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

Pierre Mérel[†] and Matthew Gammans[†]

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 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 then illustrate how the panel approach can be leveraged to simultaneously estimate short-run and long-run effects using data on French cereal yields. We extend the existing framework by allowing unobserved heterogeneity to affect both yield levels and the marginal response of yields to weather. In contrast to recent results for US crops, our results suggest significant adaptation to temperature and precipitation changes in the long run.

[†] 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 correlated with climate 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 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 introduced the panel approach to climate change impact assessment in agriculture, 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 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). Yet the extent to which such estimates should be thought of as inclusive of long-run 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.

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

[1]This 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).

[2]For 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."

results are derived for this particular specification, the insights they provide are more general. We first show that in addition to the actual extent of long-run adaptation 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 locational variation in climate dominates short-run weather fluctuations, existing panel estimates should thus be considered as already reflecting a significant share of the historical climate adaptation. 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. As a point of comparison, in France these figures translate to 86% and 53%, respectively, when using department-level data.

Of course, one cannot control the structure of the variation naturally present in the data. We therefore move on to show how the panel approach can be leveraged to simultaneously estimate the short-run and long-run effects, and propose an application to French cereal yields. Importantly, we extend the existing framework proposed by McIntosh and Schlenker (2006) by allowing unobserved heterogeneity to affect both yield levels and the marginal response of yields to weather. This refinement, made possible by the presence of climate trends in our data, makes our estimates robust to omitted variables like soils that, in addition to being likely determinants of yield levels, may be interacting with weather variables. In contrast to recent results for US crops, notably by Burke and Emerick (2016), our results suggest statistically significant and biologically relevant adaptation to temperature and precipitation changes in the long run for French cereals.

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

Assumption 2 *The DGP takes the following form:*

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

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

Assumption 3

$$\mathbb{E}[\epsilon_{it}|\alpha_i, \mu_i, x_{i1}, \ldots, x_{iT}] = 0.$$

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. We relax this assumption in our empirical application below to account for a richer model where unobserved heterogeneity enters not only additively (α_i) but also multiplicatively (i.e., as an interaction with weather). This refinement is made possible by the presence of location-specific trends in climate in our empirical setting. We leave it aside for now as the existing literature has generally not allowed for such heterogeneity and we are mainly interested in the potential inconsistency of panel estimates attributable to the lack of consideration of long-run adaptation.

In equation (1), climate μ_i enters through the penalty term $\beta_3 \left(x_{it} - \mu_i\right)^2$, where it is expected that $\beta_3 \leq 0$. The interpretation is that economic agents respond both to

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

weather shocks and climate signals, but that, consistent with the classic definition of short-run and long-run average cost 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), a one-year drought might not warrant the construction of an irrigation channel, yet it may be worth doing so in an arid climate. This parallels the idea that a firm's capital may be fixed in the short run yet variable in the long run.

Another element that may underlie different short-run and long-run responses is that it may take time and experience to 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. Of course, mean weather (μ_i) might be a poor indicator of the actual exposure of a location to a certain weather outcome. For instance, if climate is bi-modal (an alternation of very hot and very cold years), the mean weather might not be very relevant in explaining adaptation. In the empirical section, we thus propose different specifications, but in each case we aim to capture the fact that economic agents tend to adapt to what they are familiar, rather than unfamiliar, with. Finally, note that this specification is admittedly restrictive in the sense 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 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 deviations. If β_3 does not depend on γ_i, then 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:

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

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

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

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; Gourdji et al., 2013; Annan and Schlenker, 2015; Burke et al., 2015; Cooper et al., 2017; Hsiang et al., 2017),[4] 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).

We take the long-run coefficients $\boldsymbol{\beta}^{\text{LR}} = (\beta_1, \beta_2)$ 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). Also note that we

[4] Annan 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).

are only exploring 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

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

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

$$y_{it} = a_i + b_1 x_{it} + e_{it} \tag{5}$$

estimated with the fixed-effects (within) estimator.

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

We are interested in the asymptotic bias as $T \to \infty$ 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

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

and

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

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

$$\text{Bias} = \beta_3 \frac{\sum_i M^i_{\ddot{x}^3}}{\sum_i M^i_{\ddot{x}^2}}. \tag{6}$$

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 ($\sum_t \ddot{x}^3_{it} > 0$), 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 $\ddot{x}^2_{it}$, 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 $\ddot{x}^2_{it}$.[5] 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

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

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.

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

In this setting, the naive model is

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

while the DGP is 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

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

Let us further define

$$M^i_{(\ddot{x}^2 - \ddot{x}^2)^2} \equiv \text{plim}\, T^{-2} \sum_{s,t} \left(\ddot{x}_{is}^2 - \ddot{x}_{it}^2 \right)^2 .$$

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

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

with

$$\begin{aligned} N \quad = \quad & \sum_i M^i_{\ddot{x}^3} \left[\sum_i (\mu_i - \mu) M^i_{\ddot{x}^3} + 2 \sum_i \mu_i (\mu_i - \mu) M^i_{\ddot{x}^2} \right] \\ & -2 \sum_i \mu_i M^i_{\ddot{x}^3} \sum_i (\mu_i - \mu) M^i_{\ddot{x}^2} - \sum_i (\mu_i - \mu) M^i_{\ddot{x}^2} \sum_i M^i_{(\ddot{x}^2 - \ddot{x}^2)^2} \end{aligned} \tag{10}$$

and

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

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 $M^i_{\ddot{x}^3}$, 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 $M^i_{\ddot{x}^3} = 0$.

4.2 Bias when the weather distribution is not skewed

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

$$
\text{Bias}(\mu) = -2\beta_3 \frac{\sum_i\left(\mu_i-\mu\right)M^i_{\ddot{x}^2}\sum_i M^i_{(\ddot{x}^2-\ddot{x}^2)^2}}{\sum_i M^i_{\ddot{x}^2}\sum_i M^i_{(\ddot{x}^2-\ddot{x}^2)^2} + 4\sum_{i,j}\left(\mu_i-\mu_j\right)^2 M^i_{\ddot{x}^2}M^j_{\ddot{x}^2}}. \tag{12}
$$

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 long-run coefficient estimates $\boldsymbol{\beta}^{\text{LR}} = \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}$ and the location-specific short-run estimates $\boldsymbol{\beta}^{\text{SR}}_i = \begin{pmatrix}\beta_1 - 2\beta_3\mu_i\\ \beta_2+\beta_3\end{pmatrix}$:

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

where

$$
\theta = \frac{\sum_i M^i_{\ddot{x}^2}\sum_i M^i_{(\ddot{x}^2-\ddot{x}^2)^2}}{\sum_i M^i_{\ddot{x}^2}\sum_i M^i_{(\ddot{x}^2-\ddot{x}^2)^2} + 4\sum_{i,j}\left(\mu_i-\mu_j\right)^2 M^i_{\ddot{x}^2}M^j_{\ddot{x}^2}} \tag{14}
$$

and

$$
\lambda_i = \frac{M^i_{\ddot{x}^2}}{\sum_j M^j_{\ddot{x}^2}}. \tag{15}
$$

[Figure 3 about here.]

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 it is a weighted average of the short-run and long-run responses. Second, our analysis makes explicit at what particular point one should expect the tangency between the estimated and underlying long-run response to happen: 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 $\sum_i M^i_{\ddot{x}^2}$. 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 predicted of 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

$$\sum_i M^i_{\ddot{x}^2} \sum_i M^i_{(\ddot{x}^2-\ddot{x}^2)^2} << \sum_{i,j} \left(\mu_i - \mu_j\right)^2 M^i_{\ddot{x}^2} M^j_{\ddot{x}^2}. \tag{16}$$

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, $\sum_i M^i_{\ddot{x}^2} > 0$.[6] Also note that $M^i_{\ddot{x}^2} = 0 \Rightarrow M^i_{(\ddot{x}^2-\ddot{x}^2)^2} = 0$, so that if a location displays no time-series variation in weather, its index can be removed from all summations in condition (16). We can therefore limit ourselves to locations for which $M^i_{\ddot{x}^2} > 0$. Condition (16) 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 $M^i_{(\ddot{x}^2-\ddot{x}^2)^2}$, must be small relative to the cross-sectional variation in climate, captured by the terms $\left(\mu_i - \mu_j\right)^2$.

Note that if $M^i_{(\ddot{x}^2-\ddot{x}^2)^2} = 0\ \forall i$, then the identified parameter vector $\hat{\mathbf{b}}$ is consistent for the vector of long-run parameter values. Given the definition of $M^i_{(\ddot{x}^2-\ddot{x}^2)^2}$, this condition is equivalent to saying that, in a given location, weather takes on only two possible values.[7] In that case, the penalty term $\beta_3 \left(x_{it} - \mu_i\right)^2$ 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.[8] 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 ($\ddot{x}^2_{is} \neq \ddot{x}^2_{it}$).[9]

[6]Otherwise, both x_{it} and x^2_{it} 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.

[7]Because then all deviations from the mean are equal in magnitude, which implies $M^i_{(\ddot{x}^2-\ddot{x}^2)^2} = 0$. In all other instances, $M^i_{(\ddot{x}^2-\ddot{x}^2)^2} > 0$.

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

[9]McIntosh and Schlenker (2006) indicate that the naive model will yield the long-run parameter values

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.

One can revisit existing panel estimates by looking at the nature of the weather variation used. Many studies have studied how various outcomes respond to weather in the US context, often relying on county-level data over relatively long time frames (Schlenker and Roberts, 2009; Burke et al., 2015). 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 θ is Equation (14).

5 Joint estimation of short-run and long-run effects

Jointly estimating the short- and long-run responses requires regressing the outcome of interest on a vector of weather variables as well as a measure of how far the values of these variables deviate from the location-specific climate. The previous sections have assumed that climate (μ_i) is fixed over time. By defining climate as a function of past weather, naturally occurring trends in weather variables, like the ones observed in France, provide an opportunity to relax this assumption. This means that trends in climate now enter the identification of the coefficient on the penalty term, which allows for identification of a more flexible model where locations not only differ in their mean outcome through the additive fixed effect α_i but also in how outcomes may respond to weather through the introduction of location-specific slope coefficients β_{1i}.

The specification of the penalty term in Equation (1) implicitly assumes that mean weather is the signal to which agents are responding in the long run. As a result, wherever mean weather (climate) coincides with the current weather realization, the penalty term vanishes. Because mean weather may be a poor indicator of how frequently agents have been exposed to particular conditions in the past, in this section we propose an alternative formulation of the climate-deviation penalty term.

only if $\beta_3 = 0$. Here we have essentially shown that there is no bias either if $\beta_3 \neq 0$ but $\sum_i M^i_{(\ddot{x}^2 - \ddot{x}^2)^2} = 0$.

5.1 Specification A: Squared deviation from mean weather

For a given weather variable, say temperature, climate is defined as a moving average of past weather and the penalty is the squared deviation of the weather variable from climate. The DGP is:

$$y_{it} = \alpha_i + \beta_{1i} x_{it} + \beta_2 x_{it}^2 + \beta_3 (x_{it} - \mu_{it})^2 + \epsilon_{it} \tag{17}$$

where $\mu_{it} = \frac{1}{T} \sum_{k=t-T}^{t-1} x_{ik}$. This specification is similar to Equation (1) except that the climate variable varies over time. The parameter T indicates the period over which the rolling average is computed. This specification implies that at any point in time t, the agent has taken adaptation measures that minimize the yield penalty associated with a weather outcome that is equal to the average weather over the previous T years. The T observations of past weather are weighted equally in the climate calculation. A large T suggests that the agent has a "long memory," with distant observations still playing a role in defining the climate the agent has adapted to. In contrast, shorter T implies that the agent is updating climate expectations faster, with recent conditions playing a larger relative role in the perceived climate. In the application section, we estimate the model for many values of T as well as a special case where we assume $\beta_2 = 0$ (linear long-run relationship).

5.2 Specification B: Mean of squared deviations

Our second specification is defined as:

$$y_{it} = \alpha_i + \beta_{1i} x_{it} + \beta_2 x_{it}^2 + \beta_3 \frac{1}{T} \sum_{k=t-T}^{t-1} (x_{it} - x_{ik})^2 + \epsilon_{it} \tag{18}$$

In this specification, β_3 is multiplied by the average of the squared difference of the current year's weather and each the previous T observations of weather. The penalty term in this model captures the idea that agents are adapting to weather that they have been frequently exposed to, rather than simply to past mean weather. In a sense, this specification is more demanding in terms of what it requires for agents to adapt. For the penalty term to vanish, it no longer suffices that mean past weather matches the current weather occurrence: every single past weather occurrence has to have been "close enough" to the current realization. Again, we can simplify this model by assuming that $\beta_{1i} = \beta_1 \ \forall i$ and that $\beta_2 = 0$.

5.3 Multiplicative heterogeneity

To account for heterogeneity in marginal responses to weather, we allow the slope coefficient β_1 to vary across locations. This generalization is potentially important. Just as the cross-sectional approach pioneered by Mendelsohn et al. (1994) has been shown to be vulnerable to omitted variable bias (Schlenker et al., 2005; Deschênes and Greenstone, 2007), heterogeneity in slope coefficients could lead to bias if not properly accounted for. Imagine for example that soil depth affects crop yields additively but also affects the yield response to temperature. (For instance, because deep soils retain more moisture which dampens the negative effects of heat.) The additive fixed effect α_i absorbs the time-invariant soil depth and therefore removes any bias from omitting soil depth from the estimating equation. However, if soil depth also interacts with x_{it}, this interaction will end up in the error term ϵ_{it} and bias coefficient estimates. Since, in Equation (1), the adaptation coefficient β_3 is only identified from the interaction between weather x_{it} and climate μ_i, inference on the extent of adaptation will be particularly affected. Studying regions that display trends in climate provides an opportunity to control for this confounding effect.

6 Application to French cereal yields

We test for the presence of adaptation in French cereal agriculture using models (17) and (18). We use 67 years of weather and yield data across the 88 departments of mainland France to estimate the short- and long-run responses for wheat and barley. Weather variables include average growing season temperature and cumulative precipitation across the growing season, which is defined as March 1–July 31 for both crops. Data on wheat include only winter wheat yields (spring wheat is a relatively minor crop in France), while barley data represent a harvest-weighted average of winter barley and spring barley yields. It is not obvious how many previous seasons farmers may consider when forming weather expectations, therefore we estimate the model for all definitions of climate between $T = 5$ and $T = 35$ years. Since we utilize past weather to define the penalty term, our sample is restricted to all years greater than 1950+T. In all regressions, we include regional quadratic time trends. (There are 21 regions for 88 departments.)

We present results for both specifications (A and B) of the penalty term discussed in Section 5 and for models that assume either a linear long-run response (model 1, $\beta_2 = 0$ and $\beta_{1i} = \beta_1\ \forall i$), a quadratic long-run response (model 2, $\beta_{1i} = \beta_1\ \forall i$), or a quadratic long-run response that allows for multiplicative heterogeneity (model 3). Although model 3 is technically preferred as it controls for the potentially confounding effect of time-invariant unobservables that interact with weather, it is a demanding model as the purely cross-sectional variation in climate , which potentially matters in a country such as France with a variety of climates, no longer enters the identification. In fact, if climate were stationary, model 3 would not be identified. Focussing on specification A and defining $\bar{\mu}_i = \frac{\sum_t \mu_{it}}{T}$, there are two sources of identification left in model 3: (i) the residual variation in the quadratic climate μ_{it}^2, albeit conditional on the included regional quadratic trends and (ii) the interaction between weather x_{it} and the climate deviation $\mu_{it} - \bar{\mu}_i$.[10]

Following Auffhammer et al. (2013), we report standard errors that are robust to spatial correlation. These standard errors are derived using a variant of the method of Conley (1999) adapted for panel data.[11]

6.1 Evidence of adaptation

If farmers fare significantly better when the climate they are used to is closer to the current weather, this is evidence that they have adapted to climate (because then the past matters). Phrased differently, if exposure to atypical weather leads to lower yields, even after controlling for the weather itself, this is evidence of adaptation. Thus, if we find the coefficient on the penalty term for a given weather variable to be significantly negative, we can conclude that farmers have adapted to climate in that particular dimension. We calculate these adaptation coefficients for average growing season temperature and cumulative growing-season precipitation. The coefficient can be interpreted as the effect of the squared deviation between weather and climate on the logarithm of yield. For example, if the value of the temperature penalty coefficient is -0.01, a 1°C deviation is associated with a 1% decrease in yield and a 2°C deviation from climate is associated with a 4% decrease in yield.

[Figure 4 about here.]

[Figure 5 about here.]

Panels in Figures 4 and 5 show coefficients on the temperature penalty term for wheat for specifications A and B, respectively. The x-axis shows the number of years

[10]For specification B, source (ii) is replaced by the residual variation in the mean $\frac{\sum_{k=t-T}^{t-1} x_{ik}^2}{T}$ conditional on the regional trends.

[11]Specifically, we allow for spatial correlation across neighboring departments up to the third degree, using a weight function with a geometric decay rate.

used to compute climate and the y-axis displays the coefficient value. Looking at specification A, we find significant adaptation across all models for all definitions of climate longer than five years. The results across models are quite similar, with values ranging from -0.04 to slightly positive (though insignificant) depending on the climate definition length. For most definitions of climate, the coefficients lie between -0.01 and -0.03.

The results for specification B are comparable. For all models, we find significant penalty coefficients for nearly all climate definition lengths. Unsurprisingly, for both specifications, we find that model 3 has slightly wider confidence intervals than the models that do not allow for multiplicative heterogeneity.

[Figure 6 about here.]

[Figure 7 about here.]

We also find evidence of adaptation in barley yields. Figures 6 and 7 show coefficients on the temperature penalty term for barley for specifications A and B, respectively. For models 1 and 2 we find that the penalty coefficient is always negative, ranging from -0.04 to -0.01, depending on the definition of climate. For model 3, we find slightly more negative penalty coefficients, ranging from -0.05 to -0.02. Barley results for specification B are very similar, although slightly less negative.

[Figure 8 about here.]

[Figure 9 about here.]

[Figure 10 about here.]

[Figure 11 about here.]

Figures 8 through 11 show results for precipitation. For wheat, the results of both specifications are quite similar. For both specifications, models 1 and 2 are comparable with estimates ranging between -2.5e-6 and -1.0e-6. An estimate of -2.0e-6 would imply that a 100 mm deviation from the expected precipitation level is associated with a 2% decrease in yield conditional on current weather. For both specifications, model 3 results are less stable across climate definitions, with only borderline significant estimates for climate definitions less than 20 years. For longer definitions, the coefficient is negative and generally significant, with the sole exception being specification A with a climate definition length of 35 years, which is negative but insignificant. For longer climate lengths the coefficient can be as large as -1.0e-5, implying that an unexpected 100m shock in precipitation is associated with a yield decrease of 10%. Barley precipitation results are comparable to those of wheat, particularly for models 1 and 2. Model 3 results are slightly different, with some significantly positive estimates for short climate definitions, but negative and significant coefficients for all climate definitions larger than 27 years.

6.2 Yield response curves

For specification A, we also provide short- and long-run temperature response curves for all models. These responses define climate as a 15-year rolling average of weather. Figure 12 shows the responses for wheat. The long-run response is shown in red. The short-run responses for each of the 88 departments are calculated based on the department's average weather over the reference period 1992–2016 and are shown below the common long-run response. The short-run marginal effect of temperature for each department, evaluated at that department's climate during the reference period, is the slope at the tangency point between that department's short- and long-run curves. Short-run responses are colored according to their mean growing season temperature during the reference period. Departments with mean temperatures below the 25^{th} percentile of departments are green, departments with mean temperature above the 75^{th} percentile are blue, the short-run response of the department with median temperature is orange, and the remainder are black. Curves are normalized vertically so that the median department's short-run curve is tangent to the long-run curve at 0.

[Figure 12 about here.]

For model 1, the long-run curve is downward sloping, implying that the marginal effect of average growing season temperature on yield is negative. For model 2, the quadratic long-run response model, we find that the long-run response is convex, with all short-run curves' point of tangency lying on the downward-sloping portion. This finding suggests that the generalization from model 1 to model 2 does not lead to meaningful changes in terms of inference, and explains why the estimates of the adaptation parameter $\hat{b}_3$ were generally comparable between these models.

The representation of model 3 is somewhat more involved. Since β_2 is not specific to a department, the overall shape of all departments' long-run response curve is identical. However, as β_{1i} is unique to each department, these curves no longer lie atop each other because a change in β_1 also shifts the long-run curve along the x-axis. In order to represent how the marginal effect of temperature varies across departments, we've shifted back all of the departments' long-run curves both vertically and horizontally, so that they are one and the same. Thus, the x-axis labels are useful only for drawing inference about the median department (indicated in orange). However, the distribution of the departments' short-run response relative to its long-run response is preserved. Note that in model 3, unlike models 1 and 2, it is not necessarily the case that departments with a warmer climate have a more negative marginal response to temperature.

For wheat, we find that the long-run response is concave, with the majority of departments having a negative marginal effect of average growing season temperature. For barley, shown in Figure 13, we again find that the linear long-run response estimated in model 1 is downward sloping and the quadratic long-run response estimated in model 2 is convex. In both cases, the marginal effect of average growing season temperature is negative for all departments. For model 3, our estimates imply a convex long-run curve, with a majority of departments lying on the downward-sloping region.

[Figure 13 about here.]

6.3 Climate warming impacts

We calculate short- and long-run impact estimates for a uniform 3°C warming for both crops, both specifications, and all models for many definitions of climate length. All impacts are relative to a climate defined as the average weather over the period 1992–2016, to which we assume agents have fully adapted. We construct a counterfactual climate where every department is 3°C warmer than in the reference climate.[12] For the short-run impact calculation, we use all of the temperature estimates, along with the hypothetical weather observations, to compute the change in yield for each department. For the penalty term, the coefficient (β_3) is multiplied by 9 (the square of 3°C). For the long-run case, the 3°C change is assumed to be wholly expected, and thus the penalty term drops out.

[12]Due to the nonlinearity of the long-run relationship, the reference climate matters for the calculation of the predicted yield change in a given department. Total impacts are computed by calculating the predicted value of yield in each department under both reference and projection climates and multiplying by the department's average crop area over the reference period to get predicted output.

[Figure 14 about here.]

[Figure 15 about here.]

Figures 14 and 15 show the impact calculations for wheat for specifications A and B respectively. For specification A models 1 and 2, both effects are significantly negative in nearly all cases, with the long-run impacts generally significantly less negative than the short-run impact. On average, the short-run impact for model 1 is -18% compared to an average long-term impact of -7%. For model 2, the average short-run impact is -18% compared to an average long-term impact of -5%. These values are nearly identical to those found using specification B. For model 3, for both specifications we find that the long-run impact is significantly negative for shorter climate definitions, but is generally not statistically distinguishable from zero for longer lengths under either specification. The short-run impact is significantly negative for all climate definitions, but is not always significantly different from the long-run impact. Averaging across climate definition lengths, the short-run impact for specification A (resp. specification B) is -17% (resp. -16%) and the long-impact is -3% (resp. -4%).

[Figure 16 about here.]

[Figure 17 about here.]

Figures 16 and 17 show the impacts for barley. Under both specifications, we find the model 1 and 2 results to be very similar. For climate definition lengths less than 30 years, both impacts are significantly negative, although the short-run impacts are far more severe. In the long run, impacts range from 0% to -14%, depending on specification and model. In contrast, short-run impacts range from -15% to -32%.

For model 3, for specification A (resp. B), the long-run result is indistinguishable from zero for nearly all climate definitions less than 28 (resp. 25) years, and significantly positive for the longest climate definitions. The short-run impacts range from -6% to -31%.

7 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. Our application to French cereal yields extends the existing framework to allow for richer unobserved heterogeneity and suggests long-run adaptation by farmers, crop breeders, and institutions that is both statistically and biologically significant.

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.

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

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.

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*.

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. (2005). Will U.S. Agriculture Really Benefit from Global Warming? Accounting for Irrigation in the Hedonic Approach. *The American Economic Review*, 95(1):395–406.

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.

Appendices

A Derivation of the asymptotic bias

A.1 Useful expressions

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

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

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

A.2 Linear model

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

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

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

Our strong exogeneity assumption (Assumption 3) implies that $\operatorname{plim} T^{-1}\sum_t \ddot{x}_{it}\ddot{\epsilon}_{it} = 0$. In addition, $\operatorname{plim} \bar{x}_i = \mu_i$, therefore we have

$$\operatorname{plim} \hat{b}_1 = \beta_1 + \beta_3\frac{\sum_i M^i_{\ddot{x}^3}}{\sum_i M^i_{\ddot{x}^2}}.$$

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

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

Denote $\mathbf{M}_{\mathbf{XX}} \equiv \operatorname{plim} T^{-1}\ddot{\mathbf{X}}'\ddot{\mathbf{X}}$. Assumption 3 implies that $\operatorname{plim} T^{-1}\ddot{\mathbf{X}}'\ddot{\boldsymbol{\varepsilon}} = 0$. Therefore, we have

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

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

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

with

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

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

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

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

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

and

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

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

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

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

$$\begin{aligned} N \;=\; & \left[\sum_i M^i_{\ddot{x}^3}\right]\left[\sum_i \left(\mu_i-\mu\right) M^i_{\ddot{x}^3} + 2\sum_i \mu_i \left(\mu_i-\mu\right) M^i_{\ddot{x}^2}\right] \\ & -2\left[\sum_i \mu_i M^i_{\ddot{x}^3}\right]\left[\sum_i \left(\mu_i-\mu\right) M^i_{\ddot{x}^2}\right] - \left[\sum_i \left(\mu_i-\mu\right) M^i_{\ddot{x}^2}\right]\left[\sum_i \left(M^i_{\ddot{x}^4} - \left(M^i_{\ddot{x}^2}\right)^2\right)\right]. \end{aligned}$$

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

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

where the last equality obtains because each term $\ddot{x}^4_{it}$ appears $T-1$ times in the summation (each time index t is paired with one of the $T-1$ remaining indices). Defining $M^i_{(\ddot{x}^2-\ddot{x}^2)^2} \equiv \operatorname{plim} T^{-2}\sum_{s,t}\left(\ddot{x}^2_{is} - \ddot{x}^2_{it}\right)^2$ leads to expression (10) in the main text. Using a similar argument, we can rewrite

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

It is straightforward to show that the last two terms in the previous expression simplify to $4\sum_{i,j}\left(\mu_i-\mu_j\right)^2 M^i_{\ddot{x}^2} M^j_{\ddot{x}^2}$, 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 $M^i_{\ddot{x}^3} = 0\ \forall i$

Assume that $M^i_{\ddot{x}^3} = 0\ \forall i$. 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 $\boldsymbol{\beta}^{\text{SR}}_i = \begin{pmatrix} \beta_1 - 2\beta_3\mu_i \\ \beta_2 + \beta_3 \end{pmatrix}$. We seek to write $\hat{\mathbf{b}}$ as a convex combination of the underlying parameters, that is, $\hat{\mathbf{b}} = (1-\theta)\boldsymbol{\beta}^{\text{LR}} + \theta \sum_i \lambda_i \boldsymbol{\beta}^{\text{SR}}_i$ with $0 \leq \theta \leq 1$, $0 \leq \lambda_i \leq 1$, and $\sum_i \lambda_i = 1$. Such decomposition must satisfy $\theta \sum_i \lambda_i \mu_i = \frac{N_1}{D}$ and $\theta = 1 - 2\frac{N_2}{D}$. Using $M^i_{\ddot{x}^3} = 0\ \forall i$, we obtain Equations (14) and (15).

Although the decomposition also holds in the general case where $M^i_{\ddot{x}^3}$ may be nonzero, the decomposition is meaningless as it is no longer possible to ensure that the weights lie between zero and one.

List of Figures

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

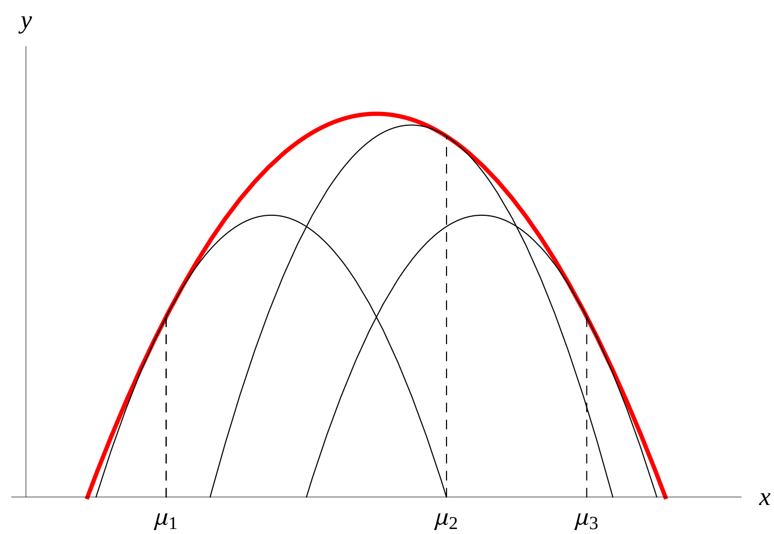
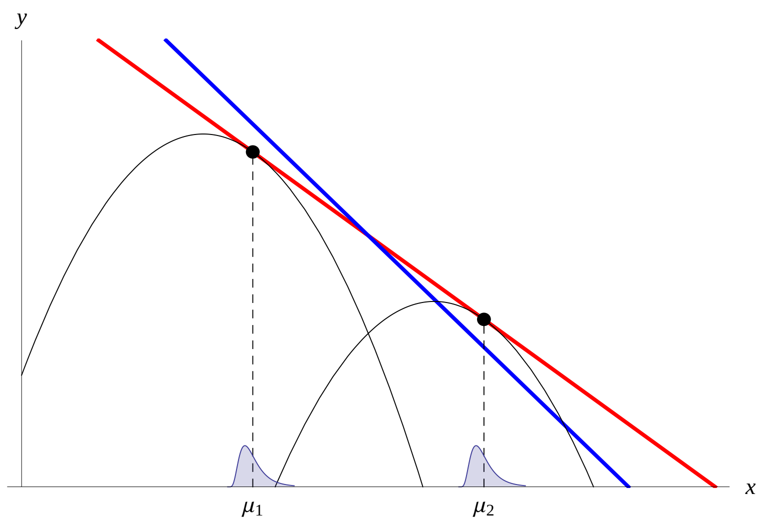
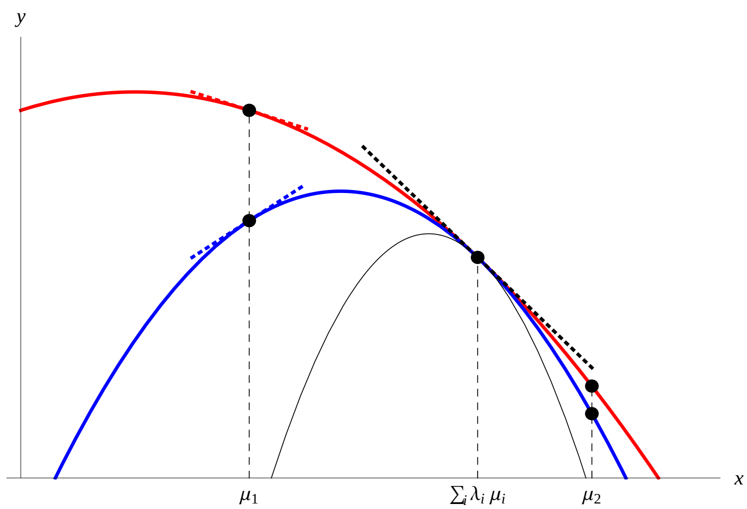


Figure 2 Asymptotic bias when weather fluctuations are right-skewed



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 Bias on counterfactual impact of climate change



Note: Red curve: underlying long-run response. Blue curve: estimated response. Black curve: underlying short-run response for climate $\bar{\mu} = \sum_i \lambda_i \mu_i$.

Figure 4 Wheat (Specification A): Temperature penalty coefficient

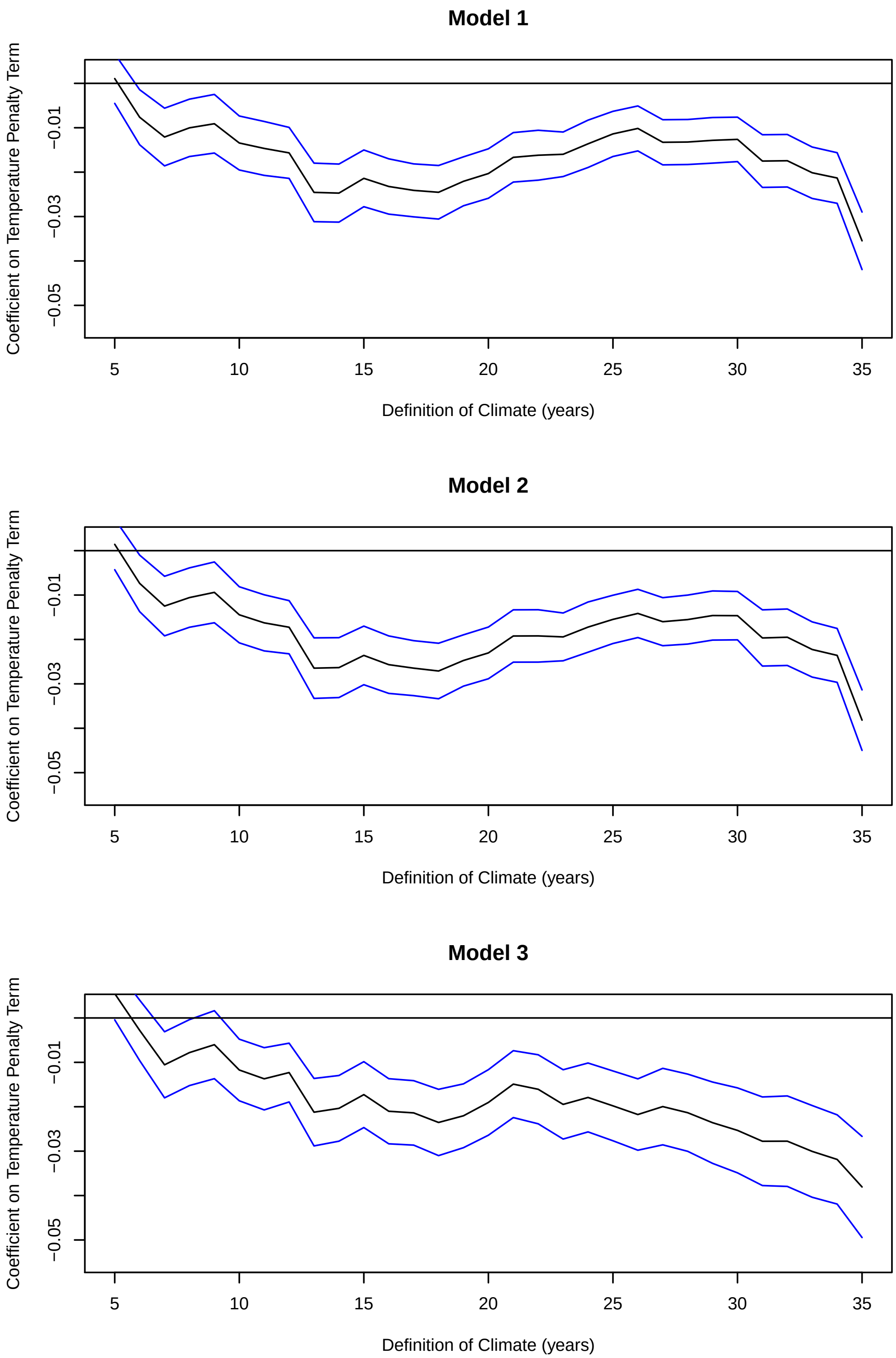


Figure 5 Wheat (Specification B): Temperature penalty coefficient

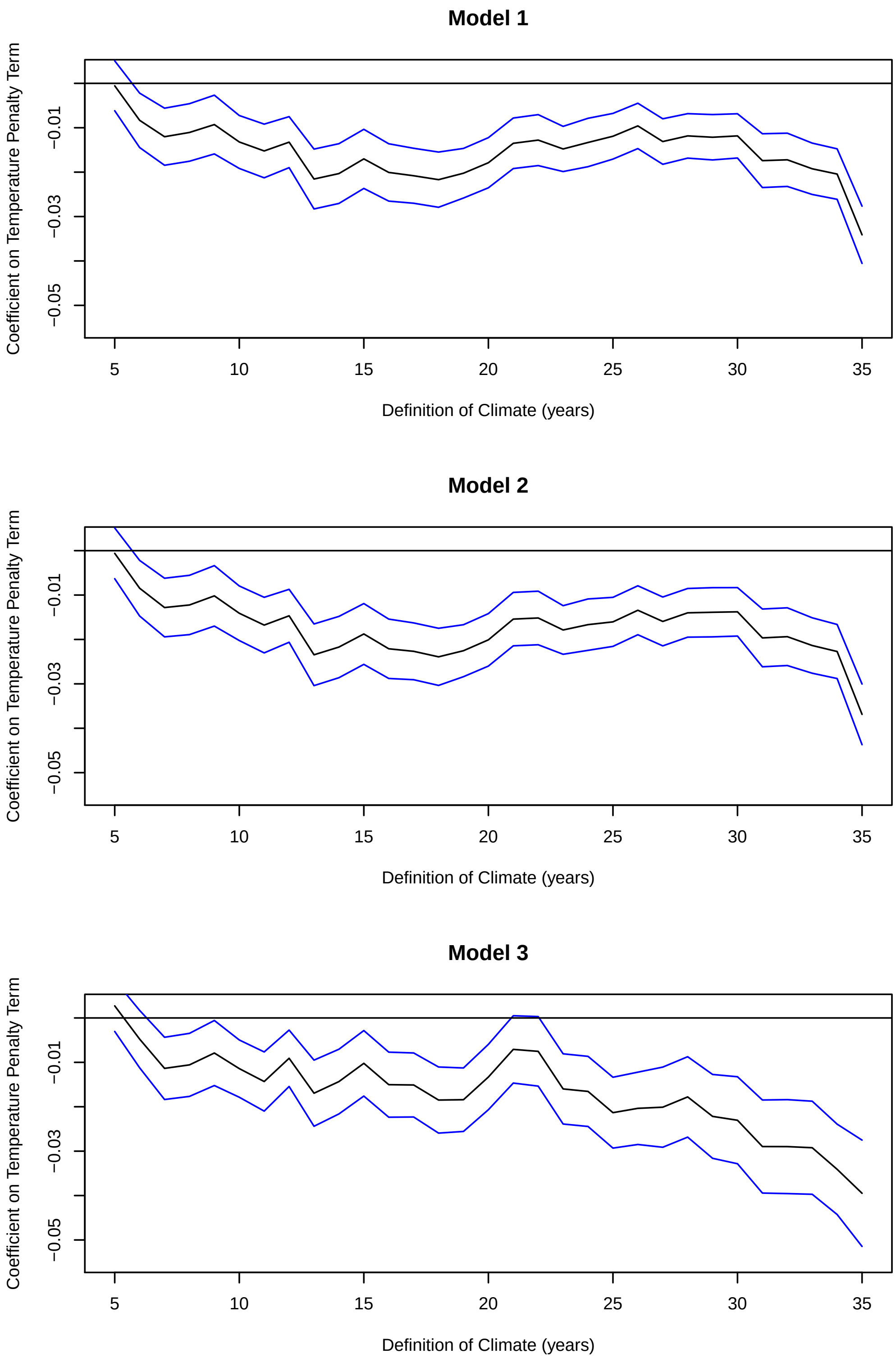


Figure 6 Barley (Specification A): Temperature penalty coefficient

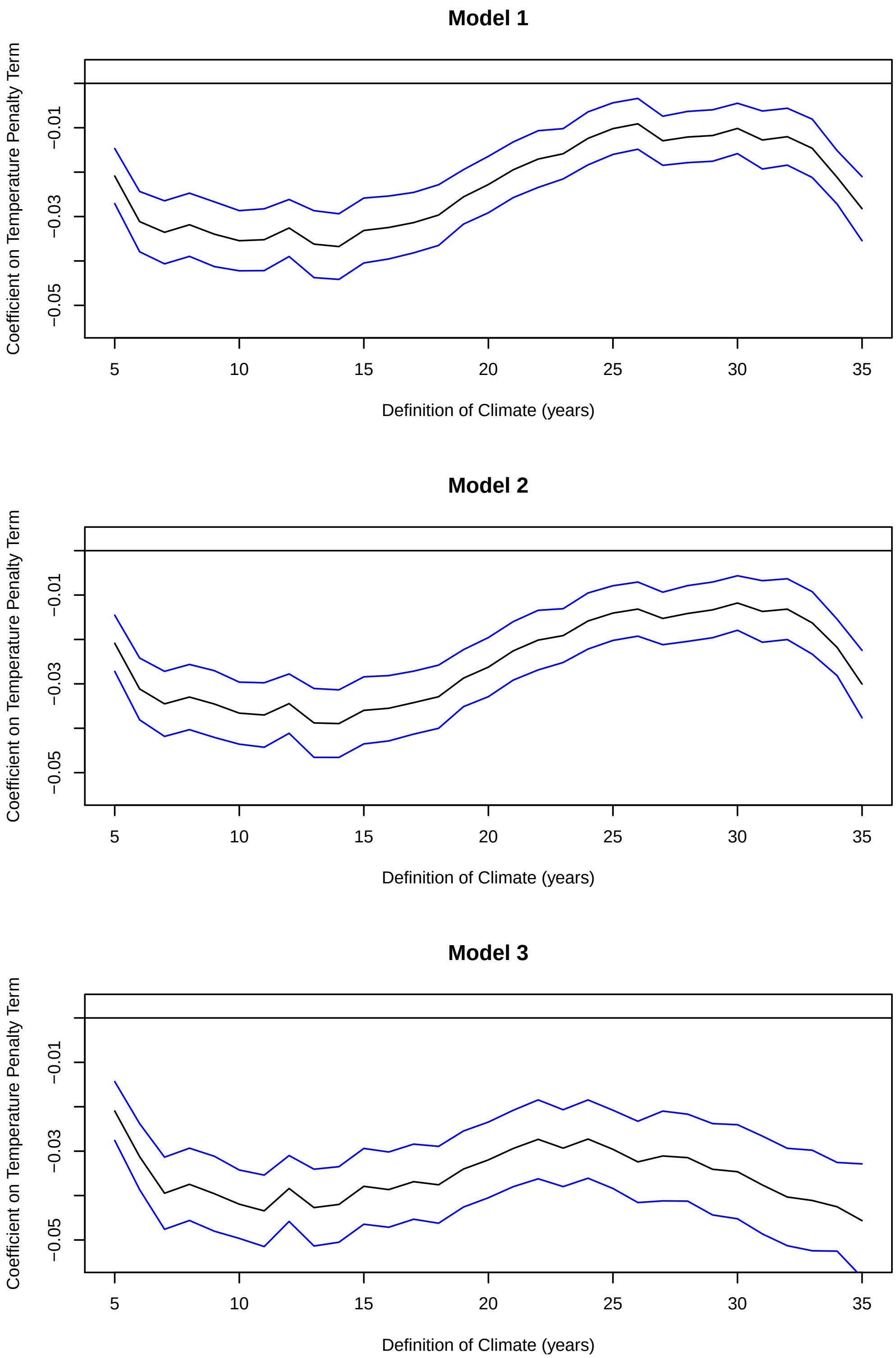


Figure 7 Barley (Specification B): Temperature penalty coefficient

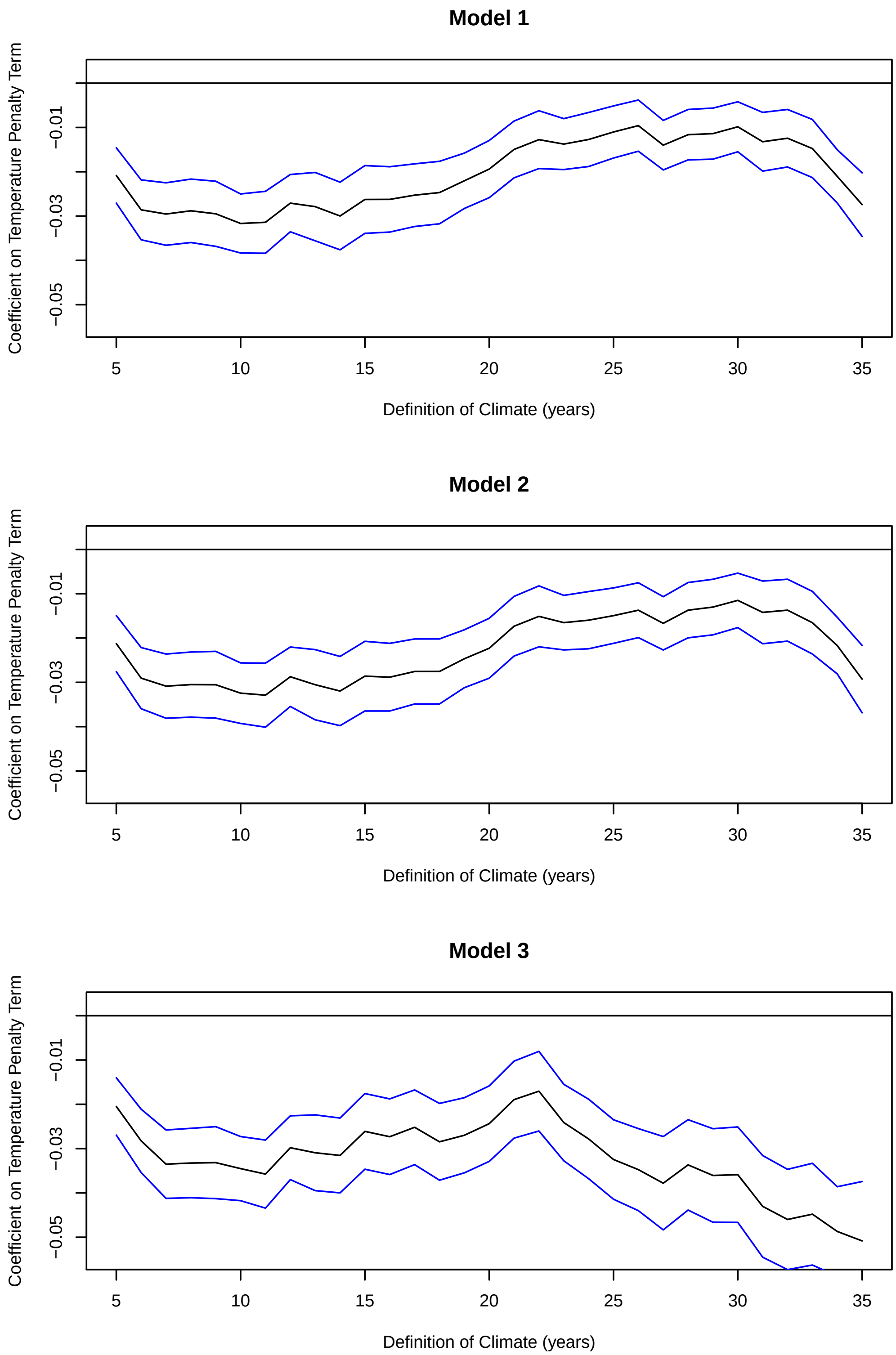


Figure 8 Wheat (Specification A): Precipitation penalty coefficient

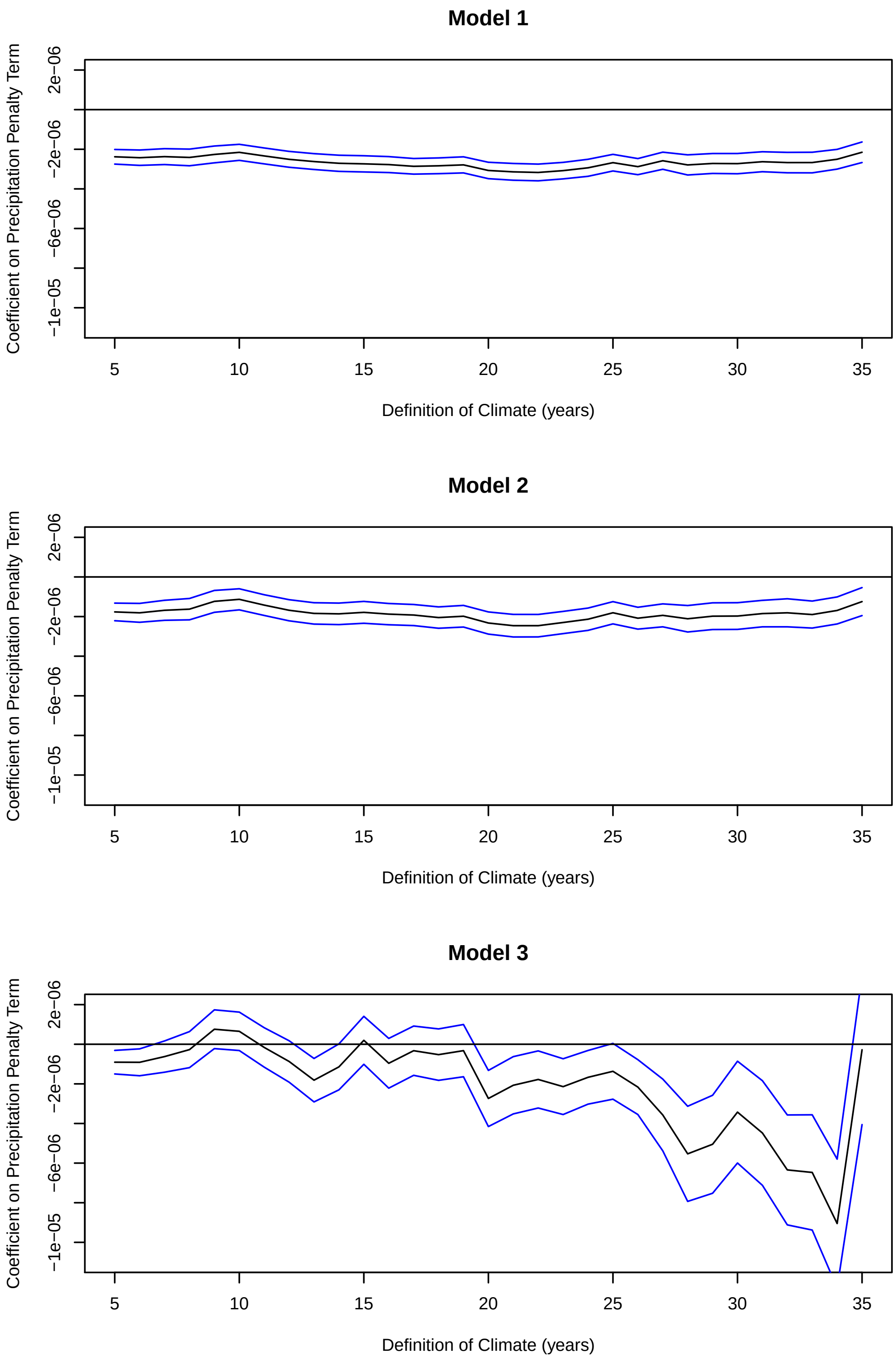


Figure 9 Wheat (Specification B): Precipitation penalty coefficient

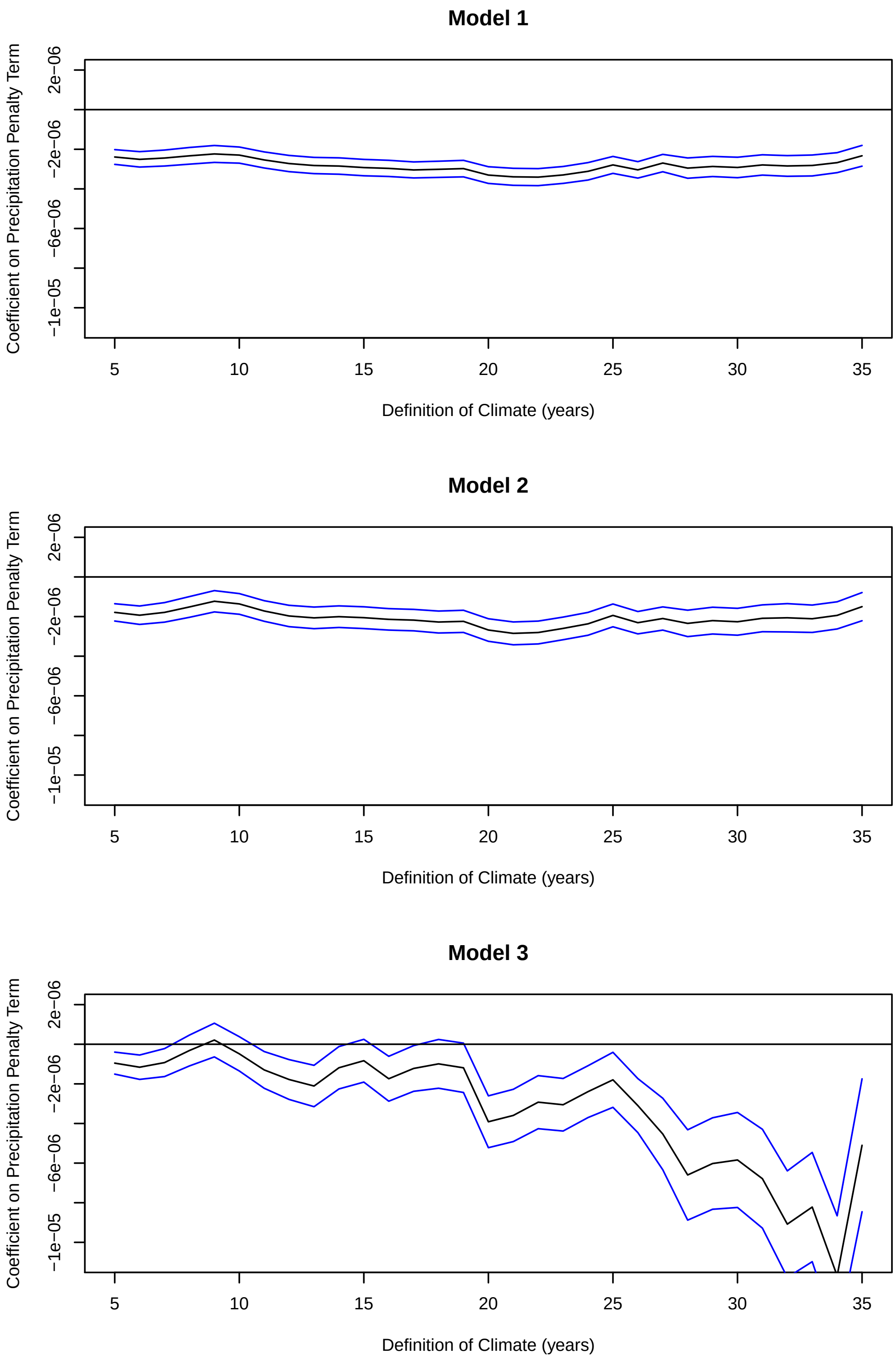


Figure 10 Barley (Specification A): Precipitation penalty coefficient

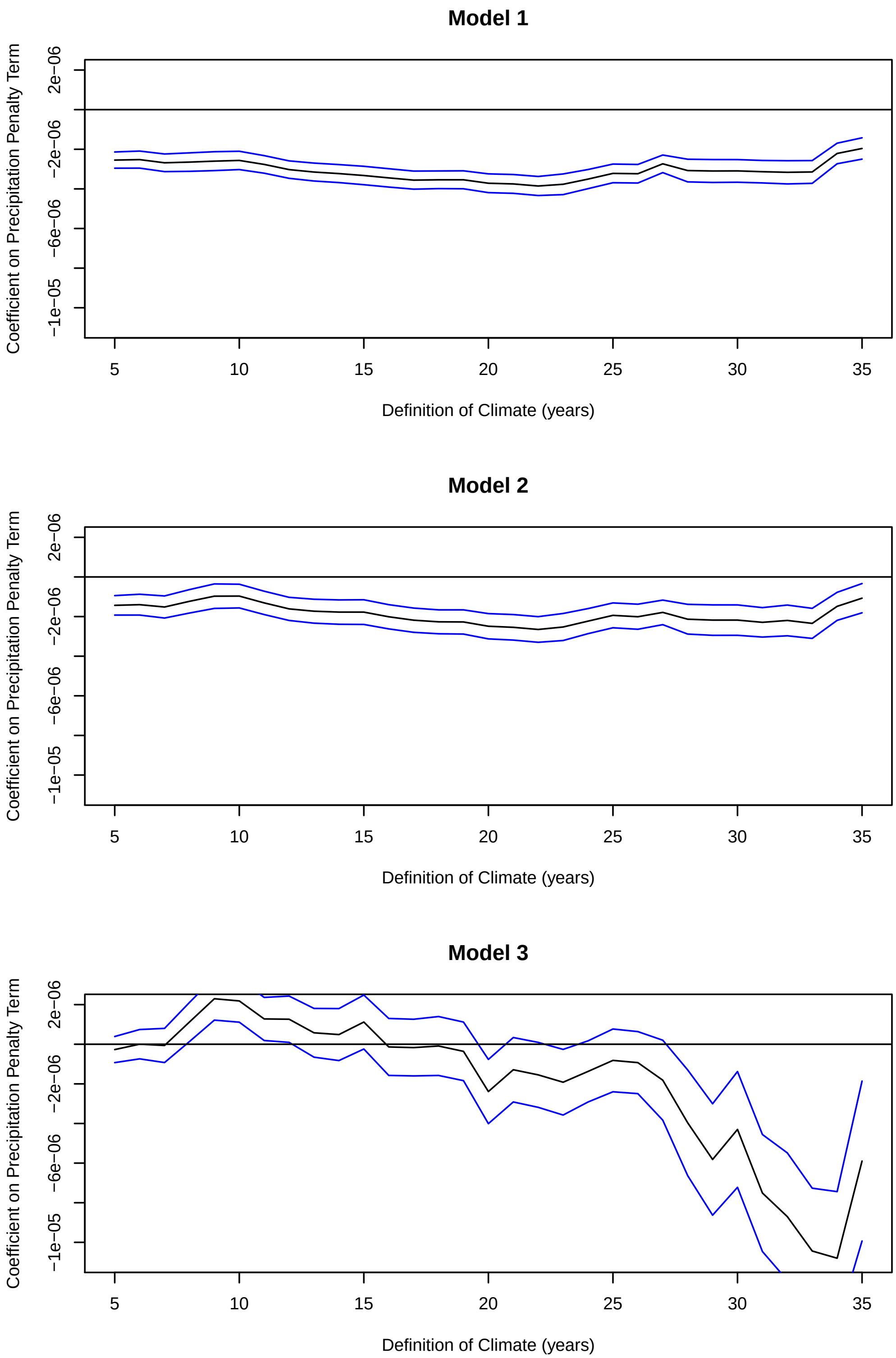


Figure 11 Barley (Specification B): Precipitation penalty coefficient

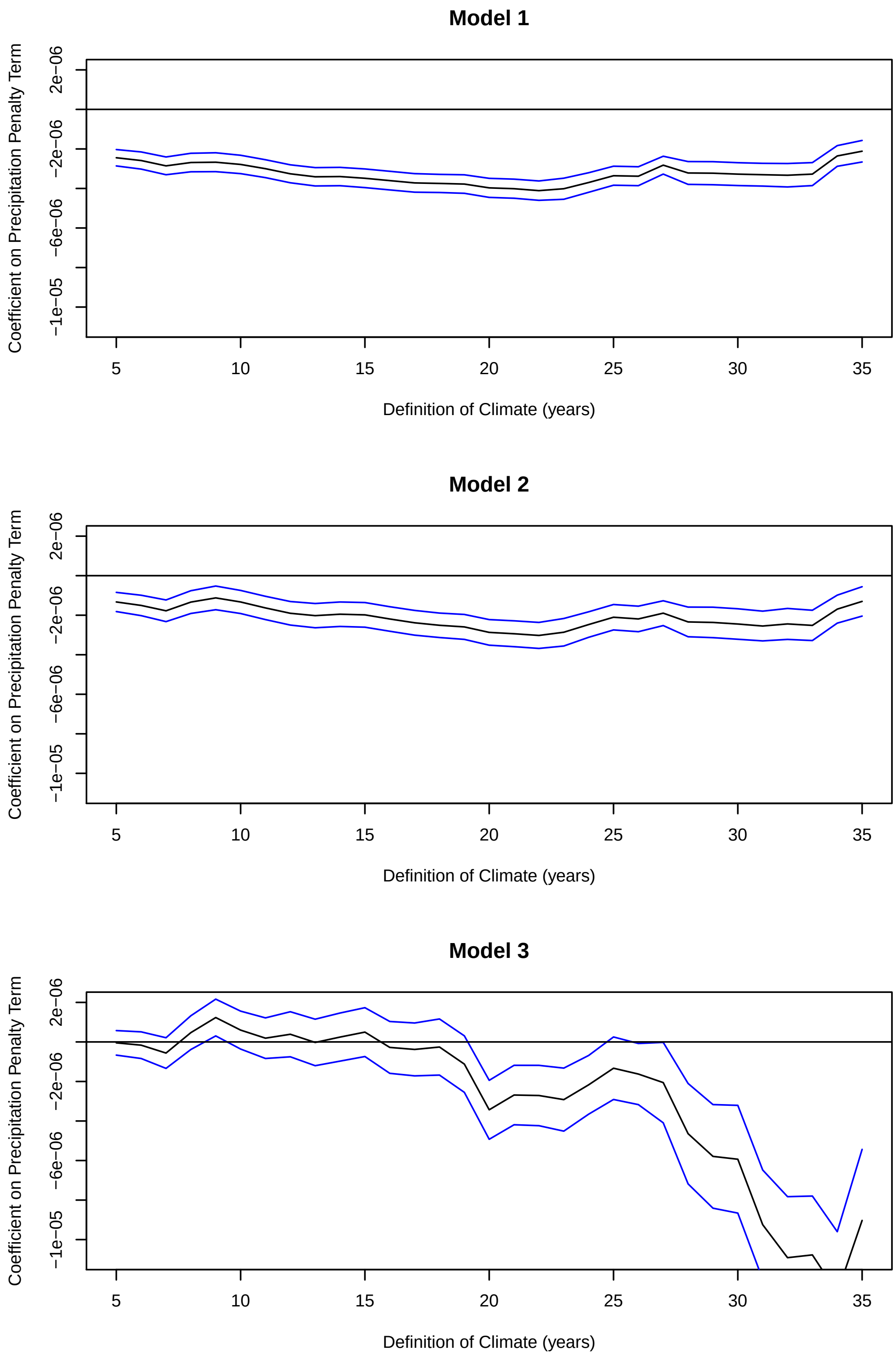


Figure 12 Wheat responses to average temperature

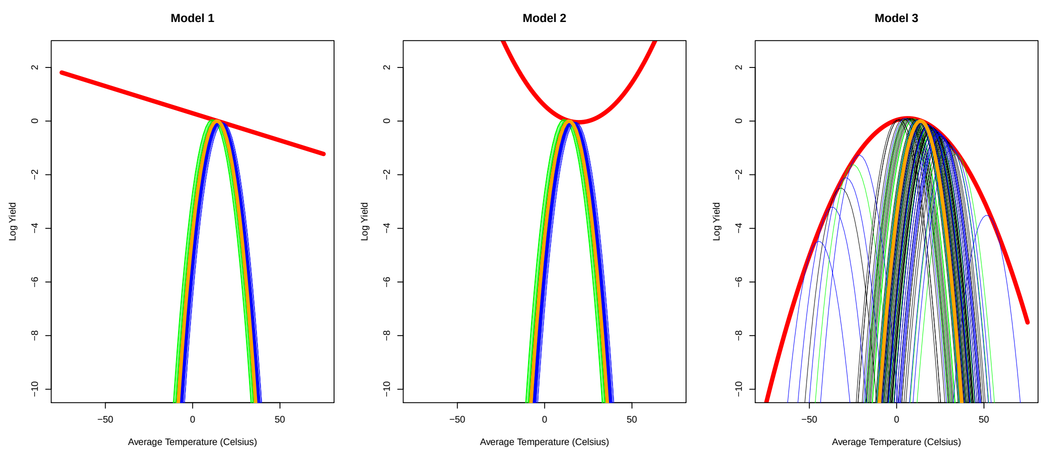


Figure 13 Barley responses to average temperature

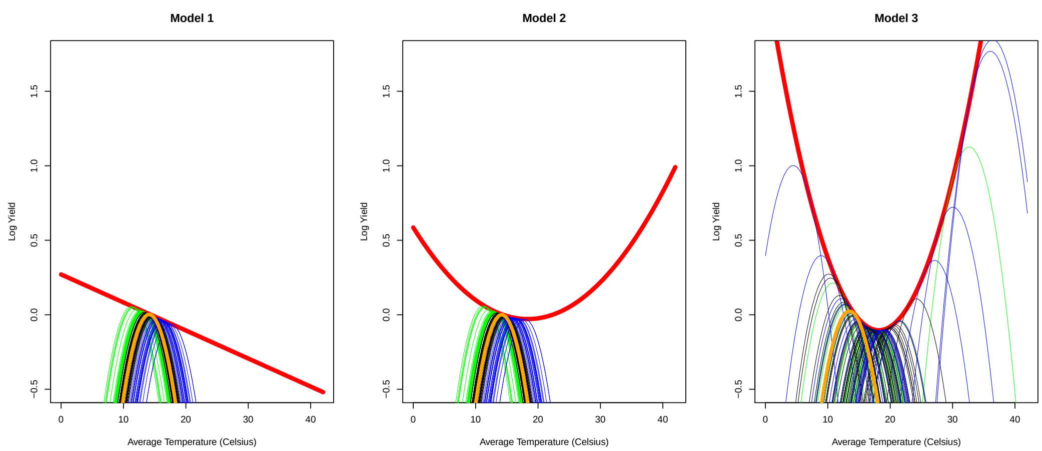


Figure 14 Wheat (Specification A): Impact of a 3°C warming

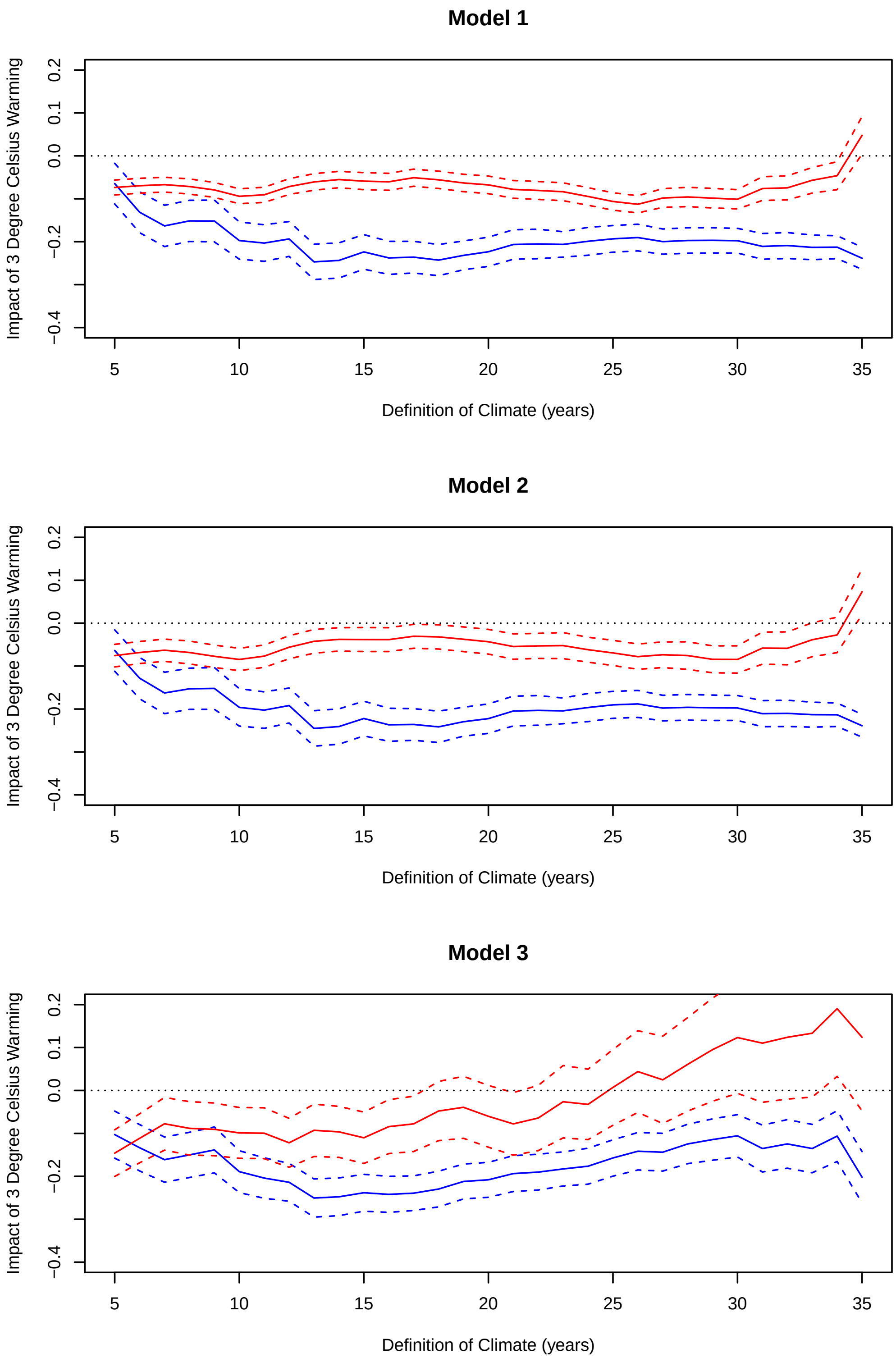


Figure 15 Wheat (Specification B): Impact of a 3°C warming

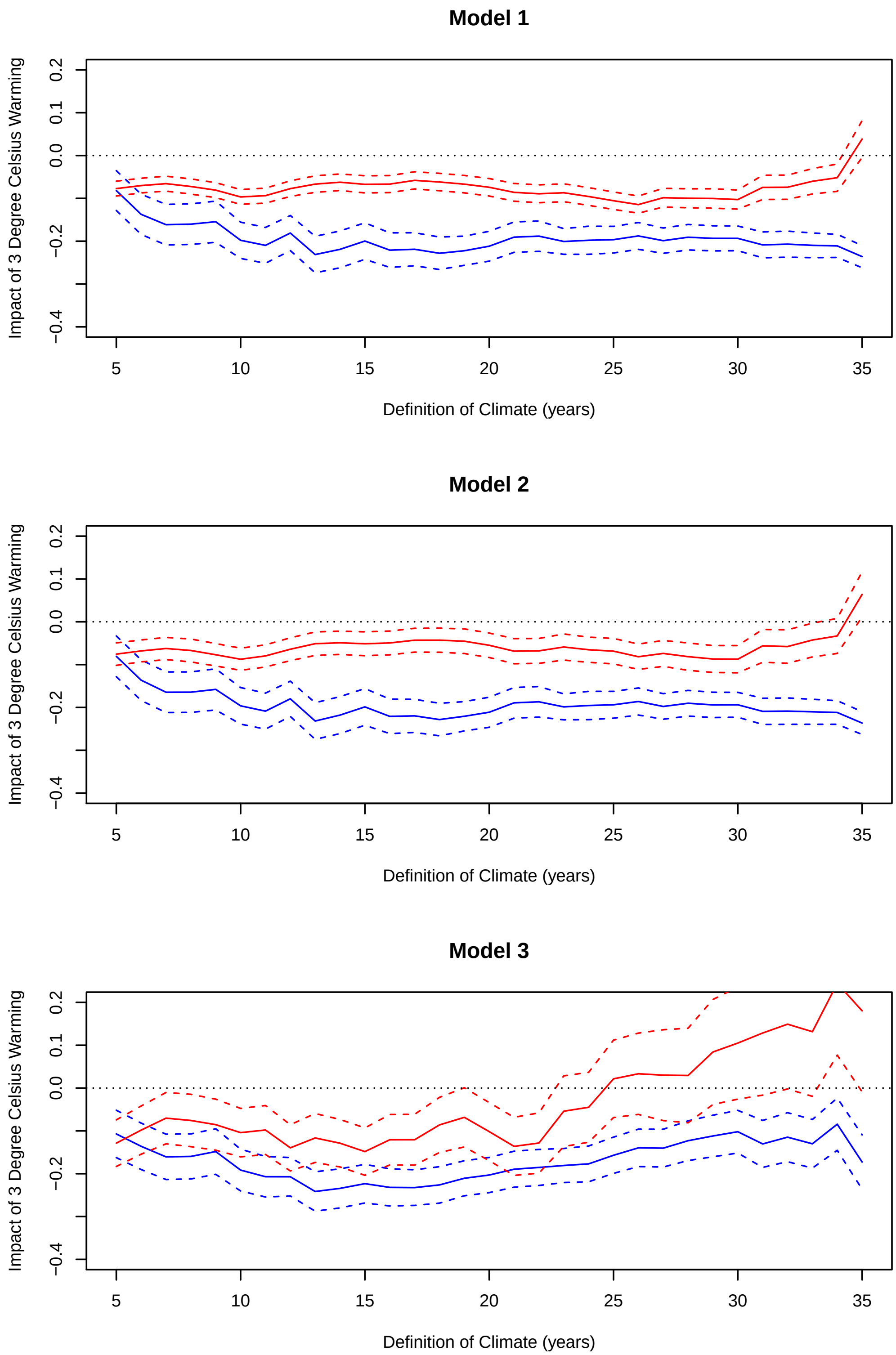


Figure 16 Barley (Specification A): Impact of 3°C warming

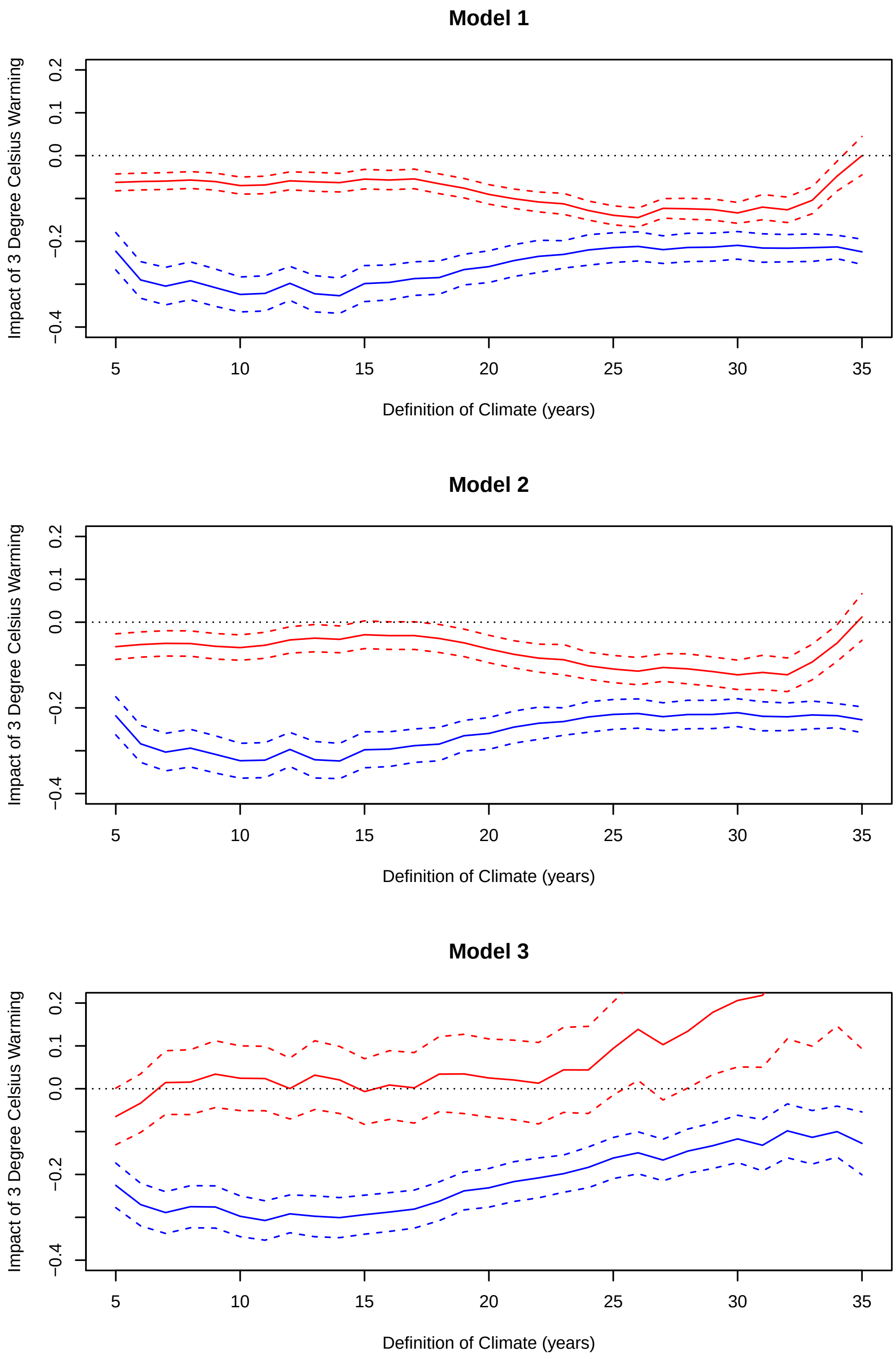


Figure 17 Barley (Specification B): Impact of 3°C warming