Estimating the Consequences of Climate Change from Variation in Weather*

Derek Lemoine

University of Arizona and NBER
McClelland Hall 401, 1130 E Helen St, Tucson, AZ, 85721-0108, USA
dlemoine@email.arizona.edu

University of Arizona Working Paper 18-09

November 2019

First version: August 2018

I formally relate the consequences of climate change to the time series variation in weather extensively explored by recent empirical literature. I show that reduced-form fixed effects estimators can recover the effects of climate if agents are myopic, if agents' payoff functions belong to a particular class, or if the actions agents take in each period do not depend on actions taken in previous periods. I also show how to recover structural estimates of climate change impacts from reduced-form weather regressions. Applying this new method, I find that an additional 2°C of global warming would reduce eastern U.S. agricultural profits by around 50% under the median estimates.

JEL: C23, Q12, Q51, Q54

Keywords: climate, weather, adaptation, forecasts, expectations, adjustment, agriculture

*I thank seminar participants at the 2019 Spring Meeting of the National Bureau of Economic Research Energy and Environment Program, the Bank of Canada Climate Change and Central Banking Workshop, the Queen's Institute of Intergovernmental Relations State of the Federation Workshop, Arizona State University, Auburn University, University of Alabama, University of Arizona, University of California Berkeley, University of Chicago, University of Massachusetts, University of Southern California, and University of Wisconsin for helpful comments. I thank Wolfram Schlenker for kindly providing data and Olivier Deschênes for helpful conversations. An early version circulated as "Sufficient Statistics for the Cost of Climate Change."

1 Introduction

A pressing research agenda seeks to estimate the economic costs of climate change. Ignorance of these costs has severely hampered economists' ability to evaluate policy. Recognizing that different locations have different climates, many economists have hoped to estimate the effects of climate change from the spatial correlation between climate and outcomes of interest (e.g., Mendelsohn et al., 1994; Schlenker et al., 2005; Nordhaus, 2006). However, any two locations differ along many dimensions, leading to concerns about omitted variables bias.[1] Fortunately, though, the same location does experience different weather at different times. Stimulated by Deschênes and Greenstone (2007), an explosively growing empirical literature estimates the consequences of a location happening to experience cooler-than-average or hotter-than-average weather.[2] These researchers project the consequences of climate change by combining their credibly estimated effects of weather with scientists' predictions about how climate change will alter the distribution of weather. Whether the estimated weather treatment is in fact a good proxy for the unobserved climate treatment has been the subject of much debate but little analysis.[3]

I here undertake the first formal analysis that precisely delineates what and how we can learn about climate impacts from weather impacts. A change in climate differs from a weather shock in being repeated period after period and in affecting expectations of weather far out into the future. Linking weather to climate therefore requires analyzing a dynamic model that captures the distinction between transient and permanent changes in weather. I study an agent (equivalently, firm) who is exposed to stochastic weather outcomes. The agent chooses actions (equivalently, investments) that suit the weather. Actions can be responses to realized weather ("ex-post adaptation") or can be proactive investments against future weather ("ex-ante adaptation"). The actions chosen in different periods may be complements or substitutes: when actions are intertemporal complements, choosing a high action in the previous period reduces the cost of choosing a high action today, but when actions are intertemporal substitutes, choosing a high action in the previous period increases the cost of choosing a high action today. The first case is consistent with adjustment costs, and the second case is consistent with actions that require scarce resources.[4] When choosing actions,

[1] See Dell et al. (2014) and Auffhammer (2018b) for expositions and Massetti and Mendelsohn (2018) for a review.

[2] For recent reviews, see Dell et al. (2014), Carleton and Hsiang (2016), and Heal and Park (2016). Blanc and Schlenker (2017) discuss the strengths and weaknesses of relying on panel variation in weather. Few doubt that weather shocks are as-good-as-randomly assigned. For instance, Dell et al. (2014, 741) write that "the primary advantage of the new literature is identification", and Blanc and Schlenker (2017, 262) describe "weather anomalies" as "ideal right-hand side variables" because "they are random and exogenous".

[3] For instance, Dell et al. (2014, 771–772) emphasize that "short-run changes over annual or other relatively brief periods are not necessarily analogous to the long-run changes in average weather patterns that may occur with climate change." And Mendelsohn (2019, 272) observes, "An important failing of current weather panel studies is that they lack a clear theoretical model."

[4] Both types of stories exist in the literature. For instance, in studies of the agricultural impacts of climate

the agent knows the current weather, has access to forecasts of future weather, and relies on knowledge of the climate to generate forecasts of weather at longer horizons. A change in the climate alters the distribution of potential weather outcomes as well as the agent's expectation of future weather outcomes.

I derive the effects of climate change in terms of model primitives and express reduced-form fixed effects estimators in terms of these same model primitives. I show that reduced-form estimates of weather impacts can exactly recover the theory-implied effects of climate change on payoffs in a few special cases. Two are of particular interest. First, if actions are neither intertemporal substitutes nor intertemporal complements (so that current decisions are not directly affected by previous decisions), then empirical researchers can recover the effects of climate on actions by combining the estimated effects on actions of current weather, lagged weather, and forecasts. And because empirical researchers can recover the effects of climate on actions, they can also recover the effect of climate on payoffs. Second, researchers do not need to recover effects on actions if agents' payoff functions satisfy a particular condition. Empirical researchers can then recover the effects of climate from especially simple regressions that are consistent with standard practice to date.

However, many applications will not satisfy the special cases. I therefore also extend conventional regression frameworks to recover structural estimates of climate impacts. Because I formally derive the reduced-form regression coefficients in terms of model primitives, I can recover combinations of model primitives from these coefficients and then calculate the theory-implied effects of climate change.[5] The identification is exactly the same as in the recent reduced-form literature (relying on within-unit variation in weather), but the regression specification and the use of the estimated coefficients both differ from standard practice.

I apply this new method to the seminal analysis of climate and agriculture from Deschênes and Greenstone (2007). The results suggest that the special cases do not hold in this application. I find that adaptation offsets some of the costs of extreme heat in the short run, but because adaptation imposes its own costs, adaptation adds to the costs of extreme heat in the long run. In contrast, adaptation to non-extreme heat imposes costs in the short run in exchange for adding to the long-run benefits of additional days with non-extreme

change, Deschênes and Greenstone (2007) conjecture that long-run adjustments to changes in climate should be greater than short-run adjustments to weather shocks because there may be costs to adjusting crops, whereas Fisher et al. (2012) and Blanc and Schlenker (2017) conjecture that constraints on storage and groundwater pumping, respectively, could make short-run adjustments exceed long-run adjustments.

[5]This approach is in the spirit of Marschak's Maxim. Heckman (2010, 359) writes, "Marschak's Maxim suggests that economists should solve well-posed economic problems with minimal assumptions. All that is required to conduct many policy analyses or to answer many well-posed economic questions are policy invariant combinations of the structural parameters that are often much easier to identify than the individual parameters themselves and that do not require knowledge of individual structural parameters." It is also related to sufficient statistics approaches (see Chetty, 2009) and to price theory (see Weyl, 2019). Throughout, I use "reduced-form" and "structural" in the way now common in empirical work rather than in accord with their original usage.

heat. Most adaptation is ex post, but there is evidence of ex-ante adaptation to extreme heat in more recent years. In total, the costs of additional days with extreme heat outweigh the benefits of more days with non-extreme heat: the median estimates suggest that the current century's warming (in the RCP 4.5 scenario of stabilized emissions) would reduce agricultural profits by 56% in the absence of adaptation and by 50% if agents adapt as they do to annual weather shocks.

But will agents display more or less adaptation to climate than to weather? The estimated regression coefficients imply that actions are intertemporal substitutes, as in resource scarcity stories. This finding is consistent with recent empirical results in agricultural economics (Hendricks et al., 2014; Kim and Moschini, 2018) and with implications of crop rotation dynamics (Eckstein, 1984) but is contrary to widespread intuition based on Le Châtelier's principle. According to the theoretical analysis, agents will undertake more adaptation to short-run weather shocks than to long-run climate change. I can therefore bound the effects of climate change by the no-adaptation and full-adaptation estimates. The median estimates suggest losses of 50–56% ($20–22 billion annually, in year 2017 dollars) from warming over the century.

There has been remarkably little formal analysis of the economic link between weather and climate, despite the importance of empirically estimating the costs of climate change and the sharpness of informal debates around the relevance of the recent empirical literature to climate change. The primary exception is an argument given in Hsiang (2016) and repeated in Deryugina and Hsiang (2017). The argument stipulates that a change in climate differs from a change in weather only by affecting beliefs about future weather. This difference in beliefs can matter for payoffs only if it affects an agent's chosen actions. However, the envelope theorem tells us that an optimizing agent's actions cannot have first-order consequences for payoffs. Therefore the effects of weather on payoffs exactly—and generically—identify the effects of climate on payoffs.

By formalizing the distinction between climate and weather in a dynamic environment, the present analysis highlights two weak points in this argument.[6] First, it is true that a change in climate alters beliefs about future weather, but it is also true that a change in climate alters past weather and past actions. Past actions are predetermined variables from the perspective of an optimizing agent and thus do not drop out through the envelope theorem. Even myopic agents can respond differently to weather and climate. Second, in a dynamic model, the envelope theorem applies to the intertemporal value function, not to the per-period payoff function investigated by much empirical work. Optimized current actions can have first-order effects on current payoffs when those are offset by first-order effects on

[6] In an earlier expositional analysis, I showed how envelope theorem arguments can fail in a three-period model (Lemoine, 2017). The present work precisely analyzes the consequences of climate change in an infinite-horizon model, constructively shows which types of empirical estimates can be informative about the climate, and develops and applies a new approach to structurally estimating the consequences of climate change.

expected future payoffs. Only in a special class of payoff functions will optimized current actions not have first-order effects on future payoffs.

A few other lines of research are related. First, calibrated numerical simulations have shown that dynamic responses are critical to the effects of climate on timber markets (Sohngen and Mendelsohn, 1998; Guo and Costello, 2013) and to the cost of increased cyclone risk (Bakkensen and Barrage, 2018). I develop a general analytic setting that precisely disentangles several types of dynamic responses and relates them to widely used fixed effects estimators. Second, empirical work has shown that agents use forecasts of future weather, even at seasonal scales. In particular, Shrader (2017) and Taraz (2017) use variation in seasonal forecasts and in past years' weather outcomes, respectively, to identify ex-ante adaptation to weather events. I formally demonstrate that estimating responses to forecasts and lagged weather is critical to recovering the consequences of climate change. Finally, Kelly et al. (2005) and Kala (2017) study learning about the climate from observed weather. I here abstract from learning in order to focus on mechanisms more relevant to the recent empirical literature.

The challenge of attempting to estimate long-run effects from short-run variation is a common one in empirical economics. To get around this challenge, environmental economists have found policy-induced variation in long-run pollution exposure that is plausibly exogenous to health outcomes (e.g., Chen et al., 2013; Anderson, 2015; Barreca et al., 2017; Bishop et al., 2018). Unfortunately, this type of variation may not be available to researchers interested in the consequences of changing the climate. Labor economists desire the long-run consequences of changing the minimum wage, but inflation converts observed minimum wage increases into short-run shocks (Sorkin, 2015).[7] And macroeconomists formerly hoped to learn about long-run output-inflation tradeoffs by estimating distributed lag models, but Lucas (1972) argued that, when agents have rational expectations, the lagged response to a transient inflation shock is not informative about the long-run effects of permanently changing inflation policy. Here we desire the long-run effect of changing the policy rule used by nature to generate weather.

The next section describes the setting. Section 3 derives the theory-implied effect of climate. Section 4 establishes conditions under which the effect of climate can be recovered from reduced-form estimates of weather impacts. Section 5 develops the new method of structurally estimating climate impacts and applies it to U.S. agriculture. The final section describes potential extensions. The appendix contains empirical details, proofs, and

[7] Three other papers are related to both Sorkin (2015) and the present paper's project. First, I here formalize analogues to arguments in Hamermesh (1995) about why the pre- and post-periods around a minimum wage increase are not true pre- and post-periods. Second, in a model of dynamic stock accumulation, Hennessy and Strebulaev (2019) show that estimated responses to transient shocks can differ substantially from the theory-implied causal effects that empirical researchers seek to test. The present paper is similar in deriving sufficient conditions for estimated effects to match theory-implied effects. Third, Keane and Wolpin (2002) describe tradeoffs between cross-sectional and time series variation when estimating the effects of welfare benefits. These tradeoffs are similar to those that motivate the present paper.

additional results, including formal analysis of recently popular long difference estimators.

2 Setting

An agent is repeatedly exposed to stochastic weather outcomes and takes actions based on realized weather and information about future weather. The realized weather in period t is w_t and the agent's chosen action is A_t. This action may be interpreted as a level of activity (e.g., time spent outdoors, energy used for heating or cooling, irrigation applied to a field) or as a stock of capital (e.g., outdoor gear, size or efficiency of furnace, number or efficiency of irrigation lines). The agent's time t payoffs are given by the twice-differentiable function $\pi(A_t, A_{t-1}, w_t, w_{t-1}) : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.[8] Letting subscripts indicate partial derivatives, I assume $\pi_{11} < 0$ and $\pi_{22} \leq 0$, implying declining marginal benefits of current and past actions.

I interpret actions as adaptations that become more valuable with high weather outcomes ($\pi_{13}, \pi_{23} \geq 0$). Following terminology from the literature on climate adaptation (e.g., Fankhauser et al., 1999; Mendelsohn, 2000), a case with $\pi_{13} > 0$ reflects adaptation that can occur after weather is realized ("reactive" or "ex-post" adaptation) and a case with $\pi_{23} > 0$ reflects adaptation that can occur before weather is realized ("anticipatory" or "ex-ante" adaptation).[9] I allow adaptation to play both roles at once. The possibility that $\pi_4 \neq 0$ reflects potential delayed impacts from the previous period's weather, with π_{14} and π_{24} capturing the potential for ex-post adaptation to alter these delayed impacts. Consistent with the normalizations above, I assume $\pi_{14}, \pi_{24} \geq 0$. Finally, observe that the actions could reflect a firm's production responses to price signals rather than responses to weather per se. In this interpretation, the normalizations imply that "high" weather outcomes increase the price of a firm's output or reduce the cost of its input.

I allow π_{12} to be positive or negative, with its magnitude constrained as described below. When $\pi_{12} < 0$, actions are "intertemporal substitutes", so that choosing a higher level of past actions increases the cost of choosing higher actions today. I describe this case as a resource scarcity story.[10] For instance, pumping groundwater today raises the cost of pumping groundwater tomorrow (see Blanc and Schlenker, 2017) or rescheduling activities around today's weather makes it hard to reschedule activities around tomorrow's weather (see Graff Zivin and Neidell, 2009). When $\pi_{12} > 0$, actions are "intertemporal complements",

[8] I generalize to vector-valued actions and multidimensional weather in the appendix. Doing so yields little new insight at the expense of exposition.

[9] When interpreting actions as the choice of capital stock, the payoff function is consistent with standard models of depreciation. If we restrict the payoff function to allow only ex-ante adaptation, then the setting corresponds to a time-to-build model with a one-period lag. Section 6 discusses the implications of capital stocks that take longer to build.

[10] Relating to the literature on resource extraction, the case with $\pi_{12} < 0$ can be seen as reflecting stock-dependent extraction costs (Heal, 1976).

so that choosing a higher level of past actions increases the benefit from choosing higher actions today. I describe this case as an adjustment cost story.[11] For instance, small changes to cropping practices or activity schedules may be easier to implement than large changes. The magnitude of π_{12} affects the agent's preferred timing of adaptation. As $|\pi_{12}|$ becomes large, the agent prefers to begin adapting before the weather event arrives, but when $|\pi_{12}|$ is small, the agent may wait to undertake most adaptation only once the weather event has arrived.[12]

The agent observes time t weather before selecting her time t action. The agent also understands the climate C, which controls the distribution of weather. We can interpret weather as realized temperature and climate as a location's long-run average temperature. At all times before $t-1$, the agent's only information about time t weather consists in knowledge of the climate. However, at time $t-1$ the agent receives a forecast f_{t-1} of time t weather: $f_{t-1} = C + \zeta\nu_{t-1}$, where the innovation ν_{t-1} is a mean-zero, serially uncorrelated random variable with variance $\tau^2 > 0$. The forecast is an unbiased predictor of time t weather: $w_t = f_{t-1} + \zeta\epsilon_t$, where ϵ_t is a mean-zero, serially uncorrelated random variable with variance $\sigma^2 > 0$.[13] The parameter $\zeta \geq 0$ is a perturbation parameter that will be useful for analysis (see Judd, 1996). The covariance between ϵ_t and ν_t is ρ. The covariance between w_t and w_{t-1} is then $\zeta^2\rho$. The agent incorporates knowledge of such serial correlation in her forecasts.

The agent maximizes the present value of payoffs over an infinite horizon:

$$\max_{\{A_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t E_0 \left[\pi(A_t, A_{t-1}, w_t, w_{t-1})\right],$$

where $\beta \in [0, 1)$ is the per-period discount factor, A_{-1} is given, and E_0 denotes expectations

[11] The benchmark quadratic adjustment cost model has $\pi_{12} = k$ for some $k > 0$ (see Hamermesh and Pfann, 1996).

[12] The magnitude of π_{12} is related to the distinction between ex-post and ex-ante adaptation insofar as it affects the agent's preferred timing of adaptation actions. However, π_{12} incentivizes early adaptation only to reduce the costs of later adaptation, not because early adaptation provides protection from weather events. I reserve the terms ex-ante and ex-post adaptation to refer to the effects of actions on the marginal benefit of weather, captured by π_{13}, π_{23}, π_{14}, and π_{24}.

[13] Consistent with much previous literature, climate here controls average weather. One might wonder about the dependence of higher moments of the weather distribution on climate. In fact, the effects of climate change on the variance of the weather are poorly understood and likely to be spatially heterogeneous (e.g., Huntingford et al., 2013; Lemoine and Kapnick, 2016). Further, for economic analysis, we need to know not just how climate change affects the variance of realized weather but how it affects the forecastability of weather: the variance of the weather more than one period ahead is $\zeta^2(\sigma^2 + \tau^2)$, so we need to apportion any change in variance between σ^2 and τ^2. I leave such an extension to future work.

at the time 0 information set. The solution satisfies the following Bellman equation:

$$\begin{aligned} V(A_{t-1}, w_t, f_t, w_{t-1}; \zeta) =& \max_{A_t} \left\{ \pi(A_t, A_{t-1}, w_t, w_{t-1}) + \beta E_t \left[V(A_t, w_{t+1}, f_{t+1}, w_t; \zeta) \right] \right\} \\ \text{s.t.} \quad w_{t+1} =& f_t + \zeta \epsilon_{t+1} \\ f_{t+1} =& C + \zeta \nu_{t+1}. \end{aligned}$$

Time t payoffs do depend indirectly on actions and weather prior to time $t-1$ because earlier actions affect time $t-1$ actions.

The setting is sufficiently general to describe many applications of interest. For instance, much empirical literature has studied the effects of weather on energy use. The agent could then be choosing indoor temperature in each period, where payoffs depend on current actions through energy use and depend on weather through thermal comfort. Habituation to outdoor temperatures is captured by π_{14}. Much empirical work has also studied the effect of weather on labor productivity. The decision variable could be effort, the dependence of payoffs on weather could reflect current thermal stress as well as the effects of the previous day's weather via sleep and physiological functioning, the resource scarcity is one of tasks needing to be done, and forecasts allow the agent to plan tasks and vacation time around weather outcomes. Finally, many researchers have studied the effects of weather on agricultural outcomes. In that case, payoffs are profits, actions include planting decisions, and weather affects yields.

I will often impose one of the following two assumptions:

Assumption 1. ζ^2 *is small.*

Assumption 2. π *is quadratic.*

Either assumption will limit the consequences of stochasticity for optimal policy, whether by limiting the variance of weather outcomes (Assumption 1) or by making the policy function independent of that variance (Assumption 2).[14]

I will be interested in empirical researchers' ability to estimate the consequences of altering C from observable responses to time series variation in w_t and f_t. It is important to be clear about the climate experiment. I study the effects of a change in climate on an agent who has had time to adapt to the new climate. This climate change treatment is consistent with the exercise common in the empirical literature, which calculates the effect of replacing today's distribution of weather with a distribution projected to hold by the end of the century. I will not study how the transition from one climate to another interacts with agents' decisions,[15] and I will not study how expectations of a future change in climate

[14]Note that when applying Assumption 2, the chosen policy is affected by the variance of weather (through the realized weather) even though the policy rule is independent of that variance.

[15]Kelly et al. (2005) frame the cost of learning as an adjustment cost. Quiggin and Horowitz (1999, 2003) discuss broader costs of adjusting to a change in climate. These papers' adjustment costs are conceptually

affect agents today.[16] These are both important questions but are beyond the scope of the present analysis—and thus far largely beyond the empirical literature that this analysis seeks to inform.

3 Theory-Implied Effect of Climate Change

I now derive the exact effect of climate change on long-run payoffs and actions within this model. I later explore how to estimate these effects from observable variation in weather.

The analysis approximates the solution to the full, stochastic model around the steady state of the deterministic model, which sets $\zeta = 0$ (Judd, 1996). The first-order condition for the deterministic model is:

$$0 = \pi_1(A_t, A_{t-1}, C, C) + \beta V_1(A_t, C, C, C; 0).$$

The envelope theorem yields:

$$V_1(A_{t-1}, C, C, C; 0) = \pi_2(A_t, A_{t-1}, C, C).$$

Advancing this forward by one timestep and substituting into the first-order condition, we have the Euler equation:

$$0 = \pi_1(A_t, A_{t-1}, C, C) + \beta\pi_2(A_{t+1}, A_t, C, C). \tag{1}$$

A steady state $\bar{A}$ of the deterministic system is implicitly defined by

$$0 = \pi_1(\bar{A}, \bar{A}, C, C) + \beta\pi_2(\bar{A}, \bar{A}, C, C). \tag{2}$$

Define $\bar{\pi} \triangleq \pi(\bar{A}, \bar{A}, C, C)$. The following lemma describes the uniqueness and stability of the steady state.[17]

Lemma 1. *For $\bar{\pi}_{12} \neq 0$ and $\zeta = 0$, $\bar{A}$ is locally saddle-path stable if and only if $(1+\beta)|\bar{\pi}_{12}| < -\bar{\pi}_{11} - \beta\bar{\pi}_{22}$, in which case $\bar{A}$ is unique. For $\bar{\pi}_{12} = 0$ and $\zeta = 0$, the agent chooses $A_t = \bar{A}$ at all times t.*

distinct from the adjustment costs studied here. The present use of "adjustment costs" follows much other economics literature in referring to the cost of changing decisions from their previous levels. I study how these adjustment costs hinder estimation of the consequences of climate change from weather impacts, not how they affect the cost of transitioning from one climate to another.

[16] Severen et al. (2018) show that land markets capitalize expectations of future climate change and correct cross-sectional analyses in the tradition of Mendelsohn et al. (1994) for this effect. I here study responses to widely available, shorter-run forecasts in a time series context and show how to use them to improve panel analyses in the tradition of Deschênes and Greenstone (2007).

[17] The steady state exists if $\sup_{\bar{A}}[\pi_1(\bar{A}, \bar{A}, C, C) + \beta\pi_2(\bar{A}, \bar{A}, C, C)] > 0$ and $\inf_{\bar{A}}[\pi_1(\bar{A}, \bar{A}, C, C) + \beta\pi_2(\bar{A}, \bar{A}, C, C)] < 0$. These conditions will hold in most applications.

Proof. See appendix. □

I henceforth assume that $(1+\beta)|\bar{\pi}_{12}| < -\bar{\pi}_{11} - \beta\bar{\pi}_{22}$, so that the deterministic steady state is unique and saddle-path stable.

Now consider optimal actions in the stochastic system. The first-order condition is:

$$0 = \pi_1(A_t, A_{t-1}, w_t, w_{t-1}) + \beta E_t[V_1(A_t, w_{t+1}, f_{t+1}, w_t; \zeta)].$$

The envelope theorem yields:

$$V_1(A_{t-1}, w_t, f_t, w_{t-1}; \zeta) = \pi_2(A_t, A_{t-1}, w_t, w_{t-1}).$$

Advancing this forward by one timestep and substituting into the first-order condition, we have the stochastic Euler equation:

$$0 = \pi_1(A_t, A_{t-1}, w_t, w_{t-1}) + \beta E_t[\pi_2(A_{t+1}, A_t, w_{t+1}, w_t)]. \tag{3}$$

The following lemma describes the evolution of $E_0[A_t]$.

Lemma 2. *Let either Assumption 1 or 2 hold, and let $E_0[(A_1 - \bar{A})^2]$ be small. Then $\lim_{t\to\infty} E_0[A_t] = \bar{A}$.*

Proof. See appendix. □

When the conditions of Lemma 2 hold, applying the implicit function theorem to equation (2) yields:

$$\lim_{t\to\infty} \frac{\mathrm{d}E_0[A_t]}{\mathrm{d}C} = \frac{\mathrm{d}\bar{A}}{\mathrm{d}C} = \frac{\overbrace{\bar{\pi}_{13} + \bar{\pi}_{14} + \beta\bar{\pi}_{24}}^{\text{ex-post}} + \overbrace{\beta\bar{\pi}_{23}}^{\text{ex-ante}}}{-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}} \geq 0. \tag{4}$$

This is the average long-run effect of climate change on actions. Expected future actions increase in the climate index because I normalize high actions to be more beneficial when the weather index is high. Equation (4) captures how climate change alters weather in all periods: the past, the present, and the future. We see the various forms of ex-post adaptation captured by $\bar{\pi}_{13}$, $\bar{\pi}_{14}$, and $\beta\bar{\pi}_{24}$. We also see the possibility of ex-ante adaptation, controlled by $\bar{\pi}_{23}$ and arising because the agent understands that the altered climate affects weather in subsequent periods. Finally, observe that $\bar{\pi}_{12}$ enters through the denominator in (4). When actions are intertemporal substitutes ($\bar{\pi}_{12} < 0$), this term reduces the magnitude of the response to climate change, as when resource scarcity makes long-run responses smaller than short-run responses. However, when actions are intertemporal complements ($\bar{\pi}_{12} > 0$), this term increases the magnitude of the response to climate change, as when adjustment costs allow long-run responses to exceed short-run responses.

Approximating the payoff function around the steady state, $w_t = w_{t-1} = C$, and $\zeta = 0$ and using either Assumption 1 or Assumption 2, we have:

$$\begin{aligned} E_0[\pi(A_t, A_{t-1}, w_t, w_{t-1})] =& \bar{\pi} + \bar{\pi}_1(E_0[A_t] - \bar{A}) + \bar{\pi}_2(E_0[A_{t-1}] - \bar{A}) \\ &+ \frac{1}{2}\bar{\pi}_{11}E_0[(A_t - \bar{A})^2] + \frac{1}{2}\bar{\pi}_{22}E_0[(A_{t-1} - \bar{A})^2] + \frac{1}{2}(\bar{\pi}_{33} + \bar{\pi}_{44})\zeta^2(\sigma^2 + \tau^2) \\ &+ \bar{\pi}_{12}E_0[(A_t - \bar{A})(A_{t-1} - \bar{A})] + \bar{\pi}_{13}Cov_0[A_t, w_t] + \bar{\pi}_{23}Cov_0[A_{t-1}, w_t] \\ &+ \bar{\pi}_{14}Cov_0[A_t, w_{t-1}] + \bar{\pi}_{24}Cov_0[A_{t-1}, w_{t-1}] + \bar{\pi}_{34}\zeta^2\rho, \end{aligned} \tag{5}$$

for $t > 1$. Differentiating equation (5) with respect to C and applying either Assumption 1 or Assumption 2 again, we find that

$$\lim_{t\to\infty} \frac{\mathrm{d}E_0[\pi(A_t, A_{t-1}, w_t, w_{t-1})]}{\mathrm{d}C} = \bar{\pi}_3 + \bar{\pi}_4 + [\bar{\pi}_1 + \bar{\pi}_2]\frac{\mathrm{d}\bar{A}}{\mathrm{d}C}. \tag{6}$$

The marginal effect of climate on long-run payoffs is composed of the direct effect of a larger weather index, in both the present ($\bar{\pi}_3$) and the past ($\bar{\pi}_4$), and the effects of changing long-run actions, including both present actions ($\bar{\pi}_1$) and past actions ($\bar{\pi}_2$). Equation (2) implies $\bar{\pi}_1 = -\beta\bar{\pi}_2$. Therefore,

$$\lim_{t\to\infty} \frac{\mathrm{d}E_0[\pi(A_t, A_{t-1}, w_t, w_{t-1})]}{\mathrm{d}C} = \bar{\pi}_3 + \bar{\pi}_4 + (1-\beta)\bar{\pi}_2\frac{\mathrm{d}\bar{A}}{\mathrm{d}C}. \tag{7}$$

Whether economic responses increase or decrease payoffs depends on the sign of $\bar{\pi}_2$. A case with $\bar{\pi}_2 > 0$ is a case in which higher actions impose costs today but provide benefits tomorrow, as when undertaking adaptation investments that take time to build. A case with $\bar{\pi}_2 < 0$ is a case in which higher actions provide benefits today but impose costs tomorrow, as when borrowing money or selling from storage. Undertaking more actions because of climate change increases payoffs if and only if actions are of the former type.

Equation (6) shows that changes in average payoffs depend on changes in actions. However, some previous literature has relied on envelope theorem intuition to argue that changes in actions are in fact irrelevant to effects on payoffs. Each solid curve in Figure 1 plots time t payoffs as a function of time t weather holding current actions, past actions, and past weather fixed. The three solid curves represent a continuum obtained by varying the agent's current actions. The dotted line gives the effect of time t weather on maximized time t payoffs, with points (a) and (b) indicating maximized payoffs at two particular weather realizations. The slope of the dotted line does not depend on how actions change with weather: because it is tangent to the solid lines, it matches the partial derivative of payoffs with respect to weather. This is the content of the envelope theorem. If a change in climate only affected current weather, then we could ignore how actions respond to the change in climate.

However, a change in climate also affects past and future weather. First consider the consequences of affecting past weather. Interpret weather as temperature. Imagine that

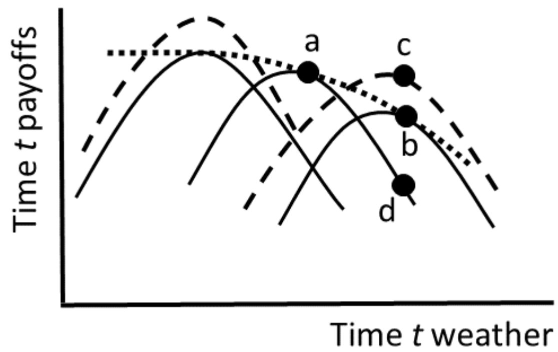

Figure 1: Per-period payoffs π_t against weather w_t. Each solid curve holds actions fixed. The dotted curve through points (a) and (b) gives the effect of time t weather shocks on myopically maximized payoffs. Point (c) shows the effect of weather shocks repeated in times $t-1$ and t, and point (d) illustrates dynamically optimizing agents.

changing inputs imposes adjustment costs (i.e., $\pi_{12} > 0$) and that the solid curves are conditioned on previous actions having been chosen in response to experiencing the temperature at point (a). The dashed curves in Figure 1 plot payoffs conditional on past actions having been chosen in response to experiencing the indicated temperature, not the temperature at point (a). Payoffs increase because the preferred current actions do not impose as many adjustment costs when past actions are better-matched to current weather. Whereas a transient weather shock induces a change from point (a) to point (b), an enduring weather shock induces a change from point (a) to point (c). This figure illustrates intuition based on Le Châtelier's principle for why the effects of transient weather shocks on payoffs could overstate the effects of climate change (e.g., Deschênes and Greenstone, 2007). What does envelope theorem intuition miss? Changes in actions chosen at time $t-1$ can have first-order effects on time t payoffs. Responses to transient weather shocks can fail to identify the effects of climate because they do not capture how climate affects the trajectory of previous actions.

Now consider the implications of climate affecting future weather. An agent who expects a transient weather shock to occur yet again in period $t+1$ may choose actions that fail to maximize short-run payoffs if these actions increase payoffs in subsequent periods, as when $\pi_{23} \neq 0$. As an example, point (d) could be optimal even though its action maximizes short-run payoffs only at point (a). Further, if $\pi_{12} \neq 0$, agents may choose not to maximize short-run payoffs even when not expecting a time $t+1$ weather shock (see equation (2)). What does envelope theorem intuition miss? Dynamic optimization typically requires trading off short-run and long-run effects. The envelope theorem no longer applies to per-period payoffs in the presence of such tradeoffs, instead applying only to expected present discounted payoffs.

4 Estimating the Effect of Climate Change from Reduced-Form Weather Regressions

I have derived the theory-implied long-run effect of climate change, but researchers do not know all of the structural parameters required to calculate this effect. Instead, empirical researchers have sought to estimate the effect of climate from reduced-form regressions with observed weather. I now consider whether and how such reduced-form regressions can recover the effect of climate.[18]

4.1 Estimating Effects on Actions

We have seen that the effects of climate on payoffs are closely related to its effects on actions. Further, much empirical research has sought to estimate the consequences of climate change for decision variables or functions of decision variables, including productivity (Heal and Park, 2013; Zhang et al., 2018), time allocation (Graff Zivin and Neidell, 2014), and energy use (Auffhammer and Aroonruengsawat, 2011; Deschênes and Greenstone, 2011; Auffhammer, 2018a). I therefore begin by considering the potential to estimate the effect of climate on actions from time series variation in weather.

First consider the determinants of time t actions. The proof of Lemma 2 shows that if either Assumption 1 or 2 holds and $(A_{t-1} - \bar{A})^2$ is small, then

$$\begin{aligned} A_t =& \bar{A} + \overbrace{\frac{\bar{\pi}_{14}}{\chi_2}(w_{t-1} - C) + \frac{\bar{\pi}_{12}}{\chi_2}(A_{t-1} - \bar{A})}^{\text{effects of past weather}} + \overbrace{\frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}(w_t - C)}^{\text{effects of current weather}} \\ &+ \underbrace{\frac{\beta\bar{\pi}_{23} + \beta\left(\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}(f_t - C)}_{\text{effects of future weather}}, \end{aligned} \tag{8}$$

where each $\chi_i > |\pi_{12}|$. We see time t actions determined by past, present, and future weather. Figure 2 illustrates the main relationships identified by this expression.

Actions depend on present weather in three ways. First, actions respond to current weather as a means of mitigating its immediate harm or amplifying its immediate benefits. This channel is controlled by $\bar{\pi}_{13}$. Second, actions respond to current weather when current actions can mitigate the harm or amplify the benefits incurred by current weather in future periods. This channel is controlled by $\bar{\pi}_{24}$ and arises only for forward-looking agents. As an

[18] I here consider only the internal validity of estimated effects. Equations (4) and (6) imply that the effect of climate change will vary with the current climate unless weather enters π only linearly. Empirical researchers should therefore take care when extrapolating estimated effects across locations and when pooling data across locations.

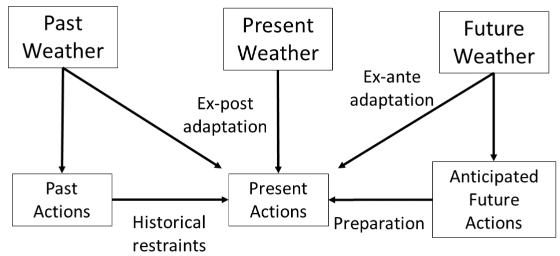

Figure 2: The determinants of present actions, from equation (8).

example of the distinction between the two channels, an agent may avoid going outside on a cold day both to minimize discomfort from the current temperature and to avoid getting sick in the near future. Both of these channels are forms of ex-post adaptation. Third, when $\bar{\pi}_{14} \neq 0$, current weather will affect the agent's chosen action in the next period (not pictured in Figure 2), leading a forward-looking agent to adjust her current action in preparation for that choice. This channel vanishes when $\bar{\pi}_{12} = 0$ because today's actions then do not directly interact with subsequent actions.

Actions also depend on forecasts of future weather. When there is the possibility of ex-ante adaptation ($\bar{\pi}_{23} > 0$), the agent chooses today's actions in order to directly mitigate the consequences (or enhance the benefits) of expected future weather. Further, expected future weather also affects desired future actions. The agent takes preparatory actions today that help her achieve her desired future actions. When $\bar{\pi}_{12} > 0$, a high forecast leads the agent to choose high actions today as a means of reducing future adjustment costs, but when $\bar{\pi}_{12} < 0$, a high forecast leads the agent to choose low actions today as a means of conserving resources for the future.

Past weather also affects actions in two ways. First, past weather affects the marginal payoffs from current actions directly when $\bar{\pi}_{14} \neq 0$. This is a form of ex-post adaptation. Second, past weather affects past actions, which impose historical restraints on current actions when $\bar{\pi}_{12} \neq 0$. When actions are intertemporal complements ($\bar{\pi}_{12} > 0$), high past actions justify higher present actions as a way to reduce adjustment costs, but when actions are intertemporal substitutes ($\bar{\pi}_{12} < 0$), high past actions justify lower present actions by depleting the resources needed to maintain a high action.

Empirical researchers hope to recover (4) from time series variation in weather. Let there be J agents (equivalently, firms) observed in each of T periods. Index these agents by j. In order to focus on the issue at hand, imagine that they are in the same climate C with the same payoff function π and the same stochastic process driving forecasts and weather, though each agent draws its own sequence of weather and forecasts. Consider the following fixed effects regression:

$$A_{jt} = \alpha_j + \Gamma_1 w_{jt} + \Gamma_2 w_{j(t-1)} + \Gamma_3 f_{jt} + \Gamma_4 A_{j(t-1)} + \eta_{jt}, \tag{9}$$

where α_j is a fixed effect for unit j and η_{jt} is an error term that is uncorrelated with the covariates.[19] I use a hat to denote the probability limit of each estimator. The following proposition relates the estimated coefficients to the effect of climate change.

Proposition 1. *Let either Assumption 1 or 2 hold, and let* $(A_{j(t-1)} - \bar{A})^2$ *be small for all observations. Then* $\hat{\Gamma}_4 \propto \bar{\pi}_{12}$ *and*

$$\hat{\Gamma}_1 + \hat{\Gamma}_2 + \hat{\Gamma}_3 = \omega \left(\frac{\mathrm{d}\bar{A}}{\mathrm{d}C} + \beta \bar{\pi}_{12} \Omega \right), \tag{10}$$

where $\bar{\pi}_{12} > 0$ *implies* $\omega \in (0, 1)$, $\bar{\pi}_{12} < 0$ *implies* $\omega > 1$, $\bar{\pi}_{12} = 0$ *implies* $\omega = 1$, *and* $\Omega \propto \bar{\pi}_{13} + \bar{\pi}_{14} + \beta \bar{\pi}_{24} \geq 0$.

Proof. See appendix. □

The three coefficients capture the three temporal relationships altered by climate change: $\hat{\Gamma}_1$ recovers consequences of altering current weather, $\hat{\Gamma}_2$ recovers consequences of altering past weather, and $\hat{\Gamma}_3$ recovers consequences of altering expectations of future weather. However, we cannot in general recover the response to a permanent change in climate from the estimated response to transient weather shocks. The reason for this failure is the possibility that $\bar{\pi}_{12} \neq 0$, which occurs if and only if $\hat{\Gamma}_4 \neq 0$.

Relationships of intertemporal substitutability or complementarity drive two types of wedges between the estimator in (10) and the effect of climate change in (4). The second term in parentheses in (10) reflects preparatory actions that are undertaken in response to forecasts but are not relevant to the long-run effects of climate. The fixed effects estimator is identified from shocks to forecasts and weather. As described above, a high forecast increases present actions both through the possibility of ex-ante adaptation and through preparatory actions. The former are important components of the effect of climate but the latter are not: an increase in the climate index C does increase forecasts, but because it also increases current and past weather, preparatory actions are not relevant to its long-run effects. When $\bar{\pi}_{12} > 0$, preparatory actions make the fixed effects estimator overstate responses to climate as observed agents are motivated by expectations of temporary adjustment costs, but when $\bar{\pi}_{12} < 0$, preparatory actions make the fixed effects estimator understate responses to climate as observed agents temporarily conserve resources.

The second wedge in (10) arises from ω. This term reflects the difference between the historical restraints on current actions imposed by transient weather shocks and those imposed by a change in climate that affects all past weather realizations. When $\bar{\pi}_{12} > 0$, historical restraints prevent an agent from adjusting too much to a transient weather shock, but when

[19] I do not explicitly model the unobservable characteristics that motivate the fixed effects specification. I am here not interested in identification but in what we learn from a well-identified weather regression. See Dell et al. (2014) and Auffhammer (2018b), among others, for expositions of identification in the climate-economy literature. I assume that the only possible sources of omitted variables bias are the failure to control for variables such as forecasts and lagged actions that are defined within the theoretical model.

that shock has been repeated many times in the past (as eventually happens following a change in climate), the many small adjustments eventually add up to much greater adjustment. The $\omega < 1$ captures how responses to transient shocks overstate historical restraints in this case. Consistent with conjectures in Deschênes and Greenstone (2007), observable short-run responses are smaller than long-run responses. In contrast, when $\bar{\pi}_{12} < 0$, an agent experiences more severe historical restraints following a change in climate than following a transient weather shock. When actions depend on scarce resources, actions can be more extreme when they are maintained for only a short period of time. The $\omega > 1$ captures how responses to transient shocks understate historical restraints in this case. Consistent with conjectures in Fisher et al. (2012) and Blanc and Schlenker (2017), short-run responses are larger than long-run responses.

The wedges introduced by Ω and ω conflict, making it impossible to sign the bias in general. However, we can make progress in two special cases. First, when $\bar{\pi}_{12} = 0$, both wedges vanish. In this case, the fixed effects estimator exactly recovers the effect of climate. Second, when $\beta = 0$, the wedge introduced by preparatory actions vanishes because myopic agents are not concerned about future actions. The sign of the bias then depends only on the wedge ω induced by historical restraints, as even myopic agents respond to their own past decisions (see also Keane and Wolpin, 2002).[20]

Now consider the following distributed lag regression, which matches most literature in not controlling for lagged actions:[21]

$$A_{jt} = \alpha_j + \sum_{i=0}^{I+1} \Gamma_{w_{t-i}} w_{j(t-i)} + \sum_{i=0}^{I+1} \Gamma_{f_{t-i}} f_{j(t-i)} + \eta_{jt}, \tag{11}$$

where $I \geq 0$. The following proposition relates the estimated coefficients to the effect of climate change.[22]

Proposition 2. *Let either Assumption 1 or 2 hold, and let $(A_{j(t-1)} - \bar{A})^2$ be small for all observations. Then*

$$\lim_{I \to \infty} \sum_{i=0}^{I} \left[\hat{\Gamma}_{w_{t-i}} + \hat{\Gamma}_{f_{t-i}} \right] = \tilde{\omega} \left(\frac{\mathrm{d}\bar{A}}{\mathrm{d}C} + \beta \bar{\pi}_{12} \Omega \right),$$

[20]The wedge introduced by preparatory actions also vanishes if there is no ex-post adaptation, but this is an artifact of modeling forecasts as existing only one period ahead. In this environment, there are no time t shocks that affect expectations of time $t+2$ weather. If there were longer-horizon forecasts, then time t shocks could affect those expectations and thereby induce preparation for ex-ante adaptation anticipated to be undertaken at time $t+1$.

[21]This omission is driven in part by concern for Nickell (1981) omitted variables bias.

[22]The proof also shows that, beyond the first lag, the coefficients alternate signs as the lags increase if and only if $\bar{\pi}_{12} < 0$. The appendix shows that $\bar{pi}_{12} = 0$ is not sufficient to recover the effects of climate if the regression does not control for forecasts and there is scope for ex-ante adaptation.

where $\beta\bar{\pi}_{12} > 0$ *implies* $\tilde{\omega} \in (\omega, 1)$*,* $\beta\bar{\pi}_{12} < 0$ *implies* $\tilde{\omega} \in (1, \omega)$*, and* $\beta\bar{\pi}_{12} = 0$ *implies* $\tilde{\omega} = 1$*, with* ω *and* Ω *from Proposition 1. If* $\bar{\pi}_{12} = 0$*, then* $\hat{\Gamma}_{w_{t-i}} = 0$ *for* $i > 1$ *and* $\hat{\Gamma}_{f_{t-i}} = 0$ *for* $i > 0$*. If* $\beta = 0$*, then* $\hat{\Gamma}_{f_{t-i}} = 0$ *for all* $i \geq 0$*.*

Proof. See appendix. □

This estimator is subject to the same bias from preparatory actions, but by using a long history of transient shocks, it reduces the bias introduced by historical restraints. If the latter bias is the dominant one, then this estimator may reduce the overall bias in estimated effects on actions. Further, this estimator recovers the effects of climate in a new case: when agents are myopic. Myopic agents are never subject to the bias induced by preparatory actions, and we now lose the bias induced by historical restraints because myopic agents respond to a long sequence of transient weather shocks in exactly the same way as they respond to living in a world with an altered climate.

4.2 Estimating Effects on Payoffs

We now consider the possibility of recovering effects on payoffs from observations of payoffs and weather. For instance, empirical research studies how variation in weather affects agricultural profits (e.g., Deschênes and Greenstone, 2007) or affects macroeconomic variables such as gross output or income that are potentially related to payoffs (e.g., Dell et al., 2012; Burke et al., 2015; Deryugina and Hsiang, 2017; Colacito et al., 2019).

The class of payoff functions defined by the following assumption will yield especially interesting results:

Assumption 3. $\pi_2(A_t, A_{t-1}, w_t, w_{t-1}) = K\pi_1(A_t, A_{t-1}, w_t, w_{t-1})$ *if* $A_{t-1} = A_t$*, for* $K \neq -\beta$*.*

Consider a few members of this class. First, adjustment cost models yield $K = 0$: if $\pi = g(A_t, (A_t - A_{t-1})^z, w_t, w_{t-1})$ for $z > 1$, then $\pi_2 = z(A_t - A_{t-1})^{z-1} g_2(A_t, (A_t - A_{t-1})^z, w_t, w_{t-1})$ and thus is equal to 0 when $A_t = A_{t-1}$. Second, a model in which the returns to resource extraction decline in previous extraction can yield $K = -1$: if $\pi = g(A_t/A_{t-1}, w_t, w_{t-1})$, then $\pi_1 = g_1(A_t/A_{t-1}, w_t, w_{t-1})/A_{t-1}$ and $\pi_2 = -A_t g_1(A_t/A_{t-1}, w_t, w_{t-1})/A_{t-1}^2$. Third, a model in which ex-post adaptation and ex-ante adaptation form a constant elasticity of substitution (CES) aggregate with distribution parameter κ yields $K = (1-\kappa)/\kappa$: $\pi = g(h(A_t, A_{t-1}), w_t, w_{t-1})$ where $h(A_t, A_{t-1}) = (\kappa A_t^\sigma + (1-\kappa)A_{t-1}^\sigma)^{1/\sigma}$ for $\sigma < 1, \neq 0$ and $h(A_t, A_{t-1}) \to A_t^\kappa A_{t-1}^{1-\kappa}$ as $\sigma \to 0$. Finally, a model without dynamic linkages has $\pi_2(\cdot,\cdot,\cdot,\cdot) = 0$ and thus $K = 0$.

Empirical researchers hope to recover (6) from time series variation in weather. They will not generally observe the full set of actions available to agents or firms. As a result, empirical researchers may estimate the following regression:

$$\pi_{jt} = \alpha_j + \sum_{i=0}^{I+1} \theta_{w_{t-i}} w_{j(t-i)} + \sum_{i=0}^{I+1} \theta_{f_{t-i}} f_{j(t-i)} + \eta_{jt}, \tag{12}$$

where I again label units as j, α_j is a fixed effect for agent j, and η_{jt} is an error term (see footnote 19). We are interested in the vector of coefficients θ. As before, I use a hat to denote the probability limit of each coefficient.

Proposition 3. *Let Assumption 1 hold, or let Assumption 2 hold with the ϵ and ν normally distributed. Also let $(A_{j(t-1)} - \bar{A})^2$ and $(A_{jt} - \bar{A})^2$ be small for all observations and let each agent's average actions be $\bar{A}$.*

1. *If $\bar{\pi}_{12} = 0$ and $I > 1$, then $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \hat{\theta}_{w_t} + \hat{\theta}_{w_{t-1}} + \hat{\theta}_{w_{t-2}} + \hat{\theta}_{f_t} + \hat{\theta}_{f_{t-1}}$ and all other coefficients are equal to 0.*
2. *If $\beta = 0$, then $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \lim_{I\to\infty} \left[\sum_{i=0}^{I} \hat{\theta}_{w_{t-i}} + \sum_{i=0}^{I} \hat{\theta}_{f_{t-i}}\right]$.*
3. *$\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \lim_{\beta\to 1} \lim_{I\to\infty} \left[\sum_{i=0}^{I} \hat{\theta}_{w_{t-i}} + \sum_{i=0}^{I} \hat{\theta}_{f_{t-i}}\right]$.*
4. *If Assumption 3 holds and $I \geq 0$, then $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \hat{\theta}_{w_t} + \hat{\theta}_{w_{t-1}}$ and all other coefficients are equal to zero.*

Proof. See appendix. □

The proposition describes four cases in which we can recover the effect of climate from time series variation in weather (and the appendix further describes cases in which we can unambiguously bound the effect). The first two cases follow directly from the analysis in Section 4.1. There we saw that we can recover the effect of climate on actions if either $\bar{\pi}_{12} = 0$ or $\beta = 0$. In the former case, we can recover the effect on current actions from the coefficient on weather, its lag, and forecasts, and the first result in Proposition 3 follows from recognizing that we need to recover effects on both current and lagged actions and that the coefficients on weather and its lag also capture the direct effects of weather in equation (6). If $\beta = 0$, we recover effects on actions only as the lags become very long, in which case we also recover effects on payoffs.

The other two cases are ones in which we do not need to recover the effect of climate on actions. The proof shows that the bias from estimating the effect of climate on payoffs from the combination of infinite lags is proportional to $\beta\bar{\pi}_{12}(\bar{\pi}_1 + \bar{\pi}_2)$. The bias vanishes as $\beta \to 1$ because agents' responses equalize the marginal value of past and current actions, without discounting the former. Alternately, Assumption 3 and equation (2) imply that $\bar{\pi}_2 = \bar{\pi}_1 = 0$: an optimizing agent sets the marginal benefit of actions to zero around a steady state. In this case, the consequences of marginal climate change are independent of changes in actions and the estimated coefficients do not include any effects of weather or forecasts on actions. Summing $\hat{\theta}_{w_t}$ and $\hat{\theta}_{w_{t-1}}$ now captures only the direct effects of weather and fully captures the effects of climate on payoffs. We can test whether Assumption 3 holds by examining the magnitude of the coefficient on forecasts: because forecasts matter for current payoffs

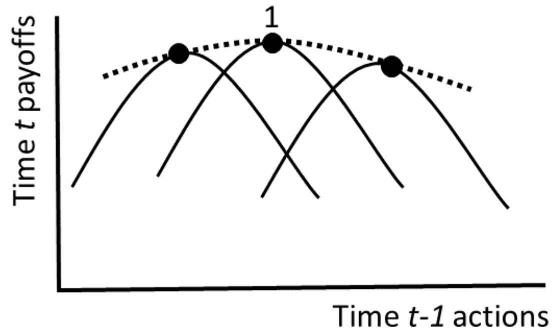

Figure 3: Per-period payoffs π_t against previous actions A_{t-1}. Each solid curve holds time t actions and weather fixed. The dotted curve gives the effect of time $t-1$ actions on myopically maximized payoffs.

only through their effects on actions, they cannot affect these payoffs if Assumption 3 indeed holds and agents are near a steady state.[23,24]

Assumption 3 leads us to the same point as envelope theorem intuition proposed in previous literature. But Section 3 explained why these arguments fail. So what does Assumption 3 do? Figure 3 plots payoffs as a function of previous actions, conditional on current actions and weather. The dotted curve gives the effect of previous actions on myopically optimized payoffs. Point (1) identifies the previous action that maximizes current payoffs. Around this point, previous actions do not have first-order effects on current payoffs. Why is that important? The dashed curves in Figure 1 reflected how past actions can shift current payoffs. If these past actions do not have first-order effects, then the dashed curves are not shifted out and point (c) converges to point (b) in Figure 1. Further, if past actions do not have first-order effects on current payoffs, then current actions do not have first-order effects on future payoffs. In that case, the myopically optimal action is also dynamically optimal and point (d) converges to point (b) in Figure 1. Assumption 3 ensures that point (1) occurs at the steady-state action. Therefore, if Assumption 3 holds and agents are near a steady state, then points (c) and (d) converge to point (b) in Figure 1 and the treatment effect of a transient weather shock indeed recovers the effect of permanently changing the weather.[25]

[23]Much literature has studied dependent variables such as crop yields (e.g., Schlenker and Roberts, 2009), mortality (e.g., Deschênes and Moretti, 2009; Deschênes and Greenstone, 2011), and health (e.g., Deschenes, 2014) that are functions of actions but are not payoff functions. If we consider recovering the effects of climate on such dependent variables from a fixed effects regression on weather, then the final two parts of Proposition 3 no longer apply because the Euler equation (1) holds only for payoffs, not for other functions of actions.

[24]Empirical researchers have usually not controlled for forecasts. The appendix shows that such regressions can still recover the full effect of climate on payoffs if either Assumption 3 holds or agents are myopic. However, the other cases now also require the absence of ex-ante adaptation.

[25]As described earlier, one of the special cases of Assumption 3 is a model with no dynamic linkages

Proposition 3 assumed that each agent's average actions are $\bar{A}$.[26] The following corollary establishes how relaxing this assumption changes the results.

Corollary 4. *Let the conditions given in Proposition 3 hold, except let each agent's average actions be different from $\bar{A}$. In addition, let at least one of $\bar{\pi}_{13}$, $\bar{\pi}_{23}$, $\bar{\pi}_{14}$, or $\bar{\pi}_{24}$ be strictly positive. Then, in each part of Proposition 3, $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C$ is strictly less (greater) than the indicated combination of coefficients if and only if $\bar{A}$ is strictly less (greater) than each agent's average actions.*

Proof. See appendix. □

The corollary establishes that the special cases that formerly sufficed to identify climate impacts from weather impacts now merely bound the effect of climate on payoffs. In particular, we obtain an upper bound if agents are approaching their steady-state actions from above and a lower bound otherwise. Intuitively, if climate shifts the steady-state action farther from the agent's current action, then weather shocks incorporate transition costs that vanish from the effect of climate on long-run payoffs.

5 Structurally Estimating Climate Impacts in U.S. Agriculture

I have thus far explored the potential for reduced-form estimates to recover climate impacts. I now show how to recover the combinations of structural parameters necessary to calculate climate impacts. I use these parameters to disentangle weather effects from adaptation and to sign the bias in the adaptation estimate. This new approach maintains precisely the same credible identification from the reduced-form specifications. As we will see, these specifications suffice because we do not need to specify or recover every underlying structural parameter in order to undertake the calculations of interest.

I demonstrate this new approach by extending the seminal analysis of agricultural impacts from Deschênes and Greenstone (2007). In order to be consistent with common regression specifications, generalize the foregoing analysis to allow for K types of weather variables, which can be correlated with each other. Let there be M actions chosen in each period, so that time t payoffs are now $\pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}) : \mathbb{R}^M \times \mathbb{R}^M \times \mathbb{R}^K \times \mathbb{R}^K \to \mathbb{R}$, with bold script indicating vectors. Superscripts will indicate elements of these vectors. The following assumption is useful for the structural calculations:

$(\pi_2(\cdot,\cdot,\cdot,\cdot) = 0)$, in which case the agent solves a series of independent, static decision problems. Appeals to the envelope theorem therefore can end up with the correct result in the types of static settings discussed by previous literature (Hsiang, 2016; Deryugina and Hsiang, 2017).

[26] Sorkin (2015) imposes an analogous restriction when relating short-run and long-run variation.

Assumption 4. *Either (i) there exist* $g : \mathbb{R}^M \to \mathbb{R}$, $h : \mathbb{R}^M \to \mathbb{R}$, *and* $\pi^0 : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^K \to \mathbb{R}$ *such that* $\pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}) = \pi^0(g(\boldsymbol{A}_t), h(\boldsymbol{A}_{t-1}), \boldsymbol{w}_t)$, *or (ii) there exist* $\pi^k : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *for* $k \in \{1, ..., K\}$ *and* $\pi^{K+1} : \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}$ *such that* $M \geq K$ *and*

$$\pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}) = \sum_{k=1}^{K} \pi^k(A_t^k, A_{t-1}^k, w_t^k) + \pi^{K+1}(\boldsymbol{A}_t^{\sim k}, \boldsymbol{A}_{t-1}^{\sim k}),$$

where $\boldsymbol{A}_t^{\sim k}$ *indicates the* $(M-K)$*-dimensional vector of actions* A_t^{K+1} *through* A_t^M.

This assumption does two things. First, it rules out delayed effects of weather, which is plausible in the below application to annual agricultural data. Some assumption about delayed effects is necessary for identification of structural parameters. Second, it requires either that the vector of actions be reducible to a composite action (trivially true for $M = 1$) or that payoffs be separable in the dimensions of weather. Either restriction ensures that the terms controlling whether actions are intertemporal substitutes or complements are scalar.

Consider the following regression:

$$\pi_{ct} = \alpha_c + \psi_{rt} + \sum_{k=1}^{K} \left[\Phi_{w_{t-2}}^k w_{c(t-2)}^k + \Phi_{w_{t-1}}^k w_{c(t-1)}^k + \Phi_{w_t}^k w_{ct}^k + \Phi_{w_{t+1}}^k w_{c(t+1)}^k \right] + \delta_{ct}, \tag{13}$$

where c indicates counties, t indicates years, π_{ct} is agricultural profits, the α_c are county fixed effects, the ψ_{rt} are region-year fixed effects,[27] and superscript k indexes weather variables of interest. The following lemma expresses the coefficients in terms of model primitives:[28]

Lemma 3. *Let Assumption 4 and the conditions of Proposition 3 hold, and let* $\boldsymbol{\epsilon}_t$ *be uncor-*

[27] In the preferred specification, the regions are USDA Farm Resource Regions (see also Deschênes and Greenstone, 2012). The appendix provides further details, reports the variance explained by the weather variables (see Fisher et al., 2012), and assesses sensitivity to instead defining regions as individual states (as in Deschênes and Greenstone, 2007) or as the whole country (as in Fisher et al., 2012).

[28] The lemma requires that weather be serially uncorrelated. This assumption seems an acceptable starting point: over all U.S. counties from 1972 to 2017, the correlation between locally demeaned growing season degree days and its lag is 0.10, the correlation between locally demeaned extreme growing season degree days and its lag is 0.074, and the correlation between locally demeaned growing season precipitation and its lag is -0.029.

related with $\boldsymbol{\nu}_t$. Then:

$$\begin{aligned}
\hat{\Phi}^k_{w_{t+1}} &= -\beta\bar{\pi}^k_2\hat{\Gamma}^k_3\frac{(\tau^k)^2}{(\tau^k)^2+(\sigma^k)^2},\\
\hat{\Phi}^k_{w_t} &= \bar{\pi}^k_3 - \beta\bar{\pi}^k_2\hat{\Gamma}^k_1 + \left(1-\beta\frac{\bar{\pi}^k_{12}}{\chi^k_2}\right)\bar{\pi}^k_2\frac{(\tau^k)^2}{(\tau^k)^2+(\sigma^k)^2}\hat{\Gamma}^k_3,\\
\hat{\Phi}^k_{w_{t-1}} &= \left(1-\beta\frac{\bar{\pi}^k_{12}}{\chi^k_2}\right)\bar{\pi}^k_2\hat{\Gamma}^k_1 + \left(1-\beta\frac{\bar{\pi}^k_{12}}{\chi^k_2}\right)\frac{\bar{\pi}^k_{12}}{\chi^k_2}\bar{\pi}^k_2\frac{(\tau^k)^2}{(\tau^k)^2+(\sigma^k)^2}\hat{\Gamma}^k_3,\\
\hat{\Phi}^k_{w_{t-2}} &= \frac{\bar{\pi}^k_{12}}{\chi^k_2}\left(1-\beta\frac{\bar{\pi}^k_{12}}{\chi^k_2}\right)\bar{\pi}^k_2\hat{\Gamma}^k_1 + \left(\frac{\bar{\pi}^k_{12}}{\chi^k_2}\right)^2\left(1-\beta\frac{\bar{\pi}^k_{12}}{\chi^k_2}\right)\bar{\pi}^k_2\frac{(\tau^k)^2}{(\tau^k)^2+(\sigma^k)^2}\hat{\Gamma}^k_3.
\end{aligned}$$

If case (i) of Assumption 4 holds, then we replace the k superscripts on the right-hand side with 0.

Proof. See appendix. □

The $\hat{\Gamma}$ were defined in regression (11) and analyzed in Proposition 1. The proof of Proposition 1 defines these in terms of model primitives. They here gain a superscript k to indicate the corresponding dimension of weather. Solving the system of equations, we find:

$$\begin{aligned}
\frac{\bar{\pi}^k_{12}}{\chi^k_2} &= \frac{\hat{\Phi}^k_{w_{t-2}}}{\hat{\Phi}^k_{w_{t-1}}},\\
\bar{\pi}^k_2\hat{\Gamma}^k_3\frac{(\tau^k)^2}{(\tau^k)^2+(\sigma^k)^2} &= -\frac{\hat{\Phi}^k_{w_{t+1}}}{\beta},\\
\bar{\pi}_2\hat{\Gamma}^k_1 &= \frac{\hat{\Phi}^k_{w_{t-1}} - \left(1-\beta\frac{\bar{\pi}^k_{12}}{\chi^k_2}\right)\frac{\bar{\pi}^k_{12}}{\chi^k_2}\bar{\pi}^k_2\frac{(\tau^k)^2}{(\tau^k)^2+(\sigma^k)^2}\hat{\Gamma}^k_3}{1-\beta\frac{\bar{\pi}^k_{12}}{\chi^k_2}},\\
\bar{\pi}^k_3 &= \hat{\Phi}^k_{w_t} + \beta\bar{\pi}^k_2\hat{\Gamma}^k_1 - \left(1-\beta\frac{\bar{\pi}^k_{12}}{\chi^k_2}\right)\bar{\pi}^k_2\frac{(\tau^k)^2}{(\tau^k)^2+(\sigma^k)^2}\hat{\Gamma}^k_3.
\end{aligned}$$

$\bar{\pi}^k_{12}/\chi^k_2$ is identified from the first and second lags of weather. Recalling that χ^k_2 must be positive for a saddle-path stable steady-state, the sign of this term matches the sign of $\bar{\pi}^k_{12}$. The lead of weather identifies $\bar{\pi}^k_2\hat{\Gamma}^k_3\frac{(\tau^k)^2}{(\tau^k)^2+(\sigma^k)^2}$, which the proof of Proposition 1 connects to ex-ante adaptation.[29] Given these two terms, the residual effects of lagged weather identify $\bar{\pi}^k_2\hat{\Gamma}^k_1$, which the proof of Proposition 1 connects to ex-post adaptation. Finally, the residual effects of contemporary weather identify the direct effects $\bar{\pi}^k_3$. Intuitively, lagged weather affects current payoffs only through actions, so we can identify ex-post adaptation from

[29] I calibrate β to the annual discount rate of 34% obtained in Duquette et al. (2012). The appendix shows that results are not sensitive to lower discount rates.

lagged weather and can identify intertemporal links between actions by comparing lags of weather. And once we have identified the scope of adaptation, we can recover the direct effects of weather from the response of payoffs to contemporary weather.

Equation (7) shows that calculating climate impacts requires $\bar{\pi}_3$ and $(1-\beta)\bar{\pi}_2 \,\mathrm{d}\bar{A}/\,\mathrm{d}C$. The above steps recover $\bar{\pi}_3$ directly. From equation (10),

$$\frac{\mathrm{d}\bar{A}}{\mathrm{d}C} = \frac{1}{\omega}\left[\hat{\Gamma}_1 + \hat{\Gamma}_3\right] - \beta\bar{\pi}_{12}\Omega,$$

with $\omega > 1$ if and only if $\bar{\pi}_{12} < 0$. If $\bar{\pi}_{12} \approx 0$, then

$$\frac{\mathrm{d}\bar{A}}{\mathrm{d}C} \approx \hat{\Gamma}_1 + \hat{\Gamma}_3.$$

I calculate the effects of ex-post adaptation using $(1-\beta)\bar{\pi}_2\hat{\Gamma}_1$ and the effects of ex-ante adaptation using $(1-\beta)\bar{\pi}_2\hat{\Gamma}_3$.[30] As described below, I use the estimated sign of $\bar{\pi}_{12}^k$ to convert these estimates into bounds.

For comparison, I also undertake two reduced-form calculations of the effects of climate change. Either can be justified from an analogue of Proposition 3. A first calculation estimates (13) without any leads or lags on the right-hand side and multiplies each weather variable's coefficient by the projected change in that weather variable. This calculation recovers the theory-implied effects of climate if Assumptions 3 and 4 hold. It matches the calculations undertaken in previous literature. A second calculation estimates (13) without the second lag on the right-hand side and multiplies the sum of each weather index's three coefficients by the projected change in that weather variable. This calculation recovers the theory-implied effects of climate if $\bar{\pi}_{12} = 0$ and Assumption 4 holds.

The appendix describes the data, details sample construction, and undertakes robustness checks. The construction of the data follows an updated version of the methodology in Deschênes and Greenstone (2007) and Fisher et al. (2012). I have observations of county-level agricultural profits and acreage every 5 years from 1987 through 2012. I follow previous literature in studying a measure of growing season degree days (i.e., accumulated heat within a temperature range favorable to plant growth), a measure of extreme growing season degree days (i.e., accumulated extreme heat, generally harmful to plant growth), and growing season precipitation. The preferred specification includes USDA Farm Resource Region-by-year fixed effects (as in Deschênes and Greenstone, 2012), weights counties by average acreage (as in Deschênes and Greenstone, 2007), clusters standard errors by state (as in Fisher et al., 2012), and restricts the sample to counties east of the 100th meridian, which are less likely to be irrigated (Schlenker et al., 2005; Fisher et al., 2012).

[30] These calculations set $(\tau^k)^2/((\tau^k)^2+(\sigma^k)^2)$ equal to 1. This fraction reflects the fraction of the variation in weather that is already realized one period ahead (i.e., that is reflected in forecasts). Estimated ex-ante adaptation is biased towards zero if the fraction is in fact less than 1. Replacing the lead of realized weather in regression (13) with forecasts would eliminate this bias.

Table 1: Top panel: Estimated coefficients and standard errors from regression (13). Bottom panel: Theory-implied structural parameters, reported as the median and lower/upper quartiles of the distribution implied by the reduced-form estimates.

	GDD	Extreme GDD	Precip
Reduced-Form Coefficients			
Current	18	-98	-7.9
	(9.1)	(54)	(3.9)
Lag 1	5.4	-60	-7.1
	(6.2)	(22)	(3.9)
Lag 2	-15	32	0.51
	(9.9)	(23)	(3.8)
Lead	-3.8	-1.4	1.3
	(4.1)	(20)	(1.5)
Theory-Implied Parameters			
$\bar{\pi}_3$	15	-130	-12
	(8.9,22)	(-160,-110)	(-16,-7.8)
$\bar{\pi}_2\hat{\Gamma}_1$	13	-46	-7.6
	(3.4,28)	(-60,-31)	(-11,-5)
$\bar{\pi}_2\hat{\Gamma}_3\frac{\tau^2}{\tau^2+\sigma^2}$	5.0	1.9	-1.8
	(1.3,8.8)	(-16,20)	(-3.2,-0.43)
$\bar{\pi}_{12}/\chi_2$	-1.7	-0.53	-0.065
	(-3.4,-0.41)	(-0.74,-0.31)	(-0.46,0.31)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. There are 13944 county-year observations and 37 state observations. Profits in thous. year 2002 dollars, GDD in °C-days, and precip in mm.

The top panel of Table 1 reports the reduced-form coefficients from regression (13). Profits increase in same-year and previous-year growing degree days between 10°C and 29°C ("GDD"), but profits decrease in same-year and previous-year growing degree days above 29°C ("Extreme GDD"). The signs of the central estimates alternate from the first to the second lag. The lead of weather does not have a statistically significant effect on profits. Additional same-year and previous-year precipitation reduces profits.

The lower panel of Table 1 reports the medians and, in parentheses, lower and upper quartiles for the theory-implied structural parameters.[31] The signs of the direct terms $\bar{\pi}_3$ and the ex-post adaptation terms $\bar{\pi}_2\hat{\Gamma}_1$ are consistent with the signs of same-year and previous-year impacts on profits. The nonzero ex-post adaptation term suggests that Assumption 3 does not hold in this application, which motivates a structural approach to recovering climate impacts. The ex-ante adaptation terms are small and noisy, as were the coefficients on the leads of weather.

The final row of Table 1 reports $\bar{\pi}_{12}/\chi_2$. We have a case of intertemporal substitutes (complements) if this term is negative (positive). The median values are all negative, and even the 75th percentile is negative for the growing degree day variables.[32] The negative $\bar{\pi}_{12}$ are identified off the result that lagged weather affects payoffs much like current weather but the second lag of weather has opposite effects to current weather. Within the economic model, these opposite effects imply that adaptive actions taken two years ago increase current payoffs by constraining the actions taken last year. Finding $\bar{\pi}_{12} < 0$ is contrary to Le Châtelier's principle but consistent with recent empirical work in agricultural economics (Hendricks et al., 2014; Kim and Moschini, 2018). Following Eckstein (1984), these researchers attribute their results to soil nitrogen and pest dynamics inducing farmers to rotate their crops over time.[33]

Table 2 reports the projected effects of climate change. It uses the RCP 4.5 trajectory of stabilized emissions from 21 downscaled CMIP5 models. In this scenario, global mean surface temperature increases by around 2 degrees Celsius over the century. The top panel reports the two reduced-form calculations. Both approaches project substantial costs from climate change. The projected increase in conventional growing degree days is estimated to

[31] The lower panel does not report means and standard errors because the distributions can be skewed due to $\bar{\pi}_{12}/\chi_2$ being the ratio of two reduced-form coefficients. The appendix reports standard errors from method of moments estimators of the structural parameters and reports results for the 10th and 90th percentiles.

[32] Reassuringly, these estimates are consistent with $|\bar{\pi}_{12}| < \chi_2$, which the proof of Lemma 2 shows follows from saddle-path stability. One might consider estimating a geometric lag structure in regression (13), with the geometric term equal to $\bar{\pi}_{12}/\chi_2$. I find that coefficient estimates can be unstable when adding longer lags, likely due to collinearity problems that commonly arise with lagged predictors. When using three lags, the estimated $\bar{\pi}_{12}/\chi_2$ for extreme growing degree days is extremely noisy, the estimated $\bar{\pi}_{12}/\chi_2$ for conventional growing degree days in a one- (two-) step GMM estimator is -1.18 (-0.56) with a standard error of 0.43 (0.34), and the estimated $\bar{\pi}_{12}/\chi_2$ for precipitation in a one- (two-) step GMM estimator is -1.47 (-0.17) with a standard error of 0.56 (0.31).

[33] Huang and Moore (2019) show that U.S. farmers do adjust crop acreage in response to pre-planting precipitation.

increase agricultural profits, but the projected increase in extreme growing degree days is projected to reduce profits to a greater degree. Projected costs are 50% greater under the assumption that $\bar{\pi}_{12} = 0$ than under Assumption 3. However, these numbers are meaningless if Assumption 3 does not hold or $\bar{\pi}_{12} \neq 0$, and we have already seen that Assumption 3 does not appear to hold and that $\bar{\pi}_{12} < 0$.

The lower panel reports the new, theory-based estimates of climate impacts. It decomposes the effects of climate change into direct effects (driven by $\bar{\pi}_3$), ex-post adaptation (driven by $\bar{\pi}_{13}$ via $\bar{\pi}_2\hat{\Gamma}_1$), and ex-ante adaptation (driven by $\bar{\pi}_{23}$ via $\bar{\pi}_2\hat{\Gamma}_3$). The median direct effects of projected changes in each measure of growing degree days are smaller than either of the reduced-form estimates, but the median combined direct effect is in between the two reduced-form estimates.

Ex-post adaptation increases the benefits from conventional growing degree days but increases the costs from extreme growing degree days. From equation (6), adaptive changes in actions affect steady-state payoffs as $\bar{\pi}_1 + \bar{\pi}_2$, and from equation (2), $\bar{\pi}_1 + \bar{\pi}_2$ is opposite in sign to $\bar{\pi}_1$. As described following equation (7), actions that provide short-run benefits (costs) in exchange for long-run costs (benefits) have negative (positive) effects on steady-state payoffs when agents are not perfectly patient. Adaptation to increases in conventional growing degree days therefore imposes short-run costs, but agents reap the benefits from past adaptation in the long run. In contrast, adaptation to increases in extreme growing degree days provides short-run benefits, but agents reap these benefits only while paying the larger costs of past adaptation, which dominate in the long run. Ex-ante adaptation is not clearly important.[34] Accounting for adaptation, projected changes in conventional growing degree days increase profits by 54% in the median estimate and projected changes in extreme growing degree days eliminate profits in the median estimate. The median total effect of climate change is a 50% reduction in profits.

The median effect of climate happens to be somewhat close to the value that would be obtained by applying previous literature's methods (which are identical to applying Assumption 3). However, the interpretation is quite different. The structural estimates approximate $\mathrm{d}\bar{A}/\,\mathrm{d}C$ by setting $\bar{\pi}_{12} = 0$, which implies that $\omega = 1$ and that Ω is irrelevant in equation (10). In contrast to reduced-form approaches, these structural calculations have a clear interpretation even if that assumption is violated. We know from Proposition 1 that the bias from $\bar{\pi}_{12} \neq 0$ reflects preparatory actions (through Ω) and historical restraints (through ω). In the present context, it is reasonable to assume that the wedge induced by preparatory actions is small relative to the wedge induced by historical restraints, an intuition reinforced by the small effects of ex-ante adaptation. Therefore $\bar{\pi}_{12} > 0$ implies that the present calculations

[34] Recall, however, that the estimated effects of ex-ante adaptation are biased towards zero (see footnote 30). The appendix reports that ex-ante adaptation to extreme growing degree days becomes clearer if we omit the earlier years from the sample, which is consistent with the reduced skill and availability of seasonal forecasts prior to the mid-1990s (see Klemm and McPherson, 2017). Takle et al. (2013) and Klemm and McPherson (2017) describe the various seasonal forecasts of interest to agriculture.

Table 2: The percentage change in eastern U.S. agricultural profits due to predicted end-of-century changes in growing degree days, extreme growing degree days, and precipitation. The reduced-form estimates report central estimates and standard errors. The theory-implied estimates report the median and lower/upper quartiles.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	51	-97	-1.0	-47
	(25)	(37)	(0.4)	(24)
Using $\bar{\pi}_{12} = 0$	44	-1.1e+02	-1.6	-71
	(29)	(35)	(0.66)	(28)
Theory-Implied				
Direct Effects	38	-93	-1.2	-56
	(22,54)	(-1.1e+02,-74)	(-1.7,-0.83)	(-71,-41)
Ex-Post Adaptation	8.3	-8.0	-0.21	0.56
	(2.1,17)	(-11,-5.4)	(-0.29,-0.13)	(-5.7,9.2)
Ex-Ante Adaptation	3.2	0.33	-0.048	3.4
	(0.84,5.5)	(-2.8,3.4)	(-0.085,-0.012)	(0.026,6.8)
Total	54	-1e+02	-1.5	-50
	(32,74)	(-1.2e+02,-79)	(-2,-1)	(-69,-31)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 13944 county-year observations and 37 state observations.

underestimate adaptation to climate (because $\omega < 1$) and $\bar{\pi}_{12} < 0$ implies that the present calculations overestimate adaptation to climate (because $\omega > 1$).

We saw in Table 1 that $\bar{\pi}_{12} < 0$, implying that we observe more adaptation to short-run weather shocks than would occur in response to long-run changes in climate. We can therefore bound the effects of climate by the estimated total effects that include projected adaptation and by the estimated direct effects that exclude adaptation. If I were following the conventional approach in the weather-climate literature, I would undertake calculations like those relying on Assumption 3 and apply intuition based on Le Châtelier's principle to conclude that climate change reduces profits by 0–47% in the median estimates. Instead, here the median estimates and the finding that $\bar{\pi}_{12} < 0$ imply that climate change reduces agricultural profits by 50–56% ($20–22 billion annually at the year 2017 price level), with the 75th percentile estimates yielding a reduction in profits of 31–41% and the 90th percentile estimates (see appendix) yielding a reduction in profits of 9–28%.

6 Potential Extensions

I have demonstrated how to estimate the effects of climate change from time series variation in weather. I conclude by discussing the primary restrictions in the present setting and describing other aspects of climate change that should be the subject of future analysis.

The theoretical model is fairly general. The three notable restrictions are that past weather and actions can directly affect payoffs with only a one-period lag (although they do indirectly affect payoffs arbitrarily far into the future) and that agents have access to specialized forecasts only one period in advance of a realized weather outcome. Longer-run delayed weather impacts are probably important to some applications, but they would not substantially change the theoretical results. In particular, distributed lag models will recover the effects of climate in exactly the same cases as analyzed here. Allowing for longer-horizon forecasts also does not change the theoretical results unless actions can have direct effects over those horizons.

Allowing actions to have longer-run direct consequences can have interesting consequences. Such an extension is attractive if we interpret actions as the choice of capital stock and we want to study capital that can be built only with a lag of more than one period. If longer-run forecasts are available, then the present results extend in a natural way, implying that it is important to control for these longer-run forecasts. But if weather forecasts do not exist over the whole horizon over which today's actions will directly affect payoffs, then empirical researchers may be unable to estimate the full effect of climate: changing the climate can lead agents to undertake actions that pay off only in the distant future, but observable variation in forecasts will not identify this adaptation margin. In the empirical application of Section 5, the existence of such actions would imply that long-run adaptation could be greater than short-run adaptation even though I estimate $\bar{\pi}_{12} < 0$ from the actions that vary in the data.

The present setting successfully captures the distinction between transient and permanent changes in weather. Future work should consider other aspects of climate change. First, global climate change differs from weather shocks not only in its temporal structure but also in its spatial structure. A change in global climate affects weather in every location and thus will have general equilibrium consequences. In the empirical application, general equilibrium channels change the prices of land and crops. The present setting has followed most empirical work in abstracting from such effects, but some recent empirical work has begun exploring the implications of changing the weather in many locations simultaneously (e.g., Costinot et al., 2016; Gouel and Laborde, 2018; Dingel et al., 2019). Future work should extend the present setting to account for general equilibrium effects.

Second, the present analysis has held the payoff function fixed over time. However, climate change should induce innovations that alter how weather affects payoffs, and many such innovations will arise even in the absence of climate change. Some types of innovation can be interpreted as actions within the present framework, but the potential for future innovation may be inherently unobservable. Historical studies have begun exploring the interaction between climate and agricultural innovation (e.g., Olmstead and Rhode, 2008, 2011; Roberts and Schlenker, 2011; Bleakley and Hong, 2017). Future work should consider approaches to bounding the scope for future innovation.

Third, the present analysis has considered only marginal changes in climate, but climate change over the next century is likely to be nonmarginal.[35] The present analysis implies that the marginal effect of climate will not generally be constant. Some recent work (Mérel and Gammans, 2019) explores the types of variation captured by quadratic regression terms. Future work should explore whether nonlinear specifications might inform estimates of the impacts from nonmarginal climate change in a dynamic setting such as the present one.

Fourth, the present analysis assumes that actions can be adjusted continuously. In the presence of fixed costs, an agent may choose to change an action only when the agent expects a change in weather to endure, and in the presence of constraints imposed by policy, actions may not respond smoothly to weather shocks, changing the interpretation of reduced-form coefficients. Future work should explore the conditions under which aggregating over many agents' fixed-cost decisions makes actions appear continuous. Future work should also explore whether responses to weather events of varying durations can identify how fixed-cost actions respond to a change in climate and whether responses to weather events of varying magnitudes can identify the role of policy constraints.

Finally, the present analysis has focused on identifying the long-run consequences of climate change, abstracting from the transition costs induced by climate change. In this regard, the present analysis matches the calculations undertaken by nearly all empirical work but omits a potentially critical aspect of climate change (see Quiggin and Horowitz, 1999, 2003; Kelly et al., 2005). Future work should consider whether imposing stronger assumptions on the decision-making environment can allow for credible simulation of counterfactual climate trajectories and thereby estimate transition costs.

References

Anderson, Michael L. (2015) "As the wind blows: The effects of long-term exposure to air pollution on mortality," Working Paper 21578, National Bureau of Economic Research.

Auffhammer, Maximilian (2018a) "Climate adaptive response estimation: Short and long run

[35] Estimating the consequences of nonmarginal climate change is critical to the damage functions required by climate-economy integrated assessment models (see Nordhaus, 2013). However, there is an argument that the consequences of marginal climate change might be especially policy relevant: if we accept climate scientists' views that the potentially nonquantifiable risks imposed by nonmarginal climate change are likely to exceed the cost of avoiding them, then the effects of marginal climate change become critical to policy choices.

impacts of climate change on residential electricity and natural gas consumption using big data," Working Paper 24397, National Bureau of Economic Research.

——— (2018b) "Quantifying economic damages from climate change," *Journal of Economic Perspectives*, Vol. 32, No. 4, pp. 33–52.

Auffhammer, Maximilian and Anin Aroonruengsawat (2011) "Simulating the impacts of climate change, prices and population on California's residential electricity consumption," *Climatic Change*, Vol. 109, No. 1, pp. 191–210.

Bakkensen, Laura and Lint Barrage (2018) "Climate shocks, cyclones, and economic growth: Bridging the micro-macro gap," Working Paper 24893, National Bureau of Economic Research.

Barreca, Alan I., Matthew Neidell, and Nicholas J. Sanders (2017) "Long-run pollution exposure and adult mortality: Evidence from the Acid Rain Program," Working Paper 23524, National Bureau of Economic Research.

Bishop, Kelly C., Jonathan D. Ketcham, and Nicolai V. Kuminoff (2018) "Hazed and confused: The effect of air pollution on dementia," Working Paper 24970, National Bureau of Economic Research.

Blanc, Elodie and Wolfram Schlenker (2017) "The use of panel models in assessments of climate impacts on agriculture," *Review of Environmental Economics and Policy*, Vol. 11, No. 2, pp. 258–279.

Bleakley, Hoyt and Sok Chul Hong (2017) "Adapting to the weather: Lessons from U.S. history," *The Journal of Economic History*, Vol. 77, No. 3, pp. 756–795.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel (2015) "Global non-linear effect of temperature on economic production," *Nature*, Vol. 527, pp. 235–239.

Carleton, Tamma A. and Solomon M. Hsiang (2016) "Social and economic impacts of climate," *Science*, Vol. 353, No. 6304, p. aad9837.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li (2013) "Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy," *Proceedings of the National Academy of Sciences*, Vol. 110, No. 32, pp. 12936–12941.

Chetty, Raj (2009) "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annual Review of Economics*, Vol. 1, No. 1, pp. 451–488.

Colacito, Riccardo, Bridget Hoffmann, and Toan Phan (2019) "Temperature and growth: A panel analysis of the United States," *Journal of Money, Credit and Banking*, Vol. 51, No. 2-3, pp. 313–368.

Costinot, Arnaud, Dave Donaldson, and Cory Smith (2016) "Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world," *Journal of Political Economy*, Vol. 124, No. 1, pp. 205–248.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken (2012) "Temperature shocks and economic growth: Evidence from the last half century," *American Economic Journal: Macroeconomics*, Vol. 4, No. 3, pp. 66–95.

——— (2014) "What do we learn from the weather? The new climate-economy literature," *Journal of Economic Literature*, Vol. 52, No. 3, pp. 740–798.

Deryugina, Tatyana and Solomon Hsiang (2017) "The marginal product of climate," Working Paper 24072, National Bureau of Economic Research.

Deschenes, Olivier (2014) "Temperature, human health, and adaptation: A review of the empirical literature," *Energy Economics*, Vol. 46, pp. 606–619.

Deschênes, Olivier and Michael Greenstone (2007) "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather," *American Economic Review*, Vol. 97, No. 1, pp. 354–385.

——— (2011) "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US," *American Economic Journal: Applied Economics*, Vol. 3, No. 4, pp. 152–185.

——— (2012) "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Reply," *American Economic Review*, Vol. 102, No. 7, pp. 3761–3773.

Deschênes, Olivier and Enrico Moretti (2009) "Extreme weather events, mortality, and migration," *The Review of Economics and Statistics*, Vol. 91, No. 4, pp. 659–681.

Dingel, Jonathan I., Kyle C. Meng, and Solomon M. Hsiang (2019) "Spatial correlation, trade, and inequality: Evidence from the global climate," Working Paper 25447, National Bureau of Economic Research.

Duquette, Eric, Nathaniel Higgins, and John Horowitz (2012) "Farmer discount rates: Experimental evidence," *American Journal of Agricultural Economics*, Vol. 94, No. 2, pp. 451–456.

Eckstein, Zvi (1984) "A rational expectations model of agricultural supply," *Journal of Political Economy*, Vol. 92, No. 1, pp. 1–19.

Fankhauser, Samuel, Joel B. Smith, and Richard S. J. Tol (1999) "Weathering climate change: some simple rules to guide adaptation decisions," *Ecological Economics*, Vol. 30, No. 1, pp. 67–78.

Fisher, Anthony C., W. Michael Hanemann, Michael J. Roberts, and Wolfram Schlenker (2012) "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Comment," *American Economic Review*, Vol. 102, No. 7, pp. 3749–3760.

Gouel, Christophe and David Laborde (2018) "The crucial role of international trade in adaptation to climate change," Working Paper 25221, National Bureau of Economic Research.

Graff Zivin, Joshua and Matthew Neidell (2009) "Days of haze: Environmental information disclosure and intertemporal avoidance behavior," *Journal of Environmental Economics and Management*, Vol. 58, No. 2, pp. 119–128.

——— (2014) "Temperature and the allocation of time: Implications for climate change," *Journal of Labor Economics*, Vol. 32, No. 1, pp. 1–26.

Guo, Christopher and Christopher Costello (2013) "The value of adaption: Climate change and timberland management," *Journal of Environmental Economics and Management*, Vol. 65, No. 3, pp. 452–468.

Hamermesh, Daniel S. (1995) "Comment," *ILR Review*, Vol. 48, No. 4, pp. 835–838.

Hamermesh, Daniel S. and Gerard A. Pfann (1996) "Adjustment costs in factor demand," *Journal of Economic Literature*, Vol. 34, No. 3, pp. 1264–1292.

Heal, Geoffrey (1976) "The relationship between price and extraction cost for a resource with a backstop technology," *The Bell Journal of Economics*, Vol. 7, No. 2, pp. 371–378.

Heal, Geoffrey and Jisung Park (2013) "Feeling the heat: Temperature, physiology and the wealth of nations," Working Paper 19725, National Bureau of Economic Research.

——— (2016) "Temperature stress and the direct impact of climate change: A review of an emerging literature," *Review of Environmental Economics and Policy*, Vol. 10, No. 2, pp. 347–362.

Heckman, James J. (2010) "Building bridges between structural and program evaluation approaches to evaluating policy," *Journal of Economic Literature*, Vol. 48, No. 2, pp. 356–398.

Hendricks, Nathan P., Aaron Smith, and Daniel A. Sumner (2014) "Crop supply dynamics and the illusion of partial adjustment," *American Journal of Agricultural Economics*, Vol. 96, No. 5, pp. 1469–1491.

Hennessy, Christopher A and Ilya A Strebulaev (2019) "Beyond random assignment: Credible inference of causal effects in dynamic economies," *Journal of Finance*, Vol. forthcoming.

Hsiang, Solomon M. (2016) "Climate econometrics," *Annual Review of Resource Economics*, Vol. 8, pp. 43–75.

Huang, Hsing-Hsiang and Michael R. Moore (2019) "Farming under weather risk: Adaptation, moral hazard, and selection on moral hazard," in Wolfram Schlenker ed. *Understanding Productivity Growth in Agriculture*: University of Chicago Press, pp. 77–124.

Huntingford, Chris, Philip D. Jones, Valerie N. Livina, Timothy M. Lenton, and Peter M. Cox (2013) "No increase in global temperature variability despite changing regional patterns," *Nature*, Vol. 500, No. 7462, pp. 327–330.

Judd, Kenneth L. (1996) "Approximation, perturbation, and projection methods in economic analysis," in Hans M. Amman, David A. Kendrick, and John Rust eds. *Handbook of Computational Economics*, Vol. 1: Elsevier, pp. 509–585.

Kala, Namrata (2017) "Learning, adaptation, and climate uncertainty: Evidence from Indian agriculture," working paper.

Keane, Michael P. and Kenneth I. Wolpin (2002) "Estimating welfare effects consistent with forward-looking behavior. Part I: Lessons from a simulation exercise," *The Journal of Human Resources*, Vol. 37, No. 3, pp. 570–599.

Kelly, David L., Charles D. Kolstad, and Glenn T. Mitchell (2005) "Adjustment costs from environmental change," *Journal of Environmental Economics and Management*, Vol. 50, No. 3, pp. 468–495.

Kim, Hyunseok and GianCarlo Moschini (2018) "The dynamics of supply: U.S. corn and soybeans in the biofuel era," *Land Economics*, Vol. 94, No. 4, pp. 593–613.

Klemm, Toni and Renee A. McPherson (2017) "The development of seasonal climate forecasting for agricultural producers," *Agricultural and Forest Meteorology*, Vol. 232, pp. 384–399.

Lemoine, Derek (2017) "Expect above average temperatures: Identifying the economic impacts of climate change," Working Paper 23549, National Bureau of Economic Research.

Lemoine, Derek and Sarah Kapnick (2016) "A top-down approach to projecting market impacts of climate change," *Nature Climate Change*, Vol. 6, No. 1, pp. 51–55.

Lucas, Robert E., Jr. (1972) "Econometric testing of the natural rate hypothesis," in Otto Eckstein ed. *The Econometrics of Price Determination*: Board of Governors of the Federal Reserve System, pp. 50–59.

Massetti, Emanuele and Robert Mendelsohn (2018) "Measuring climate adaptation: Methods and evidence," *Review of Environmental Economics and Policy*, Vol. 12, No. 2, pp. 324–341.

Mendelsohn, Robert (2000) "Efficient adaptation to climate change," *Climatic Change*, Vol. 45, No. 3-4, pp. 583–600.

——— (2019) "An examination of recent revealed preference valuation methods and results," *Review of Environmental Economics and Policy*, Vol. 13, No. 2, pp. 267–282.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw (1994) "The impact of global warming on agriculture: A Ricardian analysis," *The American Economic Review*, Vol. 84, No. 4, pp. 753–771.

Mérel, Pierre and Matthew Gammans (2019) "Climate econometrics: Can the panel approach account for long-run adaptation?" working paper, University of California Davis.

Nickell, Stephen (1981) "Biases in dynamic models with fixed effects," *Econometrica*, Vol. 49, No. 6, pp. 1417–1426.

Nordhaus, William (2013) "Integrated economic and climate modeling," in Peter B. Dixon and Dale W. Jorgenson eds. *Handbook of Computable General Equilibrium Modeling*, Vol. 1: North Holland, Elsevier B.V. pp. 1069–1131.

Nordhaus, William D. (2006) "Geography and macroeconomics: New data and new findings," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 103, No. 10, pp. 3510–3517.

Olmstead, Alan L. and Paul W. Rhode (2008) *Creating Abundance: Biological Innovation and American Agricultural Development*, New York, NY: Cambridge University Press.

——— (2011) "Responding to climatic challenges: Lessons from U.S. agricultural development," in Gary D. Libecap and Richard H. Steckel eds. *The Economics of Climate Change: Adaptations Past and Present*, pp. 169–194.

Quiggin, John and John Horowitz (2003) "Costs of adjustment to climate change," *Australian Journal of Agricultural and Resource Economics*, Vol. 47, No. 4, pp. 429–446.

Quiggin, John and John K. Horowitz (1999) "The impact of global warming on agriculture: A Ricardian analysis: Comment," *The American Economic Review*, Vol. 89, No. 4, pp. 1044–1045.

Roberts, Michael J. and Wolfram Schlenker (2011) "The evolution of heat tolerance of corn: Implications for climate change," in GaryD. Libecap and Richard H. Steckel eds. *The Economics of Climate Change: Adaptations Past and Present*, pp. 225–251.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher (2005) "Will U.S. agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach," *American Economic Review*, Vol. 95, No. 1, pp. 395–406.

Schlenker, Wolfram and Michael J. Roberts (2009) "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change," *Proceedings of the National Academy of Sciences*, Vol. 106, No. 37, pp. 15594–15598.

Severen, Christopher, Christopher Costello, and Olivier Deschnes (2018) "A forward-looking Ricardian approach: Do land markets capitalize climate change forecasts?" *Journal of Environmental Economics and Management*, Vol. 89, pp. 235–254.

Shrader, Jeffrey (2017) "Expectations and adaptation to environmental risks," working paper.

Sohngen, Brent and Robert Mendelsohn (1998) "Valuing the impact of large-scale ecological change in a market: The effect of climate change on U.S. timber," *The American Economic Review*, Vol. 88, No. 4, pp. 686–710.

Sorkin, Isaac (2015) "Are there long-run effects of the minimum wage?" *Review of Economic Dynamics*, Vol. 18, No. 2, pp. 306–333.

Takle, Eugene S., Christopher J. Anderson, Jeffrey Andresen, James Angel, Roger W. Elmore, Benjamin M. Gramig, Patrick Guinan, Steven Hilberg, Doug Kluck, Raymond Massey, Dev Niyogi, Jeanne M. Schneider, Martha D. Shulski, Dennis Todey, and Melissa Widhalm (2013) "Climate forecasts for corn producer decision making," *Earth Interactions*, Vol. 18, No. 5, pp. 1–8.

Taraz, Vis (2017) "Adaptation to climate change: historical evidence from the Indian monsoon," *Environment and Development Economics*, Vol. 22, No. 5, pp. 517–545.

Weyl, E. Glen (2019) "Price theory," *Journal of Economic Literature*, Vol. 57, No. 2, pp. 329–384.

Zhang, Peng, Olivier Deschenes, Kyle Meng, and Junjie Zhang (2018) "Temperature effects on productivity and factor reallocation: Evidence from a half million Chinese manufacturing plants," *Journal of Environmental Economics and Management*, Vol. 88, pp. 1–17.

Appendix

Appendix A establishes conditions under which reduced-form weather regressions can recover the effects of climate even without controlling for forecasts. Appendix B analyzes estimators that aggregate over multiple timesteps, including recently popular "long difference" estimators. Appendix C establishes cases in which we can use weather variation to unambiguously bound how climate affects payoffs. Appendix D provides details of the empirical implementation and additional results. Appendix E contains proofs. Appendix F generalizes the primary analysis to the case of vector-valued actions and multidimensional weather.

A Reduced-Form Weather Regressions Without Forecasts

The main text explores regressions that control for forecasts. However, the empirical literature has, almost without exception, not controlled for forecasts. Now consider such a regression:

$$A_{jt} = \alpha_j + \sum_{i=0}^{I} \gamma_{w_{t-i}} w_{j(t-i)} + \delta_{jt}. \tag{A-1}$$

The following proposition considers whether this regression can succeed in cases where previous ones could succeed:

Proposition 5. *Let either Assumption 1 or 2 hold, and let $(A_{j(t-1)} - \bar{A})^2$ be small for all observations.*

1. *If $\bar{\pi}_{12} = \beta\bar{\pi}_{23} = \bar{\pi}_{14} = 0$, then $\hat{\gamma}_{w_t} = \mathrm{d}\bar{A}/\,\mathrm{d}C$ for $I = 0$ or $I = 1$.*
2. *If $\bar{\pi}_{12} = \beta\bar{\pi}_{23} = 0$, then $\hat{\gamma}_{w_t} + \hat{\gamma}_{w_{t-1}} = \mathrm{d}\bar{A}/\,\mathrm{d}C$ for $I = 1$.*
3. *If $\rho = 0$ and either $\beta = 0$ or $\bar{\pi}_{12} = \bar{\pi}_{23} = 0$, then $\lim_{I\to\infty} \sum_{i=0}^{I} \hat{\gamma}_{w_{t-i}} = \mathrm{d}\bar{A}/\,\mathrm{d}C$.*

Proof. See Appendix E. □

Regression (A-1) does not control for forecasts or for past actions, so these affect the estimated $\hat{\gamma}$ as omitted variables. The first result establishes what we can learn from $\hat{\gamma}_{w_t}$, which is the coefficient of interest in much previous empirical literature. $\hat{\gamma}_{w_t}$ can capture part of the effect of time t forecasts when ν_t is positively correlated with ϵ_t (i.e., when $\rho > 0$); however, the proof shows that $\hat{\gamma}_{w_t}$ can never capture the total effect of forecasts. Omitted variables bias helps, but it cannot replace explicitly controlling for forecasts. Further, $\hat{\gamma}_{w_t}$ also misses the interaction between time t actions and past weather. Putting these pieces together, $\hat{\gamma}_{w_t}$ can fully recover climate impacts only if, in addition to the restriction from Propositions 1 and 2 that $\bar{\pi}_{12} = 0$, there is also no ex-ante adaptation that would use forecasts ($\beta\bar{\pi}_{23} = 0$) and past weather shocks do not affect actions directly ($\bar{\pi}_{14} = 0$). The second result shows that also using $\hat{\gamma}_{w_{t-1}}$ allows the estimator to succeed when past weather affects current choices, and the third result establishes that we can recover the effects of climate without using forecasts when agents are myopic.

Now consider a regression with payoffs as the dependent variable:

$$\pi_{jt} = \alpha_j + \sum_{i=0}^{I} \Phi_{w_{t-i}} w_{j(t-i)} + \delta_{jt}. \tag{A-2}$$

The error term δ_{jt} now includes not only actions but also current and past forecasts. Most empirical research sets $I = 0$. We are interested in the vector of coefficients Φ.[36]

Proposition 6. *Let Assumption 1 hold, or let Assumption 2 hold with the ϵ and ν normally distributed. Also let $(A_{j(t-1)} - \bar{A})^2$ and $(A_{jt} - \bar{A})^2$ be small for all observations and let each agent's average actions be $\bar{A}$.*

1. *If $\bar{\pi}_{12} = 0$ and $\beta\bar{\pi}_{23} = 0$, then:*
   - *(a)* $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \hat{\Phi}_{w_t} + \hat{\Phi}_{w_{t-1}} + \hat{\Phi}_{w_{t-2}}$ *for $I = 2$.*
   - *(b)* $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \sum_{i=0}^{I} \hat{\Phi}_{w_{t-i}}$ *for $I \geq 2$ and $\rho = 0$.*
   - *(c) If $\bar{\pi}_{14} = 0$, then* $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \hat{\Phi}_{w_t} + \hat{\Phi}_{w_{t-1}}$ *for $I = 1$ or $I = 2$.*
2. *If $\beta = 0$, then* $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \lim_{I\to\infty} \sum_{i=0}^{I} \hat{\Phi}_{w_{t-i}}$ *for $\rho = 0$.*
3. *If $\bar{\pi}_{23} = 0$, then* $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \lim_{\beta\to 1} \lim_{I\to\infty} \sum_{i=0}^{I} \hat{\Phi}_{w_{t-i}}$ *for $\rho = 0$.*
4. *If Assumption 3 holds, then:*
   - *(a)* $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \hat{\Phi}_{w_t} + \hat{\Phi}_{w_{t-1}}$ *for $I = 1$ or $I = 2$.*
   - *(b)* $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C = \sum_{i=0}^{I} \hat{\Phi}_{w_{t-i}}$ *for $\rho = 0$ and $I \geq 1$.*

Proof. See Appendix E. □

We can still recover the full effect of climate on payoffs if either Assumption 3 holds or agents are myopic. However, the other cases now require the absence of ex-ante adaptation.[37] Some

[36] The matrix of regressors is symmetric, tridiagonal, and Toeplitz. For general I and ρ, the inverse is not analytically convenient (Hu and O'Connell, 1996). Parts of the proposition assume $\rho = 0$ in order to simplify this inverse.

[37] Including a lead of weather in regression (A-2) (i.e., summing from $i = -1$ to I) would allow the sum of the $\hat{\Phi}$ to recover the effects of climate under the conditions given in Proposition 3 as $\rho \to 0$ and $\tau^2/(\tau^2 + \sigma^2) \to 1$.

of this ex-ante adaptation is captured through omitted variables bias in the plausible case where weather and forecasts are positively correlated (i.e., where $\rho > 0$), but the proof shows that it can never be captured completely. When agents can directly protect themselves against future weather outcomes, forecasts provide variation that is critical to identifying these responses (compare Keane and Wolpin, 2002).

Rather than focusing on the $\hat{\Phi}$, Deryugina and Hsiang (2017) undertake a different calculation. They estimate $\pi(A_t(w_t), A_{t-1}(w_t), w_t, w_{t-1}(w_t)) - \pi(A_t(w^0), A_{t-1}(w^0), w^0, w_{t-1}(w^0))$ for each w_t, where w^0 indicates an omitted weather category and where we write $A_{t-1}(w_t)$ and $w_{t-1}(w_t)$ in order to focus on questions besides the evaluation point. Let $p(w_t; C)$ represent the probability density function for weather in climate C. They calculate the marginal effect of climate from the following expression:

$$\int_{-\infty}^{\infty} \left[\pi(A_t, A_{t-1}, w_t, w_{t-1}) - \pi(A_t, A_{t-1}, w^0, w_{t-1})\right] \frac{\mathrm{d}p(w_t; C)}{\mathrm{d}C} \, \mathrm{d}w_t \triangleq \Psi.$$

Analyzing, we find

$$\Psi = Cov\left[\pi(A_t, A_{t-1}, w_t, w_{t-1}), \frac{\frac{\mathrm{d}p(w_t;C)}{\mathrm{d}C}}{p(w_t; C)}\right].$$

If w_t is normally distributed, then

$$\Psi = \frac{Cov\left[\pi(A_t, A_{t-1}, w_t, w_{t-1}),\ w_t\right]}{Var[w_t]},$$

which, following the proof of Proposition 6, is equal to $\hat{\Phi}_{w_t}$ with $I = 0$. Proposition 6 shows that this estimator recovers the effects of climate change in only the most special of cases.

B Aggregating over Longer Timesteps, Including Long Difference Estimators

In actual empirical work, the proper timestep of analysis may be unclear, computational requirements may require using coarser timesteps, or data may be available only over coarser timesteps. I therefore now consider the implications of aggregating weather and payoffs over longer timesteps.

Assume, as in the main text, that all agents are in the same climate and that this climate is stationary. The empirical researcher averages outcomes over Δ periods.[38] Denote the averages with a ˇ and use the time subscript to indicate the beginning of the averaging interval, so that, for instance, $\check{\pi}_{jt} \triangleq \sum_{T=t}^{t+\Delta-1} \pi_{jT}/\Delta$. Consider estimating the regression

$$\check{\pi}_{jt} = \alpha_j + \Lambda \check{w}_{jt} + \check{u}_{jt},$$

[38]The results do not depend on whether the operation is averaging or summing.

where I assume that the averaging intervals do not overlap.

The following proposition establishes properties of the estimator $\hat{\Lambda}$.

Proposition 7. *Let Assumption 1 hold, or let Assumption 2 hold with the ϵ and ν normally distributed. Also let $(A_{j(t-1)} - \bar{A})^2$ and $(A_{jt} - \bar{A})^2$ be small for all observations and let each agent's average actions be $\bar{A}$. Then the following conditions are individually sufficient for* $\lim_{s\to\infty}[\,\mathrm{d}E_0[\pi_s]/\,\mathrm{d}C] = \lim_{\Delta\to\infty} \hat{\Lambda}$*:*

1. *Assumption 3 holds.*

2. $\rho, \bar{\pi}_{12}, \beta\bar{\pi}_{23} = 0$.

3. $\rho, \bar{\pi}_{12} = 0$ *and* $\sigma^2/\tau^2 = 0$.

Proof. See Appendix E. □

Let $\hat{\Phi}^0_{w_t}$ be the estimator from regression (A-2) with $I = 0$. The proof shows that, for $\Delta > 2$, the estimator is

$$\hat{\Lambda} = \hat{\Phi}^0_{w_t} + \frac{\Delta - 1}{\Delta}\Upsilon_1 + \frac{\Delta - 2}{\Delta}\Upsilon_2,$$

with $\Upsilon_1, \Upsilon_2 \geq 0$. As Δ becomes small, the long-timestep estimator $\hat{\Lambda}$ converges towards $\hat{\Phi}^0_{w_t}$, which Proposition 6 showed can approximate the effect of climate change in only the most special of cases. As Δ becomes large, the coefficients on Υ_1 and Υ_2 go to 1. This is the case considered by Proposition 7. As Δ becomes large, $\hat{\Lambda}$ recovers the effect of climate change in a broader set of cases than does $\hat{\Phi}^0_{w_t}$: the process of aggregating Δ time periods into one picks up correlations between current payoffs and lags and leads of weather within these Δ periods.[39] However, $\hat{\Lambda}$ underperforms estimators analyzed in Section 4.2 that used forecasts. In particular, for $\rho, \bar{\pi}_{12} = 0$ and Δ large, the estimator $\hat{\Lambda}$ recovers the effect of climate only in the absence of ex-ante adaptation, which was the same restriction required by the estimator $\sum_{i=0}^{I} \hat{\Phi}_{w_{t-i}}$ studied in Proposition 6.

The estimator $\hat{\Lambda}$ is closely to the panel regression (A-2). This insight has implications for a recent literature developing "long difference" estimates of climate impacts. Rather than estimating either a cross-sectional or a panel model, this method instead averages weather and outcomes over two non-overlapping periods, differences the averages, and estimates how the differenced dependent variable changes with differenced average weather (e.g., Dell et al., 2012; Burke and Emerick, 2016). To many, this approach's appeal rests in providing "plausibly credible causal estimates of climate impacts that account for adaptation" (Auffhammer,

[39] Note, however, that aggregating over more periods reduces the sample size. I show that the bias of the probability limit falls as the length of the aggregating interval increases, but, with a finite sample, the variance of the estimator will also increase as the aggregating interval increases. Thus, the choice of Δ faces a bias-variance tradeoff.

2018, 45): differencing removes the unobserved fixed factors that may covary with climate in a cross-sectional regression, and the variation induced by spatially heterogeneous rates of climate change may identify the long-run adaptations missing from standard panel regressions. On this reasoning, comparing long difference estimates to standard panel estimates indicates whether short-run adaptation differs from long-run adaptation.

In the present setting, there is no climate change (C is constant over time), yet it is easy to show that the estimator $\hat{\Lambda}$ is equivalent to a long difference estimator. Proposition 7 therefore implies that long difference estimators are in fact identified by random differences in sequences of transient weather shocks over the aggregation intervals. At best, long difference estimators conflate this variation with differential rates of climate change, but at worst, they capture nothing but the same transient weather shocks as do the estimators in (A-2).[40] Intuitively, averaging over several periods does not eliminate the old sources of variation and need not introduce new sources of variation. In fact, previous work has found that long difference and panel estimators produce similar results (Burke and Emerick, 2016). We now see that this result should be unsurprising: rather than indicating the absence of long-run adaptation, the similarity may in fact be mechanical.

C Bounding the effects of climate on payoffs

The main text describes cases in which time series regressions recover the effects of climate on actions. It also describes some cases in which we can unambiguously bound the effect of climate on actions. I now explore when time series regressions generate clearly signed bounds on the effect of climate change on payoffs.[41]

Consider regression (12).

Proposition 8. *Let Assumption 1 hold, or let Assumption 2 hold with the ϵ and ν normally distributed. Also let $(A_{j(t-1)} - \bar{A})^2$ and $(A_{jt} - \bar{A})^2$ be small for all observations and let each agent's average actions be $\bar{A}$. Then:*

1. *No Ex-Post Adaptation: If $\bar{\pi}_{13}, \bar{\pi}_{14}, \bar{\pi}_{24} = 0$ and $\beta\bar{\pi}_{23} > 0$, then $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C < \lim_{I\to\infty}\left[\sum_{i=0}^{I}\hat{\theta}_{w_{t-i}} + \sum_{i=0}^{I}\hat{\theta}_{f_{t-i}}\right]$ if and only if $\bar{\pi}_{12}\bar{\pi}_2 < 0$.*

2. *No Ex-Ante Adaptation: If $\bar{\pi}_{23}, \bar{\pi}_{14} = 0$ and $\beta\bar{\pi}_{13} > 0$, then $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\,\mathrm{d}C < \lim_{I\to\infty}\left[\sum_{i=0}^{I}\hat{\theta}_{w_{t-i}} + \sum_{i=0}^{I}\hat{\theta}_{f_{t-i}}\right]$ if and only if $[\bar{\pi}_{12}]^2\bar{\pi}_2 < 0$.*

[40] That worst case arises when the climate actually did not change differentially (as modeled here) or when agents were not aware of ongoing changes in climate. See Dell et al. (2014) and Burke and Emerick (2016) for discussion of awareness of climate change.

[41] Recall that we have normalized actions so that $\mathrm{d}\bar{A}/\,\mathrm{d}C \geq 0$. It is straightforward to adapt the conditions below to a case with $\mathrm{d}\bar{A}/\,\mathrm{d}C$ either negative or of unknown sign.

Proof. The proof of Proposition 3 showed that

$$\lim_{I\to\infty}\sum_{i=0}^{I}\left[\hat{\theta}_{w_{t-i}}+\hat{\theta}_{f_{t-i}}\right]=\lim_{s\to\infty}\frac{\mathrm{d}E_0[\pi_s]}{\mathrm{d}C}+\frac{\beta(1-\beta)}{1-\frac{\bar{\pi}_{12}}{\chi_2}}\bar{\pi}_{12}\bar{\pi}_2\left\{\omega\Omega-\frac{\mathrm{d}\bar{A}}{\mathrm{d}C}\frac{1-\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\right\}.$$

First assume that $\bar{\pi}_{13},\bar{\pi}_{14},\bar{\pi}_{24}=0$ and $\beta\bar{\pi}_{23}>0$. We then have $\Omega=0$, $\mathrm{d}\bar{A}/\mathrm{d}C>0$, and

$$\lim_{I\to\infty}\sum_{i=0}^{I}\left[\hat{\theta}_{w_{t-i}}+\hat{\theta}_{f_{t-i}}\right]=\lim_{s\to\infty}\frac{\mathrm{d}E_0[\pi_s]}{\mathrm{d}C}-\frac{\beta(1-\beta)}{1-\frac{\bar{\pi}_{12}}{\chi_2}}\bar{\pi}_{12}\bar{\pi}_2\frac{\mathrm{d}\bar{A}}{\mathrm{d}C}\frac{1-\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}.$$

The final fraction is strictly positive by the stability restriction imposed following Lemma 1 in the main text. Therefore the bias term on the right-hand side has the same sign as $-\bar{\pi}_{12}\bar{\pi}_2$ and thus

$$\lim_{s\to\infty}\frac{\mathrm{d}E_0[\pi_s]}{\mathrm{d}C}-\lim_{I\to\infty}\sum_{i=0}^{I}\left[\hat{\theta}_{w_{t-i}}+\hat{\theta}_{f_{t-i}}\right]>0\text{ iff }\bar{\pi}_{12}\bar{\pi}_2>0.$$

We have established the first part of the proposition.

Next assume that $\bar{\pi}_{23},\bar{\pi}_{14}=0$ and $\beta\bar{\pi}_{13}>0$. $\bar{\pi}_{23}=0$ implies, by the definition of Ω and equation (4), that

$$\Omega=\frac{\frac{\mathrm{d}\bar{A}}{\mathrm{d}C}}{\chi_1}.$$

$\bar{\pi}_{13}>0$ implies that $\mathrm{d}\bar{A}/\mathrm{d}C>0$. Recall that

$$\omega\triangleq\frac{-\bar{\pi}_{11}-(1+\beta)\bar{\pi}_{12}-\beta\bar{\pi}_{22}}{\chi_2}>0.$$

Substituting, we find:

$$\lim_{I\to\infty}\sum_{i=0}^{I}\left[\hat{\theta}_{w_{t-i}}+\hat{\theta}_{f_{t-i}}\right]=\lim_{s\to\infty}\frac{\mathrm{d}E_0[\pi_s]}{\mathrm{d}C}+\frac{\beta(1-\beta)}{1-\frac{\bar{\pi}_{12}}{\chi_2}}\frac{\mathrm{d}\bar{A}}{\mathrm{d}C}\frac{1}{\chi_2}\bar{\pi}_{12}\bar{\pi}_2\left\{\frac{-\bar{\pi}_{11}-\beta\bar{\pi}_{12}-\beta\bar{\pi}_{22}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{12}\frac{\bar{\pi}_{12}}{\chi_0}-\beta\bar{\pi}_{22}}-1\right\}.$$

The term in braces is <0 if $\bar{\pi}_{12}>0$ and is >0 if $\bar{\pi}_{12}<0$. The bias on the right-hand side has the same sign as $-\bar{\pi}_2[\bar{\pi}_{12}]^2$ and thus

$$\lim_{s\to\infty}\frac{\mathrm{d}E_0[\pi_s]}{\mathrm{d}C}-\lim_{I\to\infty}\sum_{i=0}^{I}\left[\hat{\theta}_{w_{t-i}}+\hat{\theta}_{f_{t-i}}\right]>0\text{ iff }\bar{\pi}_2>0.$$

We have established the second part of the proposition.

□

The first part of the proposition signs the bias in our estimate of the effect of climate on payoffs when $\bar{\pi}_{12} \neq 0$ but there is no ex-post adaptation. In this case, $\Omega = 0$ and the estimator implicitly overestimates changes in actions if and only if $\bar{\pi}_{12} < 0$. When $\bar{\pi}_2 > 0$, equation (2) implies that $\bar{\pi}_1 + \bar{\pi}_2 > 0$, so that the effect of climate on actions increases payoffs. Overestimating (underestimating) this effect on actions then overestimates (underestimates) the net benefits of climate change. But if $\bar{\pi}_2 < 0$, then equation (2) implies that $\bar{\pi}_1 + \bar{\pi}_2 < 0$, so that the effect of climate on actions decreases payoffs. Overestimating (underestimating) this effect on actions then overestimates (underestimates) the net costs of climate change. Intuitively, a case with $\bar{\pi}_2, \bar{\pi}_{12} > 0$ is one in which actions are costly to adjust and are undertaken for future benefits, as with diverting crops to storage, and a case with $\bar{\pi}_2, \bar{\pi}_{12} < 0$ is one in which actions impose long-run costs but will not be maintained for long, as may be true of groundwater withdrawals. In either case, extrapolating from responses to weather overstates the cost of climate change; in other cases, the estimator is an overly optimistic estimate of the effect of climate change.

The second part of the proposition signs the bias in our estimate of the effect of climate on payoffs when $\bar{\pi}_{12} \neq 0$ but there is no ex-ante adaptation. Now the bias introduced by Ω is relevant and is proportional to $\mathrm{d}\bar{A}/\,\mathrm{d}C$. The proof implies that the estimator always implicitly underestimates changes in actions. As a result, the estimator overstates the benefits of climate change if and only if $\bar{\pi}_2 < 0$.

D Empirical Details

D.1 Sample construction and specification details

I use sales, expense, and farmland acreage data from the 1987, 1992, 1997, 2002, 2007, and 2012 U.S. Census of Agriculture. From 1997 on, data are available for download from the official Quick Stats site. I obtain the 1987 and 1992 data from files posted by Deschênes and Greenstone (2007) via Fisher et al. (2012). I use a balanced panel, dropping counties that are missing observations in any year. I construct the weather variables from data available for download from Wolfram Schlenker's web site, which follows Schlenker and Roberts (2009). In line with supplementary analyses in both Deschênes and Greenstone (2007) and Fisher et al. (2012), I include three weather variables ($K = 3$ in regression (13)): growing season precipitation (in mm), growing season degree days, and extreme growing season degree days. I define growing season degree days using temperatures between 10°C and 29°C. Consistent with Schlenker and Roberts (2009), I define extreme growing season degree days using temperatures above 29°C. Lags and leads are defined using adjacent years. Following arguments in Schlenker et al. (2005) and Fisher et al. (2012) regarding irrigation, the base specification restricts the sample to counties east of the 100th meridian.[42] And following

[42] County longitude is weighted by cropland, following previous literature.

Deschênes and Greenstone (2007), the base specification weights observations by (the square root of) average farmland acreage in a county over time. In Section D.2, I report results for counties west of the 100th meridian and for unweighted regressions. Table D-1 summarizes weather and economic data by year.

I measure profits as sales minus expenses, following Deschênes and Greenstone (2007). Whereas they use profits per acre as the dependent variable, I use profits as the dependent variable. One of the primary actions farmers may take is to choose their cultivated acreage (e.g., Scott, 2014). I do not normalize profits by acreage because I am interested in estimating such adaptation margins. Fisher et al. (2012) argue that market value may be a better dependent variable since it does not conflate storage decisions, but for present purposes, profits are the correct dependent variable because the theoretical analysis requires the dependent variable to be flow payoffs and because the theoretical analysis can account for storage as a type of action undertaken in response to the weather. The effects of storage should be captured by the estimates of adaptation.

Deschênes and Greenstone (2007) favor state-by-year fixed effects to account for unobservables such as local price shocks.[43] Fisher et al. (2012) raise concerns about the weather variation remaining once state-by-year fixed effects and county fixed effects combine to restrict the identifying variation to deviations from average weather that are not shared by nearby counties. In their Table A3, they report that three weather variables analogous to the ones used here explain around 1.5% of the variance with year fixed effects (column 1e) but explain only around 0.3% of the variance with state-by-year fixed effects (column 2e).[44] As a result of this pattern, they prefer year fixed effects. In my preferred specification, the variance explained by my twelve weather variables (which include two lags, a lead, and the contemporary value for each of three weather indexes) is around 12% with year fixed effects but only 2% with state-by-year fixed effects. Per weather variable, year fixed effects explain twice as much variance in my data as in Fisher et al. (2012) and state-by-year fixed effects explain over one-half more variance in my data than in Fisher et al. (2012).

However, some of the variation explained by weather in the case with year fixed effects could be "bad" variation due to unobservables such as local shocks to prices, costs, or productivity that are correlated with local weather shocks (Deschênes and Greenstone, 2007, 2012). I therefore also consider USDA Farm Region-by-year fixed effects (also explored in Deschênes and Greenstone, 2012). The USDA Farm Resource Regions cover geographic regions that are broader than states while also better reflecting patterns in crop production related to unobservables such as local price shocks (USDA, 2000).[45] There are nine Farm

[43] Carter et al. (2018) argue that interpreting marginal effects on profits in terms of social welfare requires holding prices constant.

[44] Variance explained by weather is calculated as 1 minus the ratio of residual variance from a specification with all weather variables over residual variance from a specification without any of the weather variables.

[45] I use the crosswalk between Farm Resource regions and counties available at `https://www.ers.usda.gov/data-products/arms-farm-financial-and-crop-production-practices/documentation.aspx`.

Resource Regions in the U.S., with eight of them including counties east of the 100th meridian (as opposed to 37 states that include these counties). I find that the variance explained by weather is 4.5% with these fixed effects, over two times greater than with state-by-year fixed effects. On a per-variable basis with Farm Region-by-year fixed effects, weather explains three-quarters as much variance as explained by weather with year fixed effects in Fisher et al. (2012) but nearly four times more variance than explained by weather with state-by-year fixed effects in Fisher et al. (2012). I use these Farm Resource Region-by-year fixed effects in my preferred specifications because these fixed effects navigate a tradeoff between absorbing omitted variables bias while leaving variation for weather to explain. In Section D.2, I report results for specifications that instead include either state-by-year fixed effects or year fixed effects.

I estimate covariance matrices that are cluster-robust at the state level, which accounts both for arbitrary serial correlation within a county and for arbitrary spatial correlation within a state (see also Fisher et al., 2012). The county fixed effects are nested within clusters when clustering either by state or by county. The degrees of freedom adjustment follows Cameron and Miller (2015).

I project climate change from the suite of 21 downscaled CMIP5 general circulation climate model projections from the NASA Earth Exchange (NEX) database, kindly provided by Wolfram Schlenker as county averages (weighted by measures of crop acreage). The base specification uses the RCP 4.5 trajectory as that trajectory of stabilized emissions is most consistent with the theoretical analysis of marginal climate change. (An alternate specification in Section D.2 considers the higher-warming RCP 8.5 trajectory.) I calculate each model's estimate of climate change by differencing average weather over 2075–2095 with average weather over 1985–2005. I then average over models' estimates to obtain a single climate change projection. I calculate the percentage change in profits due to climate change by multiplying the theory-implied marginal effects of changing each climate variable by the projected change in the weather variable and dividing by (acreage-weighted) average profits over the sample. Table D-2 summarizes county-level climate projections. Climate change increases both growing degree day variables but has heterogeneous effects on precipitation.

It may seem desirable to directly estimate the theory-implied parameters through the method of moments: one can replace the estimated coefficients in the ordinary least squares moment equations with their expressions in terms of theory-implied parameters. The problem is that the estimated covariance matrix is unreliable because calculating it can, depending on the data, require inverting a poorly conditioned (nearly-singular) matrix. The poorly conditioned matrix arises when the theory-implied parameters rely on division of reduced-form parameters that are close to zero. This happens to not pose a problem in the preferred specification (results given in Section D.2), but it does pose a problem in other specifications. Further, the distributions will be skewed in all specifications. I therefore instead obtain the theory-implied parameters by sampling from the joint distribution of the reduced-form parameters (as defined by their estimated means and the estimated covariance matrix). In this case, the same underlying problem leads the standard deviation and the mean to be unreliable statistics for the theory-implied parameters. I instead report the median and the lower and upper quartiles for the theory-implied results. In Section D.2, I also report the 10th and 90th percentiles. All reported results use 1 million samples. Results are robust to using 10 million samples.

Deschênes and Greenstone (2007) use sales, expenses, and farmland acreage data from the Census of Agriculture for 1987, 1992, 1997, and 2002, which overlaps with data currently available online from the Census of Agriculture only in the latter two years. Both Fisher et al. (2012) and Deschênes and Greenstone (2012) use the same sales, expense, and acreage data as Deschênes and Greenstone (2007). The three variables I downloaded exactly match the data used in those papers for 2002 but none of them ever matches the data used in those papers for 1997. On average, those papers' data underestimate sales by 3%, underestimate expenses by 6%, and underestimate farmland acreage by 5%, with substantial variation around these averages and with many observations overestimating these variables. The source of the discrepancy appears to be that the USDA changed its methodology for the 2002 Census of Agriculture. It had previously adjusted its data for non-response, but in 2002 it began also adjusting for coverage. The 1997 data currently available online include a coverage adjustment, but the data originally published for 1997 (and presumably used in those prior papers) do not. Further, there is no coverage adjustment available for the pre-1997 data, so there is no way to make them perfectly consistent with the more recent data. I assess robustness to this data issue in Section D.3. There I report specifications that do not use any of the data from Deschênes and Greenstone (2007) (so dropping 1987 and 1992), and I report specifications that use economic and acreage data only from Deschênes and Greenstone (2007) (so dropping 2007 and 2012 and replacing 1997 with their data), with and without the year 2002. For consistency with those previous papers' results, these last specifications project climate change using Scenario B2 from the Hadley III model (see Fisher et al., 2012), define growing season degree days using the interval 8–32°C, and define extreme growing season degree days as the square root of growing degree days above 34°C.

D.2 Additional results and robustness checks

Table D-3 reports the same results as in the main text except now using the 10th and 90th percentiles. Even the 90th percentile for the combined effect is negative. It is therefore unlikely that climate change will help agriculture on net. Also, even the 90th percentile for $\bar{\pi}_{12}/\chi_2$ (not shown) is negative in the case of extreme growing degree days, reinforcing the conclusion that $\bar{\pi}_{12}/\chi_2$ is likely to be negative.

Instead of reporting medians and quartiles, Table D-4 reports central estimates and standard errors obtained through method of moments estimation, with the caveats (from Section D.1) that the estimated standard errors can be computationally imprecise (although they do not appear to be in this specification) and may be misleading because the distri-

Table D-1: Summary statistics for the sample used in the preferred specification.

	1987	1992	1997	2002	2007	2012
Mean and standard deviation						
Profit (million $2002)	12.3	11.8	13.6	7.8	14.8	17.2
	(16.5)	(14.4)	(18.6)	(15.8)	(23.6)	(28.8)
GDD (thous °C-days)	2.02	1.76	1.78	2.02	2.01	2.04
	(0.415)	(0.501)	(0.468)	(0.452)	(0.422)	(0.440)
Extreme GDD (°C-days)	74.2	34.4	53.7	74.4	69.4	95.3
	(51.1)	(43.2)	(53.9)	(50.7)	(48.8)	(70.7)
Prec (mm)	550	616	591	597	569	534
	(114)	(126)	(132)	(136)	(152)	(175)
Weighted average						
Profit (million $2002)	15.0	14.5	16.8	9.1	18.7	22.4
GDD (thous °C-days)	2.03	1.76	1.79	2.02	2.01	2.05
Extreme GDD (°C-days)	78.0	37.6	59.3	79.0	68.5	107.7
Prec (mm)	542	601	579	583	590	499

The sample includes only counties east of the 100th meridian. Weights are the square root of a county's average acreage. There are 2324 counties and 6 years.

Table D-2: Projected effects of 21st century climate change for the sample used in the preferred specification.

	RCP 4.5	RCP 8.5
Mean and standard deviation		
GDD (thous °C-days)	0.394	0.709
	(0.0432)	(0.0884)
Extreme GDD (°C-days)	106	257
	(51.2)	(101)
Prec (mm)	23.3	21.3
	(26.4)	(41.0)
Weighted average		
GDD (thous °C-days)	0.397	0.714
Extreme GDD (°C-days)	111	266
Prec (mm)	17.0	11.1

Climate projections use the RCP 4.5 scenario from the NEX database. The sample includes only counties east of the 100th meridian. Weights are the square root of a county's average acreage. There are 2222 counties, slightly fewer than in the estimation sample.

butions may be highly skewed. Nonetheless, the broad story is consistent with the results we have already seen. The net effect of climate change is to reduce agricultural profits by 52–58% according to the central estimates, with a standard error of just over 20%. The central estimate (standard error) for $\bar{\pi}_{12}/\chi_2$ is -2.73 (3.21) in the case of conventional growing degree days and -0.531 (0.305) in the case of extreme growing degree days.

The remaining tables undertake robustness checks. All specifications include county fixed effects. First, Table D-5 assesses sensitivity to a lower discount rate, which can affect only the theory-based estimates. We see only small changes relative to the main text. The most interesting change is mechanical: reducing the discount rate reduces the degree of ex-ante adaptation required to explain the coefficient on the reduced-form lead of weather. The median estimates of $\bar{\pi}_{12}/\chi_2$ are essentially unchanged.

Table D-6 replaces the Farm Region-by-year fixed effects with year fixed effects. The estimated $\bar{\pi}_{12}/\chi_2$ are not too different from before. I point out a few differences. First, the median estimate of the direct effects of conventional growing degree days is now negative, which is cause for concern about specification error. Second, the estimates suggest a much larger role for both types of adaptation. If we believe that geographically differentiated fixed effects do in fact absorb local shocks to prices, costs, and productivity, then it may not be surprising that estimation of direct effects and adaptation is especially sensitive to including these fixed effects. Third, the interquartile ranges tend to become much larger than in the preferred specification.

Table D-7 instead replaces the Farm Region-by-year fixed effects with state-by-year fixed effects. The inclusion of state-by-year fixed effects shrinks the central estimates of the net effect of climate change towards zero and makes the sign unclear. This effect is consistent with Fisher et al. (2012). The estimated $\bar{\pi}_{12}/\chi_2$ is still clearly negative in the case of extreme growing degree days, but the median estimate is now much closer to zero in the case of conventional growing degree days.

Table D-8 replicates the main specification but without weighting by average farmland acreage. It therefore estimates effects for the average county rather than for the average acre of farmland. The results are broadly similar (including for $\bar{\pi}_{12}/\chi_2$), although ex-ante adaptation now plays a larger role in offsetting the harm from extreme growing degree days.

Table D-9 repeats the climate change calculation with the high-warming RCP 8.5 trajectory. This trajectory amplifies the benefit from additional growing degree days but also amplifies the cost of additional extreme growing degree days. The latter effect dominates, so that the results now suggest much greater losses from climate change. In fact, the median estimate suggests more-than-complete elimination of agricultural profits in the eastern U.S. This extreme result suggests that extrapolating estimated impacts of marginal climate change to this climate change scenario may be inappropriate.

Table D-10 studies the other side of the contiguous United States: those counties west of the 100th meridian. Schlenker et al. (2005) argue that these counties tend to be irrigated, whereas counties east of the 100th meridian tend to be rainfed. Consistent with intuition, the results suggest that the western counties are broadly less exposed to climate change and could even plausibly benefit from climate change. In particular, projected changes in extreme growing degree days are now plausibly beneficial and projected changes in conventional growing degree days are now plausibly harmful. The 75th percentile estimates of $\bar{\pi}_{12}/\chi_2$ are negative for both growing degree day variables. One difficulty with these results is that there are only 17 states in this region, raising concerns about the reliability of clustering by state. Table D-11 repeats the analysis, except clustering by county. The only notable difference in the results is that the median theory-implied estimate for total effects from extreme growing degree days now suggests negative effects, driven by costly adaptation.

Additional experiments (not reported) split the sample by USDA Farm Region. The "Heartland", "Northern Great Plains", "Eastern Uplands", and "Prairie Gateway" regions demonstrate results broadly consistent with the aggregated results, being helped by changes in conventional growing degree days, harmed by changes in extreme growing degree days, and harmed by climate change overall. The most interesting exception is the "Mississippi Portal", which appears to be harmed by changes in conventional growing degree days, to benefit from changes in extreme growing degree days, and to benefit from climate change overall.

D.3 Robustness to adjustments in USDA economic and acreage data

As described in Section D.1, the USDA changed their census methodology in 2002, with corrections available only back to 1997. I begin by assessing the robustness of my results to maintaining an internally consistent sample of years. I then assess the robustness of prior literature's results to maintaining an internally consistent sample of years. In the course of the latter, I also assess whether my results are much affected by using prior literature's slightly different weather variable definitions.[46]

Table D-12 does not use observations for 1987 and 1992, instead sticking to the years that are currently available online from the USDA and that adjust for both coverage and non-response. The direct effects of extreme growing degree days are now smaller, and ex-ante adaptation appears more important.[47] Otherwise, estimates are largely similar, albeit noisier. The median estimates suggest smaller total impacts from climate change, but even this much noisier estimate suggests that total impacts are likely to be negative. The estimated $\bar{\pi}_{12}/\chi_2$ are quite noisy. The estimates for conventional growing degree days and precipitation could plausibly have either sign, but the 75th percentile for $\bar{\pi}_{12}/\chi_2$ is still clearly negative (-0.5)

[46] The effects of climate are calculated as a percentage of average profits over the sample. Changing the sample therefore also affects the denominator of this calculation.

[47] The skill and availability of seasonal forecasts improved dramatically from the early years of the primary specification's sample to the years used in this table (see Klemm and McPherson, 2017, among others). It is possible that ex-ante adaptation actually did become more important over time.

in the case of extreme growing degree days.

The remaining tables replicate the current approach using only the economic and acreage data from Deschênes and Greenstone (2007), which means not using any years after 2002 and changing data values for 1997. Section D.1 describes differences in weather variable definitions and climate change calculations between these specifications and all previous ones. Table D-13 repeats the preferred specification with these alterations. The estimated net effect of climate change from Assumption 3 is nearly identical to the most comparable previous estimates, in column (1e) of Table A3 in Fisher et al. (2012). That specification differs in using year fixed effects (instead of the current Farm Region-year fixed effects) and in using profits per acre as the dependent variable rather than profits (which also creates some minor differences in the climate change calculation). The most notable difference with respect to the present paper's results is that estimated ex-post adaptation to growing degree days becomes rather noisy, which makes the net effect of climate change rather noisy. The median estimate for $\bar{\pi}_{12}/\chi_2$ is now positive in the case of conventional growing degree days but still negative in the case of extreme growing degree days. Neither interquartile range includes zero.

Table D-14 modifies the previous paragraph's specification to also drop all observations for the year 2002, so that the economic and acreage time series are now internally consistent. The reduced-form estimates using Assumption 3 shrink a bit. The theory-based estimates change in more interesting ways. The estimate of ex-post adaptation to growing degree days is now negative and, despite the sample being 25% smaller, is not nearly as noisy. As a result, the total effect of climate change changes fairly dramatically. In fact, it is now not too far from the main text's results, despite the different years used, the different weather variable definitions, and the different climate model projections. Finally, the interquartile range for $\bar{\pi}_{12}/\chi_2$ is still well within negative values in the case of extreme growing degree days, and it also now includes negative values in the case of conventional growing degree days.

Table D-15 combines economic and acreage data from Deschênes and Greenstone (2007) with the state-by-year fixed effects favored by Deschênes and Greenstone (2007, 2012). The estimated net effect of climate change from Assumption 3 is again nearly identical to the most comparable previous estimates, in column (2e) of Table A3 in Fisher et al. (2012). That specification differs only in using profits per acre as the dependent variable rather than profits. The theory-implied effects are broadly consistent with those in Table D-7 (which also used state-by-year fixed effects), with the exception that the signs of the adaptation channels flip for extreme growing degree days. The median estimates for $\bar{\pi}_{12}/\chi_2$ are negative for both growing degree day variables. Table D-16 drops all observations for the year 2002. The reduced-form estimates from Assumption 3 do not change much, suggesting that the results in Deschênes and Greenstone (2007, 2012) are not too sensitive to the USDA's change in variable construction between 1997 and 2002. However, the theory-based total effect of climate change is now clearly negative, driven by much smaller direct benefits from additional growing degree days. The median estimates for $\bar{\pi}_{12}/\chi_2$ are still negative for both growing degree day variables.

Table D-3: Like Table 2, except reporting the 10th and 90th percentiles (in parentheses) instead of the 25th and 75th percentiles.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	51	-97	-1	-47
	(25)	(37)	(0.4)	(24)
Using $\bar{\pi}_{12} = 0$	44	-1.1e+02	-1.6	-71
	(29)	(35)	(0.66)	(28)
Theory-Implied				
Direct Effects	38	-93	-1.2	-56
	(6.9,69)	(-1.3e+02,-58)	(-2.1,-0.44)	(-85,-28)
Ex-Post Adaptation	8.3	-8	-0.21	0.56
	(-12,36)	(-13,-2.6)	(-0.4,-0.061)	(-20,28)
Ex-Ante Adaptation	3.2	0.33	-0.048	3.4
	(-1.3,7.6)	(-5.6,6.2)	(-0.12,0.022)	(-3,9.9)
Total	54	-1e+02	-1.5	-50
	(2.2,95)	(-1.4e+02,-60)	(-2.5,-0.6)	(-92,-9.4)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 13944 county-year observations and 37 state observations.

Table D-4: Like Table 2, except reporting the central estimates and standard errors from method of moments estimation of the structural parameters.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	50.9	-97.3	-1.03	-47.4
	(25.4)	(36.6)	(0.4)	(23.8)
Using $\bar{\pi}_{12} = 0$	44	-114	-1.59	-71.5
	(29.4)	(34.6)	(0.659)	(28.4)
Theory-Implied				
Direct Effects	35.5	-91.8	-1.17	-57.5
	(23.6)	(27)	(0.567)	(21.1)
Ex-Post Adaptation	9.75	-7.42	-0.185	2.15
	(13.4)	(3.79)	(0.111)	(12.8)
Ex-Ante Adaptation	3.16	0.325	-0.0484	3.44
	(3.45)	(4.6)	(0.0547)	(5.06)
Total	48.4	-98.9	-1.41	-51.9
	(23.8)	(31)	(0.636)	(22.6)

All specifications include county and Farm Region-year fixed effects. Standard errors, in parentheses, are clustered at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 13944 county-year observations and 37 state observations.

Table D-5: Like Table 2, except using a lower annual discount rate of 12%.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	51	-97	-1	-47
	(25)	(37)	(0.4)	(24)
Using $\bar{\pi}_{12} = 0$	44	-1.1e+02	-1.6	-71
	(29)	(35)	(0.66)	(28)
Theory-Implied				
Direct Effects	41	-96	-1.4	-57
	(25,57)	(-1.2e+02,-77)	(-1.9,-0.91)	(-72,-42)
Ex-Post Adaptation	3.1	-3.2	-0.084	-0.043
	(0.83,6.2)	(-4.2,-2.2)	(-0.13,-0.052)	(-2.4,3)
Ex-Ante Adaptation	1.1	0.12	-0.017	1.2
	(0.3,1.9)	(-0.98,1.2)	(-0.03,-0.0041)	(0.009,2.4)
Total	47	-99	-1.5	-55
	(29,65)	(-1.2e+02,-79)	(-2,-0.99)	(-71,-38)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 13944 county-year observations and 37 state observations.

Table D-6: Like Table 2, except using year fixed effects instead of Farm Region-by-year fixed effects.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	53	-1.1e+02	-1.9	-61
	(22)	(27)	(0.57)	(35)
Using $\bar{\pi}_{12} = 0$	2.1	-1.2e+02	-2.3	-1.2e+02
	(53)	(25)	(0.91)	(54)
Theory-Implied				
Direct Effects	-19	-69	-1.7	-91
	(-49,10)	(-85,-54)	(-2.3,-1.2)	(-1.2e+02,-61)
Ex-Post Adaptation	39	-18	-0.29	21
	(-38,1e+02)	(-22,-15)	(-0.37,-0.22)	(-54,80)
Ex-Ante Adaptation	17	-6	-0.056	11
	(13,22)	(-9.5,-2.5)	(-0.12,0.0053)	(7.7,15)
Total	59	-94	-2.1	-39
	(-36,99)	(-1.1e+02,-76)	(-2.7,-1.5)	(-1.3e+02,9.6)

All specifications include county and year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 13944 county-year observations and 37 state observations.

Table D-7: Like Table 2, except using state-by-year fixed effects instead of Farm Region-by-year fixed effects.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	81	-89	-0.093	-8
	(33)	(38)	(0.3)	(37)
Using $\bar{\pi}_{12} = 0$	89	-82	-0.36	6.6
	(40)	(30)	(0.58)	(32)
Theory-Implied				
Direct Effects	89	-80	-0.42	9.8
	(58,1.2e+02)	(-99,-61)	(-0.98,0.04)	(-19,37)
Ex-Post Adaptation	-4.5	-4.9	-0.14	-9.5
	(-13,5.2)	(-17,3.5)	(-0.3,0.029)	(-28,6.3)
Ex-Ante Adaptation	-6	-2.6	0.048	-8.6
	(-10,-1.8)	(-6.5,1.2)	(-0.0065,0.1)	(-14,-3.3)
Total	79	-86	-0.51	-5.3
	(43,1.1e+02)	(-1.2e+02,-58)	(-1.2,0.095)	(-49,32)

All specifications include county and state-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 13944 county-year observations and 37 state observations.

Table D-8: Like Table 2, except not weighting the observations by average farmland acreage.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	45	-89	-0.72	-46
	(23)	(30)	(0.29)	(18)
Using $\bar{\pi}_{12} = 0$	34	-1.1e+02	-1.7	-81
	(24)	(28)	(0.55)	(24)
Theory-Implied				
Direct Effects	36	-99	-1.3	-65
	(22,50)	(-1.2e+02,-82)	(-1.6,-1.1)	(-80,-51)
Ex-Post Adaptation	4.5	-3.3	-0.21	2.1
	(-10,23)	(-6.4,0.69)	(-0.27,-0.14)	(-12,21)
Ex-Ante Adaptation	3.1	3.7	0.042	6.8
	(0.81,5.3)	(1.4,6.1)	(0.0056,0.079)	(3.9,9.8)
Total	48	-98	-1.5	-53
	(14,76)	(-1.2e+02,-78)	(-1.8,-1.2)	(-79,-28)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 13944 county-year observations and 37 state observations.

Table D-9: Like Table 2, except using RCP8.5.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	91	-2.3e+02	-0.67	-1.4e+02
	(46)	(87)	(0.26)	(60)
Using $\bar{\pi}_{12} = 0$	79	-2.7e+02	-1	-1.9e+02
	(53)	(83)	(0.43)	(65)
Theory-Implied				
Direct Effects	68	-2.2e+02	-0.81	-1.5e+02
	(40,98)	(-2.6e+02,-1.8e+02)	(-1.1,-0.54)	(-1.9e+02,-1.2e+02)
Ex-Post Adaptation	15	-19	-0.13	-3.2
	(3.8,31)	(-25,-13)	(-0.19,-0.087)	(-15,13)
Ex-Ante Adaptation	5.7	0.78	-0.032	6.4
	(1.5,9.9)	(-6.6,8.2)	(-0.056,-0.0075)	(-1.2,14)
Total	98	-2.4e+02	-0.97	-1.5e+02
	(58,1.3e+02)	(-2.9e+02,-1.9e+02)	(-1.3,-0.68)	(-1.9e+02,-1e+02)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 8.5 scenario averaged across 21 CMIP5 models. There are 13944 county-year observations and 37 state observations.

Table D-10: Like Table 2, except using only counties west of the 100th meridian.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	-40	5.3	0.31	-35
	(20)	(14)	(0.22)	(24)
Using $\bar{\pi}_{12} = 0$	-26	26	0.88	0.3
	(33)	(22)	(0.49)	(44)
Theory-Implied				
Direct Effects	-35	48	0.91	15
	(-59,-14)	(23,72)	(0.15,1.5)	(-26,53)
Ex-Post Adaptation	16	-28	0.19	-15
	(9.8,23)	(-65,-10)	(0.033,0.36)	(-52,1.7)
Ex-Ante Adaptation	8.6	-15	-0.044	-6.4
	(3.8,14)	(-23,-7)	(-0.12,0.029)	(-13,0.58)
Total	-7.9	1.4	1.1	-9.2
	(-31,13)	(-29,22)	(0.21,1.7)	(-63,31)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties west of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 3240 county-year observations and 17 state observations.

Table D-11: Like Table 2, except using only counties west of the 100th meridian and clustering by county.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	-40	5.3	0.31	-35
	(21)	(14)	(0.17)	(17)
Using $\bar{\pi}_{12} = 0$	-26	26	0.88	0.3
	(36)	(30)	(0.33)	(31)
Theory-Implied				
Direct Effects	-35	47	1.1	13
	(-59,-12)	(25,69)	(0.45,1.5)	(-12,38)
Ex-Post Adaptation	15	-27	0.23	-12
	(10,21)	(-62,-13)	(0.085,0.39)	(-50,3)
Ex-Ante Adaptation	8.6	-15	-0.044	-6.4
	(4.8,12)	(-21,-9.1)	(-0.08,-0.0074)	(-13,-0.14)
Total	-11	-5	1.3	-9.9
	(-34,13)	(-29,11)	(0.57,1.8)	(-54,22)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the county level. The sample includes only counties west of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 3240 county-year observations and 540 county observations.

Table D-12: Like Table 2, except using only data from 1997–2012.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	64	-1e+02	-0.96	-39
	(32)	(41)	(0.53)	(31)
Using $\bar{\pi}_{12} = 0$	18	-91	-0.3	-73
	(50)	(31)	(0.76)	(49)
Theory-Implied				
Direct Effects	37	-64	0.28	-27
	(0.59,73)	(-81,-47)	(-0.16,0.75)	(-58,3)
Ex-Post Adaptation	-1.5	-10	-0.21	-14
	(-24,14)	(-23,-3.9)	(-0.74,0.56)	(-49,10)
Ex-Ante Adaptation	3.6	-6.2	-0.32	-2.9
	(-1.1,8.5)	(-11,-1.4)	(-0.4,-0.25)	(-9.2,3.4)
Total	45	-80	-0.32	-38
	(-9.8,85)	(-1.1e+02,-56)	(-1.2,0.94)	(-1e+02,4)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the RCP 4.5 scenario averaged across 21 CMIP5 models. There are 9380 county-year observations and 37 state observations.

Table D-13: Like Table 2, except using years and weather variable definitions from Deschênes and Greenstone (2007).

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	-0.9	-68	-0.1	-69
	(19)	(18)	(0.12)	(22)
Using $\bar{\pi}_{12} = 0$	-4	-91	-0.27	-95
	(34)	(24)	(0.13)	(36)
Theory-Implied				
Direct Effects	48	-82	-0.047	-41
	(25,71)	(-98,-67)	(-0.29,0.14)	(-67,-13)
Ex-Post Adaptation	26	21	-0.019	49
	(-29,54)	(9.4,38)	(-0.11,0.11)	(-16,94)
Ex-Ante Adaptation	-12	12	-0.017	0.063
	(-16,-7.4)	(9,14)	(-0.034,-0.00024)	(-3.8,3.9)
Total	65	-47	-0.064	16
	(-32,1.2e+02)	(-69,-24)	(-0.42,0.21)	(-94,87)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the B2 scenario from the Hadley III model. There are 9528 county-year observations and 37 state observations.

Table D-14: Like Table D-13, except using only data from 1987–1997.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	6.4	-50	-0.23	-44
	(22)	(15)	(0.15)	(24)
Using $\bar{\pi}_{12} = 0$	-9.1	-89	-0.49	-99
	(31)	(27)	(0.17)	(24)
Theory-Implied				
Direct Effects	18	-96	-0.51	-80
	(-17,44)	(-1.1e+02,-77)	(-0.65,-0.36)	(-1e+02,-60)
Ex-Post Adaptation	-10	17	-0.054	10
	(-20,3.8)	(6.7,35)	(-0.11,-0.0097)	(-7.7,44)
Ex-Ante Adaptation	-11	16	0.055	4.5
	(-14,-7.8)	(13,18)	(0.036,0.073)	(0.82,8.2)
Total	-11	-58	-0.51	-66
	(-48,38)	(-81,-34)	(-0.68,-0.34)	(-1e+02,-20)

All specifications include county and Farm Region-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the B2 scenario from the Hadley III model. There are 7146 county-year observations and 37 state observations.

Table D-15: Like Table 2, except using years and weather variable definitions from Deschênes and Greenstone (2007) and using their preferred state-year fixed effects.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	43	-50	0.077	-7.8
	(16)	(13)	(0.11)	(18)
Using $\bar{\pi}_{12} = 0$	54	-57	-0.097	-3.6
	(21)	(23)	(0.15)	(24)
Theory-Implied				
Direct Effects	71	-79	-0.11	-7.6
	(56,86)	(-94,-63)	(-0.24,0.0086)	(-26,10)
Ex-Post Adaptation	-8.5	8.3	-0.044	-1.3
	(-17,4.2)	(-0.083,14)	(-0.17,0.11)	(-12,11)
Ex-Ante Adaptation	-11	9.5	0.049	-1.4
	(-16,-5.9)	(6.3,13)	(0.024,0.074)	(-5.3,2.6)
Total	50	-60	-0.077	-10
	(32,71)	(-79,-43)	(-0.37,0.17)	(-36,16)

All specifications include county and state-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the B2 scenario from the Hadley III model. There are 9528 county-year observations and 37 state observations.

Table D-16: Like Table D-15, except using only data from 1987–1997.

	GDD	Extreme GDD	Precip	Combined
Reduced-Form				
Using Assumption 3	37	-39	0.036	-1.7
	(16)	(15)	(0.12)	(11)
Using $\bar{\pi}_{12} = 0$	33	-61	-0.19	-28
	(27)	(24)	(0.19)	(20)
Theory-Implied				
Direct Effects	42	-68	-0.19	-27
	(25,60)	(-86,-51)	(-0.37,-0.011)	(-46,-7.3)
Ex-Post Adaptation	-9.1	0.99	-0.082	-6.8
	(-16,1.7)	(-18,16)	(-0.25,0.0034)	(-21,6)
Ex-Ante Adaptation	-6.1	7.3	0.065	1.2
	(-10,-2.1)	(5,9.5)	(0.034,0.097)	(-2.5,5)
Total	27	-55	-0.22	-31
	(8.5,48)	(-94,-30)	(-0.55,0.07)	(-59,-6.8)

All specifications include county and state-year fixed effects. The reduced-form estimates' standard errors and the theory-implied results derive from covariance matrices that are robust to clustering at the state level. The sample includes only counties east of the 100th meridian. Observations are weighted by (the square root of) a county's average farmland acreage. Climate projections use the B2 scenario from the Hadley III model. There are 7146 county-year observations and 37 state observations.

E Proofs

E.1 A Useful Lemma

The following lemma is critical to deriving the regression coefficients. Define $\tilde{X}_K$ as the $JT \times 2(K+1)$ matrix with rows

$$\begin{bmatrix} w_{jt} - C & f_{jt} - C & w_{j(t-1)} - C & f_{j(t-1)} - C & \dots & w_{j(t-K)} - C & f_{j(t-K)} - C \end{bmatrix},$$

where the ϵ and ν are uncorrelated over time and across units j (though allowing for the possibility that $\rho \triangleq Cov[\epsilon_{jt}, \nu_{jt}] \neq 0$). And define $\mathbf{0}_R$ as the $1 \times R$ row vector of zeros if $R > 0$ and as an empty element element if $R \leq 0$. We then have:

Lemma 4. *For $K > 0$ and $i \leq 2K$, row i of $E[\tilde{X}_K^\intercal \tilde{X}_K]^{-1}$ obeys the following rules.*

- *If i is odd, then*

$$\frac{1}{JT\zeta^2[\sigma^2\tau^2 - \rho^2]} \begin{bmatrix} \mathbf{0}_{i-1} & \tau^2 & -\rho & 0 & -\tau^2 & \mathbf{0}_{2(K+1)-(i+3)} \end{bmatrix}.$$

- *If $i = 2$, then*

$$\frac{1}{JT\zeta^2[\sigma^2\tau^2 - \rho^2]} \begin{bmatrix} -\rho & \sigma^2 & 0 & \rho & \mathbf{0}_{2(K+1)-4} \end{bmatrix}.$$

- *If i is even and $i > 2$, then*

$$\frac{1}{JT\zeta^2[\sigma^2\tau^2 - \rho^2]} \begin{bmatrix} \mathbf{0}_{i-4} & -\tau^2 & \rho & -\rho & \sigma^2 + \tau^2 & 0 & \rho & \mathbf{0}_{2(K+1)-(i+2)} \end{bmatrix}.$$

Proof. Observe that, for $K > 0$,

$$E[\tilde{X}_K^\intercal \tilde{X}_K] = \begin{bmatrix} E[\tilde{X}_{K-1}^\intercal \tilde{X}_{K-1}] & C_{K-1} \\ C_{K-1}^\intercal & D \end{bmatrix},$$

where C_{K-1} is a $2K \times 2$ matrix with the only nonzero entries being in row $2K - 1$, which is $[JT\zeta^2\rho \quad JT\zeta^2\tau^2]$, and where

$$D \triangleq JT\zeta^2 \begin{bmatrix} \sigma^2 + \tau^2 & \rho \\ \rho & \tau^2 \end{bmatrix}.$$

Define

$$B_K \triangleq \begin{bmatrix} E[\tilde{X}_{K-1}^\intercal \tilde{X}_{K-1}] & C_{K-1} \\ C_{K-1}^\intercal & B_0 \end{bmatrix}$$

for $K > 0$, where

$$B_0 \triangleq JT\zeta^2 \begin{bmatrix} \sigma^2 & \rho \\ \rho & \tau^2 \end{bmatrix}.$$

The only nonzero entries in the $2K \times 2$ matrix $C_{K-1}D^{-1}$ are in row $2K-1$, which is $[0 \quad 1]$. Therefore[48]

$$B_{K-1} = E[\tilde{X}_{K-1}^\intercal \tilde{X}_{K-1}] - C_{K-1}D^{-1}C_{K-1}^\intercal.$$

Using standard results for block matrix inversion and substituting B_{K-1},

$$E[\tilde{X}_K^\intercal \tilde{X}_K]^{-1} = \begin{bmatrix} B_{K-1}^{-1} & -B_{K-1}^{-1}C_{K-1}D^{-1} \\ -D^{-1}C_{K-1}^\intercal B_{K-1}^{-1} & D^{-1} + D^{-1}C_{K-1}^\intercal B_{K-1}^{-1} C_{K-1} D^{-1} \end{bmatrix}. \tag{E-3}$$

Also, we have, for $K > 0$,[49]

$$B_K = E[\tilde{X}_K^\intercal \tilde{X}_K] - C_K B_0^{-1} C_K^\intercal$$

and thus

$$B_K^{-1} = \begin{bmatrix} B_{K-1}^{-1} & -B_{K-1}^{-1}C_{K-1}B_0^{-1} \\ -B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1} & B_0^{-1} + B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1} C_{K-1} B_0^{-1} \end{bmatrix}. \tag{E-4}$$

We proceed by induction, first directly proving the result for $K = 1$ and $K = 2$. Consider $K = 0$. Direct calculations yield:

$$\begin{aligned} E[\tilde{X}_0^\intercal \tilde{X}_0] =& JT\zeta^2 \begin{bmatrix} \sigma^2 + \tau^2 & \rho \\ \rho & \tau^2 \end{bmatrix} \\ \Rightarrow E[\tilde{X}_0^\intercal \tilde{X}_0]^{-1} =& \frac{1}{JT\zeta^2[\sigma^2\tau^2 - \rho^2]} \begin{bmatrix} \tau^2 & -\rho \\ -\rho & \sigma^2 + \tau^2 \end{bmatrix}. \end{aligned}$$

Now consider $K = 1$. We have:

$$\begin{aligned} B_0^{-1} =& \frac{1}{JT\zeta^2[\sigma^2\tau^2 - \rho^2]} \begin{bmatrix} \tau^2 & -\rho \\ -\rho & \sigma^2 \end{bmatrix}, \\ -B_0^{-1}C_0D^{-1} =& -B_0^{-1} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \frac{1}{JT\zeta^2[\sigma^2\tau^2 - \rho^2]} \begin{bmatrix} 0 & -\tau^2 \\ 0 & \rho \end{bmatrix}. \end{aligned}$$

Therefore the first 2 rows of $E[\tilde{X}_1^\intercal \tilde{X}_1]^{-1}$ are:

$$\frac{1}{JT\zeta^2[\sigma^2\tau^2 - \rho^2]} \begin{bmatrix} \tau^2 & -\rho & 0 & -\tau^2 \\ -\rho & \sigma^2 & 0 & \rho \end{bmatrix}.$$

[48] The only nonzero entry in the $2K \times 2K$ matrix $C_{K-1}D^{-1}C_{K-1}^\intercal$ is entry $(2K-1, 2K-1)$, which is $JT\zeta^2\tau^2$. Subtracting this entry from $E[\tilde{X}_{K-1}^\intercal \tilde{X}_{K-1}]$ yields the result.

[49] The only nonzero entries in the $2K \times 2$ matrix $C_{K-1}B_0^{-1}$ are in row $2K-1$, which is $[0 \quad 1]$, and thus the only nonzero entry in the $2K \times 2K$ matrix $C_{K-1}B_0^{-1}C_{K-1}^\intercal$ is entry $(2K-1, 2K-1)$, which is $JT\zeta^2\tau^2$. Subtracting this entry from $E[\tilde{X}_{K-1}^\intercal \tilde{X}_K]$ yields the result.

This is consistent with Lemma 4.

Now consider $K = 2$. Note that

$$-B_0^{-1}C_0B_0^{-1} = -B_0^{-1}\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}0 & -\tau^2\\ 0 & \rho\end{bmatrix},$$

$$-B_0^{-1}C_0^{\intercal}B_0^{-1} = -\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}B_0^{-1} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}0 & 0\\ -\tau^2 & \rho\end{bmatrix}$$

$$\begin{aligned}B_0^{-1} + B_0^{-1}C_0^{\intercal}B_0^{-1}C_0B_0^{-1} =& B_0^{-1} + \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}B_0^{-1}C_0^{\intercal}\begin{bmatrix}0 & \tau^2\\ 0 & -\rho\end{bmatrix} = B_0^{-1} + \frac{\tau^2}{[\sigma^2\tau^2-\rho^2]}B_0^{-1}\begin{bmatrix}0 & \rho\\ 0 & \tau^2\end{bmatrix}\\ =& B_0^{-1}\begin{bmatrix}1 & \rho\frac{\tau^2}{[\sigma^2\tau^2-\rho^2]}\\ 0 & 1+\tau^2\frac{\tau^2}{[\sigma^2\tau^2-\rho^2]}\end{bmatrix}\\ =& \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}\tau^2 & -\rho\\ -\rho & \sigma^2\end{bmatrix}\begin{bmatrix}1 & \rho\frac{\tau^2}{[\sigma^2\tau^2-\rho^2]}\\ 0 & 1+\tau^2\frac{\tau^2}{[\sigma^2\tau^2-\rho^2]}\end{bmatrix}\\ =& \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}\tau^2 & -\rho\\ -\rho & \sigma^2+\tau^2\end{bmatrix}.\end{aligned}$$

Putting these pieces together, we have:

$$B_1^{-1} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}\tau^2 & -\rho & 0 & -\tau^2\\ -\rho & \sigma^2 & 0 & \rho\\ 0 & 0 & \tau^2 & -\rho\\ -\tau^2 & \rho & -\rho & \sigma^2+\tau^2\end{bmatrix}.$$

And we have:

$$-B_1^{-1}C_1D^{-1} = -B_1^{-1}\begin{bmatrix}0 & 0\\ 0 & 0\\ 0 & 1\\ 0 & 0\end{bmatrix} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}0 & 0\\ 0 & 0\\ 0 & -\tau^2\\ 0 & \rho\end{bmatrix}.$$

So the first four rows of $E[\tilde{X}_2^{\intercal}\tilde{X}_2]^{-1}$ are:

$$\frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}\tau^2 & -\rho & 0 & -\tau^2 & 0 & 0\\ -\rho & \sigma^2 & 0 & \rho & 0 & 0\\ 0 & 0 & \tau^2 & -\rho & 0 & -\tau^2\\ -\tau^2 & \rho & -\rho & \sigma^2+\tau^2 & 0 & \rho\end{bmatrix}.$$

This is consistent with Lemma 4. This result with $K = 2$ is the basis step.

The induction hypothesis is that Lemma 4 is satisfied for some $K > 1$. We seek to show it is then also satisfied for $K + 1$.

The induction hypothesis and equation (E-3) imply that all rows of B_{K-1}^{-1} have the form defined in Lemma 4. Now use equation (E-4) to get B_K^{-1}. Consider $-B_{K-1}^{-1}C_{K-1}B_0^{-1}$. Because the only nonzero entries in the $2K \times 2$ matrix $C_{K-1}B_0^{-1}$ are in row $2K-1$ and this row is $[0 \quad 1]$, the first column of $-B_{K-1}^{-1}C_{K-1}B_0^{-1}$ contains only zeros and the second column of $-B_{K-1}^{-1}C_{K-1}B_0^{-1}$ selects the second-to-last element in each row of B_{K-1}^{-1}, which is equivalent to the fourth-to-last element in each of the first $2K$ rows of $E[\tilde{X}_K^\intercal \tilde{X}_K]^{-1}$. Using the induction hypothesis and temporarily ignoring the constant factored out of the matrix, this fourth-to-last element is τ^2 in row $2K-1$, is $-\rho$ in row $2K$, and is zero in all other rows. Therefore, under the induction hypothesis,

$$-B_{K-1}^{-1}C_{K-1}B_0^{-1} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}\mathbf{0}_{2K-2}^\intercal & \mathbf{0}_{2K-2}^\intercal \\ 0 & -\tau^2 \\ 0 & \rho\end{bmatrix}.$$

Now consider $-B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1}$. Because the only nonzero entries in the $2 \times 2K$ matrix $B_0^{-1}C_{K-1}^\intercal$ are in column $2K-1$ and this column is $[0 \quad 1]^\intercal$, the first row of $-B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1}$ contains only zeros and the second row of $-B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1}$ selects the second-to-last element in each column of B_{K-1}^{-1}, which is equivalent to element $2K-1$ in each of the first $2K$ columns of $E[\tilde{X}_K^\intercal \tilde{X}_K]^{-1}$. Using the induction hypothesis and temporarily ignoring the constant factored out of the matrix, row $2K-1$ of $E[\tilde{X}_K^\intercal \tilde{X}_K]^{-1}$ has $2(K+1)-(i+3) = 2(K+1)-(2K+2) = 0$ zeros at the end of it, so in using only the first $2K$ columns of $E[\tilde{X}_K^\intercal \tilde{X}_K]^{-1}$, we drop $[0 \quad -\tau^2]$ from row $2K-1$ and are left with $[\mathbf{0}_{2K-2} \quad \tau^2 \quad -\rho]$. Then,

$$-B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}\mathbf{0}_{2K-2} & 0 & 0 \\ \mathbf{0}_{2K-2} & -\tau^2 & \rho\end{bmatrix}.$$

Now consider $B_0^{-1} + B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1}C_{K-1}B_0^{-1}$. Using previous results, we know that

$$B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1}C_{K-1}B_0^{-1} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}\mathbf{0}_{2K-2} & 0 & 0 \\ \mathbf{0}_{2K-2} & 1 & 0\end{bmatrix}\begin{bmatrix}\mathbf{0}_{2K-2}^\intercal & \mathbf{0}_{2K-2}^\intercal \\ 0 & \tau^2 \\ 0 & -\rho\end{bmatrix}.$$

Therefore

$$B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1}C_{K-1}B_0^{-1} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}0 & 0 \\ 0 & \tau^2\end{bmatrix},$$

and so

$$B_0^{-1} + B_0^{-1}C_{K-1}^\intercal B_{K-1}^{-1}C_{K-1}B_0^{-1} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}\tau^2 & -\rho \\ -\rho & \sigma^2+\tau^2\end{bmatrix}.$$

Putting the pieces together, the induction hypothesis implies that

$$B_K^{-1} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix} JT\zeta^2[\sigma^2\tau^2-\rho^2]B_{K-1}^{-1} & \begin{bmatrix}\mathbf{0}_{2K-2}^\intercal & \mathbf{0}_{2K-2}^\intercal \\ 0 & -\tau^2 \\ 0 & \rho\end{bmatrix} \\ \begin{bmatrix}\mathbf{0}_{2K-2} & 0 & 0 \\ \mathbf{0}_{2K-2} & -\tau^2 & \rho\end{bmatrix} & \begin{bmatrix}\tau^2 & -\rho \\ -\rho & \sigma^2+\tau^2\end{bmatrix}\end{bmatrix}.$$

In order to complete the first $2(K+1)$ rows of $E[\tilde{X}_{K+1}^\intercal \tilde{X}_{K+1}]^{-1}$, we need $-B_K^{-1}C_K D^{-1}$. Because the only nonzero entries in the $(2K+2)\times 2$ matrix $C_K D^{-1}$ are in row $2K+1$ and this row is $[0 \quad 1]$, the first column of $-B_K^{-1}C_K D^{-1}$ contains only zeros and the second column of $-B_K^{-1}C_K D^{-1}$ selects the second-to-last element in each row of B_K^{-1}. Therefore:

$$-B_K^{-1}C_K D^{-1} = \frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix}\mathbf{0}_{2K}^\intercal & \mathbf{0}_{2K}^\intercal \\ 0 & -\tau^2 \\ 0 & \rho\end{bmatrix}.$$

The first $2(K+1)$ rows of $E[\tilde{X}_{K+1}^\intercal \tilde{X}_{K+1}]^{-1}$ are then

$$\frac{1}{JT\zeta^2[\sigma^2\tau^2-\rho^2]}\begin{bmatrix} JT\zeta^2[\sigma^2\tau^2-\rho^2]B_{K-1}^{-1} & \begin{bmatrix}\mathbf{0}_{2K-2}^\intercal & \mathbf{0}_{2K-2}^\intercal \\ 0 & -\tau^2 \\ 0 & \rho\end{bmatrix} & \begin{bmatrix}\mathbf{0}_{2K}^\intercal & \mathbf{0}_{2K}^\intercal\end{bmatrix} \\ \begin{bmatrix}\mathbf{0}_{2K-2} & 0 & 0 \\ \mathbf{0}_{2K-2} & -\tau^2 & \rho\end{bmatrix} & \begin{bmatrix}\tau^2 & -\rho \\ -\rho & \sigma^2+\tau^2\end{bmatrix} & \begin{bmatrix}0 & -\tau^2 \\ 0 & \rho\end{bmatrix}\end{bmatrix}.$$

We already established that the top left $2K \times 2K$ submatrix satisfies Lemma 4. Begin by checking the remaining portions of the top $2K$ odd rows and as well as row $2K+1$. For odd rows up to $2K-3$ inclusive, Lemma 4 posits that there are at least $2(K+1+1)-(2K-3+3) = 4$ zeros. We have that here. Lemma 4 also posits that the odd row $2K-1$ has two zeros and a $[0 \quad -\tau^2]$ leading up to them. We have that here. And Lemma 4 posits that row $2K+1$ has no trailing zeros, with a concluding $[\tau^2 \quad -\rho \quad 0 \quad -\tau^2]$. We also have that here. The odd rows are therefore consistent with Lemma 4.

Now consider the even rows, noting that $K+1 > 2$. For even rows up to $2K-2$ inclusive, Lemma 4 posits that there are at least $2(K+1+1)-(2K-2+2) = 4$ zeros. We have that here. Lemma 4 also posits that row $2K$ has two trailing zeros and a $[0 \quad \rho]$ leading up to them. We have that here. And Lemma 4 posits that row $2K+2$ has no trailing zeros and a concluding $[-\tau^2 \quad \rho \quad -\rho \quad \sigma^2+\tau^2 \quad 0 \quad \rho]$ with zeros leading up to that. We have that here.

We have therefore established that Lemma 4 holds for $E[\tilde{X}_{K+1}^\intercal \tilde{X}_{K+1}]^{-1}$ when it holds for $E[\tilde{X}_K^\intercal \tilde{X}_K]^{-1}$ (with $K > 1$), which was the induction step we sought.

□

E.2 Proof of Lemma 1

Begin by considering the uniqueness of the steady state. The right-hand side of equation (2) monotonically decreases in $\bar{A}$ if and only if $(1+\beta)\bar{\pi}_{12} < -\bar{\pi}_{11} - \beta\bar{\pi}_{22}$. Thus, the steady state is unique if $(1+\beta)\bar{\pi}_{12} < -\bar{\pi}_{11} - \beta\bar{\pi}_{22}$, which is satisfied for all $\bar{\pi}_{12} \leq 0$.

If $\bar{\pi}_{12} = 0$, then the Euler equation is a function of A_t but not of A_{t+1} or A_{t-1}. The agent selects $A_t = \bar{A}$ for any A_{t-1}.

Now consider the stability of the steady state for $\bar{\pi}_{12} \neq 0$. Define $A^*_{t+1}(A_t, A_{t-1})$ from the Euler equation. Linearizing around $\bar{A}$ gives a first-order difference equation:

$$A_{t+1} - \bar{A} \approx \frac{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}{\beta\bar{\pi}_{12}}(A_t - \bar{A}) - \frac{1}{\beta}(A_{t-1} - \bar{A}).$$

We have:

$$\begin{bmatrix} A_{t+1} - \bar{A} \\ A_t - \bar{A} \end{bmatrix} \approx \begin{bmatrix} \frac{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}{\beta\bar{\pi}_{12}} & -\frac{1}{\beta} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} A_t - \bar{A} \\ A_{t-1} - \bar{A} \end{bmatrix}.$$

The product of the linearized system's eigenvalues is $\frac{1}{\beta} > 1$, and the sum of the linearized system's eigenvalues is $\frac{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}{\beta\bar{\pi}_{12}}$, which is positive if and only if $\bar{\pi}_{12} > 0$.

First, assume that $\bar{\pi}_{12} > 0$. Both eigenvalues are positive and at least one is greater than 1. The characteristic equation is

$$z^2 - \frac{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}{\beta\bar{\pi}_{12}} z + \frac{1}{\beta},$$

where z defines the eigenvalues. The smaller eigenvalue is less than 1 if and only if the characteristic equation is negative at $z = 1$, and therefore if and only if

$$-\bar{\pi}_{11} - \beta\bar{\pi}_{22} > (1 + \beta)\bar{\pi}_{12}.$$

In this case, the linearized system is saddle-path stable.

Now assume that $\bar{\pi}_{12} < 0$. Both eigenvalues are negative and at least one is less than -1. The characteristic equation is as before. The larger eigenvalue is greater than -1 if and only if the characteristic equation is negative at $z = -1$, and therefore if and only if

$$\begin{aligned} \frac{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}{\bar{\pi}_{12}} + 1 + \beta &< 0 \\ \Leftrightarrow -\bar{\pi}_{11} - \beta\bar{\pi}_{22} &> -(1 + \beta)\bar{\pi}_{12}. \end{aligned}$$

In this case, the linearized system is saddle-path stable.

The proposition follows from a standard application of the Hartman-Grobman theorem and from noticing that the conditions for saddle-path stability imply the condition for uniqueness.

E.3 Proof of Lemma 2

I first describe A_t, under the assumption that $(A_{t-1} - \bar{A})^2$ is small. Write A_{t+1} as $A(A_t, w_{t+1}, f_{t+1}, w_t; \zeta)$. Expanding the stochastic Euler equation around $\zeta = 0$ and noting that all terms of order ζ^2

or larger depend on at least the third derivative of π, either Assumption 1 or 2 ensures that we can drop all terms of order ζ^2 or larger. We therefore have:

$$\begin{aligned}
0 =&\pi_1(A_t, A_{t-1}, w_t, w_{t-1}) + \beta E_t \left[\pi_2(\tilde{A}_{t+1}, A_t, f_t, w_t) + \pi_{23}(\tilde{A}_{t+1}, A_t, f_t, w_t)\epsilon_{t+1}\zeta\right] \\
&+ \beta E_t \left[\pi_{12}(\tilde{A}_{t+1}, A_t, f_t, w_t) \left(\left.\frac{\partial A_{t+1}}{\partial \zeta}\right|_{\zeta=0} + \left.\frac{\partial A_{t+1}}{\partial w_{t+1}}\right|_{\zeta=0} \epsilon_{t+1} + \sum_{i=1}^{N} \left.\frac{\partial A_{t+1}}{\partial f_{i(t+1)}}\right|_{\zeta=0} \epsilon_{i(t+1)} \right) \zeta \right] \\
=&\pi_1(A_t, A_{t-1}, w_t, w_{t-1}) + \beta\pi_2(\tilde{A}_{t+1}, A_t, f_t, w_t) + \beta\pi_{12}(\tilde{A}_{t+1}, A_t, f_t, w_t) \left.\frac{\partial A_{t+1}}{\partial \zeta}\right|_{\zeta=0} \zeta, \quad \text{(E-5)}
\end{aligned}$$

where $\tilde{A}_{t+1} \triangleq A(A_t, f_t, C, w_t; 0)$.

I next establish two lemmas. The first one shows that uncertainty does not have a first-order effect on policy:

Lemma 5. $\left.\frac{\partial A_{t+1}}{\partial \zeta}\right|_{(\bar{A},C,C,C;0)} = 0.$

Proof. Equation (E-5) defines A_t as a function of A_{t-1}, w_t, f_t, and ζ. Note that

$$\begin{aligned}
\left.\frac{\partial A_t}{\partial \zeta}\right|_{(\bar{A},C,C,C;0)} =& \frac{\beta\bar{\pi}_{12} \left(\left.\frac{\partial A_{t+1}}{\partial \zeta}\right|_{(\bar{A},C,C,C;0)} + \left.\frac{\partial^2 A_{t+1}}{\partial \zeta^2}\right|_{(\bar{A},C,C,C;0)} \zeta \right)}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{122} \left.\frac{\partial A_{t+1}}{\partial \zeta}\right|_{\zeta=0} \zeta - \beta\bar{\pi}_{12} \left.\frac{\partial^2 A_{t+1}}{\partial \zeta \partial A_t}\right|_{\zeta=0} \zeta} \\
=& \frac{\beta\bar{\pi}_{12}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \left.\frac{\partial A_{t+1}}{\partial \zeta}\right|_{(\bar{A},C,C,C;0)},
\end{aligned}$$

where the second equality applies $\zeta = 0$. Forward-substituting, we have:

$$\left.\frac{\partial A_t}{\partial \zeta}\right|_{(\bar{A},C,C,C;0)} = \left(\frac{\beta\bar{\pi}_{12}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \right)^j \left.\frac{\partial A_{t+j}}{\partial \zeta}\right|_{(\bar{A},C,C,C;0)}$$

for $j \in \mathcal{Z}^+$. The term in parentheses is < 1 by the condition imposed following Lemma 1. Because A_{t+j} evaluated around $A_{t+j-1} = \bar{A}$, $w_t = C$, $f_t = C$, and $\zeta = 0$ must be $\bar{A}$, we know that A_{t+j} is not infinite. The derivative on the right-hand side must also be finite, in which case the right-hand side goes to 0 as j becomes large. Therefore:

$$\left.\frac{\partial A_t}{\partial \zeta}\right|_{(\bar{A},C,C,C;0)} = 0.$$

Because the choice of t was arbitrary, we have established the lemma. □

The second lemma solves for $\tilde{A}_{t+1}$:

Lemma 6. *If either Assumption 1 or 2 holds and $(A_t - \bar{A})^2$ is small, then there exists λ such that $|\lambda| < 1$, $sign(\lambda) = sign(\bar{\pi}_{12})$, and*

$$\begin{aligned}\tilde{A}_{t+1} =& \bar{A} + \frac{\bar{\pi}_{12}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{12}\frac{\bar{\pi}_{12}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}-\beta\bar{\pi}_{12}\lambda}}(A_t - \bar{A}) + \frac{\bar{\pi}_{14}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{12}\frac{\bar{\pi}_{12}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}-\beta\bar{\pi}_{12}\lambda}}(w_t - C)\\ &+ \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{12}\frac{\bar{\pi}_{14}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}-\beta\bar{\pi}_{12}\lambda}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{12}\frac{\bar{\pi}_{12}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}-\beta\bar{\pi}_{12}\lambda}}(f_t - C).\end{aligned}$$

Proof. For $\zeta = 0$, the weather in period $t+1$ matches the forecast in f_t, and the weather is always C after period $t+1$. By period $t+3$, even lagged weather is just C and we are back to the deterministic system. Begin by solving for policy in period $t+4$. The characteristic equation given in the proof of Lemma 1 implies the following two eigenvalues:

$$\lambda, \mu = \frac{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}{2\beta\bar{\pi}_{12}} \pm \sqrt{\left(\frac{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}{2\beta\bar{\pi}_{12}}\right)^2 - \frac{1}{\beta}}.$$

The proof of Lemma 1 showed that the two eigenvalues have the same sign and that this sign matches that of $\bar{\pi}_{12}$. Define λ as the eigenvalue that is smallest in absolute value. We seek the eigenvectors corresponding to λ, which is the stable manifold. These eigenvectors have $\tilde{A}_{t+3} - \bar{A} = \lambda(\tilde{A}_{t+2} - \bar{A})$ and thus are proportional to

$$\begin{bmatrix}\lambda(\tilde{A}_{t+2} - \bar{A})\\ \tilde{A}_{t+2} - \bar{A}\end{bmatrix}.$$

Therefore, along the saddle path,

$$\begin{bmatrix}\tilde{A}_{t+4} - \bar{A}\\ \tilde{A}_{t+3} - \bar{A}\end{bmatrix} = c\lambda \begin{bmatrix}\lambda(\tilde{A}_{t+2} - \bar{A})\\ \tilde{A}_{t+2} - \bar{A}\end{bmatrix}$$

for some $c \neq 0$. It must be true that $\tilde{A}_{t+3} - \bar{A} = c\lambda(\tilde{A}_{t+2} - \bar{A})$ and that $\tilde{A}_{t+3} - \bar{A} = \lambda(\tilde{A}_{t+2} - \bar{A})$. Therefore $c = 1$.

Now consider policy at time $t+2$. The relevant Euler equation is:

$$0 = \pi_1(\tilde{A}_{t+2}, \tilde{A}_{t+1}, C, f_t) + \beta\pi_2(\bar{A} + \lambda(\tilde{A}_{t+2} - \bar{A}), \tilde{A}_{t+2}, C, C),$$

where we recognize that $w_{t+1} = f_t$. A first-order approximation to $\tilde{A}_{t+2}$ around $\bar{A}$ is exact when either Assumption 1 or 2 holds and $(\tilde{A}_{t+1} - \bar{A})^2$ is small. We thereby obtain $\tilde{A}_{t+2}$ as a function of $\tilde{A}_{t+1}$ and f_t:

$$\tilde{A}_{t+2} = \bar{A} + \frac{\bar{\pi}_{12}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{12}\lambda}(\tilde{A}_{t+1} - \bar{A}) + \frac{\bar{\pi}_{14}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{12}\lambda}(f_t - C).$$

Now consider policy at time $t+1$. The relevant Euler equation is:

$$0 = \pi_1(\tilde{A}_{t+1}, A_t, f_t, w_t) + \beta\pi_2(\tilde{A}_{t+2}(\tilde{A}_{t+1}, f_t), \tilde{A}_{t+1}, C, f_t),$$

where we recognize that $w_{t+1} = f_t$. A first-order approximation to $\tilde{A}_{t+1}$ around $\bar{A}$ is exact when either Assumption 1 or 2 holds and $(A_t - \bar{A})^2$ is small. We then have the expression in the lemma.

□

Applying Lemmas 5 and 6 to equation (E-5), we have:

$$0 = \pi_1(A_t, A_{t-1}, w_t, w_{t-1}) + \beta\pi_2(\tilde{A}_{t+1}(A_t, f_t, w_t), A_t, f_t, w_t). \tag{E-6}$$

We now have A_t implicitly defined as $A(A_{t-1}, w_t, f_t, w_{t-1}; 0)$. If $(A_{t-1} - \bar{A})^2$ is small and either Assumption 1 or 2 holds, then we have:

$$\begin{aligned}
A_t =& \bar{A} + \left.\frac{\partial A_t}{\partial A_{t-1}}\right|_{(\bar{A},C,C,C;0)} (A_{t-1} - \bar{A}) + \left.\frac{\partial A_t}{\partial w_t}\right|_{(\bar{A},C,C,C;0)} (w_t - C) + \left.\frac{\partial A_t}{\partial f_t}\right|_{(\bar{A},C,C,C;0)} (f_t - C) \\
&+ \left.\frac{\partial A_t}{\partial w_{t-1}}\right|_{(\bar{A},C,C,C;0)} (w_{t-1} - C) + \left.\frac{\partial A_t}{\partial \zeta}\right|_{(\bar{A},C,C,C;0)} \zeta \\
=& \bar{A} + \frac{\bar{\pi}_{12}}{\chi_2}(A_{t-1} - \bar{A}) + \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}(w_t - C) \\
&+ \frac{\beta\bar{\pi}_{23} + \beta\left(\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}(f_t - C) \\
&+ \frac{\bar{\pi}_{14}}{\chi_2}(w_{t-1} - C),
\end{aligned}$$

where we use Lemma 5 in the first equality and where χ_i is defined recursively:

$$\begin{aligned}
\chi_0 &\triangleq -\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{12}\lambda, \\
\chi_i &\triangleq -\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{12}\frac{\bar{\pi}_{12}}{\chi_{i-1}} \quad \text{for } i \text{ a strictly positive integer.}
\end{aligned}$$

The condition imposed following Lemma 1 and the fact that $|\lambda| < 1$ together ensure that each $\chi_i > |\bar{\pi}_{12}|$.

Now use this result to analyze $E_0[A_t]$. If either Assumption 1 or 2 holds and $E_0[(A_1 - \bar{A})^2]$ is small, then

$$E_0[A_2] = \bar{A} + \frac{\bar{\pi}_{12}}{\chi_2}(E_0[A_1] - \bar{A}).$$

$E_0[(A_2 - \bar{A})^2]$ must also be small because $|\bar{\pi}_{12}|/\chi_2 < 1$. Iterating forward, we find, for $t > 1$,

$$E_0[A_t] =\bar{A} + \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{t-1} (E_0[A_1] - \bar{A}).$$

As $t \to \infty$, we have:

$$E_0[A_t] \to \bar{A}.$$

We have proved the desired result.

E.4 Proof of Proposition 1

Using equation (8) and standard regression properties, we have:

$$\begin{aligned}
\hat{\Gamma}_1 =& \omega \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}},\\
\hat{\Gamma}_2 =& \omega \frac{\bar{\pi}_{14}}{-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}},\\
\hat{\Gamma}_3 =& \omega \frac{\beta\bar{\pi}_{23} + \beta\left(\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}},\\
\hat{\Gamma}_4 =& \omega \frac{\bar{\pi}_{12}}{-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}},
\end{aligned}$$

where

$$\omega \triangleq \frac{-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}}{\chi_2} > 0$$

and χ_2 was defined in the proof of Lemma 2. Because $1 + \beta > \beta\bar{\pi}_{12}/\chi_1$, we have $\omega > 1$ if $\bar{\pi}_{12} < 0$, $\omega = 1$ if $\bar{\pi}_{12} = 0$, and $\omega < 1$ if $\bar{\pi}_{12} > 0$. The proposition follows from defining

$$\Omega \triangleq \frac{\bar{\pi}_{13} + \bar{\pi}_{14} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}}{\left(-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}\right)\chi_1} \geq 0.$$

E.5 Proof of Proposition 2

The vector of estimated coefficients is

$$\hat{\Gamma} = E[X_{I+1}^{\intercal} X_{I+1}]^{-1} E[X_{I+1}^{\intercal} \boldsymbol{A}],$$

where $\boldsymbol{A}$ is a $JT \times 1$ vector with rows A_{jt} and X_{I+1} is a $JT \times J + 2(I+2)$ matrix with the final $2(I+2)$ columns of each row being

$$\begin{bmatrix} w_{jt} & f_{jt} & w_{j(t-1)} & f_{j(t-1)} & \dots & w_{j(t-(I+1))} & f_{j(t-(I+1))} \end{bmatrix}.$$

By the Frisch-Waugh Theorem,

$$\hat{\Gamma} = E[\tilde{X}_{I+1}^{\intercal}\tilde{X}_{I+1}]^{-1}E[\tilde{X}_{I+1}^{\intercal}\tilde{\boldsymbol{A}}],$$

where $\tilde{X}_{I+1}$ is a $JT \times 2(I+2)$ matrix with rows

$$\begin{bmatrix} w_{jt}-C & f_{jt}-C & w_{j(t-1)}-C & f_{j(t-1)}-C & \dots & w_{j(t-(I+1))}-C & f_{j(t-(I+1))}-C \end{bmatrix}$$

and $\tilde{\boldsymbol{A}}$ is demeaned $\boldsymbol{A}$. Lemma 4 establishes the first $2(I+1)$ rows of $E[\tilde{X}_{I+1}^{\intercal}\tilde{X}_{I+1}]^{-1}$, which are the ones that are relevant for $\hat{\Gamma}_{w_t}$ through $\hat{\Gamma}_{w_{t-I}}$ and for $\hat{\Gamma}_{f_t}$ through $\hat{\Gamma}_{f_{t-I}}$. Observe that:

$$E[\tilde{X}_{I+1}^{\intercal}\tilde{\boldsymbol{A}}] = JT \begin{bmatrix} Cov[w_{jt}-C, A_{jt}] \\ Cov[f_{jt}-C, A_{jt}] \\ \vdots \\ Cov[w_{j(t-(I+1))}-C, A_{jt}] \\ Cov[f_{j(t-(I+1))}-C, A_{jt}] \end{bmatrix}.$$

From here, drop the j subscript to save on unnecessary notation.

Equation (8) holds under the given assumptions. Using that equation, we find:

$$\begin{aligned}\frac{1}{\zeta^2}Cov[w_t-C, A_t] =& \frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\sigma^2+\tau^2+\frac{\bar{\pi}_{12}}{\chi_2}\rho\right)+\frac{\bar{\pi}_{14}}{\chi_2}\rho \\ &+\frac{\beta\bar{\pi}_{23}+\beta\left(\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\rho+\frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right),\end{aligned}$$

$$\frac{1}{\zeta^2}Cov[f_t-C, A_t] = \frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\rho+\frac{\beta\bar{\pi}_{23}+\beta\left(\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\tau^2,$$

$$\begin{aligned}\frac{1}{\zeta^2}Cov[w_{t-1}-C, A_t] =& \frac{\bar{\pi}_{14}}{\chi_2}(\sigma^2+\tau^2)+\frac{\bar{\pi}_{12}}{\chi_2}\frac{\bar{\pi}_{14}}{\chi_2}\rho \\ &+\frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left[\rho+\frac{\bar{\pi}_{12}}{\chi_2}\left(\sigma^2+\tau^2+\frac{\bar{\pi}_{12}}{\chi_2}\rho\right)\right] \\ &+\frac{\bar{\pi}_{12}}{\chi_2}\left(\rho+\frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right)\frac{\beta\bar{\pi}_{23}+\beta\left(\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2},\end{aligned}$$

$$
\begin{aligned}
\frac{1}{\zeta^2} Cov[f_{t-1} - C, A_t] =& \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\tau^2 + \frac{\bar{\pi}_{12}}{\chi_2}\rho\right) \\
&+ \frac{\beta\bar{\pi}_{23} + \beta\left(\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\frac{\bar{\pi}_{12}}{\chi_2}\tau^2 + \frac{\bar{\pi}_{14}}{\chi_2}\rho.
\end{aligned}
$$

And for $i \geq 2$,

$$
\begin{aligned}
\frac{1}{\zeta^2} Cov[w_{t-i} - C, A_t] =& \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-2}\frac{\bar{\pi}_{14}}{\chi_2}\left\{\rho + \frac{\bar{\pi}_{12}}{\chi_2}\left[(\sigma^2 + \tau^2) + \rho\frac{\bar{\pi}_{12}}{\chi_2}\right]\right\} \\
&+ \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-1}\frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left[\rho + \frac{\bar{\pi}_{12}}{\chi_2}(\sigma^2 + \tau^2) + \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^2\rho\right] \\
&+ \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i}\frac{\beta\bar{\pi}_{23} + \beta\left(\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left[\rho + \frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right],
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{\zeta^2} Cov[f_{t-i} - C, \pi_t] =& \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-2}\frac{\bar{\pi}_{14}}{\chi_2}\left\{\tau^2 + \frac{\bar{\pi}_{12}}{\chi_2}\rho\right\} \\
&+ \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-1}\frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left[\tau^2 + \frac{\bar{\pi}_{12}}{\chi_2}\rho\right] \\
&+ \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i}\frac{\beta\bar{\pi}_{23} + \beta\left(\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\tau^2.
\end{aligned}
$$

Direct calculations then yield the following coefficients:

$$
\begin{aligned}
\hat{\Gamma}_{w_t} =& \hat{\Gamma}_1, \\
\hat{\Gamma}_{f_t} =& \hat{\Gamma}_3, \\
\hat{\Gamma}_{w_{t-1}} =& \frac{\bar{\pi}_{12}}{\chi_2}\hat{\Gamma}_1 + \hat{\Gamma}_2, \\
\hat{\Gamma}_{f_{t-1}} =& \frac{\bar{\pi}_{12}}{\chi_2}\hat{\Gamma}_3, \\
\hat{\Gamma}_{w_{t-i}} =& \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i}\hat{\Gamma}_1 + \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-1}\hat{\Gamma}_2 \quad \text{for } i \geq 2, \\
\hat{\Gamma}_{f_{t-i}} =& \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i}\hat{\Gamma}_3 \quad \text{for } i \geq 2.
\end{aligned}
$$

Therefore

$$\lim_{I\to\infty}\sum_{i=0}^{I}[\hat{\Gamma}_{w_{t-i}}+\hat{\Gamma}_{f_{t-i}}]=\frac{\hat{\Gamma}_1+\hat{\Gamma}_2+\hat{\Gamma}_3}{1-\frac{\bar{\pi}_{12}}{\chi_2}}=\tilde{\omega}\left(\frac{\mathrm{d}\bar{A}}{\mathrm{d}C}+\beta\bar{\pi}_{12}\Omega\right),$$

with Ω defined as in the proof of Proposition 1 and

$$\tilde{\omega}\triangleq\frac{\omega}{1-\frac{\bar{\pi}_{12}}{\chi_2}}=\frac{-\bar{\pi}_{11}-(1+\beta)\bar{\pi}_{12}-\beta\bar{\pi}_{22}}{-\bar{\pi}_{11}-\left(1+\beta\frac{\bar{\pi}_{12}}{\chi_1}\right)\bar{\pi}_{12}-\beta\bar{\pi}_{22}}.$$

The proposition follows straightforwardly.

E.6 Proof of Proposition 3

The vector of estimated coefficients is

$$\hat{\theta}=E[X_{I+1}^{\intercal}X_{I+1}]^{-1}E[X_{I+1}^{\intercal}\boldsymbol{\pi}],$$

where $\boldsymbol{\pi}$ is a $JT\times 1$ vector with rows π_{jt} and X_{I+1} is a $JT\times J+2(I+2)$ matrix with the final $2(I+2)$ columns of each row being

$$\begin{bmatrix} w_{jt} & f_{jt} & w_{j(t-1)} & f_{j(t-1)} & \dots & w_{j(t-(I+1))} & f_{j(t-(I+1))}\end{bmatrix}.$$

By the Frisch-Waugh Theorem,

$$\hat{\theta}=E[\tilde{X}_{I+1}^{\intercal}\tilde{X}_{I+1}]^{-1}E[\tilde{X}_{I+1}^{\intercal}\tilde{\boldsymbol{\pi}}],$$

where $\tilde{X}_{I+1}$ is a $JT\times 2(I+2)$ matrix with rows

$$\begin{bmatrix} w_{jt}-C & f_{jt}-C & w_{j(t-1)}-C & f_{j(t-1)}-C & \dots & w_{j(t-(I+1))}-C & f_{j(t-(I+1))}-C\end{bmatrix}$$

and $\tilde{\boldsymbol{\pi}}$ is demeaned $\boldsymbol{\pi}$. Lemma 4 establishes the first $2(I+1)$ rows of $E[\tilde{X}_{I+1}^{\intercal}\tilde{X}_{I+1}]^{-1}$, which are the ones that are relevant for $\hat{\theta}_{w_t}$ through $\hat{\theta}_{w_{t-I}}$ and for $\hat{\theta}_{f_t}$ through $\hat{\theta}_{f_{t-I}}$. Observe that:

$$E[\tilde{X}_{I+1}^{\intercal}\tilde{\boldsymbol{\pi}}]=JT\begin{bmatrix} Cov[w_{jt}-C,\pi_{jt}]\\ Cov[f_{jt}-C,\pi_{jt}]\\ \vdots\\ Cov[w_{j(t-(I+1))}-C,\pi_{jt}]\\ Cov[f_{j(t-(I+1))}-C,\pi_{jt}]\end{bmatrix}.$$

From here, drop the j subscript to save on unnecessary notation.

Consider $Cov[w_t - C, \pi_t]$. Expanding π around $A_t = \bar{A}$, $A_{t-1} = \bar{A}$, $w_t = C$, and $w_{t-1} = C$, applying either Assumption 1 or 2, and assuming that $(A_t - \bar{A})^2$ and $(A_{t-1} - \bar{A})^2$ are small, we have:

$$\begin{aligned}
\pi(A_t, A_{t-1}, w_t, w_{t-1}) =& \bar{\pi} + \bar{\pi}_1(A_t - \bar{A}) + \bar{\pi}_2(A_{t-1} - \bar{A}) + \bar{\pi}_3(w_t - C) + \bar{\pi}_4(w_{t-1} - C) \\
&+ \frac{1}{2}\bar{\pi}_{33}(w_t - C)^2 + \bar{\pi}_{13}(A_t - \bar{A})(w_t - C) + \bar{\pi}_{23}(A_{t-1} - \bar{A})(w_t - C) \\
&+ \frac{1}{2}\bar{\pi}_{44}(w_{t-1} - C)^2 + \bar{\pi}_{14}(A_t - \bar{A})(w_{t-1} - C) + \bar{\pi}_{24}(A_{t-1} - \bar{A})(w_{t-1} - C) \\
&+ \bar{\pi}_{34}(w_t - C)(w_{t-1} - C).
\end{aligned}$$

As a result,

$$\begin{aligned}
Cov[w_t - C, \pi_t] =& \bar{\pi}_1 Cov[A_t, w_t] + \bar{\pi}_2 Cov[A_{t-1}, w_t] + \bar{\pi}_3 Var[w_t] + \bar{\pi}_4 Cov[w_t, w_{t-1}] \\
&+ \frac{1}{2}\bar{\pi}_{33} Cov[w_t - C, (w_t - C)^2] + \frac{1}{2}\bar{\pi}_{44} Cov[w_t - C, (w_{t-1} - C)^2] \\
&- C\bar{\pi}_{13} Cov[A_t, w_t] - \bar{A}\bar{\pi}_{13} Var[w_t] + \bar{\pi}_{13} Cov[w_t, A_t w_t] \\
&- C\bar{\pi}_{23} Cov[A_{t-1}, w_t] - \bar{A}\bar{\pi}_{23} Var[w_t] + \bar{\pi}_{23} Cov[w_t, A_{t-1} w_t] \\
&- C\bar{\pi}_{14} Cov[w_t, A_t] - \bar{A}\bar{\pi}_{14} Cov[w_t, w_{t-1}] + \bar{\pi}_{14} Cov[w_t, A_t w_{t-1}] \\
&- C\bar{\pi}_{24} Cov[w_t, A_{t-1}] - \bar{A}\bar{\pi}_{24} Cov[w_t, w_{t-1}] + \bar{\pi}_{24} Cov[w_t, A_{t-1} w_{t-1}] \\
&- C\bar{\pi}_{34} Var[w_t] - C\bar{\pi}_{34} Cov[w_t, w_{t-1}] + \bar{\pi}_{34} Cov[w_t, w_t w_{t-1}].
\end{aligned}$$

If the ϵ and ν are normally distributed, then $Cov[w_t - C, (w_t - C)^2] = 0$, or if Assumption 1 holds, then $Cov[w_t - C, (w_t - C)^2] \approx 0$. Using results from Bohrnstedt and Goldberger (1969), we have:

$$Cov[w_t - C, (w_{t-1} - C)^2] = E[(w_t - C)(w_{t-1} - C)^2],$$

which is zero if either the ϵ and ν are normally distributed or Assumption 1 holds. Again using results from Bohrnstedt and Goldberger (1969), we also have:

$$Cov[w_t, A_t w_t] = E[A_t]\, Var[w_t] + C\, Cov[A_t, w_t] + E[(w_t - C)^2(A_t - E[A_t])].$$

If either the ϵ and ν are normally distributed or Assumption 1 holds, then this becomes :

$$Cov[w_t, A_t w_t] = E[A_t]\, Var[w_t] + C\, Cov[A_t, w_t].$$

Analogous derivations yield:

$$\begin{aligned}
Cov[w_t, A_{t-1} w_t] =& E[A_{t-1}]\, Var[w_t] + C\, Cov[w_t, A_{t-1}], \\
Cov[w_t, A_t w_{t-1}] =& E[A_t]\, Cov[w_t, w_{t-1}] + C\, Cov[w_t, A_t], \\
Cov[w_t, A_{t-1} w_{t-1}] =& E[A_{t-1}]\, Cov[w_t, w_{t-1}] + C\, Cov[w_t, A_{t-1}]
\end{aligned}$$

if either the ϵ and ν are normally distributed or Assumption 1 holds. Substituting these results in, we find:

$$\begin{aligned}
Cov[w_t - C, \pi_t] =& \bar{\pi}_1 Cov[A_t, w_t] + \bar{\pi}_2 Cov[A_{t-1}, w_t] + \bar{\pi}_3 Var[w_t] + \bar{\pi}_4 Cov[w_t, w_{t-1}] \\
&+ \left(E[A_t] - \bar{A}\right)\left(\bar{\pi}_{13} Var[w_t] + \bar{\pi}_{14} Cov[w_t, w_{t-1}]\right) \\
&+ \left(E[A_{t-1}] - \bar{A}\right)\left(\bar{\pi}_{23} Var[w_t] + \bar{\pi}_{24} Cov[w_t, w_{t-1}]\right).
\end{aligned}$$

The assumption that actions are on average around $\bar{A}$ implies $E[A_t] = \bar{A}$ and $E[A_{t-1}] = \bar{A}$. Using that and equation (8), we obtain:

$$\begin{aligned}
\frac{1}{\zeta^2} Cov[w_t - C, \pi_t] =& (\sigma^2 + \tau^2)\bar{\pi}_3 + \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}}{\chi_2}\left((\sigma^2 + \tau^2)\bar{\pi}_1 + \bar{\pi}_2\rho + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\rho\right) \\
&+ \bar{\pi}_4\rho + \frac{\bar{\pi}_{14}}{\chi_2}\bar{\pi}_1\rho + \frac{\beta\bar{\pi}_{23} + \beta(\bar{\pi}_{13} + \beta\bar{\pi}_{24})\frac{\bar{\pi}_{12}}{\chi_0}}{\chi_2}\left(\bar{\pi}_1\rho + \bar{\pi}_2\tau^2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right).
\end{aligned}$$

Analogous derivations yield:

$$\begin{aligned}
\frac{1}{\zeta^2} Cov[f_t - C, \pi_t] =& \frac{1}{\zeta^2}\left(\bar{\pi}_1 Cov[A_t, f_t] + \bar{\pi}_3 Cov[w_t, f_t]\right) \\
=& \bar{\pi}_3\rho + \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\bar{\pi}_1\rho \\
&+ \frac{\beta\bar{\pi}_{23} + \beta\left(\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\bar{\pi}_1\tau^2,
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\zeta^2} Cov[w_{t-1} - C, \pi_t] =& \frac{1}{\zeta^2}\left(\bar{\pi}_1 Cov[A_t, w_{t-1}] + \bar{\pi}_2 Cov[A_{t-1}, w_{t-1}] + \bar{\pi}_3 Cov[w_t, w_{t-1}] + \bar{\pi}_4 Var[w_{t-1}]\right) \\
=& \bar{\pi}_3\rho + \bar{\pi}_4(\sigma^2 + \tau^2) + \bar{\pi}_1\frac{\bar{\pi}_{14}}{\chi_2}(\sigma^2 + \tau^2) + \left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\frac{\bar{\pi}_{14}}{\chi_2}\rho \\
&+ \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left[\bar{\pi}_1\rho + \left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left(\sigma^2 + \tau^2 + \frac{\bar{\pi}_{12}}{\chi_2}\rho\right)\right] \\
&+ \left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left(\rho + \frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right)\frac{\beta\bar{\pi}_{23} + \beta\left(\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2},
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\zeta^2}Cov[f_{t-1}-C,\pi_t] =& \frac{1}{\zeta^2}\left(\bar{\pi}_1 Cov[A_t,f_{t-1}] + \bar{\pi}_2 Cov[A_{t-1},f_{t-1}] + \bar{\pi}_3 Cov[w_t,f_{t-1}] + \bar{\pi}_4 Cov[w_{t-1},f_{t-1}]\right)\\
=&\bar{\pi}_3\tau^2 + \bar{\pi}_4\rho + \frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\bar{\pi}_1\tau^2 + \bar{\pi}_2\rho + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\rho\right)\\
&+\frac{\beta\bar{\pi}_{23}+\beta\left(\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\bar{\pi}_2\tau^2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right)\\
&+\bar{\pi}_1\frac{\bar{\pi}_{14}}{\chi_2}\rho,
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\zeta^2}Cov[w_{t-2}-C,\pi_t] =& \frac{1}{\zeta^2}\left(\bar{\pi}_1 Cov[A_t,w_{t-2}] + \bar{\pi}_2 Cov[A_{t-1},w_{t-2}] + \bar{\pi}_3 Cov[w_t,w_{t-2}] + \bar{\pi}_4 Cov[w_{t-1},w_{t-2}]\right)\\
=&\bar{\pi}_4\rho + \bar{\pi}_1\frac{\bar{\pi}_{14}}{\chi_2}\rho + \frac{\bar{\pi}_{14}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left[(\sigma^2+\tau^2)+\rho\frac{\bar{\pi}_{12}}{\chi_2}\right]\\
&+\frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left[\rho+\frac{\bar{\pi}_{12}}{\chi_2}(\sigma^2+\tau^2)+\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^2\rho\right]\\
&+\frac{\beta\bar{\pi}_{23}+\beta\left(\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\frac{\bar{\pi}_{12}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left[\rho+\frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right],
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\zeta^2}Cov[f_{t-2}-C,\pi_t] =& \frac{1}{\zeta^2}\left(\bar{\pi}_1 Cov[A_t,f_{t-2}] + \bar{\pi}_2 Cov[A_{t-1},f_{t-2}] + \bar{\pi}_3 Cov[w_t,f_{t-2}] + \bar{\pi}_4 Cov[w_{t-1},f_{t-2}]\right)\\
=&\bar{\pi}_4\tau^2 + \bar{\pi}_1\frac{\bar{\pi}_{14}}{\chi_2}\tau^2 + \frac{\bar{\pi}_{14}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\rho\\
&+\frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left[\tau^2+\frac{\bar{\pi}_{12}}{\chi_2}\rho\right]\\
&+\frac{\beta\bar{\pi}_{23}+\beta\left(\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\frac{\bar{\pi}_{12}}{\chi_2}\tau^2.
\end{aligned}$$

For $i \geq 3$, we have:

$$\begin{aligned}
\frac{1}{\zeta^2}Cov[w_{t-i}-C,\pi_t] =&\frac{1}{\zeta^2}\left(\bar{\pi}_1 Cov[A_t,w_{t-i}]+\bar{\pi}_2 Cov[A_{t-1},w_{t-i}]+\bar{\pi}_3 Cov[w_t,w_{t-i}]+\bar{\pi}_4 Cov[w_{t-1},w_{t-i}]\right)\\
=&\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-3}\frac{\bar{\pi}_{14}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left\{\rho+\frac{\bar{\pi}_{12}}{\chi_2}\left[(\sigma^2+\tau^2)+\rho\frac{\bar{\pi}_{12}}{\chi_2}\right]\right\}\\
&+\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-2}\frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\\
&\left[\rho+\frac{\bar{\pi}_{12}}{\chi_2}(\sigma^2+\tau^2)+\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^2\rho\right]\\
&+\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-1}\frac{\beta\bar{\pi}_{23}+\beta\left(\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left[\rho+\frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right],
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\zeta^2}Cov[f_{t-i}-C,\pi_t] =&\frac{1}{\zeta^2}\left(\bar{\pi}_1 Cov[A_t,f_{t-i}]+\bar{\pi}_2 Cov[A_{t-1},f_{t-i}]+\bar{\pi}_3 Cov[w_t,f_{t-i}]+\bar{\pi}_4 Cov[w_{t-1},f_{t-i}]\right)\\
=&\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-3}\frac{\bar{\pi}_{14}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left\{\tau^2+\frac{\bar{\pi}_{12}}{\chi_2}\rho\right\}\\
&+\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-2}\frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left[\tau^2+\frac{\bar{\pi}_{12}}{\chi_2}\rho\right]\\
&+\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-1}\frac{\beta\bar{\pi}_{23}+\beta\left(\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\beta\bar{\pi}_{14}\frac{\bar{\pi}_{12}}{\chi_0}\right)\frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\left(\bar{\pi}_2+\bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\tau^2.
\end{aligned}$$

Now consider the regression coefficients, from $E[\tilde{X}_{I+1}^{\intercal}\tilde{X}_{I+1}]^{-1}E[\tilde{X}_{I+1}^{\intercal}\boldsymbol{\pi}]$ and using Lemma 4. For $I \geq 0$, the coefficient on w_{jt} is

$$\begin{aligned}
\hat{\theta}_{w_t} =&\frac{1}{\zeta^2}\left(\frac{\tau^2}{\sigma^2\tau^2-\rho^2}Cov[w_t-C,\pi_t]+\frac{-\rho}{\sigma^2\tau^2-\rho^2}Cov[f_t-C,\pi_t]-\frac{\tau^2}{\sigma^2\tau^2-\rho^2}Cov[f_{t-1}-C,\pi_t]\right)\\
=&\bar{\pi}_3+\hat{\Gamma}_1\bar{\pi}_1
\end{aligned}$$

and the coefficient on f_{jt} is

$$\begin{aligned}
\hat{\theta}_{f_t} =&\frac{1}{\zeta^2}\left(\frac{-\rho}{\sigma^2\tau^2-\rho^2}Cov[w_t-C,\pi_t]+\frac{\sigma^2}{\sigma^2\tau^2-\rho^2}Cov[f_t-C,\pi_t]+\frac{\rho}{\sigma^2\tau^2-\rho^2}Cov[f_{t-1}-C,\pi_t]\right)\\
=&\hat{\Gamma}_3\bar{\pi}_1.
\end{aligned}$$

For $I \geq 1$, the coefficient on $w_{j(t-1)}$ is

$$\hat{\theta}_{w_{t-1}} = \frac{1}{\zeta^2}\left(\frac{\tau^2}{\sigma^2\tau^2-\rho^2}Cov[w_{t-1}-C,\pi_t] - \frac{\rho}{\sigma^2\tau^2-\rho^2}Cov[f_{t-1}-C,\pi_t] - \frac{\tau^2}{\sigma^2\tau^2-\rho^2}Cov[f_{t-2}-C,\pi_t]\right)$$
$$= \bar{\pi}_4 + \hat{\Gamma}_1\left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2 + \hat{\Gamma}_2\bar{\pi}_1$$

and the coefficient on $f_{j(t-1)}$ is

$$\hat{\theta}_{f_{t-1}} = \frac{1}{\zeta^2}\Bigg(-\frac{\tau^2}{\sigma^2\tau^2-\rho^2}Cov[w_t-C,\pi_t] - \frac{\rho}{\sigma^2\tau^2-\rho^2}Cov[w_{t-1}-C,\pi_t] + \frac{\rho}{\sigma^2\tau^2-\rho^2}Cov[f_t-C,\pi_t]$$
$$+ \frac{\sigma^2+\tau^2}{\sigma^2\tau^2-\rho^2}Cov[f_{t-1}-C,\pi_t] + \frac{\rho}{\sigma^2\tau^2-\rho^2}Cov[f_{t-2}-C,\pi_t]\Bigg)$$
$$= \hat{\Gamma}_3\left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2.$$

For $i \geq 2$ and $I \geq i$, we find:[50]

$$\hat{\theta}_{w_{t-i}} = \frac{1}{\zeta^2}\left(\frac{\tau^2}{\sigma^2\tau^2-\rho^2}Cov[w_{t-i}-C,\pi_t] - \frac{\rho}{\sigma^2\tau^2-\rho^2}Cov[f_{t-i}-C,\pi_t] - \frac{\tau^2}{\sigma^2\tau^2-\rho^2}Cov[f_{t-(i+1)}-C,\pi_t]\right)$$
$$= \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-1}\hat{\Gamma}_1\left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2 + \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-2}\hat{\Gamma}_2\left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2,$$
$$\hat{\theta}_{f_{t-i}} = \frac{1}{\zeta^2}\Bigg(-\frac{\tau^2}{\sigma^2\tau^2-\rho^2}Cov[w_{t-(i-1)}-C,\pi_t] - \frac{\rho}{\sigma^2\tau^2-\rho^2}Cov[w_{t-i}-C,\pi_t]$$
$$+ \frac{\rho}{\sigma^2\tau^2-\rho^2}Cov[f_{t-(i-1)}-C,\pi_t] + \frac{\sigma^2+\tau^2}{\sigma^2\tau^2-\rho^2}Cov[f_{t-i}-C,\pi_t]$$
$$+ \frac{\rho}{\sigma^2\tau^2-\rho^2}Cov[f_{t-(i+1)}-C,\pi_t]\Bigg)$$
$$= \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-1}\hat{\Gamma}_3\left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2.$$

For $I \geq 2$, we have:

$$\sum_{i=0}^{I}\left[\hat{\theta}_{w_{t-i}} + \hat{\theta}_{f_{t-i}}\right] = \bar{\pi}_3 + \bar{\pi}_4 + \bar{\pi}_1\left[\hat{\Gamma}_1 + \hat{\Gamma}_2 + \hat{\Gamma}_3\right]$$
$$+ \left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2\left[(\hat{\Gamma}_1+\hat{\Gamma}_3)\sum_{k=0}^{I-1}\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^k + \hat{\Gamma}_2\sum_{k=0}^{I-2}\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^k\right].$$

[50]To show this, derive coefficients on $\hat{\theta}_{w_{t-2}}$, $\hat{\theta}_{w_{t-i}}$ for $i \geq 3$, $\hat{\theta}_{f_{t-2}}$, $\hat{\theta}_{f_{t-3}}$, and $\hat{\theta}_{f_{t-i}}$ for $i \geq 4$.

If Assumption 3 holds, then $\bar{\pi}_1 = \bar{\pi}_2 = 0$. In that case, equation (6) implies $\mathrm{d}\bar{A}/\mathrm{d}C = \bar{\pi}_3 + \bar{\pi}_4$. And the only nonzero coefficients are $\hat{\theta}_{w_t} = \bar{\pi}_3$ and $\hat{\theta}_{w_{t-1}} = \bar{\pi}_4$. It is easy to show that this same result holds even if $I = 0$.

If $\bar{\pi}_{12} = 0$ and $I > 1$, then

$$\sum_{i=0}^{I} \left[\hat{\theta}_{w_{t-i}} + \hat{\theta}_{f_{t-i}}\right] = \bar{\pi}_3 + \bar{\pi}_4 + [\bar{\pi}_1 + \bar{\pi}_2]\left[\hat{\Gamma}_1 + \hat{\Gamma}_2 + \hat{\Gamma}_3\right]$$

for $I \geq 2$. And using equation (6) and Proposition 1, we know that the right-hand side is equal to $\lim_{s\to\infty} \mathrm{d}E_0[\pi_s]/\mathrm{d}C$ when $\bar{\pi}_{12} = 0$. Finally, note that $\hat{\theta}_{w_{t-i}} = 0$ for $i \geq 3$ and $\hat{\theta}_{f_{t-i}} = 0$ for $i \geq 2$.

Now consider the case with $I \to \infty$, allowing $\bar{\pi}_{12} \neq 0$:

$$\lim_{I\to\infty} \sum_{i=0}^{I} \left[\hat{\theta}_{w_{t-i}} + \hat{\theta}_{f_{t-i}}\right] = \bar{\pi}_3 + \bar{\pi}_4 + \left[\hat{\Gamma}_1 + \hat{\Gamma}_2 + \hat{\Gamma}_3\right]\left[\bar{\pi}_1 + \frac{1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}\bar{\pi}_2\right].$$

Using Proposition 1 and then using equations (6) and (2), this becomes:

$$\begin{aligned}
\lim_{I\to\infty} \sum_{i=0}^{I} \left[\hat{\theta}_{w_{t-i}} + \hat{\theta}_{f_{t-i}}\right] =& \bar{\pi}_3 + \bar{\pi}_4 + \omega\left(\frac{\mathrm{d}\bar{A}}{\mathrm{d}C} + \beta\bar{\pi}_{12}\Omega\right)\left[\bar{\pi}_1 + \frac{1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}\bar{\pi}_2\right] \\
&= \lim_{s\to\infty} \frac{\mathrm{d}E_0[\pi_s]}{\mathrm{d}C} + \frac{\mathrm{d}\bar{A}}{\mathrm{d}C}\frac{\frac{\bar{\pi}_{12}}{\chi_2} - \beta\frac{\bar{\pi}_{12}}{\chi_2}}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}\bar{\pi}_2 \\
&\quad + \left((\omega - 1)\frac{\mathrm{d}\bar{A}}{\mathrm{d}C} + \omega\beta\bar{\pi}_{12}\Omega\right)\left[\bar{\pi}_1 + \bar{\pi}_2 + \frac{\frac{\bar{\pi}_{12}}{\chi_2} - \beta\frac{\bar{\pi}_{12}}{\chi_2}}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}\bar{\pi}_2\right] \\
&= \lim_{s\to\infty} \frac{\mathrm{d}E_0[\pi_s]}{\mathrm{d}C} + \frac{\mathrm{d}\bar{A}}{\mathrm{d}C}\frac{1 - \beta}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}\left(\frac{\bar{\pi}_{12}}{\chi_2} + \omega - 1\right)\bar{\pi}_2 + \omega\beta\bar{\pi}_{12}\frac{1 - \beta}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}\Omega\bar{\pi}_2.
\end{aligned}$$

Note that:

$$\omega - 1 + \frac{\bar{\pi}_{12}}{\chi_2} = -\beta\bar{\pi}_{12}\frac{1 - \frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}.$$

Using this, we have:

$$\lim_{I\to\infty} \sum_{i=0}^{I} \left[\hat{\theta}_{w_{t-i}} + \hat{\theta}_{f_{t-i}}\right] = \lim_{s\to\infty} \frac{\mathrm{d}E_0[\pi_s]}{\mathrm{d}C} + \frac{\beta(1 - \beta)}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}\bar{\pi}_{12}\bar{\pi}_2\left\{\omega\Omega - \frac{\mathrm{d}\bar{A}}{\mathrm{d}C}\frac{1 - \frac{\bar{\pi}_{12}}{\chi_1}}{\chi_2}\right\}.$$

The bias vanishes if $\beta = 0$ and also as $\beta \to 1$.

E.7 Proof of Corollary 4

Following the proof of Proposition 3, we now have:

$$
\begin{aligned}
Cov[w_t - C, \pi_t] =& \bar{\pi}_1 Cov[A_t, w_t] + \bar{\pi}_2 Cov[A_{t-1}, w_t] \\
&+ \left(\bar{\pi}_3 + \bar{\pi}_{13}(E[A_t] - \bar{A}) + \bar{\pi}_{23}(E[A_{t-1}] - \bar{A})\right) Var[w_t] \\
&+ \left(\bar{\pi}_4 + \bar{\pi}_{14}(E[A_t] - \bar{A}) + \bar{\pi}_{24}(E[A_{t-1}] - \bar{A})\right) Cov[w_t, w_{t-1}],
\end{aligned}
$$

$$
Cov[f_t - C, \pi_t] = \bar{\pi}_1 Cov[A_t, f_t] + \left(\bar{\pi}_3 + \bar{\pi}_{13}(E[A_t] - \bar{A}) + \bar{\pi}_{23}(E[A_{t-1}] - \bar{A})\right) Cov[w_t, f_t],
$$

$$
\begin{aligned}
Cov[w_{t-1} - C, \pi_t] =& \bar{\pi}_1 Cov[A_t, w_{t-1}] + \bar{\pi}_2 Cov[A_{t-1}, w_{t-1}] \\
&+ \left(\bar{\pi}_3 + \bar{\pi}_{13}(E[A_t] - \bar{A}) + \bar{\pi}_{23}(E[A_{t-1}] - \bar{A})\right) Cov[w_t, w_{t-1}] \\
&+ \left(\bar{\pi}_4 + \bar{\pi}_{14}(E[A_t] - \bar{A}) + \bar{\pi}_{24}(E[A_{t-1}] - \bar{A})\right) Var[w_{t-1}],
\end{aligned}
$$

$$
\begin{aligned}
Cov[f_{t-1} - C, \pi_t] =& \bar{\pi}_1 Cov[A_t, f_{t-1}] + \bar{\pi}_2 Cov[A_{t-1}, f_{t-1}] \\
&+ \left(\bar{\pi}_3 + \bar{\pi}_{13}(E[A_t] - \bar{A}) + \bar{\pi}_{23}(E[A_{t-1}] - \bar{A})\right) Cov[w_t, f_{t-1}] \\
&+ \left(\bar{\pi}_4 + \bar{\pi}_{14}(E[A_t] - \bar{A}) + \bar{\pi}_{24}(E[A_{t-1}] - \bar{A})\right) Cov[w_{t-1}, f_{t-1}],
\end{aligned}
$$

and so on. Following that analysis, we obtain the regression coefficients:

$$
\hat{\theta}_{w_t} = \bar{\pi}_3 + \hat{\Gamma}_1 \bar{\pi}_1 + \bar{\pi}_{13}(E[A_t] - \bar{A}) + \bar{\pi}_{23}(E[A_{t-1}] - \bar{A}),
$$

$$
\hat{\theta}_{w_{t-1}} = \bar{\pi}_4 + \hat{\Gamma}_1 \left(1 - \beta \frac{\bar{\pi}_{12}}{\chi_2}\right) \bar{\pi}_2 + \hat{\Gamma}_2 \bar{\pi}_1 + \bar{\pi}_{14}(E[A_t] - \bar{A}) + \bar{\pi}_{24}(E[A_{t-1}] - \bar{A}).
$$

The other coefficients are unchanged. Under the assumption that at least one of $\bar{\pi}_{13}, \bar{\pi}_{14}, \bar{\pi}_{23}, \bar{\pi}_{24}$ is strictly positive, we have increased $\hat{\theta}_{w_t}$ and $\hat{\theta}_{w_{t-1}}$ if average actions are above $\bar{A}$ and have decreased $\hat{\theta}_{w_t}$ and $\hat{\theta}_{w_{t-1}}$ if average actions are below $\bar{A}$. The results follow.

E.8 Proof of Lemma 3

If there exist $g : \mathbb{R}^M \to \mathbb{R}$, $h : \mathbb{R}^M \to \mathbb{R}$, and $\pi^0 : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^K \to \mathbb{R}$ such that $\pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}) = \pi^0(g(\boldsymbol{A}_t), h(\boldsymbol{A}_{t-1}), \boldsymbol{w}_t)$, then define $\breve{A}_t \triangleq g(\boldsymbol{A}_t)$ and $\breve{A}_{t-1} \triangleq h(\boldsymbol{A}_{t-1})$. We now have $\pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}) = \pi^0(\breve{A}_t, \breve{A}_{t-1}, \boldsymbol{w}_t)$, which would be a special case of the main text's setting but for $K > 1$. Section F shows that the results of Proposition 6 go through, except now indexing each $\hat{\Gamma}$ and weather derivative according to the dimension of weather under consideration. The only additional step is to demean for region-year fixed effects, as described below. The coefficient on the lead of weather follows from the analysis of $Cov[f_t - C, \pi_t]$ in the proof of Proposition 3.

Now consider the case in which

$$\pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}) = \sum_{k=1}^{K} \pi^k(A_t^k, A_{t-1}^k, w_t^k) + \pi^{K+1}(\boldsymbol{A}_t^{\sim k}, \boldsymbol{A}_{t-1}^{\sim k}).$$

Let ϵ_{t+1}^k and ν_{t+1}^k have variance $(\sigma^k)^2$ and $(\tau^k)^2$ respectively (where superscript k is, here and below, an index, not a power). Applying Assumption 4 and following the same steps as in the main text, it is easy to show that, if either ζ^2 is small or π^k is quadratic,

$$\lim_{t\to\infty} \frac{\mathrm{d}E_0[A_t^k]}{\mathrm{d}C^k} = \frac{\mathrm{d}\bar{A}^k}{\mathrm{d}C^k} = \frac{\bar{\pi}_{13}^k + \beta\bar{\pi}_{23}^k}{-\bar{\pi}_{11}^k - (1+\beta)\bar{\pi}_{12}^k - \beta\bar{\pi}_{22}^k}$$

and

$$\lim_{t\to\infty} \frac{\mathrm{d}E_0[\pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t)]}{\mathrm{d}C^k} = \bar{\pi}_3^k + [\bar{\pi}_1^k + \bar{\pi}_2^k]\frac{\mathrm{d}\bar{A}^k}{\mathrm{d}C^k},$$

with $\bar{\pi}_1^k = -\beta\bar{\pi}_2^k$.

Now consider the regression coefficients $\hat{\Phi}$. Let there be J counties, T years of observations, and R regions. The vector of estimated coefficients is

$$\hat{\Phi} = E[X^\intercal X]^{-1} E[X^\intercal \boldsymbol{\pi}],$$

where $\boldsymbol{\pi}$ is a $JT \times 1$ vector with rows π_{ct} and X is a $JT \times (J + RT + 4K)$ matrix with the final $4K$ columns of each row being

$$\begin{bmatrix} w_{j(t+1)}^1 & \cdots & w_{j(t-2)}^1 & w_{j(t+1)}^K & \cdots & w_{j(t-2)}^K \end{bmatrix}.$$

By the Frisch-Waugh Theorem,

$$\hat{\Phi} = E[\tilde{X}^\intercal \tilde{X}]^{-1} E[\tilde{X}^\intercal \tilde{\boldsymbol{\pi}}],$$

where $\tilde{\boldsymbol{\pi}}$ is demeaned $\boldsymbol{\pi}$ and $\tilde{X}$ is a $JT \times 4K$ matrix with the final $4K$ columns of each row being

$$\begin{bmatrix} w_{j(t+1)}^1 - C^1 - \bar{w}_{r(t+1)}^1 & \cdots & w_{j(t-2)}^1 - C^1 - \bar{w}_{r(t-2)}^1 & w_{j(t+1)}^K - C^K - \bar{w}_{r(t+1)}^K & \cdots & w_{j(t-2)}^K - C^K - \bar{w}_{r(t-2)}^K \end{bmatrix},$$

where C^k indicates climate index k and $\bar{w}_{rt}^k$ is the average weather dimension k observed in region r in year t. We can show by the Frisch-Waugh Theorem that controlling for all dimensions of weather means that we can analyze the regression dimension by dimension. Let $\hat{\Phi}^k$ be the portion of the vector of coefficients corresponding to weather dimension k. Then:

$$\hat{\Phi}^k = E[(\tilde{X}^k)^\intercal \tilde{X}^k]^{-1} E[(\tilde{X}^k)^\intercal \tilde{\boldsymbol{\pi}}],$$

where $\tilde{X}^k$ is a $JT \times 4$ matrix with the final 4 columns of each row being

$$\begin{bmatrix} w_{j(t+1)}^k - C^k - \bar{w}_{r(t+1)}^k & \dots & w_{j(t-2)}^k - C^k - \bar{w}_{r(t-2)}^k \end{bmatrix}.$$

The assumption that weather and forecast shocks are uncorrelated with each other then implies that $E[(\tilde{X}^k)^\intercal(\tilde{X}^k)]$ is a diagonal matrix with $JT\zeta^2[(\sigma^k)^2 + (\tau^k)^2]$ on the diagonal, so $E[(\tilde{X}^k)^\intercal(\tilde{X}^k)]^{-1}$ is a diagonal matrix with $1/[JT\zeta^2((\sigma^k)^2 + (\tau^k)^2)]$ on the diagonal. Following the proof of Proposition 6 yields the expressions in the lemma.

E.9 Proof of Proposition 5

The vector of estimated coefficients is

$$\hat{\gamma} = E[X_I^\intercal X_I]^{-1} E[X_I^\intercal \boldsymbol{A}],$$

where $\boldsymbol{A}$ is a $JT \times 1$ vector with rows A_{jt} and X_I is a $JT \times J + I + 1$ matrix with the final $I + 1$ columns of each row being

$$\begin{bmatrix} w_{jt} & \dots & w_{j(t-I)} \end{bmatrix}.$$

By the Frisch-Waugh Theorem,

$$\hat{\gamma} = E[\tilde{X}_I^\intercal \tilde{X}_I]^{-1} E[\tilde{X}_I^\intercal \tilde{\boldsymbol{A}}],$$

where $\tilde{X}_I$ is a $JT \times 2(I + 1)$ matrix with rows

$$\begin{bmatrix} w_{jt} - C & \dots & w_{j(t-I)} - C \end{bmatrix}$$

and $\tilde{\boldsymbol{A}}$ is demeaned $\boldsymbol{A}$.

First consider $I = 0$. The coefficient on w_{jt} is:

$$\hat{\gamma}_{w_t} = \frac{Cov[A_{jt}, w_{jt}]}{\zeta^2(\sigma^2 + \tau^2)} = \hat{\Gamma}_1 \left(1 + \frac{\bar{\pi}_{12}}{\chi_2} \frac{\rho}{\sigma^2 + \tau^2}\right) + \hat{\Gamma}_2 \frac{\rho}{\sigma^2 + \tau^2} + \hat{\Gamma}_3 \frac{\rho + \tau^2 \frac{\bar{\pi}_{12}}{\chi_2}}{\sigma^2 + \tau^2},$$

where the second equality uses results for the covariance in the proof of Proposition 2 that depend on the stated assumptions. If $\bar{\pi}_{14} = \beta\bar{\pi}_{23} = 0$, then $\hat{\gamma}_{w_t} = \hat{\Gamma}_1$, which the proof of Proposition 1 shows equals $\mathrm{d}\bar{A}/\mathrm{d}C$ if those same conditions hold and $\bar{\pi}_{12} = 0$.

Now let $I = 1$. Begin by considering $\hat{\gamma}_{w_t}$, via the Frisch-Waugh theorem. The residuals from regressing $w_{jt} - C$ on $w_{j(t-1)} - C$ are:

$$\tilde{w}_{jt} \triangleq w_{jt} - C - \frac{\rho}{\tau^2 + \sigma^2}(w_{j(t-1)} - C) = \zeta\epsilon_{jt} + \zeta\nu_{j(t-1)} - \zeta\frac{\rho}{\tau^2 + \sigma^2}[\epsilon_{j(t-1)} + \nu_{j(t-2)}].$$

We then have:

$$\hat{\gamma}_{w_t} = \frac{Cov[\tilde{w}_{jt}, A_{jt}]}{Var[\tilde{w}_{jt}]} = \hat{\Gamma}_1 - \hat{\Gamma}_2\frac{\bar{\pi}_{12}}{\chi_2}\frac{\frac{\rho^2}{\tau^2+\sigma^2}}{\sigma^2 + \tau^2 - \frac{\rho^2}{\tau^2+\sigma^2}} + \hat{\Gamma}_3\frac{\rho + \frac{\bar{\pi}_{12}}{\chi_2}\left(\tau^2 - \frac{\rho^2}{\tau^2+\sigma^2}\right)}{\sigma^2 + \tau^2 - \frac{\rho^2}{\tau^2+\sigma^2}},$$

where the second equality uses results for the covariance in the proof of Proposition 2 that depend on the stated assumptions. Now consider $\hat{\gamma}_{w_{t-1}}$. The residuals from regressing $w_{j(t-1)} - C$ on $w_{jt} - C$ are:

$$\tilde{w}_{j(t-1)} \triangleq w_{j(t-1)} - C - \frac{\rho}{\tau^2 + \sigma^2}(w_{jt} - C) = \zeta\epsilon_{j(t-1)} + \zeta\nu_{j(t-2)} - \zeta\frac{\rho}{\tau^2 + \sigma^2}[\epsilon_{jt} + \nu_{j(t-1)}].$$

We then have:

$$\hat{\gamma}_{w_{t-1}} = \hat{\Gamma}_1\frac{\bar{\pi}_{12}}{\chi_2} + \hat{\Gamma}_2\left(1 + \frac{\bar{\pi}_{12}}{\chi_2}\frac{\rho}{\sigma^2 + \tau^2 - \frac{\rho^2}{\tau^2+\sigma^2}}\right) + \hat{\Gamma}_3\frac{\frac{\bar{\pi}_{12}}{\chi_2}\frac{\sigma^2}{\tau^2+\sigma^2} - \frac{\rho}{\tau^2+\sigma^2}}{\sigma^2 + \tau^2 - \frac{\rho^2}{\tau^2+\sigma^2}}\rho.$$

If $\beta\bar{\pi}_{23} = \bar{\pi}_{12} = 0$, then $\hat{\gamma}_{w_t} + \hat{\gamma}_{w_{t-1}} = \hat{\Gamma}_1 + \hat{\Gamma}_2$, which the proof of Proposition 1 shows equals $\mathrm{d}\bar{A}/\mathrm{d}C$ if those same conditions hold.

Now assume $\rho = 0$ and let I be arbitrary. $E[\tilde{X}_I^\intercal\tilde{X}_I]$ is just a diagonal matrix with $JT\zeta^2(\sigma^2 + \tau^2)$ on the diagonal, so $E[\tilde{X}_I^\intercal\tilde{X}_I]^{-1}$ is a diagonal matrix with $1/[JT\zeta^2(\sigma^2 + \tau^2)]$ on the diagonal. We have:

$$\hat{\gamma}_{w_{t-i}} = \frac{Cov[w_{j(t-i)}, A_{jt}]}{\zeta^2(\sigma^2 + \tau^2)}.$$

Using results for the covariances derived in the proof of Proposition 2, we have

$$\hat{\gamma}_{w_t} = \hat{\Gamma}_1 + \hat{\Gamma}_3\frac{\bar{\pi}_{12}}{\chi_2}\frac{\tau^2}{\sigma^2 + \tau^2}$$

and, for $i > 0$,

$$\hat{\gamma}_{w_{t-i}} = \hat{\Gamma}_1\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^i + \hat{\Gamma}_2\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-1} + \hat{\Gamma}_3\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i+1}\frac{\tau^2}{\sigma^2 + \tau^2}.$$

Therefore

$$\lim_{I\to\infty}\sum_{i=0}^{I}\hat{\gamma}_{w_{t-i}} = \frac{\hat{\Gamma}_1 + \hat{\Gamma}_2 + \hat{\Gamma}_3\frac{\bar{\pi}_{12}}{\chi_2}\frac{\tau^2}{\sigma^2+\tau^2}}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}.$$

If either $\beta = 0$ or $\bar{\pi}_{23} = \bar{\pi}_{12} = 0$, then $\lim_{I\to\infty}\sum_{i=0}^{I}\hat{\gamma}_{w_{t-i}} = (\hat{\Gamma}_1 + \hat{\Gamma}_2)/(1 - \bar{\pi}_{12}/\chi_2)$, which Proposition 2 shows equals $\mathrm{d}\bar{A}/\mathrm{d}C$ if either $\beta = 0$ or $\bar{\pi}_{12}\bar{\pi}_{23} = 0$.

E.10 Proof of Proposition 6

I first derive estimators in the cases of $I = 0$, $I = 1$, and $I = 2$. I then derive estimators for general I in the special case with $\rho = 0$.

The vector of estimated coefficients is

$$\hat{\Phi} = E[X_I^\intercal X_I]^{-1} E[X_I^\intercal \boldsymbol{\pi}],$$

where $\boldsymbol{\pi}$ is a $JT \times 1$ vector with rows π_{jt} and X_I is a $JT \times J + I + 1$ matrix with the final $I + 1$ columns of each row being

$$\begin{bmatrix} w_{jt} & ... & w_{j(t-I)} \end{bmatrix}.$$

By the Frisch-Waugh Theorem,

$$\hat{\Phi} = E[\tilde{X}_I^\intercal \tilde{X}_I]^{-1} E[\tilde{X}_I^\intercal \tilde{\boldsymbol{\pi}}],$$

where $\tilde{X}_I$ is a $JT \times I + 1$ matrix with rows

$$\begin{bmatrix} w_{jt} - C & ... & w_{j(t-I)} - C \end{bmatrix}$$

and $\tilde{\boldsymbol{\pi}}$ is demeaned $\boldsymbol{\pi}$.

Begin with $I = 0$. We have:

$$\hat{\Phi}_{w_t} = \frac{Cov[w_{jt} - C, \pi_{jt}]}{\zeta^2(\sigma^2 + \tau^2)}.$$

We analyzed this covariance in the proof of Proposition 3. Using those results, we find that

$$\begin{aligned}
\hat{\Phi}_{w_t} =& \bar{\pi}_3 + \bar{\pi}_4 \frac{\rho}{\sigma^2 + \tau^2} + \hat{\Gamma}_1 \left[\bar{\pi}_1 + \left(1 - \beta \frac{\bar{\pi}_{12}}{\chi_2} \right) \frac{\rho}{\sigma^2 + \tau^2} \bar{\pi}_2 \right] \\
&+ \hat{\Gamma}_2 \bar{\pi}_1 \frac{\rho}{\sigma^2 + \tau^2} + \hat{\Gamma}_3 \left[\bar{\pi}_1 \frac{\rho}{\sigma^2 + \tau^2} + \left(1 - \beta \frac{\bar{\pi}_{12}}{\chi_2} \right) \frac{\tau^2}{\sigma^2 + \tau^2} \bar{\pi}_2 \right].
\end{aligned}$$

Now consider $I = 1$. Now,

$$E[\tilde{X}_1^\intercal \tilde{X}_1]^{-1} = \frac{1}{JT\zeta^2[(\sigma^2 + \tau^2)^2 - \rho^2]} \begin{bmatrix} \sigma^2 + \tau^2 & -\rho \\ -\rho & \sigma^2 + \tau^2 \end{bmatrix}.$$

We also have:

$$E[\tilde{X}_1^\intercal \boldsymbol{\pi}] = JT \begin{bmatrix} Cov[w_{jt} - C, \pi_{jt}] \\ Cov[w_{j(t-1)} - C, \pi_{jt}] \end{bmatrix}.$$

We analyzed these covariances in the proof of Proposition 3. Using those results, we find that

$$\begin{aligned}\hat{\Phi}_{w_t} =&\bar{\pi}_3 + \hat{\Gamma}_1 \left\{ \bar{\pi}_1 - \frac{\rho^2}{(\sigma^2+\tau^2)^2-\rho^2} \frac{\bar{\pi}_{12}}{\chi_2} \left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right) \bar{\pi}_2 \right\} - \hat{\Gamma}_2 \frac{\rho^2}{(\sigma^2+\tau^2)^2-\rho^2} \left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right) \bar{\pi}_2 \\ &+ \hat{\Gamma}_3 \left\{ \frac{\rho(\sigma^2+\tau^2)}{(\sigma^2+\tau^2)^2-\rho^2} \bar{\pi}_1 + \left(\frac{\tau^2(\sigma^2+\tau^2)-\rho^2}{(\sigma^2+\tau^2)^2-\rho^2} - \frac{\rho\tau^2}{(\sigma^2+\tau^2)^2-\rho^2} \frac{\bar{\pi}_{12}}{\chi_2} \right) \left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right) \bar{\pi}_2 \right\},\end{aligned}$$

$$\begin{aligned}\hat{\Phi}_{w_{t-1}} =&\bar{\pi}_4 + \hat{\Gamma}_1 \left(1 + \frac{\rho(\sigma^2+\tau^2)}{(\sigma^2+\tau^2)^2-\rho^2} \frac{\bar{\pi}_{12}}{\chi_2}\right) \left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right) \bar{\pi}_2 + \hat{\Gamma}_2 \left[\bar{\pi}_1 + \frac{\rho(\sigma^2+\tau^2)}{(\sigma^2+\tau^2)^2-\rho^2} \left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right) \bar{\pi}_2 \right] \\ &+ \hat{\Gamma}_3 \left\{ \frac{-\rho^2}{(\sigma^2+\tau^2)^2-\rho^2} \bar{\pi}_1 + \left(\frac{\rho\sigma^2}{(\sigma^2+\tau^2)^2-\rho^2} + \frac{\tau^2(\sigma^2+\tau^2)}{(\sigma^2+\tau^2)^2-\rho^2} \frac{\bar{\pi}_{12}}{\chi_2} \right) \left(1-\beta\frac{\bar{\pi}_{12}}{\chi_2}\right) \bar{\pi}_2 \right\}.\end{aligned}$$

Now consider $I = 2$. Now,

$$E[\tilde{X}_2^\intercal \tilde{X}_2]^{-1} = \frac{1}{JT\zeta^2[(\sigma^2+\tau^2)^2-2\rho^2]} \begin{bmatrix} \sigma^2+\tau^2-\frac{\rho^2}{\sigma^2+\tau^2} & -\rho & \frac{\rho^2}{\sigma^2+\tau^2} \\ -\rho & \sigma^2+\tau^2 & -\rho \\ \frac{\rho^2}{\sigma^2+\tau^2} & -\rho & \sigma^2+\tau^2-\frac{\rho^2}{\sigma^2+\tau^2} \end{bmatrix}.$$

We also have:

$$E[\tilde{X}_2^\intercal \boldsymbol{\pi}] =JT \begin{bmatrix} Cov[w_{jt}-C, \pi_{jt}] \\ Cov[w_{j(t-1)}-C, \pi_{jt}] \\ Cov[w_{j(t-2)}-C, \pi_{jt}] \end{bmatrix}.$$

We analyzed these covariances in the proof of Proposition 3. Using those results, we find that

$$\begin{aligned}\hat{\Phi}_{w_t} =&\bar{\pi}_3 + \hat{\Gamma}_1 \left\{ \bar{\pi}_1 + \frac{\rho}{\sigma^2+\tau^2} \frac{\rho^2}{(\sigma^2+\tau^2)^2-2\rho^2} \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^2 \left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right) \right\} \\ &+ \hat{\Gamma}_3 \frac{1}{(\sigma^2+\tau^2)^2-2\rho^2} \left\{ \left(\sigma^2+\tau^2-\frac{\rho^2}{\sigma^2+\tau^2}\right) \left(\bar{\pi}_1\rho + \bar{\pi}_2\tau^2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right) \right. \\ &\qquad \left. + \left(\frac{\rho^2}{\sigma^2+\tau^2}\frac{\bar{\pi}_{12}}{\chi_2} - \rho\right) \left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right) \left(\rho + \frac{\bar{\pi}_{12}}{\chi_2}\tau^2\right) \right\} \\ &+ \hat{\Gamma}_2 \frac{\rho^2}{(\sigma^2+\tau^2)^2-2\rho^2} \frac{\rho}{\sigma^2+\tau^2} \left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right) \frac{\bar{\pi}_{12}}{\chi_2},\end{aligned}$$

$$\begin{aligned}\hat{\Phi}_{w_{t-1}} =&\bar{\pi}_4 + \hat{\Gamma}_1\left\{1 - \frac{\rho^2}{(\sigma^2+\tau^2)^2 - 2\rho^2}\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^2\right\}\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\\
&+ \hat{\Gamma}_3\frac{1}{(\sigma^2+\tau^2)^2 - 2\rho^2}\left\{ -\rho^2\bar{\pi}_1 + \sigma^2\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\rho + [\tau^2(\sigma^2+\tau^2) - \rho^2]\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\frac{\bar{\pi}_{12}}{\chi_2}\right.\\
&\qquad\left. - \rho\tau^2\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^2\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\right\}\\
&+ \hat{\Gamma}_2\left\{\bar{\pi}_1 - \frac{\rho^2}{(\sigma^2+\tau^2)^2 - 2\rho^2}\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\frac{\bar{\pi}_{12}}{\chi_2}\right\},\end{aligned}$$

$$\begin{aligned}\hat{\Phi}_{w_{t-2}} =&\hat{\Gamma}_1\left\{\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\frac{\bar{\pi}_{12}}{\chi_2} + \frac{\rho\,(\sigma^2+\tau^2)}{(\sigma^2+\tau^2)^2 - 2\rho^2}\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^2\right\}\\
&+ \hat{\Gamma}_3\frac{1}{(\sigma^2+\tau^2)^2 - 2\rho^2}\\
&\quad\left\{\frac{\rho^2}{\sigma^2+\tau^2}\bar{\pi}_1\rho - \left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\frac{\sigma^2\rho^2}{\sigma^2+\tau^2} + \rho\left(\sigma^2 - \frac{\rho^2}{\sigma^2+\tau^2}\right)\frac{\bar{\pi}_{12}}{\chi_2}\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\right.\\
&\qquad\left. + \tau^2\left(\sigma^2+\tau^2 - \frac{\rho^2}{\sigma^2+\tau^2}\right)\left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^2\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right)\right\}\\
&+ \hat{\Gamma}_2\left\{1 + \frac{(\sigma^2+\tau^2)^2 - \rho^2}{(\sigma^2+\tau^2)^2 - 2\rho^2}\frac{\rho}{\sigma^2+\tau^2}\frac{\bar{\pi}_{12}}{\chi_2}\right\}\left(\bar{\pi}_2 + \bar{\pi}_1\frac{\bar{\pi}_{12}}{\chi_2}\right).\end{aligned}$$

Now assume $\rho = 0$ and let I be arbitrary. $E[\tilde{X}_I^\intercal\tilde{X}_I]$ is just a diagonal matrix with $JT\zeta^2(\sigma^2+\tau^2)$ on the diagonal, so $E[\tilde{X}_I^\intercal\tilde{X}_I]^{-1}$ is a diagonal matrix with $1/[JT\zeta^2(\sigma^2+\tau^2)]$ on the diagonal. Following the proof of Proposition 3, we find:

$$\hat{\Phi}_{w_t} =\bar{\pi}_3 + \hat{\Gamma}_1\bar{\pi}_1 + \left(1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2\frac{\tau^2}{\tau^2+\sigma^2}\hat{\Gamma}_3,$$

$$\hat{\Phi}_{w_{t-1}} =\bar{\pi}_4 + \bar{\pi}_1\hat{\Gamma}_2 + \left(1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2\hat{\Gamma}_1 + \left(1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\frac{\bar{\pi}_{12}}{\chi_2}\bar{\pi}_2\frac{\tau^2}{\sigma^2+\tau^2}\hat{\Gamma}_3,$$

and, for $i \geq 2$,

$$\hat{\Phi}_{w_{t-i}} = \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-2}\left(1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2\hat{\Gamma}_2 + \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i-1}\left(1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2\hat{\Gamma}_1 + \left(\frac{\bar{\pi}_{12}}{\chi_2}\right)^{i}\left(1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}\right)\bar{\pi}_2\frac{\tau^2}{\tau^2+\sigma^2}\hat{\Gamma}_3.$$

Therefore

$$\lim_{I\to\infty}\sum_{i=0}^{I}\hat{\Phi}_{w_{t-i}} =\bar{\pi}_3 + \bar{\pi}_4 + \left[\hat{\Gamma}_1 + \hat{\Gamma}_2\right]\left[\bar{\pi}_1 + \frac{1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}\bar{\pi}_2\right] + \hat{\Gamma}_3\frac{\tau^2}{\tau^2+\sigma^2}\frac{1 - \beta\frac{\bar{\pi}_{12}}{\chi_2}}{1 - \frac{\bar{\pi}_{12}}{\chi_2}}\bar{\pi}_2.$$

From the proof of Proposition 1, $\beta\bar{\pi}_{23} = 0$ implies that $\hat{\Gamma}_3 = 0$. And Proposition 1 itself then implies

$$\hat{\Gamma}_1 + \hat{\Gamma}_2 = \omega \left(\frac{\mathrm{d}\bar{A}}{\mathrm{d}C} + \beta\bar{\pi}_{12}\Omega \right).$$

The proposition follows by inspection (noting that $\bar{\pi}_{14} = 0$ implies $\hat{\Gamma}_2 = 0$ and that $\bar{\pi}_{12} = \beta\bar{\pi}_{23} = 0$ implies $\hat{\Gamma}_3 = 0$), by recalling that Assumption 3 implies $\bar{\pi}_1 = \bar{\pi}_2 = 0$, and, for the case with $I \to \infty$, by following the last part of the proof of Proposition 3.

E.11 Proof of Proposition 7

Observe that

$$\hat{\Lambda} = \frac{Cov[\check{\pi}_{jt}, \check{w}_{jt} - C]}{Var[\check{w}_{jt} - C]}.$$

Begin by considering the case in which Assumption 3 holds. Equation (2) requires $\bar{\pi}_2 = \bar{\pi}_1 = 0$. Using intermediate results in the proof of Proposition 3, we have:

$$\frac{1}{\zeta^2} Cov[w_{jt}, \pi_{jt}] = (\sigma^2 + \tau^2)\bar{\pi}_3 + \rho\bar{\pi}_4,$$

$$\frac{1}{\zeta^2} Cov[f_{jt}, \pi_{jt}] = \rho\bar{\pi}_3,$$

$$\frac{1}{\zeta^2} Cov[w_{j(t-1)}, \pi_{jt}] = \rho\bar{\pi}_3 + \bar{\pi}_4(\sigma^2 + \tau^2),$$

$$\frac{1}{\zeta^2} Cov[w_{j(t-2)}, \pi_{jt}] = \rho\bar{\pi}_4,$$

$$\frac{1}{\zeta^2} Cov[w_{j(t-i)}, \pi_{jt}] = 0 \text{ for } i > 2.$$

We then have, for $\Delta > 2$:

$$\begin{aligned}
Cov[\tilde{\pi}_{jt}, \tilde{w}_{jt}] =& \frac{1}{\Delta^2}\Bigg\{ \sum_{T=t+2}^{t+\Delta-2} \bigg(Cov[\pi_{jT}, w_{jT}] + Cov[\pi_{jT}, w_{j(T-1)}] + Cov[\pi_{jT}, w_{j(T-2)}] + Cov[\pi_{jT}, f_{jT}] \bigg) \\
& + Cov[\pi_{j(t+1)}, w_{j(t+1)}] + Cov[\pi_{j(t+1)}, w_{jt}] + Cov[\pi_{j(t+1)}, f_{j(t+1)}] \\
& + Cov[\pi_{jt}, w_{jt}] + Cov[\pi_{jt}, f_{jt}] \\
& + Cov[\pi_{j(t+\Delta-1)}, w_{j(t+\Delta-1)}] + Cov[\pi_{j(t+\Delta-1)}, w_{j(t+\Delta-2)}] + Cov[\pi_{j(t+\Delta-1)}, w_{j(t+\Delta-3)}] \Bigg\} \\
=& \frac{1}{\Delta^2}\Big\{ \Delta[\sigma^2 + \tau^2 + 2\rho][\bar{\pi}_3 + \bar{\pi}_4] - 2\rho\bar{\pi}_3 - [\sigma^2 + \tau^2 + 2\rho]\bar{\pi}_4 \Big\} \\
=& \frac{1}{\Delta^2}\Big\{ [\sigma^2 + \tau^2][\Delta\bar{\pi}_3 + (\Delta - 1)\bar{\pi}_4] + 2\rho[\Delta - 1][\bar{\pi}_3 + \bar{\pi}_4] \Big\}.
\end{aligned}$$

Note that

$$Var(\check{w}_{jt} - C) = \frac{\Delta(\sigma^2 + \tau^2) + 2\rho(\Delta - 1)}{\Delta^2}.$$

The estimator is then:

$$\begin{aligned}
\hat{\Lambda} =& \frac{[\sigma^2 + \tau^2][\Delta\bar{\pi}_3 + (\Delta - 1)\bar{\pi}_4] + 2\rho[\Delta - 1][\bar{\pi}_3 + \bar{\pi}_4]}{\Delta(\sigma^2 + \tau^2) + 2\rho(\Delta - 1)} \\
=& \bar{\pi}_3 + \bar{\pi}_4 \frac{\Delta - 1}{\Delta} \underbrace{\frac{\sigma^2 + \tau^2 + 2\rho}{\sigma^2 + \tau^2 + 2\rho\frac{\Delta-1}{\Delta}}}_{\triangleq \Upsilon_1}.
\end{aligned}$$

Therefore

$$\lim_{\Delta\to\infty} \hat{\Lambda} = \bar{\pi}_3 + \bar{\pi}_4 = \lim_{s\to\infty} \frac{\mathrm{d}E[\pi_s]}{\mathrm{d}C},$$

where the second equality applies $\bar{\pi}_1 = \bar{\pi}_2 = 0$ (from Assumption 3) to equation (6).

Now consider the case in which Assumption 3 need not hold but $\bar{\pi}_{12} = 0$ and $\rho = 0$. Using intermediate results in the proof of Proposition 3, we have:

$$\frac{1}{\zeta^2} Cov[w_{jt}, \pi_{jt}] = (\sigma^2 + \tau^2)\bar{\pi}_3 + \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}(\sigma^2 + \tau^2)\bar{\pi}_1 + \frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}\bar{\pi}_2\tau^2,$$

$$\frac{1}{\zeta^2} Cov[f_{jt}, \pi_{jt}] = \frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}\bar{\pi}_1\tau^2,$$

$$\frac{1}{\zeta^2} Cov[w_{j(t-1)}, \pi_{jt}] = \bar{\pi}_4(\sigma^2 + \tau^2) + \bar{\pi}_1\frac{\bar{\pi}_{14}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}(\sigma^2 + \tau^2) + \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}\bar{\pi}_2\left(\sigma^2 + \tau^2\right),$$

$$\frac{1}{\zeta^2}Cov[w_{j(t-2)},\pi_{jt}] =\bar{\pi}_2\frac{\bar{\pi}_{14}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}(\sigma^2+\tau^2),$$

$$\frac{1}{\zeta^2}Cov[w_{j(t-i)},\pi_{jt}] =0 \text{ for } i>2.$$

Summing these, we have:

$$\begin{aligned}&Cov[w_{jt},\pi_{jt}]+Cov[f_{jt},\pi_{jt}]+Cov[w_{j(t-1)},\pi_{jt}]+Cov[w_{j(t-2)},\pi_{jt}]\\ =&(\sigma^2+\tau^2)\left\{\bar{\pi}_3+\bar{\pi}_4+\frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\bar{\pi}_{14}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}[\bar{\pi}_1+\bar{\pi}_2]+\frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}\frac{\tau^2}{\sigma^2+\tau^2}[\bar{\pi}_1+\bar{\pi}_2]\right\},\end{aligned}$$

$$\begin{aligned}&Cov[w_{j(t+1)},\pi_{j(t+1)}]+Cov[f_{j(t+1)},\pi_{j(t+1)}]+Cov[w_{jt},\pi_{j(t+1)}]\\ =&(\sigma^2+\tau^2)\left\{\bar{\pi}_3+\bar{\pi}_4+\frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}[\bar{\pi}_1+\bar{\pi}_2]+\bar{\pi}_1\frac{\bar{\pi}_{14}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}+\frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}\frac{\tau^2}{\sigma^2+\tau^2}[\bar{\pi}_1+\bar{\pi}_2]\right\},\end{aligned}$$

$$Cov[w_{jt},\pi_{jt}]+Cov[f_{jt},\pi_{jt}] =(\sigma^2+\tau^2)\left\{\bar{\pi}_3+\frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}\bar{\pi}_1+\frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}\frac{\tau^2}{\sigma^2+\tau^2}[\bar{\pi}_1+\bar{\pi}_2]\right\},$$

$$\begin{aligned}&Cov[w_{j(t+\Delta-1)},\pi_{j(t+\Delta-1)}]+Cov[w_{j(t+\Delta-2)},\pi_{j(t+\Delta-1)}]+Cov[w_{j(t+\Delta-3)},\pi_{j(t+\Delta-1)}]\\ =&(\sigma^2+\tau^2)\left\{\bar{\pi}_3+\bar{\pi}_4+\frac{\bar{\pi}_{13}+\beta\bar{\pi}_{24}+\bar{\pi}_{14}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}[\bar{\pi}_1+\bar{\pi}_2]+\frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11}-\beta\bar{\pi}_{22}}\frac{\tau^2}{\sigma^2+\tau^2}\bar{\pi}_2\right\}.\end{aligned}$$

Note that

$$Var(\check{w}_{jt})=\frac{\sigma^2+\tau^2}{\Delta}.$$

The estimator is then, for $\Delta > 2$,

$$
\begin{aligned}
\hat{\Lambda} =& \frac{1}{\Delta^2 Var(\check{w}_{jt})} \Bigg\{ \sum_{T=t+2}^{t+\Delta-2} \left(Cov[\pi_{jT}, w_{jT}] + Cov[\pi_{jT}, w_{j(T-1)}] + Cov[\pi_{jT}, w_{j(T-2)}] + Cov[\pi_{jT}, f_{jT}] \right) \\
&\quad + Cov[\pi_{j(t+1)}, w_{j(t+1)}] + Cov[\pi_{j(t+1)}, w_{jt}] + Cov[\pi_{j(t+1)}, f_{j(t+1)}] \\
&\quad + Cov[\pi_{jt}, w_{jt}] + Cov[\pi_{jt}, f_{jt}] \\
&\quad + Cov[\pi_{j(t+\Delta-1)}, w_{j(t+\Delta-1)}] + Cov[\pi_{j(t+\Delta-1)}, w_{j(t+\Delta-2)}] + Cov[\pi_{j(t+\Delta-1)}, w_{j(t+\Delta-3)}] \Bigg\} \\
=& \bar{\pi}_3 + \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \bar{\pi}_1 + \frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \frac{\tau^2}{\sigma^2 + \tau^2} \bar{\pi}_2 \\
&+ \frac{\Delta - 1}{\Delta} \left\{ \bar{\pi}_4 + \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \bar{\pi}_2 + \bar{\pi}_1 \frac{\bar{\pi}_{14}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} + \frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \frac{\tau^2}{\sigma^2 + \tau^2} \bar{\pi}_1 \right\} \\
&+ \frac{\Delta - 2}{\Delta} \bar{\pi}_2 \frac{\bar{\pi}_{14}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \\
=& \hat{\Phi}^0_{w_t} + \frac{\Delta - 1}{\Delta} \underbrace{\left\{ \bar{\pi}_4 + \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \bar{\pi}_2 + \bar{\pi}_1 \frac{\bar{\pi}_{14}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} + \frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \frac{\tau^2}{\sigma^2 + \tau^2} \bar{\pi}_1 \right\}}_{\triangleq \Upsilon_1} \\
&+ \frac{\Delta - 2}{\Delta} \underbrace{\bar{\pi}_2 \frac{\bar{\pi}_{14}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}}}_{\triangleq \Upsilon_2} .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\lim_{\Delta \to \infty} \hat{\Lambda} =& \bar{\pi}_3 + \bar{\pi}_4 + \frac{\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \bar{\pi}_{14}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} [\bar{\pi}_1 + \bar{\pi}_2] + \frac{\beta\bar{\pi}_{23}}{-\bar{\pi}_{11} - \beta\bar{\pi}_{22}} \frac{\tau^2}{\sigma^2 + \tau^2} [\bar{\pi}_1 + \bar{\pi}_2] \\
=& \bar{\pi}_3 + \bar{\pi}_4 + [\bar{\pi}_1 + \bar{\pi}_2] \left[\hat{\Gamma}_1 + \hat{\Gamma}_2 + \frac{\tau^2}{\sigma^2 + \tau^2} \hat{\Gamma}_3 \right] .
\end{aligned}
$$

The results follow from inspection and previous results on the marginal effect of climate.

F Extension to vector-valued actions and multidimensional weather

Generalize the main text's setting to allow K weather variables, collected in a column vector $\boldsymbol{w}_t$, and M actions, collected in a column vector $\boldsymbol{A}_t$. The payoff function is now

$\pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1})$. Agents solve

$$\begin{aligned} V(\boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{f}_t, \boldsymbol{w}_{t-1}; \zeta) = &\max_{\boldsymbol{A}_t} \left\{ \pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}) + \beta E_t \left[V(\boldsymbol{A}_t, \boldsymbol{w}_{t+1}, \boldsymbol{f}_{t+1}, \boldsymbol{w}_t; \zeta) \right] \right\} \\ \text{s.t.} \quad \boldsymbol{w}_{t+1} =& \boldsymbol{f}_t + \zeta \boldsymbol{\epsilon}_{t+1} \\ \boldsymbol{f}_{t+1} =& \boldsymbol{C} + \zeta \boldsymbol{\nu}_{t+1}, \end{aligned}$$

where $\boldsymbol{f}_t$, $\boldsymbol{\epsilon}_t$, and $\boldsymbol{\nu}_t$ are column vectors of length K. Let ϵ^k_{t+1} and ν^k_{t+1} have variance $(\sigma^k)^2$ and $(\tau^k)^2$ respectively (where superscript k is, as always, an index, not a power). Now π_i represents a column vector corresponding to derivatives with respect to argument i and π_{ij} represents a matrix with columns differentiating π_i with respect to each element of argument j.

Analyze the deterministic model, which fixes $\zeta = 0$. The first-order conditions are:

$$\boldsymbol{0} = \pi_1(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{C}, \boldsymbol{C}) + \beta V_1(\boldsymbol{A}_t, \boldsymbol{C}, \boldsymbol{C}, \boldsymbol{C}; 0).$$

The envelope theorem yields:

$$V_1(\boldsymbol{A}_{t-1}, \boldsymbol{C}, \boldsymbol{C}, \boldsymbol{C}; 0) = \pi_2(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{C}, \boldsymbol{C}).$$

Advancing this forward by one timestep and substituting into the first-order condition, we have the Euler equation:

$$\boldsymbol{0} = \pi_1(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{C}, \boldsymbol{C}) + \beta \pi_2(\boldsymbol{A}_{t+1}, \boldsymbol{A}_t, \boldsymbol{C}, \boldsymbol{C}).$$

A steady state $\bar{\boldsymbol{A}}$ of the deterministic system is implicitly defined by

$$\boldsymbol{0} = \pi_1(\bar{\boldsymbol{A}}, \bar{\boldsymbol{A}}, \boldsymbol{C}, \boldsymbol{C}) + \beta \pi_2(\bar{\boldsymbol{A}}, \bar{\boldsymbol{A}}, \boldsymbol{C}, \boldsymbol{C}). \tag{F-7}$$

Define $\bar{\pi} \triangleq \pi(\bar{\boldsymbol{A}}, \bar{\boldsymbol{A}}, \boldsymbol{C}, \boldsymbol{C})$. Linearizing around $\bar{A}$ gives a first-order difference equation:

$$\boldsymbol{A}_{t+1} - \bar{\boldsymbol{A}} \approx [\beta \bar{\pi}_{12}]^{-1} [-\bar{\pi}_{11} - \beta \bar{\pi}_{22}] (\boldsymbol{A}_t - \bar{\boldsymbol{A}}) - \frac{1}{\beta}(\boldsymbol{A}_{t-1} - \bar{\boldsymbol{A}}).$$

The dynamic system is:

$$\begin{bmatrix} \boldsymbol{A}_{t+1} - \bar{\boldsymbol{A}} \\ \boldsymbol{A}_t - \bar{\boldsymbol{A}} \end{bmatrix} \approx \begin{bmatrix} [\beta \bar{\pi}_{12}]^{-1} [-\bar{\pi}_{11} - \beta \bar{\pi}_{22}] & -\frac{1}{\beta} I_M \\ I_M & 0_{MM} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_t - \bar{\boldsymbol{A}} \\ \boldsymbol{A}_{t-1} - \bar{\boldsymbol{A}} \end{bmatrix},$$

where I_M is an $M \times M$ identity matrix and 0_{MM} is an $M \times M$ matrix of zeros. It is easy to show that as each $\pi_{1_j 1_k}, \pi_{1_j 2_k}, \pi_{2_j 2_k}$ goes to zero for $j \neq k$, there are M pairs of eigenvalues like those identified in the proof of Lemma 2. More generally, using results for the determinants of block matrices, the product of the linearized system's eigenvalues is

$$det\Big([\beta \bar{\pi}_{12}]^{-1} [-\bar{\pi}_{11} - \beta \bar{\pi}_{22}] \Big) det\left(\frac{1}{\beta} \left\{ [\beta \bar{\pi}_{12}]^{-1} [-\bar{\pi}_{11} - \beta \bar{\pi}_{22}] \right\}^{-1} \right) = \left(\frac{1}{\beta} \right)^M > 1.$$

At least one eigenvalue must be greater than 1 in absolute value. If $\pi_{1_j 2_k} = 0$ for all $j \neq k$, then the sum of the linearized system's eigenvalues is $\sum_{m=1}^{M} \frac{-\bar{\pi}_{1_m 1_m} - \beta \bar{\pi}_{2_m 2_m}}{\beta \bar{\pi}_{1_m 2_m}}$, which is positive if each $\bar{\pi}_{1_m 2_m} > 0$ and is negative if each $\bar{\pi}_{1_m 2_m} < 0$.

Denote the eigenvalues that are less than 1 in absolute value as λ_i. These define the stable manifold. We seek the eigenvectors corresponding to each λ_i. These eigenvectors have $\boldsymbol{A}_t - \bar{\boldsymbol{A}} = \lambda_i(\boldsymbol{A}_{t-1} - \bar{\boldsymbol{A}})$ and thus are proportional to

$$\begin{bmatrix} \lambda(\boldsymbol{A}_{t-1} - \bar{\boldsymbol{A}}) \\ \boldsymbol{A}_{t-1} - \bar{\boldsymbol{A}} \end{bmatrix}.$$

Therefore, along the stable manifold,

$$\begin{bmatrix} \boldsymbol{A}_{t+1} - \bar{\boldsymbol{A}} \\ \boldsymbol{A}_t - \bar{\boldsymbol{A}} \end{bmatrix} = \sum_i c_i \lambda_i \begin{bmatrix} \lambda_i(\boldsymbol{A}_{t-1} - \bar{\boldsymbol{A}}) \\ \boldsymbol{A}_{t-1} - \bar{\boldsymbol{A}} \end{bmatrix}$$

with at least one $c_i \neq 0$. Because it must be true that $\boldsymbol{A}_t - \bar{\boldsymbol{A}} = \sum_i c_i \lambda_i (\boldsymbol{A}_{t-1} - \bar{\boldsymbol{A}})$ and also that $\boldsymbol{A}_t - \bar{\boldsymbol{A}} = \lambda_i(\boldsymbol{A}_t - \bar{\boldsymbol{A}})$ for each i, it must be true that one c_i equals 1 and the rest equal 0. I now assume that some such λ_i exists, as we know it must when $\pi_{1_j 1_k}$, $\pi_{1_j 2_k}$, and $\pi_{2_j 2_k}$ are small for all pairs (j, k) such that $j \neq k$. I label the λ_i corresponding to the $c_i = 1$ as λ.

Now consider optimal actions in the stochastic system. The first-order condition is:

$$\boldsymbol{0} = \pi_1(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}) + \beta E_t[V_1(\boldsymbol{A}_t, \boldsymbol{w}_{t+1}, \boldsymbol{f}_{t+1}, \boldsymbol{w}_t; \zeta)].$$

The envelope theorem yields:

$$V_1(\boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{f}_t, \boldsymbol{w}_{t-1}; \zeta) = \pi_2(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}).$$

Advancing this forward by one timestep and substituting into the first-order condition, we have the stochastic Euler equation:

$$\boldsymbol{0} = \pi_1(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1}) + \beta E_t[\pi_2(\boldsymbol{A}_{t+1}, \boldsymbol{A}_t, \boldsymbol{w}_{t+1}, \boldsymbol{w}_t)].$$

Extending Lemma 2 shows that, under its assumptions, $\lim_{t\to\infty} E_0[\boldsymbol{A}_t] = \bar{\boldsymbol{A}}$. And extending its proof shows that

$$\begin{aligned} \boldsymbol{A}_t =& \bar{\boldsymbol{A}} + \boldsymbol{\chi}_2^{-1} \bar{\pi}_{12}(\boldsymbol{A}_{t-1} - \bar{\boldsymbol{A}}) + \boldsymbol{\chi}_2^{-1} \left[\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{12}\boldsymbol{\chi}_1^{-1}\bar{\pi}_{14}\right](\boldsymbol{w}_t - \boldsymbol{C}) \\ &+ \boldsymbol{\chi}_2^{-1}\left[\beta\bar{\pi}_{23} + \beta\bar{\pi}_{12}\boldsymbol{\chi}_1^{-1}\left(\bar{\pi}_{13} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{12}\boldsymbol{\chi}_0^{-1}\bar{\pi}_{14}\right)\right](\boldsymbol{f}_t - \boldsymbol{C}) \\ &+ \boldsymbol{\chi}_2^{-1}\bar{\pi}_{14}(\boldsymbol{w}_{t-1} - \boldsymbol{C}), \end{aligned}$$

where each $\boldsymbol{\chi}_i$ is $M \times M$ and defined recursively:

$$\begin{aligned} \boldsymbol{\chi}_0 &\triangleq -\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{12}\lambda, \\ \boldsymbol{\chi}_i &\triangleq -\bar{\pi}_{11} - \beta\bar{\pi}_{22} - \beta\bar{\pi}_{12}\boldsymbol{\chi}_{i-1}^{-1}\bar{\pi}_{12} \quad \text{for } i \text{ a strictly positive integer.} \end{aligned}$$

If $\bar{\pi}_{1_j 1_k}$, $\bar{\pi}_{1_j 2_k}$, and $\bar{\pi}_{2_j 2_k}$ are small for all pairs (j,k) such that $j \neq k$, then the analogue of the condition imposed following Lemma 1 and the fact that $|\lambda| < 1$ together ensure that the absolute value of each element on the diagonal of $\boldsymbol{\chi}_i^{-1}\bar{\pi}_{12}$ is less than 1 (for i a weakly positive integer).

When the conditions of Lemma 2 hold, applying the implicit function theorem to equation (F-7) yields:

$$\lim_{t\to\infty} \frac{\mathrm{d}E_0[\boldsymbol{A}_t]}{\mathrm{d}\boldsymbol{C}} = \frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\boldsymbol{C}} = \left[-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}\right]^{-1}\left[\bar{\pi}_{13} + \bar{\pi}_{14} + \beta\bar{\pi}_{24} + \beta\bar{\pi}_{23}\right].$$

Following the main text, we then obtain:

$$\begin{aligned}\lim_{t\to\infty} \frac{\mathrm{d}E_0[\pi(\boldsymbol{A}_t, \boldsymbol{A}_{t-1}, \boldsymbol{w}_t, \boldsymbol{w}_{t-1})]}{\mathrm{d}\boldsymbol{C}} =& \bar{\pi}_3 + \bar{\pi}_4 + \left(\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\boldsymbol{C}}\right)^{\intercal}[\bar{\pi}_1 + \bar{\pi}_2] \\ =& \bar{\pi}_3 + \bar{\pi}_4 + \left(\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\boldsymbol{C}}\right)^{\intercal}\bar{\pi}_2(1-\beta).\end{aligned}$$

With these expressions and the expression for $\boldsymbol{A}_t$ in hand, it is easy to derive Proposition 1 and its underlying terms. Let either Assumption 1 or 2 hold, and let $(A_{j(t-1)} - \bar{A})^2$ be small for all observations. Consider the following regression:

$$A_{jt}^m = \alpha_j + \sum_{k=1}^{K}\left[\Gamma_1^k w_{jt}^k + \Gamma_2^k w_{j(t-1)}^k + \Gamma_3^k f_{jt}^k\right] + \sum_{m=1}^{M}\Gamma_4^m A_{j(t-1)}^m + \eta_{jt}.$$

Stack the coefficients of the M regressions in vectors $\boldsymbol{\Gamma}$. Then:

$$\begin{aligned}\hat{\boldsymbol{\Gamma}}_1^k =& \boldsymbol{\omega}\left[-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}\right]^{-1}\left[\bar{\pi}_{13_k} + \beta\bar{\pi}_{24_k} + \beta\bar{\pi}_{12}\boldsymbol{\chi}_1^{-1}\bar{\pi}_{14_k}\right], \\ \hat{\boldsymbol{\Gamma}}_2^k =& \boldsymbol{\omega}\left[-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}\right]^{-1}\bar{\pi}_{14_k}, \\ \hat{\boldsymbol{\Gamma}}_3^k =& \boldsymbol{\omega}\left[-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}\right]^{-1}\left[\beta\bar{\pi}_{23_k} + \beta\bar{\pi}_{12}\boldsymbol{\chi}_1^{-1}\left(\bar{\pi}_{13_k} + \beta\bar{\pi}_{24_k} + \beta\bar{\pi}_{12}\boldsymbol{\chi}_0^{-1}\bar{\pi}_{14_k}\right)\right],\end{aligned}$$

where each $\hat{\boldsymbol{\Gamma}}_i^k$ is $M \times 1$ and

$$\boldsymbol{\omega} \triangleq \boldsymbol{\chi}_2^{-1}\left[-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}\right] > 0.$$

If each $\bar{\pi}_{1_j 1_k}$, $\bar{\pi}_{1_j 2_k}$, and $\bar{\pi}_{2_j 2_k}$ is small for all pairs (j,k) such that $j \neq k$, then each diagonal element of $\boldsymbol{\omega}$ is > 1 if each $\bar{\pi}_{1_m 2_m} < 0$, $= 1$ if each $\bar{\pi}_{1_m 2_m} = 0$, and < 1 if each $\bar{\pi}_{1_m 2_m} > 0$. Then:

$$\begin{aligned}&\hat{\Gamma}_1^k + \hat{\Gamma}_2^k + \hat{\Gamma}_3^k \\ =& \boldsymbol{\omega}\left(\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\boldsymbol{C}^k} + \beta\left[-\bar{\pi}_{11} - (1+\beta)\bar{\pi}_{12} - \beta\bar{\pi}_{22}\right]^{-1}\bar{\pi}_{12}\boldsymbol{\chi}_1^{-1}\left[\bar{\pi}_{13_k} + \bar{\pi}_{14_k} + \beta\bar{\pi}_{24_k} + \beta\bar{\pi}_{12}\boldsymbol{\chi}_0^{-1}\bar{\pi}_{14_k}\right]\right).\end{aligned}$$

Now consider the regression:

$$\pi_{jt} = \alpha_j + \sum_{k=1}^{K} \left[\sum_{i=0}^{I+1} \theta^k_{w_{t-i}} w^k_{j(t-i)} + \sum_{i=0}^{I+1} \theta^k_{f_{t-i}} f^k_{j(t-i)} \right] + \eta_{jt}.$$

The vector of estimated coefficients is

$$\hat{\theta} = E[X_{I+1}^{\intercal} X_{I+1}]^{-1} E[X_{I+1}^{\intercal} \boldsymbol{\pi}],$$

where $\boldsymbol{\pi}$ is a $JT \times 1$ vector with rows π_{jt} and X_{I+1} is a $JT \times J + 2K(I+2)$ matrix with the final $2K(I+2)$ columns of each row being

$$\begin{bmatrix} w^1_{jt} & f^1_{jt} & \dots & w^K_{jt} & f^K_{jt} & \dots & w^1_{j(t-(I+1))} & f^1_{j(t-(I+1))} & \dots & w^K_{j(t-(I+1))} & f^K_{j(t-(I+1))} \end{bmatrix}.$$

By the Frisch-Waugh Theorem,

$$\hat{\theta} = E[\tilde{X}_{I+1}^{\intercal} \tilde{X}_{I+1}]^{-1} E[\tilde{X}_{I+1}^{\intercal} \tilde{\boldsymbol{\pi}}],$$

where $\tilde{X}_{I+1}$ is a $JT \times 2K(I+2)$ matrix with rows

$$\begin{bmatrix} w^1_{jt} - C^1 & f^1_{jt} - C^1 & \dots & w^K_{j(t-(I+1))} - C^K & f^K_{j(t-(I+1))} - C^K \end{bmatrix}$$

and $\tilde{\boldsymbol{\pi}}$ is demeaned $\boldsymbol{\pi}$. We can show by the Frisch-Waugh Theorem that controlling for all dimensions of weather means that we can analyze the regression dimension by dimension. Let $\hat{\theta}^k$ be the portion of the vector of coefficients corresponding to weather dimension k. Then:

$$\hat{\Phi}^k = E[(\tilde{X}^k)^{\intercal} \tilde{X}^k]^{-1} E[(\tilde{X}^k)^{\intercal} \tilde{\boldsymbol{\pi}}],$$

where $\tilde{X}^k$ is a $JT \times 2(I+2)$ matrix with rows

$$\begin{bmatrix} w^k_{jt} - C^k & f^k_{jt} - C^k & w^k_{j(t-1)} - C^k & f^k_{j(t-1)} - C^k & \dots & w^k_{j(t-(I+1))} - C^k & f^k_{j(t-(I+1))} - C^k \end{bmatrix}.$$

From here, we can follow the proof of Proposition 3 fairly directly. Analogous logic shows that the proof of Proposition 6 also goes through fairly directly.

References from the Appendix

Auffhammer, Maximilian (2018) "Quantifying economic damages from climate change," *Journal of Economic Perspectives*, Vol. 32, No. 4, pp. 33–52.

Bohrnstedt, George W. and Arthur S. Goldberger (1969) "On the exact covariance of products of random variables," *Journal of the American Statistical Association*, Vol. 64, No. 328, pp. 1439–1442.

Burke, Marshall and Kyle Emerick (2016) "Adaptation to climate change: Evidence from US agriculture," *American Economic Journal: Economic Policy*, Vol. 8, No. 3, pp. 106–140.

Cameron, A. Colin and Douglas L. Miller (2015) "A practitioner's guide to cluster-robust inference," *Journal of Human Resources*, Vol. 50, No. 2, pp. 317–372.

Carter, Colin, Xiaomeng Cui, Dalia Ghanem, and Pierre Mérel (2018) "Identifying the economic impacts of climate change on agriculture," *Annual Review of Resource Economics*, Vol. 10, No. 1, pp. 361–380.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken (2012) "Temperature shocks and economic growth: Evidence from the last half century," *American Economic Journal: Macroeconomics*, Vol. 4, No. 3, pp. 66–95.

——— (2014) "What do we learn from the weather? The new climate-economy literature," *Journal of Economic Literature*, Vol. 52, No. 3, pp. 740–798.

Deryugina, Tatyana and Solomon Hsiang (2017) "The marginal product of climate," Working Paper 24072, National Bureau of Economic Research.

Deschênes, Olivier and Michael Greenstone (2007) "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather," *American Economic Review*, Vol. 97, No. 1, pp. 354–385.

——— (2012) "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Reply," *American Economic Review*, Vol. 102, No. 7, pp. 3761–3773.

Fisher, Anthony C., W. Michael Hanemann, Michael J. Roberts, and Wolfram Schlenker (2012) "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Comment," *American Economic Review*, Vol. 102, No. 7, pp. 3749–3760.

Hu, G. Y. and R. F. O'Connell (1996) "Analytical inversion of symmetric tridiagonal matrices," *Journal of Physics A: Mathematical and General*, Vol. 29, No. 7, pp. 1511–1513.

Keane, Michael P. and Kenneth I. Wolpin (2002) "Estimating welfare effects consistent with forward-looking behavior. Part I: Lessons from a simulation exercise," *The Journal of Human Resources*, Vol. 37, No. 3, pp. 570–599.

Klemm, Toni and Renee A. McPherson (2017) "The development of seasonal climate forecasting for agricultural producers," *Agricultural and Forest Meteorology*, Vol. 232, pp. 384–399.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher (2005) "Will U.S. agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach," *American Economic Review*, Vol. 95, No. 1, pp. 395–406.

Schlenker, Wolfram and Michael J. Roberts (2009) "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change," *Proceedings of the National Academy of Sciences*, Vol. 106, No. 37, pp. 15594–15598.

Scott, Paul T. (2014) "Dynamic discrete choice estimation of agricultural land use," Working paper TSE-526, Toulouse School of Economics.

USDA (2000) "Farm Resource Regions," Agricultural Information Bullentin 760, U.S. Department of Agriculture Economic Research Service.