sum_t \check{x}_{it}^2 \right) + \sum_i \sum_t \check{x}_{it} \check{\epsilon}_{it} \right]. \end{aligned}$$

Our strong exogeneity assumption implies that $\text{plim } T^{-1} \sum_t \check{x}_{it} \check{\epsilon}_{it} = 0$. In addition, $\text{plim } \bar{x}_i = \mu_i$, therefore we have

$$\text{plim } \hat{b}_1 = \beta_1 + \beta_3 \frac{\sum_i M_{\check{x}^3}^i}{\sum_i M_{\check{x}^2}^i}.$$

A.3 Quadratic model

Define $\ddot{\mathbf{X}} = \begin{pmatrix} \ddot{x}_{11} & \dot{x}_{11}^2 \\ \vdots & \vdots \\ \ddot{x}_{1T} & \dot{x}_{1T}^2 \\ \vdots & \vdots \\ \ddot{x}_{I1} & \dot{x}_{I1}^2 \\ \vdots & \vdots \\ \ddot{x}_{IT} & \dot{x}_{IT}^2 \end{pmatrix}$ and $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 - 2\beta_3\mu_1 \\ \vdots \\ \beta_1 - 2\beta_3\mu_I \\ \beta_2 + \beta_3 \end{pmatrix}$. The fixed-effects estimator of $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ is

$$\begin{aligned} \hat{\mathbf{b}} &= (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \ddot{\mathbf{X}}'\ddot{\mathbf{y}} \\ &= (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \ddot{\mathbf{X}}'(\ddot{\mathbf{W}}\boldsymbol{\beta} + \ddot{\boldsymbol{\epsilon}}). \end{aligned}$$

Denote $\mathbf{M}_{\mathbf{X}\mathbf{X}} \equiv \text{plim } T^{-1}\ddot{\mathbf{X}}'\ddot{\mathbf{X}}$. Strong exogeneity implies that $\text{plim } T^{-1}\ddot{\mathbf{X}}'\ddot{\boldsymbol{\epsilon}} = 0$. Therefore, we have

$$\begin{aligned} \text{plim } \hat{\mathbf{b}} &= \mathbf{M}_{\mathbf{X}\mathbf{X}}^{-1} \text{plim } T^{-1}\ddot{\mathbf{X}}'\ddot{\mathbf{W}}\boldsymbol{\beta} \\ &= \mathbf{M}_{\mathbf{X}\mathbf{X}}^{-1} \text{plim } T^{-1} \left[\ddot{\mathbf{X}}'\ddot{\mathbf{X}} \begin{pmatrix} \beta_1 \\ \beta_2 + \beta_3 \end{pmatrix} - 2\beta_3 \begin{pmatrix} \sum_i \mu_i \sum_t \dot{x}_{it}^2 \\ \sum_i \mu_i \sum_t \ddot{x}_{it} \dot{x}_{it}^2 \end{pmatrix} \right] \\ &= \begin{pmatrix} \beta_1 \\ \beta_2 + \beta_3 \end{pmatrix} - 2\beta_3 \mathbf{M}_{\mathbf{X}\mathbf{X}}^{-1} \begin{pmatrix} \sum_i \mu_i \text{plim } T^{-1} \sum_t \dot{x}_{it}^2 \\ \sum_i \mu_i \text{plim } T^{-1} \sum_t \ddot{x}_{it} \dot{x}_{it}^2 \end{pmatrix}. \end{aligned}$$

We can write $\mathbf{M}_{\mathbf{X}\mathbf{X}} = \begin{pmatrix} \sum_i \text{plim } T^{-1} \sum_t \dot{x}_{it}^2 & \sum_i \text{plim } T^{-1} \sum_t \ddot{x}_{it} \dot{x}_{it}^2 \\ \sum_i \text{plim } T^{-1} \sum_t \ddot{x}_{it} \dot{x}_{it}^2 & \sum_i \text{plim } T^{-1} \sum_t (\dot{x}_{it}^2)^2 \end{pmatrix}$, which, defining $M_{\dot{x}^2}^i \equiv \text{plim } T^{-1} \sum_t \dot{x}_{it}^2$, implies:

$$\text{plim } \hat{\mathbf{b}} = \begin{pmatrix} \beta_1 - 2\beta_3 \frac{N_1}{D} \\ \beta_2 + \beta_3 \left(1 - 2\frac{N_2}{D}\right) \end{pmatrix}$$

with

$$\begin{aligned}
D &= \left[\sum_i M_{\ddot{x}^2}^i \right] \left[\sum_i \text{plim } T^{-1} \sum_t (\dot{x}_{it}^2)^2 \right] - \left[\sum_i \text{plim } T^{-1} \sum_t \ddot{x}_{it} \dot{x}_{it}^2 \right]^2 \\
N_1 &= \left[\sum_i \text{plim } T^{-1} \sum_t (\dot{x}_{it}^2)^2 \right] \left[\sum_i \mu_i M_{\ddot{x}^2}^i \right] - \left[\sum_i \text{plim } T^{-1} \sum_t \ddot{x}_{it} \dot{x}_{it}^2 \right] \left[\sum_i \mu_i \text{plim } T^{-1} \sum_t \ddot{x}_{it} \dot{x}_{it}^2 \right] \\
N_2 &= \left[\sum_i M_{\ddot{x}^2}^i \right] \left[\sum_i \mu_i \text{plim } T^{-1} \sum_t \ddot{x}_{it} \dot{x}_{it}^2 \right] - \left[\sum_i \text{plim } T^{-1} \sum_t \ddot{x}_{it} \dot{x}_{it}^2 \right] \left[\sum_i \mu_i M_{\ddot{x}^2}^i \right].
\end{aligned}$$

Using the expressions in Section A.1, defining $M_{\ddot{x}^3}^i \equiv \text{plim } T^{-1} \sum_t \ddot{x}_{it}^3$ and $M_{\ddot{x}^4}^i \equiv \text{plim } T^{-1} \sum_t \ddot{x}_{it}^4$, and using $\text{plim } \bar{x}_i = \mu_i$, the denominator D can be written as

$$\begin{aligned}
D &= \left[\sum_i M_{\ddot{x}^2}^i \right] \left[\sum_i M_{\ddot{x}^4}^i + 4\mu_i M_{\ddot{x}^3}^i + 4\mu_i^2 M_{\ddot{x}^2}^i - \left(M_{\ddot{x}^2}^i \right)^2 \right] - \left[\sum_i M_{\ddot{x}^3}^i + 2\mu_i M_{\ddot{x}^2}^i \right]^2 \\
&= \left[\sum_i M_{\ddot{x}^2}^i \right] \left[\sum_i M_{(\ddot{x}^2+2\mu\ddot{x})}^i - \left(M_{\ddot{x}^2}^i \right)^2 \right] - \left[\sum_i M_{\ddot{x}^3+2\mu\ddot{x}^2}^i \right]^2
\end{aligned}$$

where we have also defined $M_{(\ddot{x}^2+2\mu\ddot{x})}^i \equiv \text{plim } T^{-1} \sum_t \left(\ddot{x}_{it}^2 + 2\mu_i \ddot{x}_{it} \right)^2 = M_{\ddot{x}^4}^i + 4\mu_i M_{\ddot{x}^3}^i + 4\mu_i^2 M_{\ddot{x}^2}^i$ and $M_{\ddot{x}^3+2\mu\ddot{x}^2}^i = \text{plim } T^{-1} \sum_t \ddot{x}_{it}^3 + 2\mu_i \ddot{x}_{it}^2 = M_{\ddot{x}^3}^i + 2\mu_i M_{\ddot{x}^2}^i$. The numerator terms can then be written as

$$N_1 = \left[\sum_i M_{(\ddot{x}^2+2\mu\ddot{x})}^i - \left(M_{\ddot{x}^2}^i \right)^2 \right] \left[\sum_i \mu_i M_{\ddot{x}^2}^i \right] - \left[\sum_i M_{\ddot{x}^3+2\mu\ddot{x}^2}^i \right] \left[\sum_i \mu_i M_{\ddot{x}^3+2\mu\ddot{x}^2}^i \right]$$

and

$$N_2 = \left[\sum_i M_{\ddot{x}^2}^i \right] \left[\sum_i \mu_i M_{\ddot{x}^3+2\mu\ddot{x}^2}^i \right] - \left[\sum_i M_{\ddot{x}^3+2\mu\ddot{x}^2}^i \right] \left[\sum_i \mu_i M_{\ddot{x}^2}^i \right].$$

We can then write the bias on the marginal effect of climate at the evaluation point μ as:

$$\begin{aligned}
\text{Bias}(\mu) &= \text{plim } \hat{b}_1 - \beta_1 + 2\mu \left(\text{plim } \hat{b}_2 - \beta_2 \right) \\
&= -2\beta_3 \frac{N_1}{D} + 2\mu\beta_3 \left(1 - 2\frac{N_2}{D} \right) \\
&= \frac{2\beta_3}{D} [-N_1 + \mu(D - 2N_2)]
\end{aligned}$$

Simple algebra shows that the term in square brackets can be rewritten as

$$N = \left[\sum_i M_{\check{x}^3}^i \right] \left[\sum_i (\mu_i - \mu) M_{\check{x}^3}^i + 2 \sum_i \mu_i (\mu_i - \mu) M_{\check{x}^2}^i \right] - 2 \left[\sum_i \mu_i M_{\check{x}^3}^i \right] \left[\sum_i (\mu_i - \mu) M_{\check{x}^2}^i \right] - \left[\sum_i (\mu_i - \mu) M_{\check{x}^2}^i \right] \left[\sum_i \left(M_{\check{x}^4}^i - \left(M_{\check{x}^2}^i \right)^2 \right) \right].$$

Let us now show that $M_{\check{x}^4}^i - \left(M_{\check{x}^2}^i \right)^2 = \text{plim } T^{-2} \sum_{s,t} \left(\check{x}_{is}^2 - \check{x}_{it}^2 \right)^2$, where the summation $\sum_{s,t}$ is taken over all un-ordered bundles of indices s and t . First note that $\left(M_{\check{x}^2}^i \right)^2 = \left(\text{plim } T^{-1} \sum_t \check{x}_{it}^2 \right)^2 = \text{plim } T^{-2} \left(\sum_t \check{x}_{it}^2 \right)^2 = \text{plim } T^{-2} \left[\sum_t \check{x}_{it}^4 + 2 \sum_{\substack{s,t \\ s \neq t}} \check{x}_{is}^2 \check{x}_{it}^2 \right]$. Therefore,

$$\begin{aligned} M_{\check{x}^4}^i - \left(M_{\check{x}^2}^i \right)^2 &= \text{plim } T^{-1} \left[\sum_t \check{x}_{it}^4 - \frac{1}{T} \sum_t \check{x}_{it}^4 - \frac{2}{T} \sum_{\substack{s,t \\ s \neq t}} \check{x}_{is}^2 \check{x}_{it}^2 \right] \\ &= \text{plim } T^{-2} \left[(T-1) \sum_t \check{x}_{it}^4 - 2 \sum_{\substack{s,t \\ s \neq t}} \check{x}_{is}^2 \check{x}_{it}^2 \right] \\ &= \text{plim } T^{-2} \sum_{s,t} \left(\check{x}_{is}^2 - \check{x}_{it}^2 \right)^2 \end{aligned}$$

where the last equality obtains because each term \check{x}_{it}^4 appears $T-1$ times in the summation (each time index t is paired with one of the $T-1$ remaining indices).

Defining $M_{(\check{x}^2 - \check{x}^2)^2}^i \equiv \text{plim } T^{-2} \sum_{s,t} \left(\check{x}_{is}^2 - \check{x}_{it}^2 \right)^2$ leads to expression (10) in the main text.

Using a similar argument, we can rewrite

$$\begin{aligned} D &= \left[\sum_i M_{\check{x}^2}^i \right] \left[\sum_i M_{(\check{x}^2 - \check{x}^2)^2}^i \right] - \left[\sum_i M_{\check{x}^3}^i \right]^2 + 4 \left[\sum_i M_{\check{x}^2}^i \right] \left[\sum_i \mu_i M_{\check{x}^3}^i \right] \\ &\quad - 4 \left[\sum_i M_{\check{x}^3}^i \right] \left[\sum_i \mu_i M_{\check{x}^2}^i \right] + 4 \left[\sum_i M_{\check{x}^2}^i \right] \left[\sum_i \mu_i^2 M_{\check{x}^2}^i \right] - 4 \left[\sum_i \mu_i M_{\check{x}^2}^i \right]^2 \end{aligned}$$

It is straightforward to show that the last two terms in the previous expression simplify to $4 \sum_{i,j} (\mu_i - \mu_j)^2 M_{\check{x}^2}^i M_{\check{x}^2}^j$, where the summation is taken over the set of un-ordered bundles of indices i and j . Expression (11) follows.

A.4 Parameter estimates when $M_{\ddot{x}^3}^i = 0 \forall i$

Assume that $M_{\ddot{x}^3}^i = 0 \forall i$. Recall that $\text{plim } \hat{\mathbf{b}} = \begin{pmatrix} \beta_1 - 2\beta_3 \frac{N_1}{D} \\ \beta_2 + \beta_3 \left(1 - 2\frac{N_2}{D}\right) \end{pmatrix}$ while $\boldsymbol{\beta}^{\text{LR}} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ and $\boldsymbol{\beta}_i^{\text{SR}} = \begin{pmatrix} \beta_1 - 2\beta_3 \mu_i \\ \beta_2 + \beta_3 \end{pmatrix}$. We seek to write $\hat{\mathbf{b}}$ as a convex combination of the underlying parameters, that is, $\hat{\mathbf{b}} = (1 - \bar{\theta})\boldsymbol{\beta}^{\text{LR}} + \bar{\theta} \sum_i \lambda_i \boldsymbol{\beta}_i^{\text{SR}}$ with $\bar{\theta} \in [0, 1]$, $\lambda_i \geq 0$, and $\sum_i \lambda_i = 1$. Such decomposition must satisfy $\bar{\theta} \sum_i \lambda_i \mu_i = \frac{N_1}{D}$ and $\bar{\theta} = 1 - 2\frac{N_2}{D}$. Using $M_{\ddot{x}^3}^i = 0 \forall i$, we obtain Equations (14) and (15), which make it clear that $\bar{\theta} \in [0, 1]$, $\lambda_i \geq 0$, and $\sum_i \lambda_i = 1$.

List of Figures

Long-run and short-run responses to weather	
Asymptotic bias when weather fluctuations are right-skewed	
Asymptotic bias on counterfactual impact of climate change when $M_{x^3}^i = 0$	

Figure 1 Long-run and short-run responses to weather

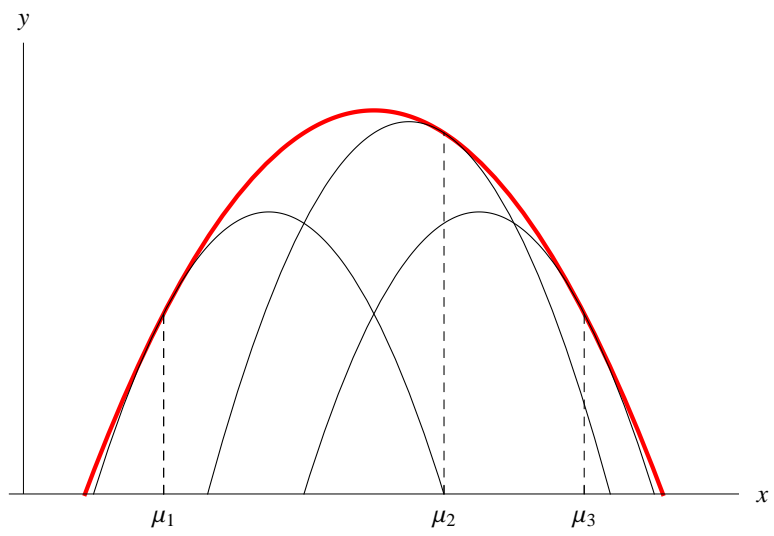
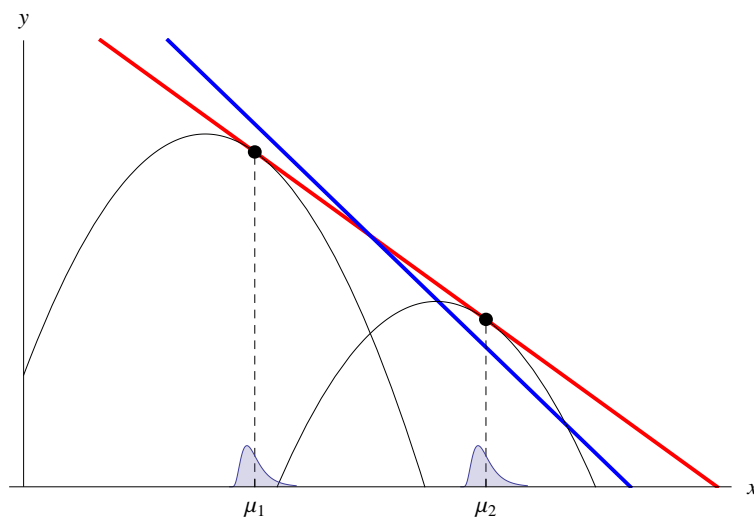
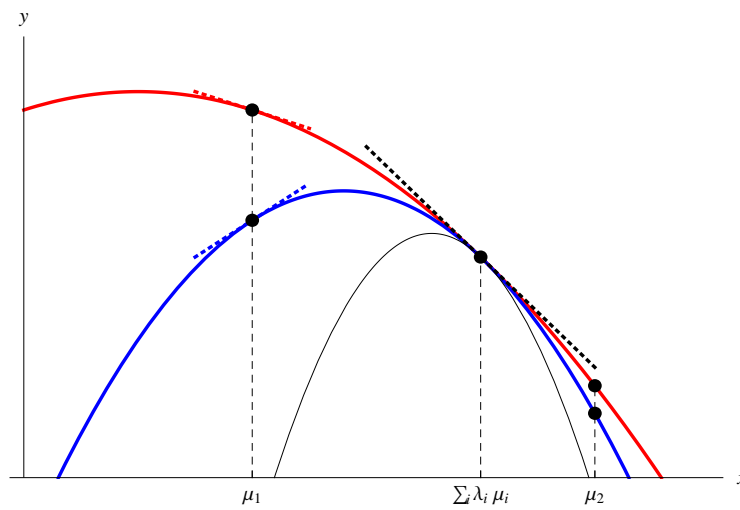


Figure 2 Asymptotic bias when weather fluctuations are right-skewed



Note: Red curve: underlying long-run response. Blue curve: estimated response. Black curve: underlying short-run responses. Weather distributions are shown on the x-axis for locations with climates μ_1 and μ_2 .

Figure 3 Asymptotic bias on counterfactual impact of climate change when $M_{x^3}^i = 0$



Note: Red curve: underlying long-run response. Blue curve: estimated response. Black curve: underlying short-run response for climate $\bar{\mu} = \sum_i \lambda_i \mu_i$. The blue curve is obtained as a convex combination of the red and black curves, with the weight on the black curve given by expression (14).

List of Tables

Bias decomposition

Table 1 Bias decomposition

Description	I	T	$\bar{\theta}$	$\tilde{\theta}$	$\bar{\mu}$	$\tilde{\mu}$
French department growing season temperature	88	67	0.14 (0.03)	0.15 (0.02)	13.76 (0.11)	13.38 (0.58)
French department growing season precipitation	88	67	0.46 (0.04)	0.45 (0.04)	324.00 (7.94)	308.32 (10.18)
US county spring-summer temperature	3037	66	0.02 (0.001)	0.03 (0.01)	17.05 (0.19)	14.02 (2.30)
US county spring-summer precipitation	3037	66	0.33 (0.03)	0.32 (0.03)	708.56 (9.07)	641.73 (12.49)
Country annual temperature	167	51	0.002 (0.0004)	0.01 (0.01)	15.32 (0.48)	15.21 (4.52)
Country annual precipitation	167	51	0.13 (0.01)	0.11 (0.02)	1894.90 (31.32)	1761.40 (109.40)