POVERTY

POVERTY AND EQUITY

EQUITABLE GROWTH, FINANCE & INSTITUTIONS INSIGHT

Background Note on Bringing Climate Change into Vulnerability Analysis

WORLD BANK GROUP

Sandra Baquié
Guillem Foucault

ABSTRACT

Weather vulnerability is often assessed using historical data, but this can be very misleading in a world of changing climate. Weather refers to short-term atmospheric conditions, while climate is the weather averaged over a long period. With climate change, some places are becoming wetter, some drier, and extreme weather events, such as heatwaves, floods, droughts, and tropical cyclones, are becoming more likely. Hence, the nature of weather risks will vary considerably. Despite the magnitude of this shift, there is currently no widely accepted method for bringing climate change into catastrophe risk modeling.

The objective of this note is to review, compare, and contrast the different techniques used in this literature to include climate change into vulnerability analysis. To do so, it summarizes recent research papers exploring how to bring climate change into catastrophe risk modeling. The note builds on this review to propose and explain a robust methodology and highlight its potential caveats. As such, this note is a first step towards unifying approaches and disseminating the analysis of climate change in vulnerability analysis. The method proposed in this note can be applied by researchers, economists, and public policy practitioners to study a wide range of topics, from the impact of climate change on diseases to stress-testing social protection programs.

Acknowledgements

The team is grateful to Benu Bidani (Practice Manager, POV), Ana Varela-Varela (Assistant Professor, University of Amsterdam), Ruth Hill (Lead Economist, POV), Alejandro De la Fuente (Senior Economist, POV), Oscar A. Ishizawa (Senior Disaster Risk Management Specialist, Urban), Rafael van der Borght (Consultant, Urban), and Joaquín Muñoz (Consultant, Urban) for their precious comments.

The corresponding author is Sandra Baquie (sbaquie@worldbank.org).

>>>

Overview of the Standard Methodology Assessing Weather Impacts

Estimating vulnerability to weather is critical to identify, ex-ante, the people who will likely need assistance in the aftermath of adverse weather or weather shocks. Policymakers can use this information to target horizontal expansions of social protection programs. The World Bank's Stress Test Tool (2021) provides a framework to assess the adaptiveness of social protection systems, in other words, their ability to respond to shocks, including extreme weather events. The first step is to evaluate the need by estimating how many people would suffer or fall into poverty in the aftermath of adverse weather. To do so, welfare is often proxied by consumption per capita, and we focus on this relationship in the rest of the document. However, the methodology described in this note can be applied to other dimensions of welfare, such as food insecurity (Blanchard et al., 2023) or health outcomes (Aguilar-Gomez et al., n.d.).

This note complements the standard methodology of vulnerability study by considering long-term changes in weather, i.e., climate change, to estimate long-term changes in needs. Although the World Bank work has mainly focused on analyzing the welfare impact of extreme weather events, particularly floods and droughts, this note considers a broader notion of weather for completeness and in agreement with the literature. In what follows, weather refers to all short-term variations[1] in atmospheric conditions, which encompassing both minor deviations to the historical mean and large ones-extreme weather events-. Atmospheric conditions can be measured by different weather indicators, such as monthly temperature, daily precipitation, and drought indexes.

The World Bank (2021)'s tool for stress testing social protection mentions three methods to assess needs in the aftermath of weather shocks: a scenario approach, a multilevel approach (Gao et al., 2020; Günther and Harttgen, 2009), and a simulation approach (Baquie and Fuje, 2020; Hill and Porter, 2017). The scenario approach focuses on analyzing selected shocks of a given magnitude, and the multilevel approach requires strong assumptions on the welfare distribution. In this note, we focus on the third method based on simulations, which avoids very strong assumptions regarding the distribution of weather shocks and welfare and is widely used in the recent literature. As explained earlier, we also broaden this method to study weather, beyond extreme weather events. **The current methodology has two steps:**

1. **Estimating the impact of weather on welfare.**

2. **Simulating the welfare distributions resulting from the observed historical distribution of weather.**

The first step relies on quantifying the relationship between welfare and weather. Most often, welfare is proxied by household consumption, and we focus on this relationship in what follows. Household consumption is often taken from household surveys and weather from reanalyzed data based on remote sensing imagery and weather stations. Matching one to the other requires having some spatial information in the household surveys, such as the household's GPS coordinates or district name. Then, one can regress log consumption on weather and use a specification similar to the one in (Hill and Porter, 2017):

$$\ln(C_{hkt}) = \beta X_{hkt} + \delta W_{kt} + \gamma W_{kt} \times X'_{hkt} + \alpha_h + \eta_t + \varepsilon_{hkt} \quad (1)$$

Where C_{hkt} is household h's consumption when living in district k at time t, X_{hkt} are household characteristics, and α_h and η_t are household and time fixed effects. W_{kt} is a vector of weather variables, that could potentially be complemented with variables describing idiosyncratic shocks at the household level. In this regression, δ and γ are the parameters of interest that quantify the impact of weather on household consumption. Results from this regression quantify the relationship between weather and consumption for the years included in the dataset, i.e., past years for which we have both weather and consumption data. Often, researchers do not have panel data because households are not interviewed over many years in nationally representative household surveys. In that case, researchers can use synthetic panel data. They would replace α_h by fixed effects at a more aggregated level, such as district fixed effects (α_k). Measurement error could be higher in this specification and bias the coefficient towards zero.

This methodology can also be applied to other dimensions of welfare, such as food consumption (Blanchard et al., 2023) or health outcomes (Aguilar-Gomez et al., n.d.). In this case, the variable on the left-hand side would not be consumption but a measure of the considered outcome. The more frequent and disaggregated the welfare data is, the more powered the regression is, and the more accurate the estimates. The choice of the weather variable(s) is also crucial in assessing the impact of weather on the outcome of interest. For instance, one may expect drought to impact crop yields significantly but to have a minor effect on infrastructure losses. Conducting a literature review is an essential first step toward selecting a set of weather variables that will likely be relevant to the considered outcome. As explained in Section 2.a, considering non-linearity in this specification is also central because the literature shows that weather usually has a non-linear effect on most outcome variables.

The second step simulates the distribution of consumption given the observed historical distribution of weather. It uses the estimated coefficients to calculate the households' consumption distribution in states of the world with alternative weather. Doing so allows us to quantify the distribution of consumption, including in years for which there is no measure of consumption. This step often relies on bootstrapping by drawing states of the world from the historical distribution of weather. In this case, the estimated consumption distribution mainly represents the impact of past weather, which may not describe future weather very well. Indeed, climate models show that the mean and variance of weather indicators are expected to shift (IPCC, 2022).

In what follows, we review the literature and explain the options to include climate change in the computation of future welfare distributions. We mention and discuss a few suggestions to tailor Step 1's specification to capture the impact of climate change on welfare. However, given the inherent tradeoff between identification and estimating long-term changes, we suggest mainly focusing on adopting a rigorous approach to the regression specification in Step 1 to capture the historical relationship as accurately as possible. We present some of the potential caveats in doing so in what follows. Then, we explain how to adapt the bootstrap approach in Step 2 to include climate projections and their uncertainty in the simulation of consumption distributions. In the conclusion, we explain the potential limitations of the suggested methodology. **Table 1 summarizes the recommendations proposed in this note and refers to their related sections.**

> > >

TABLE 1 - Summary of the Main Recommendations

Recommendation	Associated section
Step 1: Impact of weather on welfare	
• Prioritize the fixed effect specification as it captures causal effects and is easy to implement.	Section 1
• Avoid using the cross-sectional specification given the threat to the exogeneity restriction.	Section 1
• Keep in mind the subtlety of the difference between weather and climate and, if possible, test their results' robustness by implementing the long-differences specification and/or filtering weather and outcome data to isolate low frequencies.	Section 1
• Iterate on the choice of the specification and independent variables to capture non-linearity in weather extremes and improve the goodness of fit as much as possible. The quality of the results heavily depends on it.	Section 2a
• Check the robustness of the results by comparing different weather data sources. In any case, it is crucial to accurately document the data sources and the construction of the weather variable.	Section 2b
• Although there is no perfect specification to separate the direct effect of weather shocks from adaptation consequences, we suggest carrying robustness checks to shed light on this crucial distinction. Depending on data availability and the study's objective, the options range from evaluating the impact of weather on variables measuring adaptation, including interactions with variables reflecting adaptation mechanisms, or controlling for proxies of adaptation.	Section 2c
• Consider the heterogeneity of weather impact on welfare outcomes and delays and displacement.	Section 2d and 2e
Step 2: Predicting the impact of future weather by simulating consumption distributions resulting from the forecasted distribution of weather	
• Be extremely clear on the assumption one makes regarding uncertainty, particularly when considering model and regression uncertainty.	Section 3a
• If possible, use the projection of all models in the ensemble to quantify model uncertainty, even though this process may be more computationally intensive.	Section 3a
• Use the projections of the CMIP6 GCMs for their comparability.	Section 3b
• Use the most widely used approach in the Economics literature for projections. It derives the predicted changes between future years and a historical baseline for each GCM grid cell before adding it to the observed weather data in the historical baseline.	Section 3c1
• The weather variable selected in Step 1 should be the same as one predicted by the GCMs, or there needs to be a physics formula to calculate this variable from the set of variables available in the GCM data.	Section 3c2

>>>

Glossary

This note relies on the below definitions provided by the Intergovernmental Panel on Climate Change (IPCC, 2022) and Dell et al. (2014), except for the definition of vulnerability.

Adaptation: The process of adjustment to actual or expected climate and its effects, in order to moderate harm or exploit beneficial opportunities (IPCC, 2022)

Climate: Climate is usually defined as the average weather, or more rigorously, as the statistical description in terms of the mean and variability of relevant quantities over a period ranging from months to thousands or millions of years. The classical period for averaging these variables is 30 years, as defined by the World Meteorological Organization. The relevant quantities are most often surface variables such as temperature, precipitation, and wind (IPCC, 2022).

Climate model: A numerical representation of the climate system based on the physical, chemical and biological properties of its components, their interactions and *feedback* processes, and accounting for some of its known properties (IPCC, 2022).

Climate projection: A climate projection is the simulated response of the climate system to a scenario of future emission or concentration of greenhouse gases (GHGs) and aerosols, generally derived using climate models (IPCC, 2022).

Exposure: The presence of people; livelihoods; species or ecosystems; environmental functions, services, and resources; infrastructure; or economic, social, or cultural assets in places and settings that could be adversely affected (IPCC, 2022).

Extreme weather event: An extreme weather event is an event that is rare at a particular place and time of year. Definitions of rare vary, but an extreme weather event would normally be as rare as or rarer than the 10th or 90th percentile of a probability density function estimated from observations (IPCC, 2022).

Hazard: The potential occurrence of a natural or human-induced physical event or trend that may cause loss of life, health impacts, damage to property, infrastructure, livelihoods, service provision, and environmental resources (IPCC, 2022). This note focuses on weather-related hazards.

Risk: The potential for adverse consequences of a hazard on lives, livelihoods, health and well-being, ecosystems and species, economic, social and cultural assets, services, and infrastructure. Risk results from the interaction of vulnerability (of the affected system), its exposure over time (to the hazard), as well as the (weather-related) hazard and the likelihood of its occurrence (IPCC, 2022). This note focuses on the potential impact of hazards on welfare

Vulnerability: In this note, we use vulnerability and risk interchangeably, following a common abuse of terminology in the associated Economic literature. The exact definition of vulnerability is narrower and refers to the propensity or predisposition to be adversely affected, which encompasses sensitivity or susceptibility to harm and lack of capacity to cope and adapt (IPCC, 2022). In this note, we refer to the narrow definition of vulnerability as adaptive capacity.

Weather: The word climate is reserved for the distribution of outcomes, which may be summarized by averages over several decades, while weather describes a particular realization from that distribution and can provide substantial variability (Dell et al., 2014).

>>>

Contents

Figures

>>>

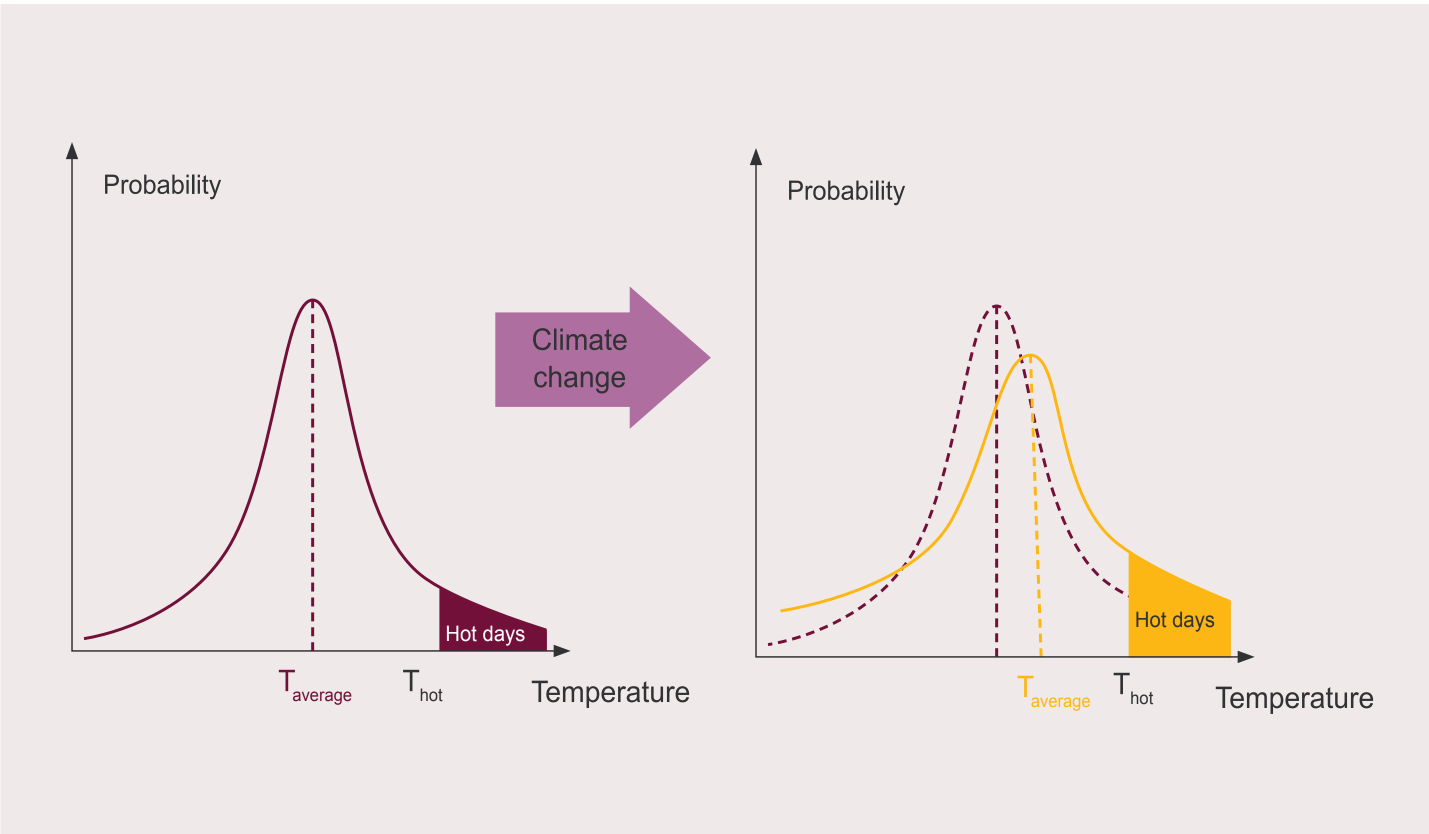
Step 1: Options to Directly Include Climate Change in the Estimation of the Impact of Weather on Welfare

Before considering some options to include climate change in Step 1's specification, let us explain the essential difference between weather and climate. These two concepts differ in the frequency of the observed changes. USGS defines it as follows: *"Weather refers to short-term atmospheric conditions while climate is the weather of a specific region averaged over a long period of time. Climate change refers to long-term changes"* (USGS, 2022). Based on this definition, monthly temperature or average temperature over a crop growing season describe weather patterns. Extreme weather events are also weather patterns since they are short-term changes in atmospheric conditions. Although the World Bank's work often focuses on extreme weather events, the literature has also applied this methodology to common weather conditions, such as the ones measured by average monthly temperature or precipitation. This note refers to both rare and common weather patterns under the term "weather".

The distribution of weather evolves with climate change, i.e., long-term changes in atmospheric conditions. Average temperatures are likely to increase in many countries due to global warming. The frequency and severity of extreme weather events are also expected to worsen because of climate change. For instance, the number of days with temperature exceeding 29°C during maize's growing season is a crucial predictor of yields. This indicator measures short-term (daily) changes in weather, but climate models also predict that, in many regions, the temperature will exceed the 29°C-degree threshold more and more frequently. Figure 1 illustrates the expected change in temperature distribution due to climate change. Not only does the average temperature increase, but the distribution's tails are also expected to be thicker. Consequently, the frequency and severity of extremely hot days ($T>T_{hot}$) are expected to increase. With this in mind, we clarify what can be estimated in Step 1 and how to do so in what follows.

> > >

FIGURE 1 - Simplified Illustration of the Impact of Climate Change on the Distribution of Temperature and Its Consequences for Weather (Average Temperature and Extreme Weather Events)



a. Research Design: The Inherent Tradeoff between Identification and Long-Run Changes

Step 1's specification (1) includes geographical fixed effects to control for temporally invariant omitted variables. They are often needed to estimate causal effects because one cannot control for any possible confounder. However, the statistical power in the specification including geographical fixed effects comes from weather deviations from the average. As a result, this specification estimates the response to short-term changes in climate, i.e., weather. This is why this specification estimates the impact of weather shocks or extreme weather events on consumption. To do so, W_{kt} is defined to capture significant deviations from the historical mean. However, this specification does not capture the impact of long-term changes in weather and their associated responses in terms of investment and adaptation (Kolstad and Moore, 2020).

When evaluating the impact of climate change, it is crucial to understand how the long-term changes in atmospheric conditions have impacted welfare. To consider the long-run changes in weather, i.e., the climate component of atmospheric patterns, one may be tempted to remove the fixed effects and use the following non-identified cross-sectional specification:

$$\ln(C_{hkt}) = \beta X_{hkt} + \delta W_{kt} + \gamma W_{kt} \times X'_{hkt} + \varepsilon_{hkt} \quad (2)$$

In equation (2), the identification of climate's impact, δ, no longer relies on weather's deviations from the mean because fixed effects are removed. As a result, this specification accounts for long-term changes in weather. However, the estimate of δ is biased because causality requires that, if two households face the same shocks, their expected consumption conditional on observables X_{hkt} is the same. This assumption is implausible in specification (2) due to potential confounders. In other words, the absence of fixed effects threatens the exogeneity restriction needed to recover causality. This is the inherent tradeoff between identification and long-run changes (Hsiang, 2016).

A middle ground between the well-identified specification (1) assessing weather's impact and the biased specification (2) considering climate change is the long-differences specification. It compromises the pros and cons of the cross-sectional and fixed-effect specifications by differentiating the data for two points far apart in time (often two to three decades). By doing so, specification (1) becomes:

$$\Delta \ln(C_{hk}) = \beta \Delta X_{hk} + \delta \Delta W_{k} + \gamma \Delta (W_{k} \times X'_{hk}) + U_{hk} \quad (3)$$

Where $\Delta \ln(C_{hk}) = \ln(C_{hk,2050})$ - $\ln(C_{hk,2020})$ for instance. In this specification, the dependent variable no longer captures yearly variation in weather, but long-term changes in weather, which is more in line with the definition of climate. Moreover, the assumption required to recover causality is weakened compared to specification (2). Here, causality requires that if two households face the same change in weather shocks, their expected *change* in consumption conditional on the change in observables ΔX_{hk} is the same. Nevertheless, the exogeneity assumption is stronger than it is in the fixed effect specification.

Given the pros and cons of all approaches, Hsiang (2016) suggests comparing the three estimators to test the results' robustness. Assuming that the identification of the three estimates stands (including the one of the cross-sectional specification), three equal estimates of δ means that the effect of short-term deviations from the mean is similar to that of long-term deviations. In other words, weather's impact is equal to climate's impact. This occurs if there is no adaptation, i.e., no belief effect. As a result, the difference in δ estimated with (1) and δ estimated with (2) or (3) can be interpreted as a measure of adaptation if the estimates of δ are unbiased. However, the equality of the estimates can also stem from the omitted variable bias compensating the adaptation effect, particularly when considering the likely biased specification (2). To shed light on this, Hsiang (2016) proposes an additional specification explained in the following section.

b. Isolating the Climate Change Component of Weather

Hsiang (2016) suggests a methodology expanding the scope of the three approaches presented in the previous subsection to further shed light on whether climate's impact equals weather's impact. We present this approach in what follows. However, although very promising, this technique is new and has not been widely applied in the literature at the time of writing.

As mentioned earlier, the difference between weather and climate lies in the periodicity of the observed changes, with climate change corresponding to long periods/low frequencies. The link between periodicity and frequencies has been studied intensively in signal Engineering. A simple mathematical transformation, the Fourier transformation, translates temporal data into a sum of sinusoids of given frequencies. Applying this formula to weather data gives the following:

$$W_{kt} = a_0^k + \sum_{\omega=1}^{\infty} [a_{\omega_k}^k \sin(\omega_k t) + b_{\omega_k}^k \cos(\omega_k t)] \quad (4)$$

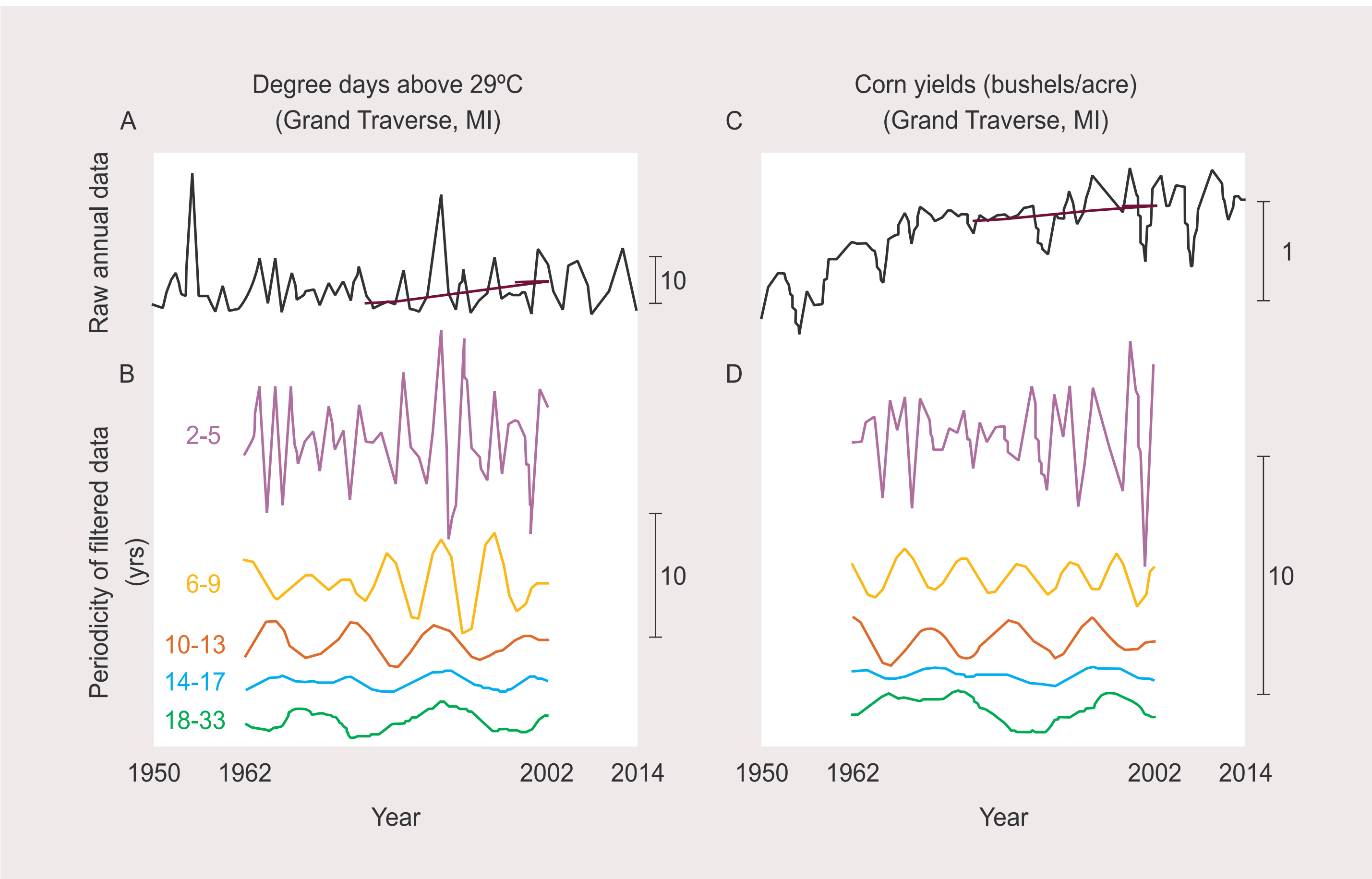
Where W_{kt} is the weather data time series, ω_k are the frequencies, and $a_{\omega_k}^k$ and $b_{\omega_k}^k$ are constants. For instance, a drought that occurs precisely every eight years would have only one term at a frequency equal to (1/8). More complex temporal patterns have more than one frequency. The Fourier transformation allows to filter the weather components according to their frequency. By doing so, one can isolate the high frequencies/rapid changes related to weather patterns from the long-term changes/low frequencies related to climate. Applying specification (2) to the resulting weather components is another way to find a middle ground between cross-sectional and fixed effect specifications.

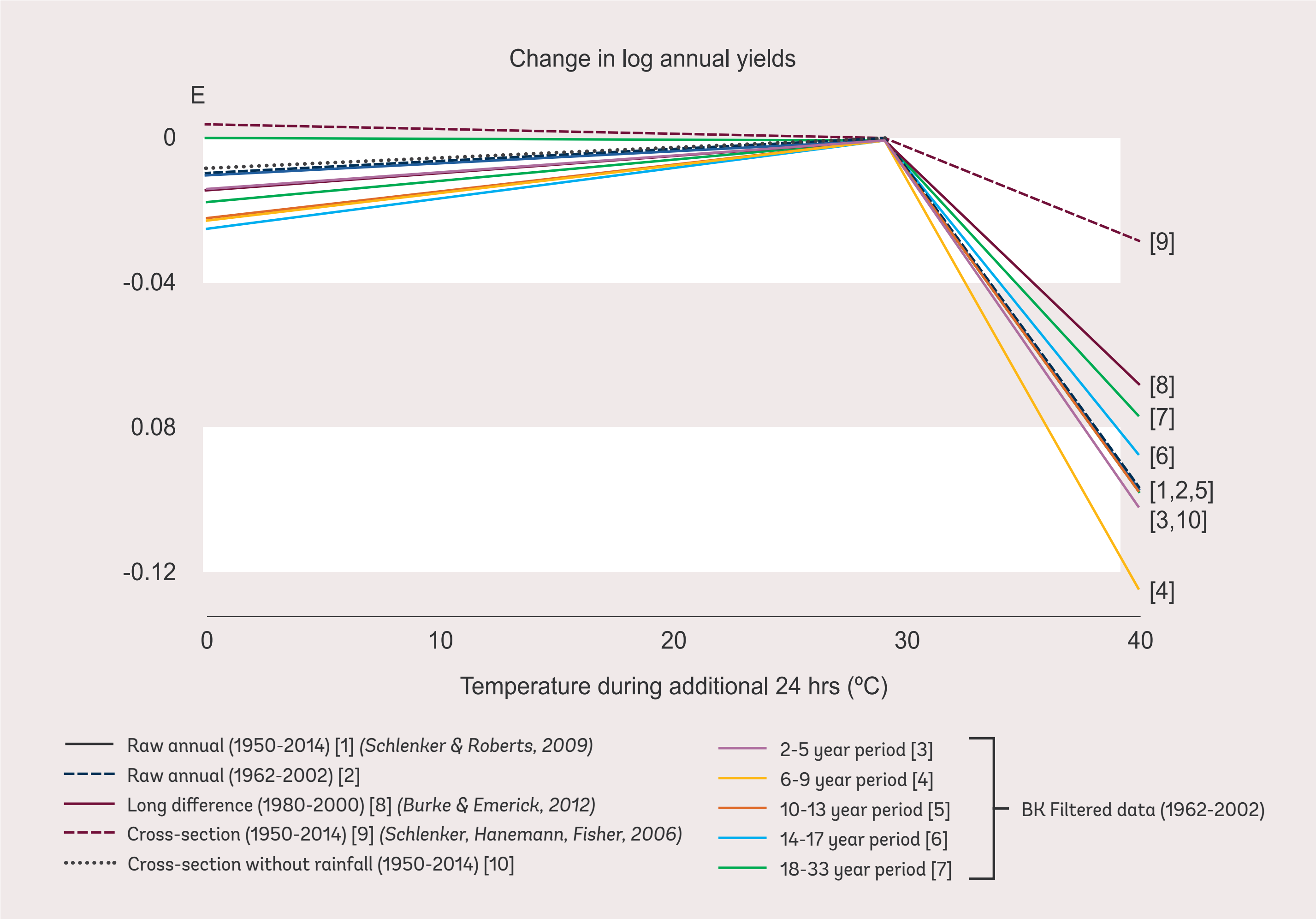
Figure 2 presents the results of this methodology used to estimate the impact of degree days above 29°C on crop yields (Hsiang, 2016). Panel (A) shows the independent variable, degree days above 29°C, and the filtered components of this weather data are presented in panel (B). Similarly, corn yield time series are presented in panel (C), and the filtered components of this output data are in panel (D). One can observe how the rapid changes (blue) are separated from the long-term changes (red). Then, the author compares all the specifications described above in panel (E): panel data applied to the raw data [1,2], cross-section [9,10], long-differences [8], and panel data with the filtered components [3,4,5,6,7]. He finds that the coefficients are not statistically different except for the cross-sectional specification. These results suggest that the effect of gradual changes may be similar to that of more rapid changes in this context. This result holds when considering other outcomes. For instance, the long-differences specification applied to the impact of climate on growth (Dell et al., 2012) and conflict (M. Burke et al., 2015) shows similar estimates as fixed-effects specifications.

Hsiang (2016) proposes a theoretical justification for the equality of the various estimates of weather and climate impacts. Under certain common conditions[2], the total effect of climate can be exactly recovered using the coefficient derived from weather variation in the panel data specification (1). Applying the Envelope Theorem and the Gradient Theorem, he shows that the marginal effect of the climate on an optimized outcome is the same as the marginal effect of the weather. This result relies on the fact that, under certain conditions, the marginal effect of adaptation and beliefs for marginal climate change is zero on an optimized outcome.

> > >

FIGURE 2 - Example of Time Series Filtered at Different Frequencies (A-B and C-D) And Comparison of the Different Specifications: Cross-Section, Long-Differences, and Panel Using Filtered Data and Raw Data (E)





Source: Hsiang (2016)

Note from Hsiang (2016): (A)-(D) Example outcome and climate time series data from Grand Traverse, Michigan filtered at different frequencies. (A) raw annual degree-days data (black) and 30-year long-difference (maroon) following Burke and Emerick (2016). (B) Same data decomposed into time series at different frequencies, where a Baxter-King band-pass filter has been applied for different periodicities. Filtering causes a loss of data at start and end of time series. (C) same as (A) but for corn yields. (D) Same as (B) but for corn yields. (E) Comparison of estimated effect of daily temperature using raw panel data sets, filtered data sets, long-differences, and cross-sectional approaches. Sample and estimation indicated by both line and bracketed numbers.

c. Takeaway

We recommend using the fixed effect specification described in Step 1 because it captures causal effects and is easy to implement. We also warn against using the cross-sectional specification given the threat to the exogeneity restriction. Finally, we encourage researchers to keep in mind the subtlety of the difference between weather and climate and test their results' robustness by implementing the long-differences specification and/or filtering weather and outcome data to isolate low frequencies. Agreement between the various estimates can be interpreted as minor adaptation to climate change. On the contrary, finding significantly different coefficients in well-identified specifications suggests a significant role of adaptation in response to climate change. We describe methods to investigate adaptation mechanisms in section 2c.

>>>

Accurately Estimating the Historical Impact of Weather on Welfare in Step 1

As mentioned in the previous section, we recommend using specification (1) in Step 1. When doing so, there are several potential caveats to remember to estimate the historical impact of weather on welfare accurately. In what follows, we focus on some of them and provide solutions to tackle them.

a. Assessing Non Linearities

Assuming that the effect of weather on consumption is linear is a strong assumption, and the literature has shown that this relationship is often non-linear. This is intuitive since an increase of one degree at a very high temperature is more likely to adversely impact yields, health, and consumption than a one-degree increase at a moderate temperature. The same reasoning applies to precipitation or wind. With climate change, tail events are expected to be more and more likely. As a result, **estimating the non-linearity of the impact of weather on welfare is essential to use the relationship estimated in specification (1) to forecast the potential impacts of climate change in Step 2.** Therefore, even when an analysis focuses on common weather patterns and not extreme events, it should describe these tail situations well enough to be relevant in the context of climate change. Similarly, studies on extreme weather events should not only capture mild disasters but also some with a severe magnitude to be useful in the context of climate change. **This is often a challenge since Step 1's estimation needs to be well specified to capture the tails, for which we have less data, relatively well.**

A non-linear relation should be captured by changing the functional form of weather to a nonlinear one (bins, splines, polynomials, or piecewise linear and stepwise functions). The first step is to regress consumption on bins of weather since it allows to be relatively agnostic on the form. Alternatively, weather indicators themselves can be transformed to capture non-linearities while maintaining the linearity in the regression equation. For example, while rainfall may improve welfare, excess rain can cause severe economic damage at an exponential rate. In this case, an appropriate indicator could include variables counting the number of days experiencing excess rainfall (Baquie and Fuje, 2020). Another classic example of such an indicator is the number of degree days above 29°C when studying crop yields (Schlenker and Roberts, 2009). Plotting the relationship between the selected climate variable and consumption is also helpful to visually investigate the relationship's shape. Stata users can easily create binned scatterplots to do so (with the *binscatter2* command, for instance). Nonlinearities can further be detected by plotting the association between the fitted values of the considered regression and its residuals. If the fit is good, residuals should not systematically depart from zero.

We encourage researchers to iterate on the choice of the specification and independent variables to capture non-linearity in weather extremes and improve the goodness of fit as much as possible. The quality of the results heavily depends on it.

b. Evaluating the Source of Weather Data

Second, researchers should carefully consider the weather data source since it may influence results due to measurement error. We explain why in what follows.

Gridded weather datasets are the most widely used data source in the literature. These datasets often combine various data inputs, such as weather stations and remote sensing imagery, to estimate weather in all grid cells. Estimations rely on two types of methods:

- **Grid interpolation** is a method dealing with the problem of missing time or station data by interpolating. The precision increases with the frequency of data points and the number of nearby weather stations.

- **Data assimilation** is an alternative method that combines observational data with a physics-based model to produce “reanalyses.” The model uses physical laws to fill in missing values.

Grid interpolation and data assimilation are powerful tools for dealing with missing data. As the sparsity of the data at hand increases, data assimilation should be favored as long-distance interpolation might miss the underlying physical laws.

To illustrate this point, Auffhammer et al. (2013) compare three traditional weather datasets: the CRU and UDEL datasets, relying on statistical interpolation, and the NCEP/NCAR reanalyzes. The authors show that although averaged weather indicators measures are all significantly correlated (>0.99 for temperature), one should be careful when studying deviations from the mean because the correlation across the datasets drops when considering deviations. When focusing on precipitation, Auffhammer et al. (2013) find a correlation of 0.70 between the grids using statistical interpolation and approximately 0.30 when comparing them to the reanalysis grid. Therefore, we should be careful about the data source when applying specification (1) because the inclusion of fixed effects may amplify measurement error and, in turn, attenuation bias.

If possible, we advise checking for the robustness of the results by comparing different weather data sources. In any case, it is crucial to accurately document the data sources and the construction of the weather variable.

c. Carefully Considering Adaptation in the Regression Specification

Researchers may want to carefully consider how adaptation enters the regression in Step 1. Indeed, Step 1 estimates the overall impact of weather on consumption, which includes the direct adverse impact of weather on consumption but also its mitigation through adaptation. For instance, farmers may diversify their crops in a bad weather year, protecting them from the adverse impact. Understanding the extent to which the estimate of δ reflects adaptation is crucial for policy implications. Indeed, it allows us to understand whether pre-existing safety nets and behaviors help mitigate the impact of the shock or not. As mentioned in Section 1b, accounting for adaptation is particularly important when the estimates from the panel and long-differences specifications differ.

The literature has attempted to disentangle the two effects with various approaches. **A first method directly evaluates the impact of weather on outcomes known to be adaptation actions**. For instance, Hornbeck (2009) study the effect of drought on the migration of agricultural households. In turn, one could estimate the impact of push migration on consumption to evaluate the increase in welfare attributable to this adaptation strategy[3]. **Alternatively, researchers can include interactions with variables that are indicators of the extent of adaptation in specification (1).** For instance, when studying the impact of drought on consumption, one may want to include the number of plots or access to irrigation as an interaction. Expert knowledge is often used to choose these interaction terms, and Stata users can easily couple it with a LASSO method to determine which one to include.

Another approach is to control for potential differences in adaptation in specification (1) to isolate the direct adverse effect of weather shocks. One may want to do so to simulate households' consumption distribution without safety nets and adaptive behaviors to identify the most vulnerable. If available, household fixed effects capture time-invariant confounders in specification (1), including household's adaptation to the climate where they live. However, when panel data is unavailable, researchers often use geographical fixed effects at a coarser level than the spatial unit of observation, such as district fixed effects for weather defined at the village level. In this case, the difference in adaptation across villages within the same district threatens the exogeneity restriction. For instance, in a given district, arid areas could be poorer and more frequently hit by droughts due to their geography. Controlling for the weather shock's historical distribution (mean or standard deviation) may partially address this concern. Indeed, the timing of weather shock is likely exogenous conditional on its probability distribution. Baquie and Fuje (2020) use this method while also building a measure of consumption that excludes transfers to assess the impact of adverse weather shocks in the absence of formal or informal safety nets.

Overall, particular attention should be given to the measure of consumption used in the regression before interpreting the coefficient. The consumption measures used to build national consumption aggregates in poverty measurement exercises rely on several assumptions regarding the inclusion or exclusion of expenses and how to treat missing values. For instance, exceptional costs such as hospitalizations are often excluded from the consumption aggregate to reflect consumption under the permanent income hypothesis. This assumption implies that the impact on consumption measured in Step 1 may be a lower bound of the effect on welfare. The reduction in the magnitude of the coefficient due to measurement error should not be mistaken to result from adaptation. Therefore, **we advise accurately documenting the measure of consumption when describing the results and performing sensitivity analysis when possible.**

Although there is no perfect specification to separate the direct effect of weather shocks from adaptation consequences, we suggest carrying robustness checks to shed light on this crucial distinction. Depending on data availability and the study's objective, the options range from evaluating the impact of weather on variables measuring adaptation, including interactions with variables reflecting adaptation mechanisms, or controlling for proxies of adaptation.

No matter the chosen specification, understanding adaptation is essential to assess whether Step 1's relationship will likely hold in the future. Indeed, since the relationship between weather shocks and welfare is estimated using historical data, it assumes that adaptation will stay constant, which may not hold in the long run. If we want to avoid making a theoretical assumption on the magnitude of future adaptation, we can restrict ourselves to medium-term climate forecasts (0-30 years). Otherwise, we can use the understanding developed by applying the methods in this subsection to make theoretical assumptions about future adaptation and run different scenarios of simulations in Step 2. Still, the longer the considered time horizon, the higher the uncertainty on the adaptation scenarios.

d. Considering the Heterogeneity of Weather Impact on Welfare Outcomes

Weather impacts on welfare may differ significantly across observables, including along dimensions related to adaptive capacity as explained above. If so, including an interaction term between the weather variable(s) and these observables could increase the predictive power of Step 1's model. For instance, the impact of drought on food consumption may be higher in rural areas if market access is lower and food imports are significant. In that case, including the interaction of weather with a dummy variable describing rural/urban could improve model's fit. However, one needs to avoid including too many interactions to prevent overfitting, and expert knowledge and/or machine learning are often used to decide which interactions to include (Blanchard et al., 2023; Hill and Porter, 2017).

When Step 1 includes an interaction term between the weather variable(s) and observable(s), additional assumptions are required in Step 2. We need to assume that the considered observables are exogenous and stay constant during the considered time horizon or make a theoretical assumption on their change over time. In the above example, we would need to assume an exogenous scenario of rural/urban migration and the resulting distribution to compute the expected measure of consumption. Similarly to adaptation action, the longer the considered time horizon, the higher the uncertainty on the scenario for the observables. As such, considering the heterogeneity of weather impact on welfare outcomes is likely to improve the predictive power of the model in Step 1, but it complexifies the implementation of Step 2[4].

e. Allowing for Delays and Displacement

We also recommend considering temporal and spatial dynamics in robustness checks when data allows. For instance, a drought at time t may have long-term consequences on consumption. If we have enough years of data, we may include lags in specification (1) to capture this persistence and potential intertemporal substitution. Falsification tests can also be run by including leads. Then, a climatic event's net effect is the sum of the lag terms. A similar idea can be applied to spatial displacement, where temporal lags are substituted with the average climate exposure of all locations at various distances of the considered household. If one worries that remote effects may be delayed, it is also possible to include both spatial and temporal lags.

3

Step 2: Predicting the Impact of Future Weather by Simulating Consumption Distributions Resulting from the Forecasted Distribution of Weather

As explained in the introduction, Step 2 uses the estimation of specification (1) to simulate households' consumption distribution given the distribution of weather. This methodology relies on bootstrapping by drawing states of the world from the assumed distribution of weather. For simplicity, most papers have assumed that the distribution of weather is equal to the observed historical distribution. In this section, **we extend this methodology to consider a distribution of weather equal to the forecasted distribution of weather in given climate scenarios (also called Shared Socioeconomic Pathways [SSPs]).** For instance, the forecasted weather distribution could be the forecast made by a climate model in a pessimistic scenario. We discuss sources of climate forecasts in what follows and potential caveats when calculating expected consumption with their projections.

a. Sources of Uncertainty

Let us introduce an essential concept when talking about climate projections: uncertainty. Indeed, forecasts are only helpful if they are somehow accurate. Uncertainty in climate models stems from three different sources: emission uncertainty, model uncertainty, and climate uncertainty (Burke et al., 2015). In what follows, we explain each source and how the scientific community assesses them.

First, **we do not have a perfect knowledge of the trajectory of all the variables that affect the climate system,** from greenhouse gas emissions to population growth. Their trends depend on many factors, including the chosen development path, future economic growth, and technological change. This source of uncertainty is called ***emission uncertainty.* It is tackled by assuming a given climate scenario -Shared Socioeconomic Path- when running climate models.** International scenarios have been developed to ensure the comparability of the models' results. SSPs range from a somewhat optimistic scenario aligned with the Paris agreement -SSP1-1.9- to a rather pessimistic one -SSP3-7.0-.

Second, ***model uncertainty* stems from discrepancies amongst researchers on modeling choices** about the underlying physical relationships and how to initialize the models. To make the models more comparable, international projects coordinate teams to run their models with common parameters. One such project is the Coupled Model Intercomparison Project, now in its 6th phase (CMIP6). Nevertheless, discrepancies exist across the different climate models composing the CMIP6 ensemble.

Many papers ignore model uncertainty by relying on the model ensemble's median or mean. However, using a small number of models or relying on the ensemble mean or median does not capture the full range of potential climate variations. As mentioned by Burke et al. (2015), proceeding like this make the *"findings seem more precise than they actually are, and as a result make them less credible among climate scientists and potentially misleading for policymakers."* As a result, **we recommend using the projection of all models in the ensemble, even though this process may be more computationally intensive.** In practice, we would draw states of the world from the forecasted weather distributions obtained using each CMIP6 model. **Doing so includes the assessment of model uncertainty in the bootstrap estimation of Step 2.** Although we could initially give equal weight to all models, one could adjust them in future work as they learn more about models' biases and their fit to the weather patterns of specific regions. **The second-best solution is to directly use the ensemble mean or median or a selected set of climate models predicting the lowest and biggest changes in weather. However, in this case, the author should state clearly that they ignore model uncertainty in their estimates.**

Third, ***climate uncertainty* reflects our imperfect knowledge of physical processes. There is no fix or estimate** for this source of uncertainty due to the impossibility of quantifying unknown processes and methods.

On top of the three sources of uncertainty associated with climate models, **regression uncertainty** stems from the finiteness of the sample and the impossibility of estimating population parameters exactly. The estimation of the uncertainty due to the estimation of the δ coefficient can be included in Step 2 as one would do to **estimate the coefficient's standard error by bootstrap** (Burke et al., 2015). In the simulations, one should draw states of the world defined by both the sampled projected weather and the sampled regression coefficient to evaluate the uncertainty of climate models and regression jointly.

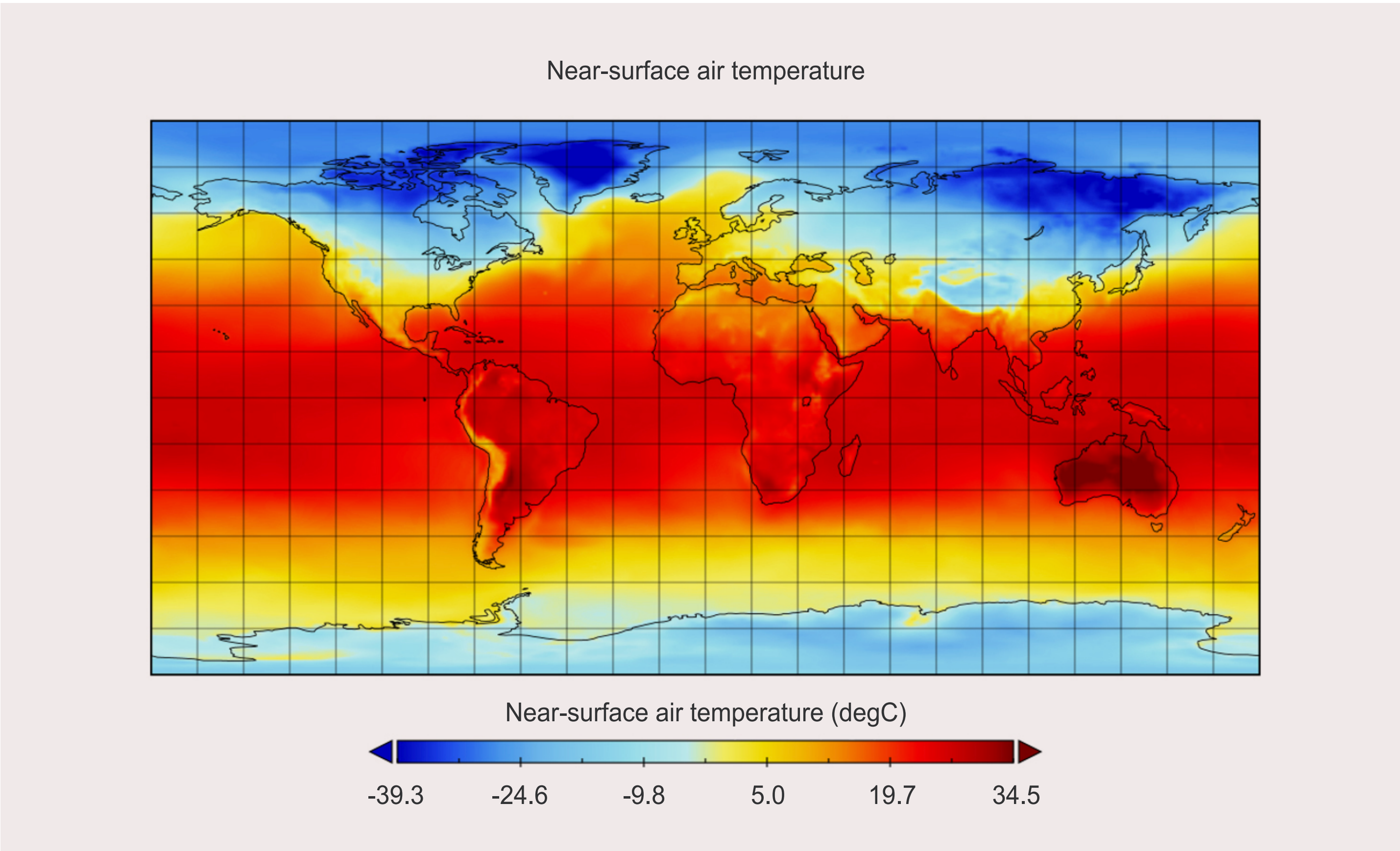
Overall, we advise being extremely clear on the assumption one makes regarding uncertainty, particularly when considering model and regression uncertainty. Reporting the resulting welfare metrics with confidence intervals is essential to communicate this uncertainty.

b. Sources of Data on Climate Change

The most widely used climate projections are the ones created by the 33 global climate models (GCM) included in the IPCC's CMIP6 coordination effort. These climate models are well-studied, comparable, and run for identical scenarios (SSPs). Their outcomes can be downloaded as NetCDF files defined for a given climate model, forecasted weather variable, region, and time-frequency. This practical file type is a 4D matrix with the following dimensions: longitude, latitude, weather, and time. It can be interpreted as a set of stacked maps representing the forecasted weather variable for each time step. Figure 3 presents the temperature forecast produced by the global climate model cams-csm-1-0 in the SSP1-1.9 climate scenario for January 1950. Ideally, the downloaded forecast describes the variable W_{kt} of Step 1 in future years. In some situations, W_{kt} is a transformation of the available forecasted variables. In this case, the researcher must calculate the forecast for W_{kt} from the available projected variables for each climate model, scenario, and time step.

> > >

FIGURE 3 - Global Temperature Forecast for January 1950 Produced by the Global Climate Model cams-csm-1-0 in the Climate Scenario SSP1-1.9



Source: World Bank Group (2022), Climate Change Knowledge Portal. File is plotted using the Panoply software.
Note: Data Min = -46.4, Max = 37.6, Mean = 13.2

The global climate models are numerical approximations of fluid's laws of motion computed by discretizing fluids such as the ocean and the atmosphere into three-dimensional grids which contain a given number of state variables. After initialization, the variables at t+1 are calculated using the fluids' law of motion, and the procedure is iterated to compute future states of the world over centuries (Auffhammer et al., 2013). Since climate change is a worldwide phenomenon, the grid must cover the whole world. As a result, it cannot have a very high spatial resolution for computational reasons.

To improve the spatial resolution, one could work with regional models, including CORDEX (Coordinated Regional Climate Downscaling Experiment). Often, regional models combine the results from GCMs with regional climate models to predict weather inside the GCM grid cells for given regions. Although these models have a better spatial resolution, they are less comparable and not widely used.

Therefore, we suggest using the projections of the CMIP6 GCMs for their comparability. They can be easily downloaded on the World Bank's Climate Change Knowledge Platform (CCKP). The CCKP team has harmonized the CMIP6 climate models' forecasts for various indicators and disaggregated them in 100*100km annual and monthly grids.

c. Aggregation Bias in Climate Data

Due to limits in computing power, global climate models can only make predictions on relatively coarse grids, usually about 2x2 degree cells or 200*200 km. As a result, climate may vary significantly within a cell and is often matched to more disaggregated data on consumption. This mismatch in spatial resolution results in measurement error, also called the "aggregation bias." This bias is more substantial when the grid cell topography is mountainous or near the ocean because climatic differences accentuate (Auffhammer et al., 2013).

Aggregation bias is problematic for studies assessing the impact of climate with the above methodology. Indeed, suppose that Step 1 is estimated with disaggregated observational data and that a GCM model's predictions are used for the counterfactual scenario in Step 2. In that case, the change in consumption forecasted by plugging the forecasted weather in the estimated equation could stem from the impact of climate change. However, it could also come from the GCM aggregation bias (Auffhammer et al., 2013). The following subsections explain how to deal with the aggregation bias.

c.1. Tackling the Aggregation Bias: Downscaling

The first solution, downscaling, uses the results from the global climate models (GCM) as an input to obtain more disaggregated climate projections. Two main downscaling techniques are used in the literature.

The first approach, that was mentioned above, is dynamical downscaling to increase spatial resolution. It extends the GCMs through regional climate models (RCM), such as CORDEX, or limited-area models (LAMs). These models are based on the same physically consistent processes but provide projections at a higher resolution, typically at the 0.5x0.5-degree resolution. They are computationally intensive, and their performance depends strongly on the biases inherited from the driving GCM and the presence and strength of regional forcings, such as orography, land-sea contrast, and vegetation cover. Over the last two decades, variable-resolution models have been developed, combining global-scale climate models with embedded regional grids. However, they have not been widely used in the literature since their computational burden is even higher than RCMs' requirements (Kotamarthi et al., 2021).

The second approach, statistical downscaling, can be implemented with methodologies ranging from simple models to artificial intelligence. Some of the most widely used ones are:

- Simple factors (Change Factor, or Delta Method): the differences between GCM historical and future projections are added to historical observations. This allows to include the climate change prediction of the model and removes the bias associated with the GCM. A second-level adjustment can be included to model changes in standard deviations.

- Regression: historical data is regressed on historical GCM predictions, and the estimated coefficients are used to predict future local weather from GCM predictions.

- Transforming the distributions by quantiles: this method focuses on estimating the extremes. Those models explicitly resolve, bias-correct, and downscale the quantiles of the distribution. One of the main approaches is Empirical Quantile Mapping. It consists in applying the Delta Method to each quantile.

- Neural Networks: conditional neural networks are increasingly used in the literature. They perform at least as well as other statistical downscaling methods and circumvent the problem of feature selection (Baño-Medina et al., 2021)

In general, those models are relatively easy to implement, but this approach does not include climate feedback, and their performance may strongly depend on the choice of predictors.

Downscaling approaches have not been discussed much in the Economics literature at the time of writing. The method that has been the most widely used so far in the literature is similar to the simple factor downscaling. We present it in the following subsection.

c.2. Recommended Solution to the Aggregation Bias: Applying Changes in Climate to the Historical Distribution

In the absence of downscaled data for the region and period of interest, the most widely used approach in the Economics literature is to derive the predicted changes between future years and a historical baseline for each GCM grid cell before adding it to the observed weather data in the historical baseline. This method, similar to simple factor downscaling, eliminates the location-specific bias and preserves the within-GCM grid variation of the historical data. Note that this approach only works if the aggregation bias is stationary in time and leaves the variance of the historical time series unchanged. If interested in adjusting the variance, one can rescale the variance of the historical data by the ratio of the variance of the GCM forecasts in future years relative to the variance of the GCM forecasts in the historical baseline (Auffhammer et al. 2013).

Finally, this method requires having the same indicator in the GCM data and the observed historical data. Therefore, before selecting the variable to include in the regression in Step 1, one should restrict oneself to indicators that can be calculated with the available GCM projection data. One may be tempted to regress the historical data indicator chosen in Step 1 on variables available in the GCM data to "estimate" their relationship and use the fitted values as projections of the historical data indicator. However, this is a false good idea. Indeed, if the available GCM indicators were not selected when working on the regression in Step 1, the linear combination of these same indicators will not capture the variation explaining the change in consumption either. The relevant source of variation will be in the residuals, which are not forecasted. Therefore, **the variable selected in Step 1 should be the same as one predicted by the GCMs, or there needs to be a physics formula to calculate this variable from the set of variables available in the GCM data.** This is crucial to applying this method in Step 2. The World Bank CCKP provides projections of a large set of weather variables, ranging from monthly temperature to cold spells to the SPEI drought index. In addition, the CCKP team can help calculate specific indicators with GCM data if these indicators rely on variables projected by GCMs.

>>>

Conclusion

In this note, we present the different techniques used in the literature to include climate change into vulnerability analysis and describe a robust methodology. Although we focus on applying this method to the impact of climate on consumption, it can be applied to study a wide range of topics, from the effects of climate change on diseases to stress-testing social protection programs. When doing so, researchers and practitioners should consider some limitations of the methodology proposed in this note.

First, the methodology presented in this note is a partial equilibrium analysis. As such, it misses factor reallocations across space or time that require a general-equilibrium approach. For instance, the identification strategy in this note's methodology assumes that the relationship between welfare and weather remains the same in the future. The longer the considered time horizon, the stronger this assumption becomes. If needed, one may relax the partial-equilibrium assumption by adopting a structural approach (Chetty, 2008), such as a Computable General Equilibrium (CGE) model associated with a microsimulations tool[5]. The method in this note and the CGE-microsimulations tool complement each other (The World Bank, n.d.). Indeed, due to the assumptions on market clearing, the general equilibrium approach is expected to describe long-term outcomes better, and its result typically represents a lower bound of effects. On the other hand, the partial equilibrium analysis described in this note is likely to perform better on short- and medium-term outcomes.

Second, tipping points and unprecedented events would not necessarily be well predicted in this type of analysis. The relationship estimated in Step 1 relies on historical data and is considered fixed over time. Like the discussion on adaptation, one would need to make theoretical assumptions about tipping points to infer how they would impact this relationship. Moreover, unprecedented events such as rising sea levels and other major climate events are not necessarily well predicted by climate models. Empirical progress on these questions is needed (Diffenbaugh et al., 2018).

Third, the bootstrapping exercise in Step 2 is relatively data and computationally intensive. One may want to contact IT to ensure that they have enough space on their drive and available memory to store all climate projections and run all the computations required for the project. In that respect, using a dedicated virtual machine or storing data on a server can help. One may also want to explore task parallelization to speed up the computations. The latter can be easily implemented in R or on a high-performance computing cluster.

>>>

Notes

1. Short-term is a relative notion, but in this context, weather usually refers to changes over periods that are shorter than a decade: year, month, or day.

2. The outcome is a solution to a maximization problem with an outcome-generating function depending on climate and adaptive actions. This function is continuous and differentiable in the space of all adaptive actions.

3. Migration is a coping mechanism in the aftermath of adverse shocks (push migration). However, other factors can drive migration decisions, including positive ones (pull migration). One must carefully design their identification strategy to isolate the component of interest.

4. This method becomes even more complex and computing-intensive if the chosen interaction term is not exogenous. For instance, if the observable at t depends on consumption at t-1. In this case, Step 2 could be implemented iteratively after assuming a relationship between consumption at t and the interaction term at t+1. Similarly, suppose the relationship in Step 1 is estimated with quantile regression. In this case, one may assume that consumption at t is informative of consumption at t+1 to assign a coefficient of weather's impact to each household. However, this assumption may not hold if the impact of weather is significant, and results may be very sensitive to the initial conditions. As a result, we recommend starting the analysis with the study of mean effects.

5. The World Bank uses several CGE models-MANAGE, MFMOD-to predict the impact of climate change on aggregate macro variables. The Equity and Policy Lab's microsimulations tool uses CGE models' predictions to estimate the distributional impacts resulting from their associated demographic changes and labor transition.

>>>

Bibliography

Aguilar-Gomez S, Baquie S and Robyn PJ (n.d.) Environmental hazards and health in Cambodia. *Cambodia CCDR Background paper.*

Auffhammer M, Hsiang SM, Schlenker W, et al. (2013) Using Weather Data and Climate Model Output in Economic Analyses of Climate Change. *Review of Environmental Economics and Policy* 7(2). The University of Chicago Press: 181–198.

Baño-Medina J, Manzanas R and Gutiérrez JM (2021) On the suitability of deep convolutional neural networks for continental-wide downscaling of climate change projections. *Climate Dynamics* 57(11): 2941–2951.

Baquie S and Fuje H (2020) Vulnerability to Poverty Following Extreme Weather Events in Malawi. World Bank, Washington, DC. DOI: 10.1596/1813-9450-9435.

Blanchard PB, Ishizawa Escudero OA, Humbert T, et al. (2023) *Struggling with the Rain : Weather Variability and Food Insecurity Forecasting in Mauritania.* 10457, 30 May. The World Bank. Available at: https://ideas.repec.org/p/wbk/wbrwps/10457.html (accessed 18 July 2023).

Burke M and Emerick K (2016) Adaptation to Climate Change: Evidence from US Agriculture. *American Economic Journal: Economic Policy* 8(3): 106–140.

Burke M., Hsiang SM and Miguel E (2015) Climate and conflict. *Annual review of economics.* annualreviews.org. Available at: https://www.annualreviews.org/doi/abs/10.1146/annurev-economics-080614-115430.

Burke M., Dykema J, Lobell DB (2015) Incorporating climate uncertainty into estimates of climate change impacts. *The review of economics and statistics* 97(2). MIT Press - Journals: 461–471.

Chetty R (2008) Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. *https://www.nber.org › papershttps://www.nber.org › papers.* Working Paper Series 14399. National Bureau of Economic Research. DOI: 10.3386/w14399.

Dell M, Jones BF and Olken BA (2012) Temperature Shocks and Economic Growth: Evidence from the Last Half Century. *American Economic Journal: Macroeconomics* 4(3): 66–95.

Dell M, Jones BF and Olken BA (2014) What Do We Learn from the Weather? The New Climate-Economy Literature. *Journal of economic literature* 52(3): 740–798.

Diffenbaugh NS, Singh D and Mankin JS (2018) Unprecedented climate events: Historical changes, aspirational targets, and national commitments. *Science advances* 4(2): eaao3354.

Gao J, Vinha KP and Skoufias E (2020) World Bank Equity Policy Lab (EPL) Vulnerability Tool to Measure Poverty Risk. *Washington, D.C. World Bank Group.*

Günther I and Harttgen K (2009) Estimating Households Vulnerability to Idiosyncratic and Covariate Shocks: A Novel Method Applied in Madagascar. *World development* 37(7): 1222–1234.

Hill RV and Porter C (2017) Vulnerability to Drought and Food Price Shocks: Evidence from Ethiopia. *World development* 96: 65–77.

Hornbeck R (2009) The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe. *The American economic review* 102(15605). American Economic Association. DOI: 10.1257/aer.102.4.1477.

Hsiang S (2016) Climate Econometrics. *Annual Review of Resource Economics* 8(1). Annual Reviews: 43–75.

IPCC (2022) *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, UK and New York, NY, USA: IPCC.

Kolstad CD and Moore FC (2020) Estimating the Economic Impacts of Climate Change Using Weather Observations. *Review of Environmental Economics and Policy* 14(1). The University of Chicago Press: 1–24.

Kotamarthi, R., Hayhoe, K., Mearns, L. O., Wuebbles, D., Jacobs, J., & Jurado, J (2021) *Downscaling Techniques for High-Resolution Climate Projections: From Global Change to Local Impacts.* Cambridge University Press.

Schlenker W and Roberts MJ (2009) Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences of the United States of America* 106(37): 15594–15598.

The World Bank (2021) *Stress Testing Social Protection: A Rapid Appraisal of the Adaptability of Social Protection Systems and Their Readiness to Scale-Up.* The World Bank. Available at: https://documents1.worldbank.org/curated/en/559321634917529231/pdf/Stress-Testing-Social-Protection-A-Rapid-Appraisal-of-the-Adaptability-of-Social-Protection-Systems-and-Their-Readiness-to-Scale-Up-A-Guide-for-Practitioners.pdf.

The World Bank (n.d.) *Cambodia Country Climate and Development Report (CCDR).* The World Bank.

USGS (2022) What is the difference between weather and climate change? Available at: https://www.usgs.gov/faqs/what-difference-between-weather-and-climate-change (accessed 14 July 2023).


TP. BẮC GIANG
0204 3854 538

WORLD BANK GROUP