Ecological Economics

Weather, Climate and Economic Outcomes: Evidence from Italy

--Manuscript Draft--

Manuscript Number:	ECOLEC_2020_890R4
Article Type:	Methodological and Ideological Options
Section/Category:	Analysis
Keywords:	Weather; Climate; Economic impact; Agriculture; Panel econometrics
Corresponding Author:	Alessandro Olper, PhD Milano, ITALY
First Author:	Alessandro Olper, PhD
Order of Authors:	Alessandro Olper, PhD
	Veronica Manara, PhD
	Maurizio Maugeri, PhD
	Valentina Raimondi, PhD
Abstract:	The economic analysis of climate change requires estimates at both aggregate and local level. Although there are numerous studies that estimate the global impact of climate change, country-level studies are still rare, particularly as far as Italy is concerned. By exploiting a panel of 110 provinces observed between 1980 and 2014, this paper investigates the impact of variation of weather variables on GDP per capita and agricultural productivity in Italy. To address issues of model uncertainty, the analysis explores to what extent these economic outcomes are affected by weather variables, linearly or non-linearly, as well as how their growth rate and levels are affected as a result. Main findings show that there is considerable model uncertainty. The most robust econometric results showing statistically significant effects of temperature for both the GDP and agricultural reaction function are from a levels specification, where temperature is included in a non-linear form while weather variables and economic outcomes enter in first differences. Projections of the impact of climate change by the end of the century show slight average effects for GDP per-capita, but important losses in agriculture, due to a persistent increase in average temperatures under both the RCPs 4.5 and 8.5.
Suggested Reviewers:	
Opposed Reviewers:	

Weather, Climate and Economic Outcomes: Evidence from Italy

Abstract. The economic analysis of climate change requires estimates at both aggregate and local level. Although there are numerous studies that estimate the global impact of climate change, country-level studies are still rare, particularly as far as Italy is concerned. By exploiting a panel of 110 provinces observed between 1980 and 2014, this paper investigates the impact of variation of weather variables on GDP per capita and agricultural productivity in Italy. To address issues of model uncertainty, the analysis explores to what extent these economic outcomes are affected by weather variables, linearly or non-linearly, as well as how their growth rate and levels are affected as a result. Main findings show that there is considerable model uncertainty. The most robust econometric results showing statistically significant effects of temperature for both the GDP and agricultural reaction function are from a levels specification, where temperature is included in a non-linear form while weather variables and economic outcomes enter in first differences. Projections of the impact of climate change by the end of the century show slight average effects for GDP per-capita, but important losses in agriculture, due to a persistent increase in average temperatures under both the RCPs 4.5 and 8.5.

JEL: O13, Q51, Q54
Keywords: Weather, Climate, Economic impact, Agriculture, Panel econometrics.

1. Introduction

An enhanced understanding of the relationship between climate and economic outcomes is crucial when it comes to projecting the impacts of global warming, as policy makers need reliable information on which to base their response to climate change (Newell et al. 2021). In the last decade, a growing body of econometric evidence linking random variation of weather variables to economic outcomes has been amassed both at the country and global level (see Dell et al. 2014 and Kolstad and Moore, 2020, for recent surveys). One important innovation of the emerging literature is the use of modern panel data econometrics (Deschênes and Greenstone 2007; Dell at al. 2014; Hsiang, 2016; Blanc and Schlenker 2017; Auffhammer 2018) together with short-run (i.e. inter-annual) variation of weather variables, so as to better isolate the (causal) temperature-productivity relationship[1]. Much of this literature focuses on the global relationship between weather variables and economic aggregates, such as Gross Domestic Product (GDP) per capita (e.g. Dell et al. 2012; Burke et al. 2015; Pretis et al. 2018; Dasgupta et al. 2017; Newell et al. 2020) or total factor productivity (Letta and Tol, 2019). These country-global estimates are essential elements when building aggregate damage functions, which are central in the evaluation of mitigation policies, such as the social cost of carbon (SSC). However, they also have certain drawbacks.

First, as recently showed by Newell et al. (2021), the lack of a clear theoretical foundation around the weather-economic relationship gives researchers plenty of room when deciding how to fit the model by assuming different functional forms of the relationship between climate and economic outcomes. This, however, can be rather problematic, as the predicted economic damage of future climate change is extremely sensitive to the underlying model specification. Second, global studies based on fixed effects models may suffer from measurement error in weather variables, particularly when developing countries are concerned, so that the predicted economic impact of climate change could be biased downward (Auffhammer and Schlenker 2014)[2]. Third, at country level, if temperatures are averaged over different areas/regions, crucial information concerning temperature anomalies may be lost in the process. That is to say, different productive units within a country could be exposed to opposing temperature shocks within a given year, particularly when a country is large and has more than one climatic zone. Thus, aggregation at country level could introduce uncertainty into the estimates and change the sign of the true weather effect (see Burke and Tanutama, 2019)[3]. Finally, given the wide income variation within countries, a cross-country study would give only a partial snapshot of the distributional effects of global warming at country level, and this might not fare well for Italy, which has considerable income differentials across population subgroups.

Building on recent contributions estimating climate impacts at global level (e.g. Dell et al. 2012; Burke et al. 2015; Kahn et al. 2019; Newell et al. 2021), this paper studies the impact of annual

[1] The panel model approach is significantly different from the so-called Ricardian model, often used to evaluate the impact of climate change on agriculture. Indeed, this hedonic approach exploits cross-sectional variability to estimate long-run (e.g. 30 years) climate impacts on economic outcomes (see Mendelsohn et al. 1994). For an in-depth discussion of the pros and cons of the panel vs. the cross-sectional approaches, see Dell et al. (2014) and Kolstad and Moore (2020).

[2] As the quality of weather data tends to be negatively correlated with the country income level, the attenuation bias induced by errors in weather variables will be more or less important depending on how many "poor" countries are involved in the dataset.

[3] See Conte et al. (2020) for a recent analysis of changing economic geography and sectoral specialization due to climate change, using a high-level spatial resolution.

variation of weather variables on GDP per-capita and agricultural gross value added (GVA) per-worker in Italy, exploiting information at NUTS3 level from year 1980 to year 2014.

Italy makes for an interesting study for a number of reasons. First, it is characterized by notable heterogeneity in climatic conditions from North to South (Cubasch et al. 1996). Second, there are clear differences in the level of development between North and South, and such information is needed when investigating the extent to which climate change disproportionately affects "poor" areas, as recently highlighted by several global studies (see Dell et al. 2012; Burke et al. 2015, 2018; Letta and Tol 2019). To a certain extent, therefore, Italy provides grounds to replicate the conditions implied by a global study and, most importantly, it requires working with a higher level of spatial granularity, at least vis-à-vis global studies.

We applied modern panel data econometrics to contrast several prominent empirical specifications used by the recent (global) literature. Our main findings confirm that selecting the "best performing" model in out-of-sample prediction through a cross validation (CV) exercise is problematic (see Newell et al. 2021). This is because, although non-linear level models are often slightly superior in terms of out-of-sample prediction, across-model differences are never statistically significant. Econometrically, the model that performs best in terms of in-sample statistical tests on weather variables is a non-linear (quadratic) in temperature *levels* specification, with economic outcomes and weather variables entering in first differences. Our projection for the year 2100 suggests that a persistent increase in average global temperature, under both the RCP 4.5 and the worst-case RCP 8.5, induces a small and insignificant effect on GDP per capita at aggregated level, but significant losses within the range of 10%-30% on agricultural GVA per-worker. At geographical level, all these effects are heterogeneous, with a clear North-South pattern.

The paper contributes to the literature along several lines. First, to our knowledge, it represents the first Italian application when it comes to panel data econometrics exploiting inter-annual variation of weather variables to infer the economic impacts of climate at NUTS3 level. Second, it provides a useful empirical application to test recent hypotheses about the impact of the variation in weather variables on per-capita GDP and agricultural productivity. Consistent with Newell et al. (2021) and Avecedo et al. (2020), our results support the notion that in general temperature shocks have non-linear *level* effects in Italy. These findings have important implications for the Italian economy, as they predict a significantly *lower* impact from global warming than estimates in recent studies.

This paper also relates to the few studies that estimate the economic impact of climate change on Italian agriculture using Ricardian models and farm-level data (see Van Passel et al. 2017; Bozzola et al. 2018). Our results are significantly less pessimistic than the findings of Van Passel et al.

(2017), and our end-of-century projections are quantitatively similar to those of Bozzola et al. (2018) [4].

The paper is organized as follows: in Section 2 we discuss the empirical approach, the model specifications analyzed, and the statistical tests put forward to assess their performances, concluding with a short discussion on the suitability of the non-linear in weather variables fixed effects model to infer the short-run vs. long-run impact of climate. Section 3 discusses the economic and weather data used in the estimations, while Section 4 first summarizes results of a cross validation exercise finalized to test the models' reliability in out-of-sample predictions, and subsequently discusses the main econometric results. Finally, Section 5 reports future projections, and Section 6 concludes the paper.

2. Method and econometric specification

In the following, we summarize the main approaches applied in the last decade to identify the impact of climate change using panel data methods and inter-annual variation of weather variables. Next, we briefly discuss the extent to which the use of the inter-annual variation in weather variables can help us identify the effect of climate.

2.1 Model specifications

In the last few years there has been an extensive use of panel data methods to estimate the relationship between inter-annual variation of weather variables and different economic outcomes, such as per-capita GDP, agricultural productivity, migration and health outcomes.

This literature uses variants of the following baseline model:

$$(1) \qquad y_{it} = \beta X_{it} + \alpha_i + \theta_t + h_i(t) + \varepsilon_{it},$$

where y_{it} is the economic outcome of interest in year *t* of the spatial unit *i* (e.g. a country, a county or a region), X_{it} is a matrix of realizations of weather variables, α_i and θ_t are spatial unit and year fixed effects, and $h_i(t)$ is a flexible linear or non-linear time trend term, specific to each unit of observation. Finally, ε_{it} is an error term.

Thus, equation (1) exploits the within-unit variation in temperature and precipitation (the two main weather variables in the regression model), after controlling for (time-invariant) heterogeneity in the units of observation, common time shocks and, eventually, unit specific time trends.

[4] See De Salvo at al. (2013), for an application to permanent crops in an Alpine region of Italy. An interesting and related agricultural productivity study by Ortiz-Bobea at al. (2018), showed that technological change and the pattern of regional specialization, led to growing climatic sensitivity in US agriculture.

Equation (1) is normally estimated without adding other controls apart from fixed effects and weather variables and, eventually, their higher order terms. The justification of this parsimonious choice is based on the idea that many traditional, time-varying determinants of growth and productivity – i.e. investments – are themselves endogenous to weather variation. Hence, variables that are endogenous to weather variation could mask, or partially offset, the true weather effects. This is the so-called over-controlling problem first stressed by Dell et al. (2014) in the econometric literature on climate[5]. Thus, we adopt the same parsimonious specification, also for comparability with previous literature,.

A key problem in applying equation (1) is that we do not have a dominant theory suggesting how weather variables should affect economic outcomes (see Dell et al. 2014; Newell et al. 2021). For example, temperature could affect the productivity level or growth rate (or both), and this could happen in a linear or a non-linear way (see Carleton and Hsiang, 2016 and Letta and Tol, 2019 for contrasting evidence). Consider, for example, agriculture. We know that temperature shocks above 29°C, during the growing season, may reduce maize yield (Schlenker and Robert, 2009). This is a standard (non-linear) level effect because, if the following year the temperature shocks disappear, then the normal yield level will be fully recovered, *ceteris paribus*. Now, take the case of a drought for a rainfed crop. If the reduction in rainfall affects the overall availability of water reserves, then the negative yield effect could persist in the following year, even if precipitations return to their normal value. This is an example of a growth effect that could persist over time.

The lack of a consolidated theory means that the true weather-economic relationship tends to become an empirical question, which is of crucial importance because model specification matters greatly for the long-run projections of climate damage. For example, Burke et al. (2015), using a specification where GDP growth is regressed on a quadratic function of level temperature (and precipitations), projected a 2100 GDP drop of about of 20-40% under RCP 8.5. By contrast, Newell et al (2021), using a specification in levels where both per-capita GDP and temperature enter in first differences, projected a 2100 GDP per-capita drop in the range of 1-3% under the same emission scenario. Similarly, Ricke et al. (2018) showed that the social cost of carbon (SCC) is extremely sensitive to the specification of the underlying damage function. These projection differences are extraordinarily large, rendering their use for policy evaluations problematic.

Current literature has proposed several different variants of equation (1). Starting from the dependent variable, the most important differences are between *specification in level* (e.g.

[5] One possible concern might be weather variables that are omitted and correlated with the included ones. This would be a problem particularly if the correlation between the included and omitted weather variables was predicted to change with climate change, because then the included variables would act as a poor proxy. We would like to thank one of the referees for drawing our attention to this problem. However, for comparability with previous literature we focus our attention only on temperature and precipitations variables.

Deryugina and Hsiang, 2017; Dasgupta et al. 2017, and many others) and *specification in growth rate* (e.g. Burke et al. 2015; Dell et al. 2012; Pretis et al. 2018). Considering weather variables (temperature and precipitation), there are several alternatives: linear in levels (Dell et al. 2012); non-linear in levels (Burke et al. 2015; Pretis et al. 2018); non-linear in first differences (Newell et al, 2021); non-linear in positive and negative deviation from the long-run climate (Kahn et al. 2019). Finally, differences also emerge in relation to how unobserved time heterogeneities are modelled, such as time fixed effects (Deryugina and Hsinag, 2017), regional-time fixed effects (Dell et al. 2012), linear and/or non-linear unit-specific time trends (Burke et al. 2015), and more recently autoregressive distributed lag models (Kahn et al. 2019).

In this study, we compare different specifications from the recent literature, to find out which model best explains the impact of weather variables on the Italian economy, focusing on both GDP per-capita and GVA per-worker in agriculture.

We tested the following models. First, a linear specification proposed in the seminal paper of Dell, Jones and Olken (2012) where the growth of an economic outcome is regressed on the level of weather variables (growth-level model), and they enter linearly and/or interacted with a "poor dummy" variable. The poor dummy, defined in the first year of the panel (1980), is equal to 1 (0 otherwise) for NUTS3 provinces with a per-capita GDP lower than the median of the sample distribution[6]. We labelled these specifications DJO1 when we assume that the impact of weather on economic outcomes is homogeneous across provinces, and DJO2 when it is assumed to be heterogeneous with respect to income, and thus the weather variables also interact with the "poor" dummy.

Second, by using a growth model close to the specification of Burke, Hsiang and Miguel (2015), where the weather variables in level enter non-linearly in a quadratic fashion, we assume that temperature in level (as well as precipitation) affects the growth rate of economic outcomes non-linearly. This specification will be labelled BHM.

Third, a specification where both the economic outcomes and weather variables enter in first-differences, so that the non-stationary nature of both GDP per-capita (or GVA per-worker) and the temperature series can be fully accounted for. This model closely mimics one of the preferred specifications of Newell, Preston and Sexton (2021), and will hence be labelled as NPS. Note that, although the dependent variable enters the NPS specification as a growth rate, we still capture level, and not growth, effects.

[6] As for the classification of the "poor" provinces, we run robustness checks by considering alternative definitions, by using GDP per-capita in 1995 or period average GDP per-capita (instead of 1980). Overall, results are similar.

All the aforementioned specifications include province and time fixed effects and are estimated both with and without province-specific time trends[7]. Although these specifications can eventually include lags of weather variables, they are static in nature. As a result, we opted for the dynamic panel model recently proposed by Kahn et al. (2019), labelled as KAHN1. Their specification is based on an autoregressive distributed lags (ARDL) model and has several useful econometric features, such as a novel way to account for non-linearity, weather variables entering in deviation from the (long-run) climate, and the possibility to measure long-run effects[8]. Following Kahn et al (2019), we also considered a specification where the weather variables effect is heterogeneous with respect to income, and thus weather variables enter also interacted with the "poor" dummy (labelled as KAHN2). Table 1 below summarizes the characteristics of such models.

For insight into which model specification best fits the data, we apply different econometric tools for model selection, from standard in-sample statistics to out-of-sample predictive performances obtained from different cross validation (CV) exercises, as recently proposed by Newell et al. (2021). This is important because, as highlighted by Hsiang (2016), climate econometric models are often used to inform projections under different climate scenarios.

Cross validation is a simple and powerful (non-parametric) statistical tool, where the available data are split into a training set and a test set. Each model is estimated in the training set, and then evaluated in terms of its predictive accuracy against the test set. We compare out-of-sample forecast accuracy, using three different CV methods, namely forecast CV, backcast CV and k-fold CV. The "best performing" model in CV is the one that minimizes root-mean-square errors (RMSE), where RMSE is obtained using the actual (y_i) and predicted ($\hat{y}_i$) values from the test set[9]. We compare the RMSE obtained from the specifications discussed above using, alternatively, no trends and (province-specific) linear trends (see Table 1).

2.2 Short-run (weather) vs long-run (climate) effects

A crucial and hotly debated issue in the recent climate econometrics literature is the trade-off between improving confidence in empirical estimates using fixed-effects to control for unobserved

[7] We also experimented specifications with province-specific quadratic time trends, omitted here to save space because they do not add significant information in term of out-of-sample prediction.

[8] In Kahn et al. (2019), the growth rate of GDP per-capita is regressed on p–1 lags of the dependent variable and the first differences of p lags of the positive and negative (absolute) deviation of weather variables from their long-run means: $\Delta y_{it} = a_i + \sum_{l=1}^{p} \varphi_l \Delta\, y_{i,t-l} + \sum_{l=0}^{p} \beta_l \Delta\, \boldsymbol{x}_{i,t-l} + \varepsilon_{it}$, where y_{it} is the log of real GDP per capita of province i in year t, $\boldsymbol{x}_{i,t}$ is the vector of positive and negative deviations of average temperature and precipitations from the historical norms (moving averages of the past 30 years). The number of p lags has been set equal to 3 in both the GDP and GVA equation.

[9] Where $RMSE_i = \sqrt{n^{-1} \sum_i e_i^2} = \sqrt{n^{-1} \sum_i (y_i - \hat{y}_i)^2}$.

omitted variables and estimating the long-run equilibrium response to climate change (Kolstad and Moore 2019). The standard view is that equation (1), by exploiting inter-annual variation of weather for identification, instead of permanent differences in climate – that are absorbed by fixed effects – can identify the short-run impact of weather variables (see Mendelsohn and Massetti 2017 for a discussion). However, to understand the economic implications of future global warming, we are particularly interested in the long-run effect of climate change. That is, the effect inclusive of adaptation representing the efforts of economic agents to contrast the negative impact of a change in climate (Hsiang, 2016).

Yet, recently, several authors have started to realize that what the panel fixed effects model really estimates is somewhat different from the above representation, at least under certain conditions. This is because, if the underlying data-generating process (DGP) implies a non-linear behavior of the quantity of interest to weather shocks, then the variation of the mean climate in the units of observation enters in the process of identification. This intuition, initially advanced by Schlenker (2006) and McIntosh and Schlenker (2006), has recently been formally extended by Mérel and Gammans (2021). More precisely, the critical reason is that the quadratic specification in temperature in a fixed-effects model does not measure non-linearity within units because, by taking the square of temperature and then demeaning it, a function of the mean temperature (i.e. the climate) has been reintroduced into the fixed-effects model.

Based on this consideration, Mérel and Gammans (2021) formally derived two results that are relevant for understanding the magnitude and the direction of the bias induced by the panel fixed-effects model that exploits (short-run) inter-annual variations of weather variables, rather than the (long-run) effect of climate.

First, they showed that the common idea that the estimated impacts of panel fixed effects are biased upward, due to failure to account for long-term adaptation, is conceptually wrong when considering a quadratic in weather specification. This is because the bias depends on the (positive or negative) skewness of the distribution of weather variables and can therefore be negative or positive, vanishing under symmetric distribution. Notably, they also showed that the estimated bias induced by the skewness of the distribution of weather variables is quantitatively negligible in applied works.

A second crucial consideration is that the estimated coefficients of a panel fixed-effects model with a quadratic in weather variables specification are a weighted average of the long-term and short-term response functions, that will be close to the long-term (short-term) true parameters, to the extent to which the time-series variation in weather variables is small (large) relative to the cross-units variation in climate[10]. Thus, in the Mérel and Gammans (2021) framework what is required to pick up the long-run effect of climate in a quadratic weather response function is the existence of substantial differences in the climate across units relative to the weather shocks within units.

A somewhat different, but related, argument is highlighted by Hsiang (2016) and Deryugina and Hsiang (2017), who exploited a combined application of the Envelope Theorem and the Gradient Theorem. In short, they argued that if we have enough (small) spatial units with a continuum of (small) differences in climate between contiguous territorial units, then the estimated marginal effect of weather will be heavily influenced by adaptation to local climate, so that the marginal effect of weather tends to coincide with the marginal effect of climate at the margin. Note that the latter result holds true only under certain restrictive conditions, such as whenever the outcome variable of interest is a maximized quantity from economic agents (see Hsiang, 2016), and if and only if the relation between outcome and weather is non-linear (see Mérel and Gammans 2021)[11].

In our application to Italy, many of the above conditions are broadly satisfied: small skewness in the distribution of weather variables; marked non-linearity in the economic-temperature relationship; cross-units variation in climate significantly larger than within-unit variation in weather; and finally, economic outcomes of interest that well approximate optimized quantity by economic agents, such as GDP per-capita and agricultural GVA per-worker. Hence, within this logic, though under some restrictive conditions, in the results reported below we are also capturing, to a certain degree, the long-term effect of climate on economic outcomes[12].

3. Data and variables

Our analysis exploits all Italian NUTS3 provinces, considering 35 yearly observations from 1980, the first year of data availability concerning economic outcomes, until 2014. During the observed period, the number of Italian NUTS3 changed from 106 (1980) to 110 (2014), as an effect of redefinition of the territorial units. However, the economic data refer to the 2010 NUTS3 classification and, as such, covered all the 110 NUTS3 provinces in force from 2008, irrespective of

---

[10] This discussion is relevant particularly when comparing linear vs non-linear panel models, because while the former only capture short-run effects of weather, the later retrieve a mix between short and long-run impacts (see Kolstad and Moore, 2020).

[11] See also the discussion in Carter et al. (2018). Furthermore, Lemoine (2018) qualifies under which specific conditions we can exploit weather variation to identify the (long-run) effect of climate, showing for example that the Envelope Theorem can fail in a dynamic framework.

[12] Mérel and Gammans (2021), using three different panel datasets related to France at NUTS3 level, US at county level and Global at country level, have calculated that estimated coefficients incorporated long-term (short-term) temperature effect with weights of 86% (14%), 98% (2%) and 100%, respectively. Considering our exercise on Italy, as more fully discussed in the results section, we show that the estimated temperature coefficients capture about 96% (4%) of the long-run (short-run) effects, *ceteris paribus*.

their effective presence in the previous years. As a result, we worked with a balanced dataset of 3,850 (=110*35) observations[13].

Following the literature (e.g. Dell et al. 2012), overall economic outcome is measured as real GDP divided by total population, obtaining a NUTS3 per-capita GDP indicator at constant prices. Agricultural outcome is measured as sectoral (real) gross value added divided by the number of agricultural workers. As such, it represents a measure of agricultural labor productivity. GDP at constant price, population, agricultural value added and agricultural workers at NUTS3 level, are all obtained from the Cambridge Econometrics' Regional Database, based on Eurostat.

Monthly weather data (temperature and precipitation) from each province over the 1980-2014 period have been set up by means of the anomaly method (New et al., 2000; Mitchell and Jones, 2005), as described in Brunetti et al. (2012). Specifically, we first obtained the monthly temperature and precipitation normals (i.e. climatologies (means) over the 1961-1990 period) for each Italian province, interpolating available Italian 30-arc-second-resolution (about 800 m) temperature (Brunetti et al. 2014) and precipitation (Crespi et al. 2018) climatologies. They were obtained by starting from high-density datasets and exploiting the assumption that the spatial distribution of temperature and precipitation normals is strictly related to geographical features, especially elevation. Then, from a different dataset, we estimated, for each province, the monthly records of deviations (i.e. anomalies) for the 1951-2014 period from the temperature and precipitation normals by means of a weighted average of the anomalies calculated for the neighboring series. In this case, we used the network of high-quality homogenized records that the Institute of Atmospheric Sciences and Climate (CNR-ISAC) of the National Research Council of Italy uses to produce its monthly climate bulletins. This network is based on updated versions of the databases of secular Italian observatories presented by Brunetti et al. (2006) and of Italian Air Force (Aeronautica Militare) synoptic stations presented by Simolo et al. (2010). Finally, by the superposition (addition for temperature and product for precipitation) of the two fields, we obtained for each Italian province and for each month of the 1951-2014 period, an absolute record of monthly temperature and precipitations values. This methodology allows us to obtain unbiased records even when the stations used for a specific point have normal values that are different from those corresponding to this point. It can be applied to any point of the grid, and average records can be obtained for any area by simply averaging the corresponding grid-point series. We decided, however, to consider the record corresponding to the main city of each province rather than the provincial average, because

[13] Cambridge Econometrics, the source of economic data, on the basis of Eurostat Statistics, estimated the main economic variables (e.g. GDP, population, sectoral value added) backward, to supply a balanced dataset. Note, all the results presented in the paper are robust to the elimination of the provinces that were reclassified in the observed period.

the former is generally more representative of the economic activities and population of the province than the latter.

Table A.6 in Appendix, reports descriptive statistics of the variables described above and additional variables used in the robustness checks discussed below.

In the baseline regressions, we used temperature and precipitations of the main city of each province for both the GDP and agricultural reaction function. This is done mainly for reasons of internal coherence and comparability with previous global-level evidence (see Burke et al. 2015; Pretis et al. 2018; Kahn et al. 2019; Acevedo et al. 2020; Newell et al. 2021)[14]. However, this may not be the best choice for agricultural outcomes, as shown by previous literature that considered average weather over farmland areas using land cover information (see Schlenker and Roberts, 2009; Ortiz-Bobea et al. 2018). As a result, besides the record corresponding to the main city in each province, we also considered an average record over each province's agricultural areas, which were identified by means of the Global Land Cover 2000 database (European Commission-JRC, 2003). Nevertheless, estimating the agricultural GVA reaction function with average temperature and precipitations based on agricultural areas delivered similar results to our baseline regressions[15].

As for future climate change projections, we rely on high resolution temperature data at 0.0715° resolution (about 8 km) obtained from the regional circulation model COSMO-CLM (see Bucchignani et al. 2015). The simulated climate data for the period 1971-2100 are forced by the CMCC-CM global climate model, under intermediate and worse-case Representative Concentration Pathways (4.5 and 8.5)[16]. Specifically, the 1971-2005 period was simulated following the CMIP5 historical experiment – that is, the two datasets contain the same climate data – while the period 2006-2100 was simulated by forcing the model with the corresponding RCP. Moreover, to remove the existing bias between observed and modeled data and to make the two datasets comparable, we downscaled the modelled scenarios by means of the observed records.

A preliminary glance at the data (Figure 1) gives the economic outcomes at NUTS3 level against temperature over the available historical period. The predicted lines are obtained through a flexible fractional polynomial regression using Stata, without the imposition of any functional form restriction. The graph allows us to study how, without additional controls, the level of economic outcomes behaves against temperature. The figure clearly shows a smoothing non-linear and concave-in-temperature relationship, with a maximum around 12°C and 14°C for the GDP per-capita and agricultural GVA per-worker, respectively. Thus, productivity increases as annual

[14] In the cited papers, indeed, the agricultural reaction function has been estimated on the same weather data used for the GDP reaction function.

[15] These additional results are available from the authors upon request.

[16] Note that the use of these data represents an improvement compared to previous studies (e.g. Van Passel et al. 2017; Bozzola et al. 2018) that based their projections on data from General Circulation Models (GCMs).

temperature increases, until the optimum. Then, productivity gradually declines with further warming. This effect is particularly evident for the agricultural sector, highlighting the greater sensitivity of agriculture to temperature. As seen in the reported scattered data points, variability in economic outcomes is particularly marked for agricultural GVA, showing important variation across provinces and over time. Furthermore, what this graphical analysis implies is that a simple test of non-linearity, against the alternative linear specification in temperature, rejects the null at 1% significance level. This represents a preliminary, albeit not conclusive, indication that a non-linear function between temperature and economic outcomes lies in the data.

4. Results

To assess the performance of alternative models in out-of-sample prediction we use different Cross Validation (CV) exercises. Our CV exercise is, however, slightly less agnostic than the one in Newell et al. (2021), because we are comparing specifications that, in one sense, have already been selected as the "best" alternative within the set of criteria used in the literature considered. We thus reassess existing literature in the Italian context to shed some light on two critical questions: the linearity vs. non-linearity effect of weather variables, and their growth vs. levels effect on economic outcomes.

4.1 Model performance in out-of-sample prediction

Table 2 reports results of the estimated RMSE of our models using forecast CV. This method accounts for the time-series nature of the dataset, dividing the data into training sets of early data and testing sets of later data[17]. For this reason, forecast CV is the standard approach for evaluating out-of-sample prediction of models with time series information.

Overall, the results from forecast CV show that models including linear time trends perform poorly compared to models without them, i.e. RMSE are minimized by excluding time trends irrespective of other modeling assumptions. Considering the per-capita GDP reaction function, the specifications with the lower RMSE, highlighted in red in Table 2, are the DJO1 and the NPS. These two specifications are equivalent in terms of out-of-sample prediction accuracy at four decimal places. Within the static models, the non-linear BHM growth model and the linear DJO2 growth specification with the "poor" interaction effect are the specifications with highest RMSE, suggesting a weak preference of forecast CV for models in level vs. models in growth. Note, however, that these differences are extremely small and emerge at only the fourth decimal place.

[17] We construct the training sample by years, using five training windows beginning in 1984 and ending in 1990, 1995, 2000, 2005 and 2010, respectively, to predict periods that begin in the year following the end of each training window and span five years, such as 1991-1995, 1996-2000, and so forth. We start in 1983, instead of 1980, to account for the dynamic structure of the Kahn et al. (2019) specification, that include 3 lags of the dependent and weather variables.

The dynamic specification recently proposed by Kahn et al. (2019) performs poorly in out-sample-prediction (KAHN1), particularly when heterogeneity with respect to income is considered (KAHN2).

Moving to the forecast CV of agricultural productivity models, the overall result is similar, in the sense that the same specifications – linear in growth DJO1 and non-linear in levels NPS – display the same (lower) RMSE while the BHM and, particularly, the KAHN1 and KAHN2 display the highest one. Though the differences in RMSE are still quite small across specifications, in agriculture the specifications that assume heterogeneity of weather effects by income level (DJO2 and KAHN2) are not preferred in out-of-sample prediction.

Table A.1 in the Appendix reports results from the backcast CV, with results that closely mimic the ones of the forecast CV[18]. Across static models and considering the GDP reaction function, results are similar to those from forecast CV, strongly favoring models that exclude time trends, and selecting the same models as equivalent in terms of RMSE (highlighted in red). As with forecast CV, backcast CV confirms that there is a very weak preference for models in levels vs. models in growth rate, i.e. the BHM and DJO2 models display the highest RMSE, but only at fourth decimal point. For models in agriculture, backcast CV shows that RMSE is minimized by the NPS non-linear level model. Again, the Kahn et al. (2019) specification performs poorly. Finally, Table A.2 in Appendix reports results from the k-fold CV. One shortcoming of this method is that it ignores the time-series nature of the data, hence in theory it is less suited to our purpose[19]. Notwithstanding this, the results from k-folder CV broadly confirm the previous findings based on forecast and backcast CV.

We formally tested models using the Welch t-test, against the null hypothesis that RMSE in 1000 replications is the same for any comparison of two models. Clearly, given the very slight differences in RMSE highlighted before, it is not surprising that for both overall and agricultural outcomes this test suggests that the NPS and DJO1 specifications, that display the lower RMSE, are not statistically superior to all other specifications (DJO2, BHM, KAHN1 and KAHN2) at 95% confidence level (see Table 3).

In sum, comparisons based on out-of-sample forecast accuracy suggest that there are several models that are statistically indistinguishable in terms of RMSE. The reason for this is that RMSE is largely

[18] In backcast CV we are using five training windows that end in 2014 (last year of our panel) and begin in 1986, 1991, 1996, 2001 and 2006, to predict periods that begin in the years preceding the beginning of each training window, and span five years, i.e. 1981-1985, 1986-1990, and so for.

[19] K-folder CV splits the data randomly into k partitions (we use k=100). Then, for each partition, it fits the specified model using the other k-1 groups and uses the resulting parameters to predict the outcome variable of interest in the unused group. The data are first centred by regressing all variables (dependents and explanatory variables) on country fixed-effects and country-specific trends. The residuals from these regressions were used for the cross validation calculating the RMSE.

invariant to the exclusion of weather variables, or to their inclusion as linear or non-linear functions (see Newell et al. 2021)[20]. As far as the GDP reaction function is concerned, the non-linear NPS specification in levels and the linear DJO1 in growth rate are always selected within the class of "best" models, although they are not statistically different from the other specifications. In agriculture things are slightly different, as the non-linear NPS specification in levels is systematically the preferred model in all CV exercises, though this model is not statistically different in out-of-sample prediction from the other models considered. Finally, our cross validation results find that the non-linear specification of BHM in growth rate is never among the best performing models in terms of RMSE, though its predictive ability is not statically different from those of the other models. Thus, similarly to the Newell et al. (2021) CV exercise conducted at global level, it emerges that it is difficult to discriminate between growth vs level models or between linear vs non-linear models in out-of-sample prediction, suggesting the existence of model uncertainty that is also significant for Italian data.

4.2 Econometric results: per-capita GDP reaction function

Table 4 reports econometric results for the per-capita GDP equation, considering all the static models of the CV exercise discussed above. We do not display precipitation coefficients because they are never statistically significant when GDP is considered. For comparison with the DJO2 (and KAHN2) specifications used in the CV exercise, we also report estimates for BHM2 and NPS2 models in Columns (4) and (6) respectively, that account for heterogeneity with respect to income[21]. We report the standard error clustered by province (in round brackets) and, following Dell et al. (2012), also clustered by province and region-year (square brackets), following the two-way clustering of Cameron et al. (2011), which also, accounts for (within region) spatial correlation.

Starting from the linear DJO1 specification (Columns 1), the estimated temperature effect is positive, though statistically significant only when the standard error is clustered at NUTS3 level. Accounting for heterogeneity with respect to income (DJO2), does not change the picture in the sense that the temperature interacted with the poor dummy is negative but small in magnitude and statistically insignificant when the more conservative two-way standard error is considered. As a result, the effect of temperature in the poor provinces reported at the bottom of column (2) is

[20] In our dataset, weather variables using the NPS specification have the following explanatory power. Considering the GDP equation, temperature and temperature-and-precipitation explain, respectively, 0.89% and 1.02% of GDP growth. In the GVA per-worker equation, the same numbers are 2.96% and 4.19%, respectively. Thus, as expected, temperature significantly more relevant than precipitations, and weather variables are significantly more important in agriculture, but still explain a small fraction of the productivity growth.

[21] Following Burke et al. (2015), this specification includes a dummy $d_i = 1$ for provinces with below-median per capita income in 1980 (and zero otherwise), interacted with the weather function, $h(X_{it}) = \beta_1 X_{it} + \beta_2 X_{it}^2 + \beta_3 (X_{it} \times d_i) + \beta_4 (X_{it}^2 \times d_i)$.

positive and statically insignificant. Hence, the data appear to reject the linear temperature specification in growth rate[22].

Column (3) reports results of the non-linear (quadratic) growth in temperature specification of BHM1. The positive linear temperature coefficient and the negative quadratic one supports a quadratic relationship that is, however, not statistically significant[23]. This result is in sharp contrast with previous (global) studies, where a significant quadratic relation in temperature is always present in the data (see Burke et al. 2015; Pretis et al. 2018; Newell et al. 2021). Testing for heterogeneity with respect to income (Column 4) we find that both the linear and quadratic interactions with the "poor" dummy (β_3 and β_4, respectively) are never statistically significant.

Column (5) reports results of the non-linear in levels specification of Newell-Preston-Sexton (NPS1), which is always among the best performing models in our CV exercises. In this specification temperature (and precipitation) enters in first differences, accounting for the trending nature of the temperature series. In this model, the linear and square temperature coefficients are estimated with great precision, irrespective of the type of standard errors considered. The resulting maximizing GDP temperature is equal to 17.5°C, with a 95% quintile range of 14.2°÷20.9°C significant at 1% statistical level[24]. Provinces with average temperature below (above) this optimum will experience an increase (decrease) of per-capita GDP with an increase of temperature.

It is useful to compare our estimate of maximizing GDP temperature of 17.5°C with previous global studies. Newell et al. (2021) report a 95% quantile range from different models of 11.6–18.4°C, with median values centered around 14°C. Thus, our NPS estimates lie within the range of current findings. However, as the average temperature in Italy was 14.9°C in the observed period (1980 to 2014), using our preferred optimal temperature of 17.5°C would imply that Italian GDP per-capita on average will increase slightly with a small rise in temperature. Differently, using the global optimal temperature estimates from previous studies centered on around 13-14°C, future projection of global warming in Italy will go in the opposite direction, pointing to a *reduction* of GDP per-capita. This contrasting evidence highlighted the importance of working at the sub-national country level to identify the "true" temperature-GDP reaction function. Finally, Column (6) looks for heterogeneity with respect to income of the NPS2 specification. As with previous results, when we

[22] Following Letta and Tol (2018), we also experiment a linear DJO specification where weather variables entered in first differences, thus transforming the DJO growth model in a levels model. The data appear to prefer this last specification to the ones reported in columns 1-2. However, several additional tests of non-linearity, suggest that the data prefers a non-linear in temperature specification instead of a linear one. These additional results are available from the authors upon request.

[23] A joint significant test of the temperature coefficients in the BHM1 specification delivered an F-statistic of 1.17 with a *p*-value of 0.31.

[24] These are obtained through a bootstrap procedure, providing a measure of variance in the optimal temperature due to sampling uncertainty.

use a two-way clustered standard error, the data cannot reject the hypothesis that the relation estimated for "rich" and "poor" provinces is the same.

It follows that the non-linear quadratic in temperature specification estimated in first differences, i.e. the NPS model in levels, performs better econometrically than the linear growth models *a la* DJO or the non-linear in temperature growth specification of BHM (2015), though they are statistically indistinguishable in out-of-sample prediction. The finding that a non-linear model appears to be preferred by the data is an important result in terms of external validity with respect to climate change. In fact, linear panel models only capture short-run elasticities, whereas non-linear ones also include a cross-sectional component which implies that the effect is a mix between short- and long-run impacts (see Kolstad and Moore, 2020; Mérel and Gammans, 2021). In an extension discussed below, we report the precise quantification of these weights.

4.3 Econometric results: Agricultural GVA per-worker reaction function

Results for the agricultural GVA per-worker are reported in Table 5 for the temperature effect. In Table 6 we consider precipitations as they are important determinants of productivity in agriculture.

Starting from the DJO linear specifications, the impact of a rise in temperature on GVA per-worker is always negative, although it is only marginally significant in "poor" provinces. Indeed, the temperature effect in the linear DJO1 specification is negative but not statistically significant when the two-way clustered standard error is considered. In column (2) the temperature interacted with the poor dummy is significant and negative, so that the magnitude of temperature effect in poor provinces, reported at the bottom of column (2), suggests that a 1°C increase in temperature reduces agricultural labor productivity by around 5 percent points, an effect significant at only 10% level[25].

Moving to the non-linear specification in growth rate of BHM, we find that also in agriculture the linear and square temperature coefficients turn out to be statistically insignificant[26]. This is hardly surprising when agriculture is concerned, as temperature directly affects output rather than output growth. The maximizing GVA temperature is equal to 7.1°C but with a very large 95% quantile range (-8.9°÷ 23.2°C). There is no evidence of heterogeneity in the temperature effect with respect to income in this non-linear growth specification, namely the β_3 and β_4 coefficients of the temperature interaction terms are never statistically significant (see Column 4).

[25] Similarly to the GDP reaction function, if we include temperature (and precipitations) in first difference in the linear DJO2 growth specification, the temperature effect in "poor" provinces becomes significantly at 1% level, suggesting that a levels specification is preferred by the data. However, additional tests to check whether this linear specification is preferred by the data, against a non-linear quadratic equation in weather variables strongly prefer the latter..

[26] A joint significant test of the temperature coefficients in the BHM1 specification, delivered an F-statistic of 1.19 with a *p*-value of 0.30.

The picture changes substantially when the non-linear specification in levels of NPS is considered (Column 5). Results point to a clear and robust non-linear relationship between temperature and agriculture GVA per-worker. The maximizing GVA temperature – equal to 12.5°C – is estimated with high precision (1% statistical level) and delivers a 95% quantile range of 10°÷15°C. Once again, the data cannot reject the hypothesis that the relation estimated for rich and poor provinces is the same (see Column 6).

Regarding precipitations, the effects on agricultural productivity are never statistically significant in the levels model of NPS (see Table 6, columns 5-6). It is worth noting, though, that the coefficient estimates on precipitations become statistically significant and precisely estimated when the growth models of DJO and BHM are employed. In Column (1) the effect of a 100 mm rise in precipitations is to increase the average growth rate of productivity by 0.6 percentage points, rising to 0.8 percentage points when "poor" provinces are considered (see bottom of Column 2), both effects are statistically significant at 1% level. A similar result can be detected from the BHM non-linear growth model, where there is a clear quadratic relationship between precipitations and agricultural GVA per-worker. The maximizing GVA precipitation level is equal to about 1,600 mm per year. However, as seen in Column (4), this effect is mainly driven by "poor" provinces. Indeed, the linear and quadratic interactions of precipitations with the "poor" dummy (β_3 and β_4) are statistically significant at 5% level. The bottom of the Table reports the precipitation effects in "poor" provinces, ($\beta_1 + \beta_3$) and ($\beta_2 + \beta_4$). They are significant at a 1% level, delivering a maximizing GVA precipitation level of about 1,300 mm per year in "poor" provinces. Overall, these findings suggest that, unlike the temperature impacts that induce mainly transitory (level) effects, precipitation appears to have important growth effects in agriculture, a result consistent with the intuition and the agronomic literature (e.g. McCallum et al. 2013) and also strongly supported by additional robustness tests discussed in the Supporting Information, Section C.

4.4 Econometric results: Kahn et al. (2019) ARDL model

Tables A.3 and A.4 in Appendix report econometric results of the Kahn et al. (2019) ARDL model. Though this model systematically displayed higher RMSE in the CV exercise, temperature variables are often estimated with precision, especially when agricultural productivity is concerned. Following Kahn et al. (2019), we used both the standard fixed effects and the half-panel Jackknife FE (XTSPJ) estimator, to deal with the possible bias induced when weather variables are not strictly exogeneous[27].

[27] Kahn et al. (2019) motivated this choice following Chudik et al (2018), according to which the Nickell bias in FE dynamic panel model (see Nickell 1981) exists regardless of whether the lags of the dependent variable are included or

Results for the GDP per-capita equation suggest that the estimated coefficients on the precipitation variables are never statistically significant. In line with Kahn et al. (2019), when the model is estimated with the XTSPJ estimator, the size of the estimated coefficients is often larger in (absolute) magnitude and slightly more significant, particularly for the GDP per-capita equation, so we shall now focus attention on the results from this estimator.

Long-run GDP per-capita growth is negatively affected when temperature deviates positively or negatively from the historical norm, and the impact of negative deviation appears large in magnitude, suggesting that also extreme cold can play a role. These results mimic the findings of Kahn et al. (2019) on US States levels. However, the magnitude of the negative temperature deviation coefficient is also the result of the few, small negative deviations in the historical data (see bottom of Table A.6). So, notwithstanding the size of the estimated coefficient, the contribution of negative temperature deviation on long-run growth is, on average, in order of magnitude smaller than the positive temperature deviation one.

Considering the specification in Column (4), which omits the (insignificant) precipitation controls, results for Italy show that a persistent annual rise of 0.01°C in temperature above the historical norm decreases real GDP per-capita growth by 0.019 percentage points per year in the long run, an effect significant at 1% level. When negative temperature deviations are considered, the same effect leads to a reduction of long-run GDP per-capita growth of 0.214 percentage points.

In Column (6), we test for the heterogeneity of the temperature deviation effects by income level. The interaction terms of positive and negative temperature deviation from historical norm with the "poor" dummy are, respectively, insignificantly positive and significantly negative, while for rich provinces the negative impact of positive temperature deviation is even larger than before and the negative temperature deviations switch from negative to positive, though statistically insignificant. Thus, overall, the evidence on the heterogeneity of the impacts in "rich" vs. "poor" provinces with regard to GDP per-capita is mixed, namely heterogeneity exists for negative, but not for positive temperature deviations, which are the ones that matter for future global warming.

Table A.4 reports the same battery of regressions for the agricultural GVA per-worker. First, the precipitation variables are not statistically significant, a result at odds with previous findings where precipitations were often significant in (static) growth models. Both positive and negative temperature deviations from the historical norm display negative estimated coefficients statistically significant at 1% and 10% level, respectively. Results from column (4) that exclude precipitations,

---

not, so long as one or more regressor is not strictly exogenous. Because the length of our time-series is 37% lower (34 vs 54 years) than in Kahn et al. (2019), in our baseline regressions we used 3 lags of both the dependent and weather variables. However, in a robustness check we test for the sensitivity of the results changing both how we measured the long-run climate (from 30 to 20 years window) and the lags structure.

suggest that a persistent 0.01°C annual increase in temperature above its historical norm reduces agricultural GVA per-worker growth by 0.117 percentage points per year in the long run, and a 0.01°C annual decrease in the temperature below its historical norm reduces agricultural GVA per-worker by 0.223 percentage points per year in the long-run, though the latter effect is only marginally significant.

In Columns (5)-(6) we test for heterogeneity of the temperature deviation effect in "poor" vs. "rich" provinces. The interaction terms of the positive and negative temperature deviation from historical norm are both negatively and statistically significant at 1% level and large in magnitude, so that we can reject the hypothesis that there are no differential effects of climate change on "poor" versus "rich" provinces in agriculture. Note, moreover, that the estimated coefficients on rich provinces though still negative for positive temperature deviations, are no longer statistically significant, suggesting that the Kahn et al. (2019) model predicts robust negative GVA growth effects of temperature deviations mainly for "poor" Italian provinces in agriculture, *ceteris paribus*[28].

To sum up, as regards the use of a long-run dynamic growth model to infer the impact of (positive/negative) deviations from the long-run climate, the results are, qualitatively, not in contradiction with our preferred non-linear NPS level model for both GDP and GVA. However, quantitatively, given the different structure of the two models, it is hard to make a direct comparison of the magnitude of the estimated temperature effects. In Section 5, we will discuss this issue in more detail.

4.5 Robustness checks and extension

We conduct a variety of robustness checks of our results. First, we test for the stationarity property of our outcome and weather variables. Im, Pesaran, and Shin (2003) panel unit root test firmly rejects the non-stationarity of GDP per-capita, agricultural GVA per-worker, temperature, and precipitations series, when they are considered in first differences. These additional results are discussed in Supporting Information, Section A (see Table S.1). However, both economic variables in levels as well as temperature, are not stationary when Hadri (2000) Lagrange multiplier stationarity test is used. This tends to suggest that models that include level temperature, such as the DJO and BHM, are not suitable from an econometric point of view due to the non-stationarity of temperature.

Second, following Burke et al. (2015) and Dell et al. (2012) we test for persistency in the weather effects, running regressions with one, two and four lags of the weather variables. Then we analysed

[28] Unlike the Kahn et al. (2019) results obtained on the US States, we never find significant effects of positive/negative precipitation deviations from long-run climate. In addition, our (positive/negative) temperature deviations effects in agriculture are larger in size, estimated with more precision, and heterogeneous with respect to income.

the results by computing marginal effects at different temperature levels, as described in the Supporting Information, Section B. Lagged temperatures are sometimes significant up to two lags, but importantly, they systematically change sign with respect to the contemporaneous effect, so that the marginal impact of additional warming shrinks or changes in sign, thus providing clear evidence that genuine level effects are captured, a result that holds true for both the GDP and particularly the GVA reaction function. In addition, when lagged precipitations are included in the agricultural GVA regressions, they are often significant and move in the direction of the contemporaneous effects. This finding represents further confirmation that precipitations, unlike temperature, have persistent effects on agricultural productivity.

Third, following Mérel and Gammans (2021), we computed the expected weight on the long-run and short-run decompositions, under the hypothesis of a non-linear long-run temperature effect, i.e. $y_{it} = a_i + b_1 x_{it} + b_2 x_{it}^2 + e_{it}$. With this approach, they showed that the estimated coefficients from a fixed-effects model is a weighted average of the short-run and long-run coefficients, with weight $\bar{\theta}$ that can be interpreted as reflecting the share of overall weather variation attributable to time-series fluctuations. Using their equation (12), the estimated weights, $\bar{\theta}$, based on our Italian weather data, have a range value of $0.02 \div 0.04$, depending on the time period considered, i.e. 1951-2016 and 1980-2014, respectively. Hence, in line with the Mérel and Gammans (2021) approach, the coefficients of our non-linear in temperature NPS model capture from 96% (4%) to 98% (2%) of the long-run (short-run) effect. Details of this calculation are reported in the Supporting Information, Section C.

Finally, we test for the sensitivity of the Kahn et al. (2019) results to the lags structure and how the long-run climate is measured. Results are quite sensitive to these modelling choices, particularly when the GDP per-capita is concerned. For example, on moving from two to four lags the estimated effect of both positive and negative temperature deviations, often switch from significant positive to significant negative (see Table S.5). A similar result can be observed when moving from a long-run climate measured with a 30- versus 20-year moving average. Results for the agricultural GVA per-worker are in general less sensitive to these modelling choices (see Table S.6). These additional results are discussed in the Supporting Information, Section D.

5. Projections of climate change impacts

In the following we discuss results from projections of the impacts of climate change. We estimated global warming projections under two different Representative Concentration Pathways: the RCP 4.5, a stabilization scenario that is characterized by an average temperature rise between 2°C and 3°C at the end of the century, relative to pre-industrial temperatures; and the (unlikely) worst-case scenario (high emissions scenario) RCP 8.5 characterized by an average increase in temperature of about 4°C (see Hausfather and Peters, 2020). The projections are based on the NPS results, accounting for sampling uncertainty through a bootstrap procedure, but not for model uncertainty, because for the DJO and BHM models the temperature coefficients are never statistically significant. For completeness, we also briefly discuss climate change projections based on the Kahn et al. (2019) model, as in our baseline results the estimated temperature coefficients are often significant, particularly in agriculture[29].

5.1 Projecting economic impact of global warming

Following Burke et al. (2015) we generate projected changes in per-capita GDP and agricultural GVA per-worker for each NUTS3 province, under the two RCPs. This approach assumes that the evolution of per-capita GDP (or GVA per-worker) in province *i* in year *t* is given by:

$$GDP_{i,t} = GDP_{i,t-1}(1 + \eta_{i,t} + \delta_{i,t})$$

where $\eta_{i,t}$ is the economic growth rate without temperature change. The "baseline" scenario is the average growth rate during the 1980-2014 period for both the per-capita GDP and agricultural GVA. Parameter $\delta_{i,t}$ is the temperature induced variation in the respective growth rate due to projected temperature change in each NUTS3, with respect to the 1980-2014 average. $\delta_{i,t} = h(T_{i,t+}) - h(\overline{T_i})$, where $T_{i,t+}$ is the projected temperature beyond 2014, and $\overline{T_i}$ is the NUTS3 specific average temperature in the period 1980-2014. Finally, $h(T_{i,t})$ is our estimated relationship based on the NPS specification. As discussed in the data section, our projected NUTS3 estimates of $T_{i,t+}$ are based on the COSMO-CLM model, forced by the CMCC-CM model (see Bucchignani et al. 2016)[30]. In addition, projected effects of temperature changes in Italy (or in most affected North and South provinces) are obtained by aggregating NUTS3 impacts, as $GDPpc_t = \sum_i w_i * GDPpc_{it}$, where w_i is the province i's observed share of GDP in 2014[31].

Because under the RCP 8.5 scenario for several Southern provinces the predicted temperature, $T_{i,t+}$, is beyond the highest observed annual average temperatures in the historical data, we produce two different sets of projections: with and without boundary on the max temperature. In the first case,

[29] The Kahn et al. (2019) projections are based on the methodology presented in their Section 4. See the Supporting Information, Section E, for calculation details and additional results.

[30] Climate projections are often based on model ensembles, to account for uncertainty. Here, our exercise is based on temperature projected by only one (high resolution) climate model. Thus, our projected economic outcomes should be considered with caution. However, note that uncertainty due to the estimated reaction function (economic model), as well as sampling uncertainty, are systematically higher than climatic model uncertainty (see Moore and Lobell, 2014). Therefore, added uncertainty in climate model predictions is unlikely to affect our main conclusions.

[31] Alternative weights for aggregation, e.g. projected changes in NUTS3 GDP based on historical trend, clearly render our projection results slightly more negative, so that our aggregation approach should be considered conservative.

$T_{i,t+}$ is equal to the estimated projected temperature without any restriction. In the second case, following Burke et al. (2015), we conservatively cap $T_{i,t+}$ at 20.2°C, which is the upper bound of the historical annual average temperature in our dataset[32]. Clearly, this strategy will understate the loss in provinces that exceed this max temperature, so our projected impacts in the hottest provinces tend to be conservative.

5.2 Projection results

Results from this exercise under both RCP 4.5 and 8.5 scenarios with their 95% bootstrap confidence interval, are reported in Figures 2 and 3, for the GDP per-capita and GVA per-worker, respectively. Note, Figure 2 considers GDP per-capita scenarios for the most affected Northern and Southern provinces, because the Italian mean effect is always near zero and insignificant (see Figure A.1 in Appendix). On the other hand, Figures 4 and 5 report Italian maps at provincial level, with end of the century projections measured as the average between 2080-2100.

The point estimates of climate change impacts vary depending on the RCP and time horizon considered. The average projected change of per-capita GDP in Italy at the end of the century, tends to be weakly positive under RCP 4.5 (+0.23%) and weakly negative under RCP 8.5 (–0.33% without bound temperature), but these effects are never statistically significant (see Figure A.1 in Appendix). The range of projected impacts follows a clear geographical pattern, with slight positive gains of around 0–2% in Northern provinces, and losses of about 1–3% and higher than 4%, under RCP 8.5, in the Southern provinces (see Figure 4). Under RCP 4.5, our per-capita GDP projections showed a significant increase of around 0.75% for the most (positively) affected North provinces that, however, becomes insignificant under the RCP 8.5 (see Figure 2). The pattern of most (negatively) affected Southern provinces is the opposite, showing a nil and insignificant effect at the end of the century under RCP 4.5, but a negative and significant effect of around 1.4% at the end of the century under RCP 8.5.

Overall, projected impacts of global warming on the Italian economy are several times lower of the (few) existing (aggregate) estimates. For example, Kahn et al. (2019), using a global reaction function, have estimated end of century per-capita GDP loss for Italy of around 7% under RCP 8.5. This significant difference is largely driven by exploiting within country heterogeneity in climate at a high granular level, and not just by model uncertainty[33].

[32] So, every time the projected temperature is $T > T_{max}$, we fix the effect in this year of additional warming at $\delta_{i,t} = h(20.2) - h(\overline{T_i})$.

[33] In fact, our projected loss of GDP per-capita in Italy based on Kahn at al. (2019), is around 60% lower to the one estimated in Kahn at al. (2019) and based on a global model, i.e. a GDP loss of 2.7% vis-à-vis a loss of 7% in 2100

Moving to the projected impacts in the agricultural sector, the results change substantially (see Figures 3 and 5). Under RCP 4.5, the end of the century real GVA per worker average loss is around 9%, with 95% of provincial losses in the interval 6‒12%. Under the RCP 8.5 scenario, this effect is roughly double, from an average loss of 22% (95% of provincial losses in the interval 17‒30%) with bounded temperature, to a loss of 35% (95% of provincial losses in the interval 28‒44%) without temperature boundary. The distribution of these impacts follows a clear North-South pattern. Less affected Northern provinces are projected to lose from 3% to 20% in agricultural GVA per-worker, depending on the RCP and time horizon considered. Instead, most affected Southern provinces are predicted to lose substantially from 10% to 40%, in the unmitigated worst-case scenario, and without temperature boundary (see Figure 5). Finally, note that RCP 8.5 projections in agriculture based on Kahn et al. (2019), support this North-South pattern, predicting GVA per-capita losses in "rich" provinces of around 5% and in "poor" provinces of around 14% by the end of the century (see Table A.5, in Appendix).

The stark discrepancies in the impact between agriculture and the overall economy may appear, at first sight, difficult to fathom. However, if we account for the agricultural share of GDP in Italy, and its distribution across (North-South) provinces, these discrepancies become less surprising. For example, an average loss in agricultural GVA at the country level of 22% under RCP 8.5, translates to a GDP loss of 0.66%, if we account for the fact that the GDP share of agriculture in Italy was equal to 3% in our sample[34]. In addition, Figures 4 and 5 clearly show that the impact of global warming on GDP and agricultural GVA, displays a clear positive correlation at the geographical level.

In comparison with existing studies, our predicted losses for Italian agriculture are similar in magnitude to the findings of Bozzola et al. (2018), but significantly less severe than in Van Passel et al. (2017). Indeed, the latter paper estimated losses in land values at the end of the century in the range of 34%-71% under RCP 8.5, losses around two times higher than our conservative estimates.

Comparison between our projections and those from Ricardian models can be somewhat problematic; nevertheless it is reassuring that our estimates tend to be close to, or lower than those in existing Ricardian studies. In fact, a recurring critique of the use of panel data to model the impact of climate change is that by failing to fully account for adaptation by farmers, these short-term weather estimates overstate the true impact. In fact, we find exactly the opposite, a result consistent with the discussion in Section 2.2.

under RCP 8.5. Similarly, we predict a small gain of 0.15% in 2100 under the RCP 4.5 scenario, in comparison to a small loss of 0.05% estimated by Kahn et al. (2019), under the same emission scenario (see Table A.5 in Appendix).

[34] Similarly, a 40% loss in agricultural GVA in the more affected Southern provinces translates into a 4% GDP loss if we account for the 10% agricultural GDP share of these provinces.

6. Conclusions

Understanding the economic impact of climate change is a complex but necessary step in the formulation of sound policy recommendations. However, current econometric tools used to investigate the economic impact of climate change are plagued by uncertainty related to both the empirical model adopted and data sampling. A related issue is the level of spatial aggregation the analysis is conducted at: when averaging temperature at country level, we lose a huge amount of useful information that could be relevant if the country's climate is heterogeneous. We have contributed to this literature by focusing on model uncertainty, working at provincial level in Italy, a country with considerable heterogeneity in both climate and levels of development.

Our analyses show that model selection based on cross-validation in out-of-sample prediction, cannot discriminate between growth vs level models or even between linear vs non-linear models, confirming the notable uncertainty of current models – a result similar to Newell et al. (2021). However, selection based on more standard econometric tools (i.e. significance of temperature variables and their linear vs. non-linear effects) suggests that the one that best approximates the economic-temperature relationship, for both GDP and agricultural GVA reaction function in Italy, is a non-linear model that relates temperature to levels of economic variables, estimated in first differences. Additional robustness checks support the notion that temperature shocks in Italy have mainly had non-linear level effects on economic outcomes. Other models, such as that of Burke et al. (2015), who relate level temperature to economic growth – or Dell et al. (2012), which is linear in temperature – are less supported by our analyses, even though they are not statistically inferior in out-of-sample prediction. Among growth models, the results obtained from the long-run dynamic specification of Kahn et al. (2019), appear more promising, despite not performing well in the cross-validation exercise, and delivering results that are sensitive to the modelling choices, particularly when the GDP per-capita equation is considered.

Climate change projections, based on our chosen NPS specification, suggest that the Italian economy should be only marginally affected by global warming at aggregated level, with losses of around 2–4% concentrated in Southern regions, *ceteris paribus*. This result is mainly the consequence of exploiting the within country heterogeneity in climate, a result that also holds true when the projections are based on the Kahn et al. (2019) model. However, at the aggregated level this does not apply to the agricultural sector where climate projections based on the NPS model, both under a stabilization RCP 4.5 scenario and under the worst RCP 8.5 scenario, predict important productivity losses in the region of 5% - 35%. Projections in agriculture based on the Kahn et al. (2019) model, are qualitatively similar, though less severe. Importantly, these effects are highly uneven across regions, with important losses concentrated in the Centre and, especially, in

Southern provinces with lower income levels. Consequently, our results suggest that the effect of global warming in Italy will disproportionally affect regions that, as an effect of their lower development level, tend to be more vulnerable to weather variability, precisely because their economy is more dependent on the (climate sensitive) agricultural sector. This does not mean that weather and climate cannot affect the Italian economy through mechanisms different from the agricultural channel. However, we find a clear overlap between the estimated climate change damage in agriculture and the findings at aggregated level.

From a policy perspective, this could have useful implications. We believe that, while policies could be required to mitigate the impact of climate change on certain vulnerable agricultural sectors, a strategy to exploit the potential adaptation of the economy in Southern regions should consider a policy mix fostering a growth of economic activities other than agriculture.

Our analyses and results have several caveats. First, and perhaps most importantly given the huge model uncertainty emerging from our cross validation analysis, the results should be treated with caution, particularly in interpreting the estimated size of the impact of future climate change. Second, the extent to which the estimated reaction functions truly capture the effect of climate, and not just short-run weather variation, remains an open question. Though, as recently argued by Mérel and Gammans (2021) and supported in this paper, non-linear specifications do not capture just short-run weather effects, the reader should be aware of the difference between weather and climate impacts. This is because the non-linear panel methods deployed in this paper are likely to still capture a mix of short- and long-run effects. In addition, our analysis focuses on the historical record and thus cannot capture potential extreme events that lie outside this range but could be triggered by unmitigated climate change in the years to come.

Future empirical analysis of the economic impact of climate change should be extended in several directions, to deal with model uncertainty, by explicitly modelling adaptation, and developing more theory-driven empirical models to better discriminate between growth vs. level effects, as well as short- vs. long-run effects. From this perspective, the Kahn et al. (2019) model that exploits deviation of weather variables from long-run climate could represent a good starting point, though- as we have also shown, when applied to GDP per-capita growth in the Italian context, this model is particularly sensitive to the modelling choices.

References

Auffhammer, M. (2018). Quantifying Economic Damages from Climate Change. *Journal of Economic Perspectives*. Volume 32, Number 4: 33–52.

Auffhammer, M., Schlenker, W. (2014). Empirical studies on agricultural impacts and adaptation. *Energy Economics* 46:555–61.

Acevedo, M.M., Novta, N., Pugacheva, E. and Topalova, P. (2020). The Effects of Weather Shocks on Economic Activity: What are the Channels of Impact? *Journal of Macroeconomics*, Volume 65, September, 103207.

Blanc, E., Schlenker, W. (2017). The Use of Panel Models in Assessments of Climate Impacts on Agriculture. *Review of Environmental Economics and Policy,* volume 11, issue 2: 258–279.

Bozzola, M., Massetti, E., Mendelsohn, R. and Capitanio, F. (2017). A Ricardian analysis of the impact of climate change on Italian agriculture. *European Review of Agricultural Economics,* 43: 217-235.

Brunetti, M., Maugeri, M., Monti, F. and Nanni, T. (2006). Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series, *International Journal Climatology*, 26, 345-381.

Brunetti, M. Lentini, G., Maugeri, M., Nanni, T. Simolo, C., and Spinoni, J. (2012). Projecting North Eastern Italy temperature and precipitation secular records onto a high resolution grid. *Physics and Chemistry of the Earth* 40–41: 9–22.

Brunetti, M., Maugeri, M., Nanni, T., Simoloa, C. and Spinoni, J. (2014). High-resolution temperature climatology for Italy: interpolation method intercomparison. *International Journal of Climatology* 34: 1278–1296

Bucchignani, E., M. Montesarchio, A. L. Zollo and Mercogliano, P. (2015). High-resolution climate simulations with COSMO-CLM over Italy: performance evaluation and climate projections for the 21st century. *International Journal of Climatology* 36: 735–756.

Burke, M., Hsiang, S. and Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, volume 527: 235–239.

Burke, M., Davis, W.M. Diffenbaugh, N.S. (2018). Large potential reduction in economic damages under UN mitigation targets. *Science*, 557: 549–553.

Burke, M. and Tanutama, V. (2019). Climate constraints on aggregate economic output. NBER Working Paper No. 25779.

Carter, C., Cui, X., Ghanem, D. and Mérel, P. (2018). Identifying the Economic Impacts of Climate Change on Agriculture, *Annual Review of Resources Economics*, 10 (25): 1–20.

Carleton, Tamma A., Hsiang, S.M. (2016). Social and Economic Impacts of Climate. *Science* 353(6304): aad9837.

Conte, B., Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2020). Local sectoral specialization in a warming world. NBER Working Paper, (w28163).

Crespi, A. Brunetti, M. Lentini, G. and Maugeri, M. (2018). 1961–1990 high-resolution monthly precipitation climatologies for Italy. *International Journal of Climatology* 38: 878–895.

Cubasch, U., von Storch, H., Waszkewitz, J., Zorita, E. (1996). Estimates of climate change in Southern Europe derived from dynamical climate model output. *Climate Research* **7**: 129–149.

Dasgupta S., Bosello F., De Cian E. and Mistry M. (2017). Global Temperature Effects on Economic Activity and Equity: A Spatial Analysis. CMCC Working paper.

Deryugina, T., & Hsiang, S. M. (2017). The Marginal Product of Climate (NBER Working Paper Series No. 24072). Cambridge, MA.

Dell, M., Jones, B. F., Olken, B.A. (2012) Temperature shocks and Economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics,* 4: 66-95

Dell M., Jones B. F., Olken, B.A. (2014). What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature*, 52: 740-798

Deschênes, O., Greenstone, M. (2007). The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather. *American Economic Review* 97 (1): 354–85.

De Salvo, M., Raffaelli, R. and Moser, R. (2013). The impact of climate change on permanent crops in an Alpine region: a Ricardian analysis. *Agricultural Systems* 118: 23–32.

European Commission, Joint Research Centre. (2003). Global Land Cover 2000 database.

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

Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *Econometrics Journal*, Vol. 3: 148–161.

Hausfather, Z., and Peters, G. P. (2020). Emissions–the 'business as usual' story is misleading. (2020). *Nature* 577, 618-620.

Hsiang, S. (2016). Climate Econometrics. *Annual Review of Resource Economics*, 8: 43-75

Kahn, M. E., Mohaddesby, K.M, Ngb, Ryan N. C., Pesaran, M. H., Raissid, Mehdi and Yange, J-C. (2019). Long-Term Macroeconomic Effects of Climate Change: A Cross-country Analysis. (NBER Working Paper 26167). Cambridge, MA.

Kolstad, C. D. and Moore, Frances C. (2020). Estimating the impact of climate change using weather observation, *Review of Environmental Economics and Policy*, Volume 14, Issue 1: 1–24.

Lemoine (2018). Sufficient Statistics for the Cost of Climate Change (NBER Working Paper No. 25008). Cambridge, MA.

Letta, M. and Tol, R. (2019). Weather, Climate and Total Factor Productivity. *Environmental Resource Economics* **73:** 283–305.

McCallum, S., Dworak, T., Prutsch, A., Kent, N., Mysiak, J., Bosello, F., Klostermann, J., Dlugolecki, A., Williams, E., Konig, M., Leitner, M., Miller, K., Harley, M., Smithers, R., Berglund, M., Glas, N., Romanovska, L., van de Sandt, K., Bachschmidt, R., Voller, S., Horrocks, L. (2013). Support to the development of the EU Strategy for Adaptation to Climate Change: Background report to the Impact Assessment, Part I – Problem definition, policy context and assessment of policy options. Environment Agency Austria, Vienna.

McIntosh, C.T., Schlenker, W. (2006). Identifying non-linearities in fixed effects models. Working Paper, Sch. Int. Relat.

Mendelsohn, R., Massetti, E. (2017). The use of cross-sectional analysis to measure climate impact on agriculture: theory and evidence. *Review of Environmental Economic Policy* 11: 280-98

Mendelsohn R., Nordhaus WD., Shaw D. (1994). The Impact of Global Warming on Agriculture: a Ricardian analysis. *American Economic Review* 84:753-771

Mérel, P, Gammans M. (2021). Climate Econometrics: Can the Panel Approach Account for Long- Run Adaptation?. *American Journal of Agricultural Economics*, https://doi.org/10.1111 /ajae.12200.

Mitchell, T.D., Jones, P.D. (2005). An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*. 25, 693–712.

Moore, F.C., Lobell, D.B. (2014). Adaptation potential of European agriculture in response to climate change. *Nature Climate Change* 4: 610.

New, M., Hulme, M., Jones, P. D. (2000). Representing twentieth-century space-time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate, *Journal of Climate.*, 13, 2217–2238.

Newell G. R., Prest B. C., Sexton E. S. (2021). The GDP Temperature Relationship: Implications for Climate Change Damages. *Journal of Environmental Economics and Management.* In press.

Ortiz-Bobea, A., Knippenberg, E., and Chambers, R.G. (2018). Growing climatic sensitivity of U.S. agriculture linked to technological change and regional specialization. *Science Advances,* 4: eaat4343.

Pretis F., Schwarz, M., Tang, K., Haustein, K., Myles R. A. (2018). Uncertain impacts on economic growth when stabilizing global temperatures at 1.5°C or 2°C warming. *Philosofical Transaction of the Royal Society, A376: 20160460.*

Ricke K., Drouet L., Caldeira K. and Tavoni M. (2018). Country-level social cost of carbon *Nature Climate Change,* Vol 8, October 2018: 895–900.

Schlenker, W. (2006). Inter-annual weather variation and crop yields. Working Paper Department of Economics, Columbia University.

Schlenker, W., Roberts. M. J. (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences of the United States of America* 106 (37):15594–98.

Simolo, C., M. Brunetti, M. Maugeri, T. Nanni, and A. Speranza (2010), Understanding climate change-induced variations in daily temperature distributions over Italy, *J. Geophys. Res.*, 115, D22110, doi:10.1029/2010JD014088.

Tol, R.S.J. (2018). Economic impacts of climate change. *Review of Environmental Economics Policy*, 12(1): 4–25.

Van Passel, S., Massetti, E., Mendelsohn, R. (2017). A Ricardian analysis of the impact of climate change on European agriculture. *Environmental and Resource Economics* 67: 725–760.

Figure 1. Temperature and economic outcomes (GDP per capita and agricultural GVA per-worker).

Notes: Fitted relationships (red line) are obtained using the Fractional Polynomial (fp) regression command in Stata with the respective 95% confidence interval (gray area). Note that the two panels show the logarithmic of the economic outcomes.

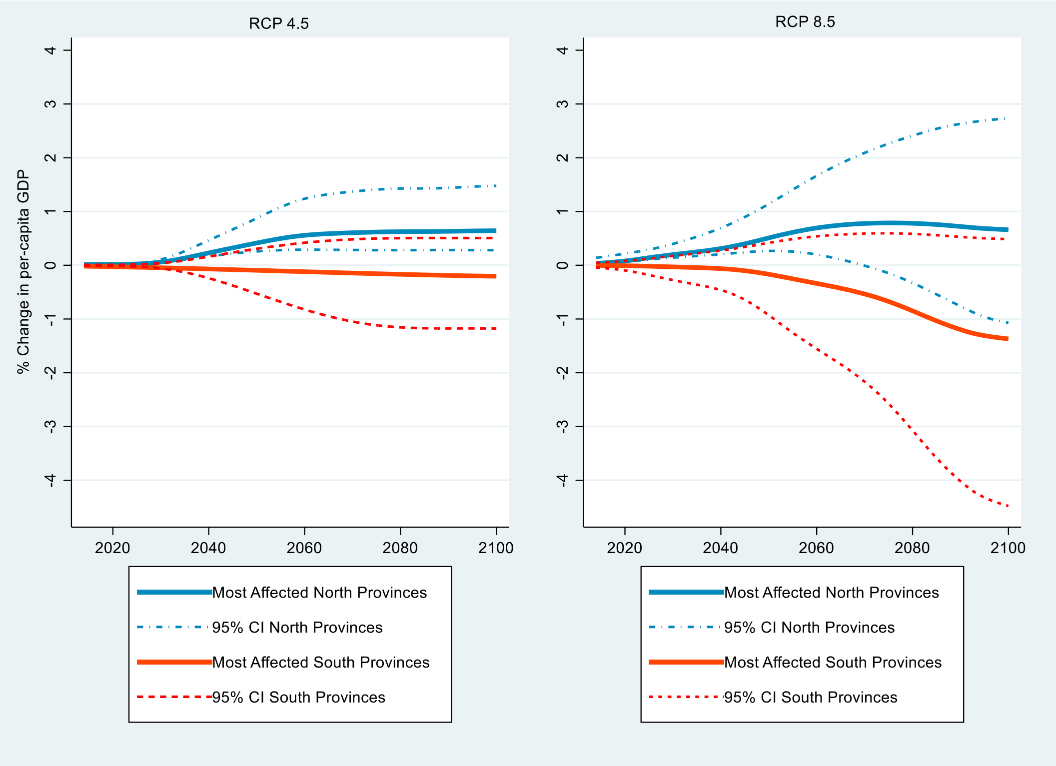


Figure 2. Projected effect of temperature change on per-capita GDP of the most affected North and South provinces.

Notes: Percentage change in GDP per-capita according to COSMO-CLM model projected temperatures. Bold lines are projections using point estimates; dot lines refer to the 95% bootstrap confidence interval. Blue lines refer to estimates of the most affected North provinces (34 provinces of the following Italian Regions: Piemonte, Lombardia, Trentino, Veneto, and Friuli); Red lines refer to estimates of the most affected South provinces (26 provinces of the following Italian Regions: Campania, Calabria, Basilicata; Puglia, Sicilia). These projections are based on the NPS econometric results reported in Table 4, column 5.

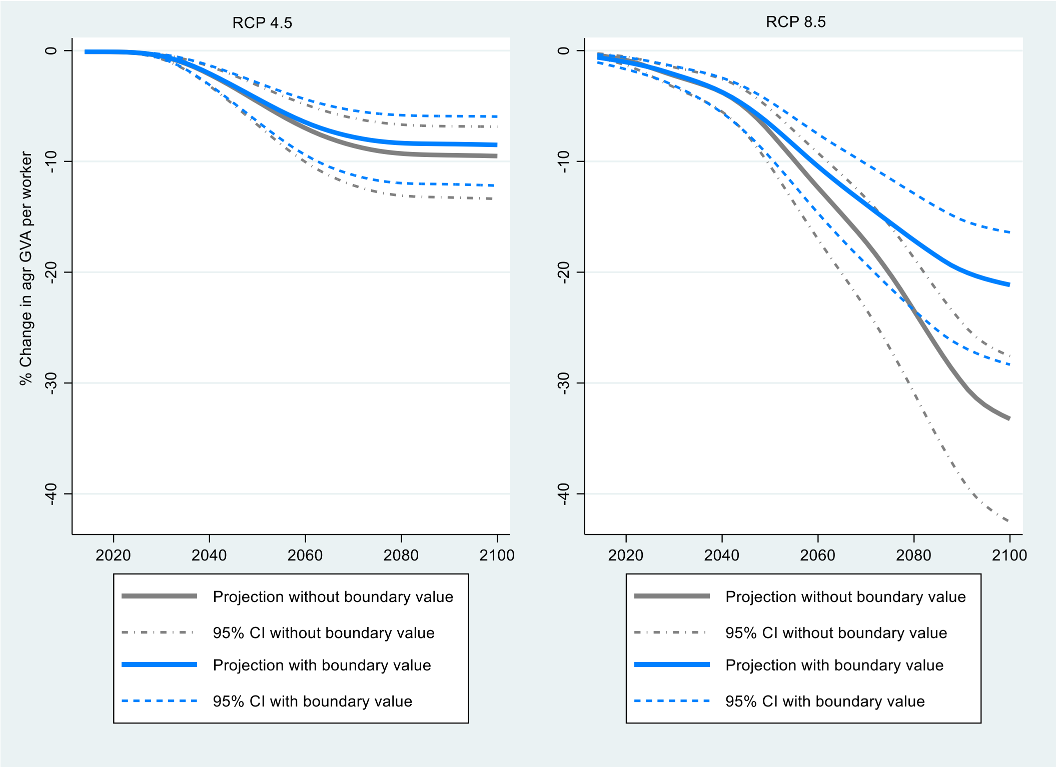
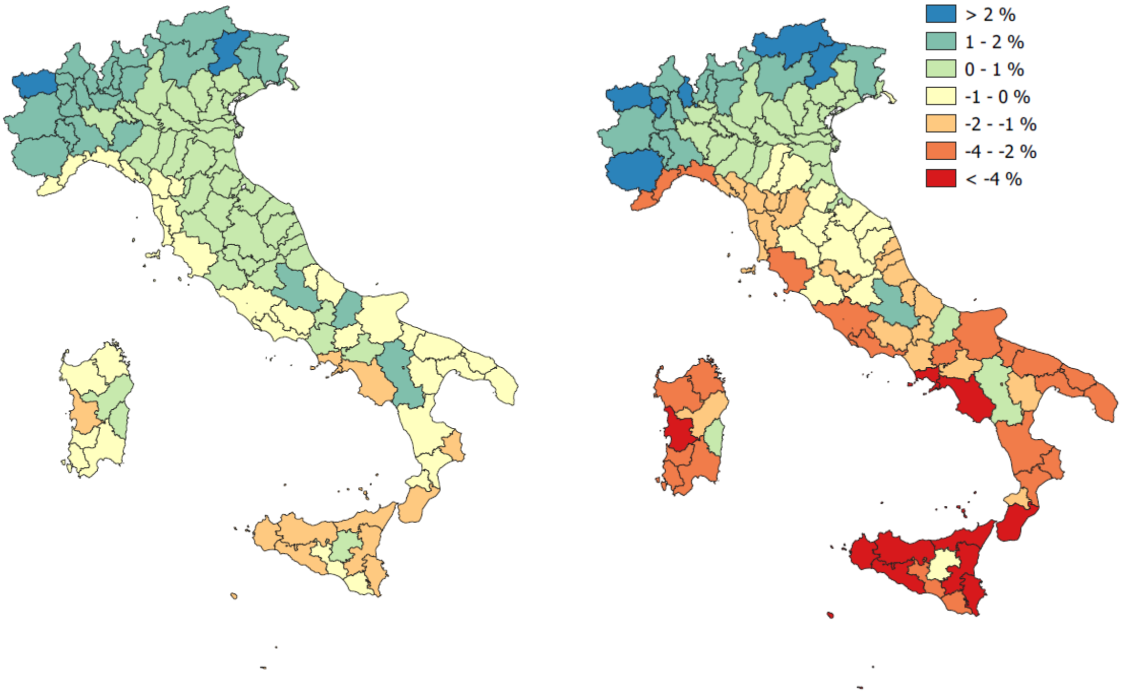


Figure 3. Projected effect of temperature change on agricultural GVA per-worker in Italy.

Notes: Percentage change in agricultural GVA per-worker according to COSMO-CLM model projected temperatures. Bold lines are projections using point estimates; dot lines refer to the 95% bootstrap confidence interval. Blue (grey) lines refer to estimates with temperature bounded (unbounded) to the maximum value of the historical distribution. (See section 5.2). These projections are based on the NPS econometric results reported in Table 5, column 5.



RCP 4.5 RCP 8.5

Figure 4. **Effects of temperature change on per-capita GDP at provincial level.**

Notes: Province (NUTS3) level estimates of the impact of temperature changes on per-capita GDP during 2080-2100 (percent change) under RCP4.5 (left panel) and RCP8.5 (right panel). These projections are based on the NPS econometric results reported in Table 4, column 5.

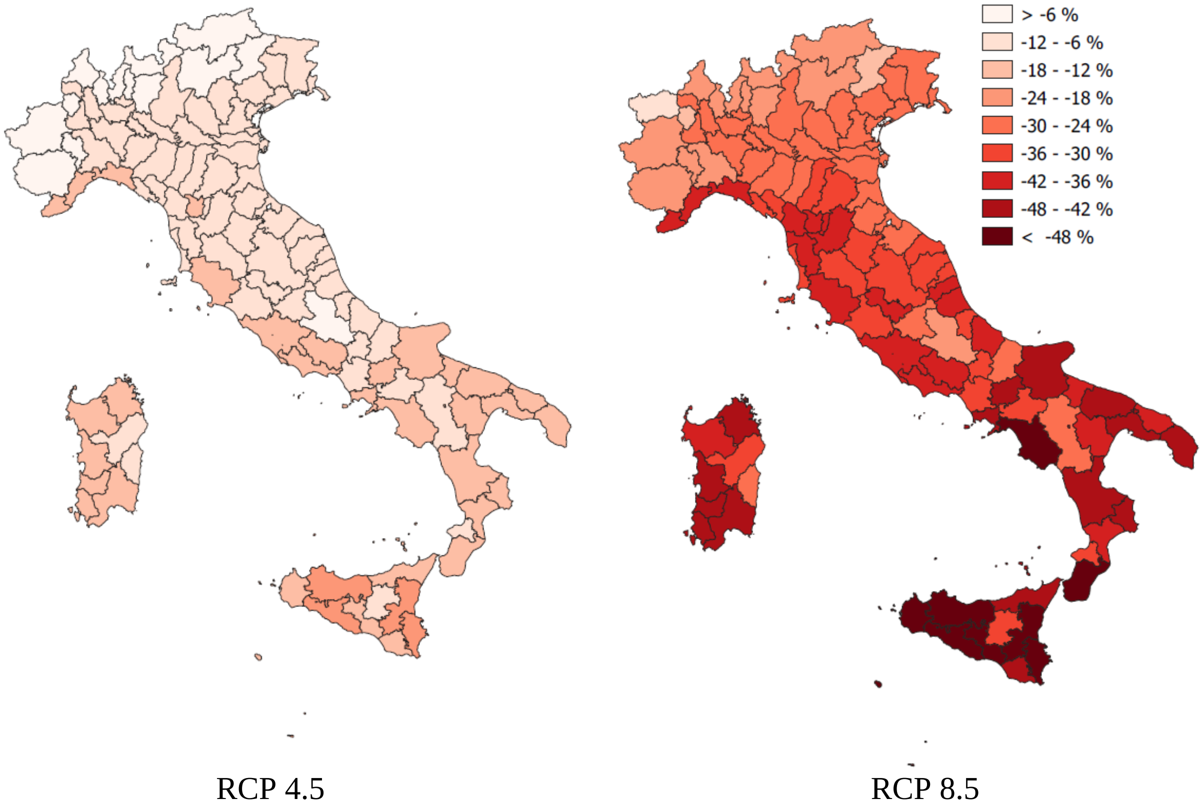


Figure 5. **Effects of temperature change on agricultural GVA per-worker at provincial level.**

Notes: Province (NUTS3) level estimates of the impact of temperature changes on GVA per-worker during 2080-2100 (percent change) under RCP4.5 (left panel) and RCP8.5 (right panel). These projections are based on the NPS econometric results reported in Table 5, column 5.

Table 1. Summary of Models Evaluated in Out-of-Sample Prediction

	Growth vs Level	Linear vs Non-linear effects	Heterogeneity with respect to income	Weather variables	Fixed effects	Trends
DJO1	Growth	Linear	No	Level	NUTS3, Time	No, Yes
DJO2	Growth	Linear	Yes	Level	NUTS3, Time	No, Yes
BHM	Growth	Non-linear (quadratic)	No	Level	NUTS3, Time	No, Yes
NPS	Levels	Non-linear (quadratic)	No	First-difference	NUTS3, Time	No, Yes
KHAN1	Growth	Non-linear (+/- dev. from historical norm)	No	First-difference*	NUTS3	No, Yes
KHAN2	Growth	Non-linear (+/- dev. from historical norm)	Yes	First-difference*	NUTS3	No, Yes

Notes: Table includes key modeling assumptions of models evaluated in the Cross Validation exercise. * In Kahn et al (2019) temperature (and precipitations) entered as first difference of positive/negative deviation from its historical norm (30 years moving average). See text.

Table 2. Forecast Cross-Validation Root Mean Square Error

	GDP per-capita		Agr. GVA per-worker	
DJO1	**0.0147**	0.0150	**0.0831**	0.0849
DJO2	0.0148	0.0150	0.0832	0.0856
BHM	0.0148	0.0151	0.0833	0.0851
NPS	**0.0147**	0.0149	**0.0831**	0.0847
KHAN1	0.0155	0.0158	0.0846	0.0862
KHAN2	0.0157	0.0160	0.0857	0.0871
Province-Time trends	No	Linear	No	Linear

Notes: In forecast CV, we used five training windows that begin in 1981 and end in 1990, 1995, 2000, 2005, and 2010, respectively. Each test period begins in the year following the end of the training window and spans five years, (e.g. 1991-1995, 1996-2000, etc.). The specifications with the lowest RMSE are highlighted in red.

Table 3: Model Comparison Tests for Out-of-Sample Prediction Accuracy

	Welch Test for Equal Forecasting Accuracy (RMSE)					
	GDP per-capita Equation					
	RMSE	DOJ2	BHM	NPS	KHAN1	KHAN2
DOJ1	0.0252	0.71	0.57	0.68	0.94	0.88
DOJ2	0.0253		0.85	0.96	0.77	0.84
BHM	0.0253			0.89	0.63	0.69
NPS	0.0252				0.74	0.80
KHAN1	0.0253					0.94
KHAN2	0.0253					
	Agriculture GVA per-worker Equation					
	RMSE	DOJ2	BHM	NPS	KHAN1	KHAN2
DOJ1	0.1409	0.89	0.96	0.96	0.94	0.98
DOJ2	0.1407		0.93	0.93	0.95	0.92
BHM	0.1405			1.00	0.98	0.99
NPS	0.1402				0.98	0.99
KHAN1	0.1404					0.97
KHAN2	0.1403					

Notes: Table compares alternative model specifications according to out-of-sample prediction accuracy. The first column reports the average root mean squared error from 1000 replications (K-Fold CV). The last five columns present pair-wise *p*-values of a Welch t-test, against the null hypothesis of equal RMSE.

Table 4: Per-capita GDP and Temperature: A Comparison Across Specifications

Dependent variable	GDP Growth (ΔLog(GDP))					
Specification	**DJO1**	**DJO2**	**BHM1**	**BHM2**	**NPS1**	**NPS2**
	(1)	(2)	(3)	(4)	(5)	(6)
Temp / ΔTemp (β_1)	0.00742	0.00823	0.01484	0.00818	0.03189	0.04125
	(0.0026)**	(0.0026)***	(0.0054)***	(0.0115)	(0.0066)***	(0.0119)***
	[0.0059]	[0.0058]	[0.0105]	[0.0164]	[0.0134]**	[0.0181]**
Temp^2 / ΔTemp^2 (β_2)			-0.00026	-0.00001	-0.00091	-0.00120
			(0.0002)	(0.0004)	(0.0002)***	(0.0004)***
			[0.0003]	[0.0005]	[0.00043]**	[0.0006]*
Temp / ΔTemp x Poor (β_3)		-0.00217		0.0053		-0.02865
		(0.0013)*		(0.0142)		(0.0151)*
		[0.0022]		[0.0171]		[0.0190]
Temp^2 / ΔTemp^2 x Poor (β_4)				-0.00023		0.00086
				(0.0005)		(0.0005)*
				[0.0006]		[0.0007]
Precipitation control	Yes	Yes	Yes	Yes	Yes	Yes
Temperature function	Linear	Linear	Quadratic	Quadratic	Quadratic	Quadratic
GDP growth or Levels effects	Growth	Growth	Growth	Growth	Levels	Levels
Max GDP temperature (°C)			28.16		**17.54**	
Max temp. 95% quantile range (°C)			(-18.6, 74.9)		(14.2, 20.9)	
Temperature effect in poor provinces						
($\beta_1 + \beta_3$)		0.0061		0.0134		0.0126
Standard error		(0.0063)		(0.0108)		(0.0137)
($\beta_2 + \beta_4$)				-0.0002		-0.0003
Standard error				(0.0004)		(0.0004)
R^2 (within)	0.441	0.441	0.441	0.442	0.444	0.445
Obs.	3740	3740	3740	3740	3740	3740

Notes: Robust standard errors clustered within NUTS3 (NUTS3 and Region-year) in round (square) brackets, respectively. Temperature (and precipitations) enters linearly in level in columns (1)-(2); non-linearly in level in columns (3)-(4); non-linearly in first difference in columns (5)-(6). Maximizing GDP growth temperatures (and their 95% CI) are obtained through bootstrap; figures in bold (*Italics*) significant at 95% (90%). DJO = Dell, Jones and Olsen (2012); BHM = Burke, Hsiang and Miguel (2015); NPS = Newell, Prest and Sexton (2018).

*, **, *** indicate significance at 90%, 95% and 99% confidence level, respectively.

Table 5: GVA per-worker and Temperature: A Comparison Across Specifications

Dependent variable	GVA per-worker growth (ΔLog(GVA))					
Specification	**DJO1**	**DJO2**	**BHM1**	**BHM2**	**NPS1**	**NPS2**
	(1)	(2)	(3)	(4)	(5)	(6)
Temp / ΔTemp (β_1)	-0.03514	-0.02760	0.03215	0.07168	0.14835	0.11374
	(0.0147)**	(0.0147)*	(0.0325)	(0.0530)	(0.0455)***	(0.0803)
	[0.0270]	[0.0274]	[0.0590]	[0.1078]	[0.0673]**	[0.1120]
Temp^2 / ΔTemp^2 (β_2)			-0.00226	-0.00344	-0.00593	-0.00454
			(0.0010)**	(0.0017)*	(0.0014)***	(0.0027)*
			[0.0019]	[0.0038]	[0.0021]***	[0.0038]
Temp / ΔTemp x Poor (β_3)		-0.02340		-0.14687		-0.02550
		(0.0071)***		(0.0698)**		(0.0939)
		[0.0109]**		[0.1097]		[0.1108]
Temp^2 / ΔTemp^2 x Poor (β_4)				0.00431		0.00016
				(0.0023)*		(0.0032)*
				[0.0039]		[0.0039]
Precipitation control	Yes	Yes	Yes	Yes	Yes	Yes
Temperature function	Linear	Linear	Quadratic	Quadratic	Quadratic	Quadratic
GVA growth or Levels effects	Growth	Growth	Growth	Growth	Levels	Levels
Max GVA temperature (°C)			7.10		**12.52**	
Max temp. 95% quantile range (°C)			(-8.9, 23.17)		(9.6, 15.4)	
Temperature effect in poor provinces						
($b_1 + b_3$)		*-0.0509*		-0.0752		0.0882
Standard error		(0.0282)		(0.0680)		(0.0730)
($b_2 + b_4$)				0.0009		*-0.0044*
Standard error				(0.0022)		(0.0023)
R^2 (within)	0.222	0.225	0.225	0.229	0.226	0.228
Obs.	3740	3740	3740	3740	3740	3740

Notes: Robust standard errors clustered within NUTS3 (NUTS3 and Region-year) in round (square) brackets, respectively. Temperature (and precipitations) enters linearly in level in columns (1)-(2); non-linearly in level in columns (3)-(4); non-linearly in first difference in columns (5)-(6). Maximizing agricultural GVA growth temperatures (and their 95% CI) are obtained through bootstrap; figures in bold (*Italics*) significant at 95% (90%). DJO = Dell, Jones and Olsen (2012); BHM = Burke, Hsiang and Miguel (2015); NPS = Newell, Prest and Sexton (2018).

*, **, *** indicate significance at 90%, 95% and 99% confidence level, respectively.

Table 6: GVA per-worker and Precipitations: A Comparison Across Specifications

Dependent variable	GVA per-worker growth (ΔLog(GDP))					
Specification	**DJO1**	**DJO2**	**BHM1**	**BHM2**	**NPS1**	**NPS2**
	(1)	(2)	(3)	(4)	(5)	(6)
Prec. / ΔPrec. (β_1)	0.00619	0.00382	0.01923	0.00851	0.00692	0.00374
	(0.0026)***	(0.0028)	(0.0056)***	(0.0074)	(0.0049)	(0.0067)
Prec.^{2} / ΔPrec.^{2} (β_2)			-0.00061	-0.00021	-0.00014	-0.00005
			(0.0002)***	(0.0002)	(0.0002)	(0.0003)
Prec. / ΔPrec. x Poor (β_3)		0.00482		0.01920**		0.00584
		(0.0035)		(0.0088)		(0.0076)
Prec.^{2} / ΔPrec.^{2} x Poor (β_4)				-0.0008		-0.000155
				(0.0004)**		(0.0003)
Temperature control	Yes	Yes	Yes	Yes	Yes	Yes
Temperature function	Linear	Linear	Quadratic	Quadratic	Quadratic	Quadratic
GVA growth or Levels effects	Growth	Growth	Growth	Growth	Levels	Levels
Max GVA precipitations (100 mm)			**15.85**		25.12	
Max prec. 95% quantile range (100 mm)			(12.1, 19.6)		(-10.9, 61.18)	
Precipitation effect in poor provinces						
($b_1 + b_3$)		**0.0086**		**0.0277**		*0.0095*
Standard error		(0.0035)		(0.0069)		(0.0056)
($b_2 + b_4$)				**-0.0010**		-0.0002
Standard error				(0.0003)		(0 .0002)
R^2 (within)	0.222	0.225	0.225	0.229	0.226	0.228
Obs.	3740	3740	3740	3740	3740	3740

Notes: Robust standard errors clustered within NUTS3 and region-year in brackets. Temperature (and precipitations) enters linearly in level in columns (1)-(2); non-linearly in level in columns (3)-(4); non-linearly in first difference in columns (5)-(6). Maximizing agricultural GVA growth precipitations (and their 95% CI) are obtained through bootstrap; figures in bold (*Italics*) significant at 95% (90%). DJO = Dell, Jones and Olsen (2012); BHM = Burke, Hsiang and Miguel (2015); NPS = Newell, Prest and Sexton (2019).

*, **, *** indicate significance at 90%, 95% and 99% confidence level, respectively.

Appendix

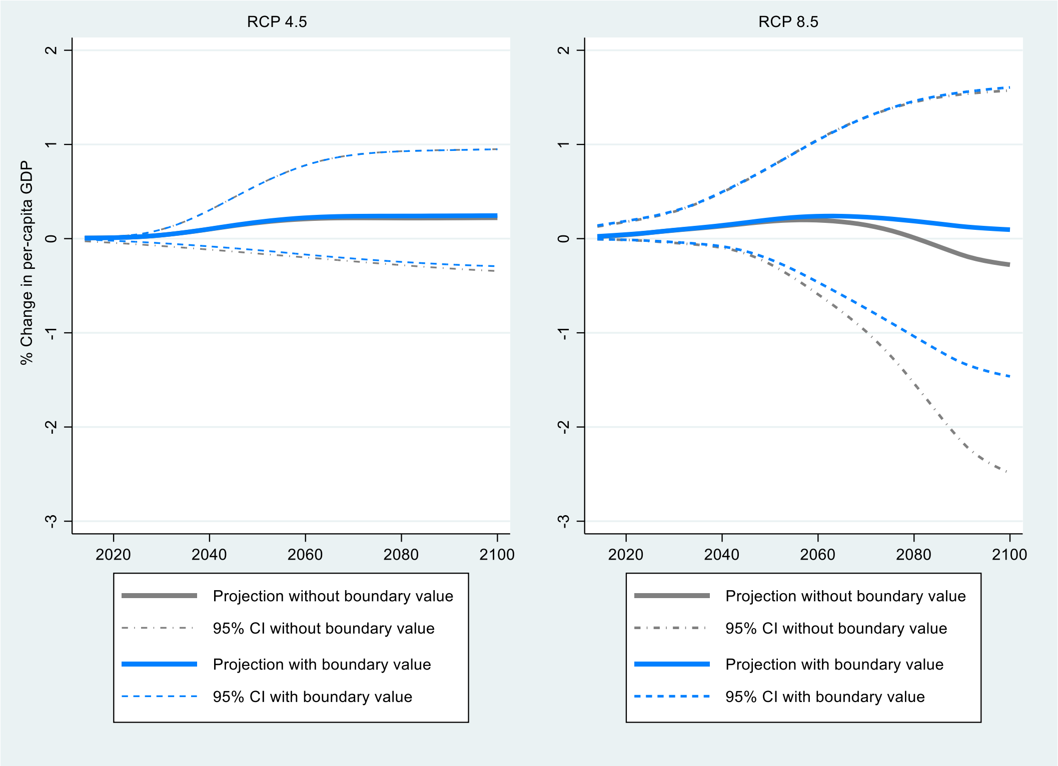


Figure A.1. Projected effect of temperature change on per-capita GDP in Italy.

Notes: Percentage change in GDP per-capita according to COSMO-CLM model projected temperatures. Bold lines are projections using point estimates; dot lines refer to the 95% bootstrap confidence interval. Blue (grey) lines refer to estimates with temperature bounded (unbounded) to the maximum value of the historical distribution. (See section 5.2). These projections are based on the NPS econometric results reported in Table 4, column 5.

Table A.1 Backcast Cross-Validation Root Mean Square Error

	GDP per-capita		Agr. GVA per-worker	
DJO1	**0.0134**	0.0136	0.0821	0.0828
DJO2	0.0135	0.0136	0.0821	0.0829
BHM	0.0135	0.0137	0.0821	0.0828
NPS	**0.0134**	0.0136	**0.0819**	0.0824
KHAN1	0.0158	0.0162	0.0941	0.0948
KHAN2	0.0160	0.0164	0.0945	0.0951
Province-Time trends	No	Linear	No	Linear

Note: We used five training windows that end in 2015 and begin in 1986, 1991, 1996, 2001, and 2006, respectively. Each test period spans the five years preceding the beginning of the training window (e.g. 1981-1985, 1986-1990, etc.). The specifications with the lowest RMSE are highlighted in red.

Table A2: K-fold Cross-Validation Root Mean Square Error

	GDP per-capita		Agr. GVA per-worker	
DJO1	**0.0252**	0.0258	0.1409	0.1434
DJO2	0.0253	0.0257	0.1407	0.1433
BHM	0.0253	0.0256	0.1405	0.1424
NPS	**0.0252**	0.0257	**0.1402**	0.1429
KHAN 1	0.0253	0.0258	0.1404	0.1430
KHAN 2	0.0253	0.0258	0.1403	0.1429
Province-Time trends	No	Linear	No	Linear

Note: in K-fold CV, the data are first centered by regressing all variables (dependents and explanatory variables) on country fixed-effects and country-specific linear and quadratic trends. The residuals from these regressions were used for the cross validation calculating the Root Mean Squared Error. The procedure splits the data randomly into k partitions, then for each partition it fits the specified model using the other k-1 groups and uses the resulting parameters to predict the dependent variable in the unused group. The specifications with the lowest RMSE are highlighted in red.

Table A3: Long-Run Effect of Weather Change on per-capita GDP Growth (Historical Norms as the Moving Averages of Past 30 Years)

	GDP per-capita Growth Rate (M=30)					
	Specification 1		**Specification 2**		**Specification 3**	
	xtreg	xtspj	xtreg	xtspj	xtreg	xtspj
Sum_delta_temp_+	-0.0102***	-0.0113*	-0.0142***	-0.0190***	-0.0204***	-0.0258***
	(0.0039)	(0.0062)	(0.0040)	(0.0057)	(0.0052)	(0.0142)
Sum_delta_temp_-	-0.0299	-0.0665**	-0.0142	-0.0444	0.0542*	0.0322
	(0.0290)	(0.0313)	(0.0293)	(0.0361)	(0.0306)	(0.0425)
Sum_delta_prec_+ (x 100)	-0.0015	-0.0012				
	(0.0015)	(0.0020)				
Sum_delta_prec_- (x 100)	0.0015	-0.0015				
	(0.0018)	(0.0024)				
Sum_delta_temp_+ x I(province i is poor)					0.0107	0.0119
					(0.0079)	(0.0118)
Sum_delta_temp_- x I(province i is poor)					-0.1628***	-0.1864***
					(0.0585)	(0.0636)
Sum_lagged_y	0.7042***	0.5394***	0.7284***	0.5682***	0.7192***	0 .4670***
	(0.0337)	(0.0266)	(0.0327)	0.0273	(0.0320)	(0.0300)
No of Provinces (N)	110	110	110	110	110	110
T	31	31	31	31	31	31
No of Observations (N x T)	3410	3410	3410	3410	3410	3410

Notes: Specification 1 is given by $\Delta y_{it} = a_i + \sum_{l=1}^{p} \varphi_l \Delta y_{i,t-l} + \sum_{l=0}^{p} \beta_l \Delta \boldsymbol{x}_{i,t-l} + \varepsilon_{it}$, where y_{it} is the log of real GDP per capita of province i in year t, $\boldsymbol{x}_{i,t} = (delta\_temp_{it}^{+}, delta\_temp_{it}^{-}, delta\_prec_{it}^{+}, delta\_prec_{it}^{-})$ with $delta\_temp_{it}$ and $delta\_prec_{it}$ the average temperature and precipitations deviations (positive and negative) from the historical norms (moving averages of the past 30 years), and p=3 the number of lags. The long-run effect, Sum_i, are calculated from the OLS estimates of the short-run coefficients: $sum = \phi^{-1} \sum_{l=0}^{p} \beta_l$ with $\phi^{-1} = 1 - \sum_{l=1}^{p} \varphi_l$. Specification 2 drops precipitation variables from the model 1. Specification 3 interacts the temperature variables with a dummy for "poor" provinces (see text). xtreg means fixed effects model; xtspj means the half-panel jackknife fixed effects estimator of Chudik et al. (2018). Standard errors clustered by provinces in parenthesis.

*, **, *** indicate significance at 90%, 95% and 99% confidence levels, respectively.

Table A4: Long-Run Effect of Weather Change on Agriculture GVA per-worker Growth (Historical Norms as the Moving Averages of Past 30 Years)

	GVA per-worker Growth Rate					
	Specification 1		Specification 2		Specification 3	
	xtreg	xtspj	xtreg	xtspj	xtreg	xtspj
Sum_delta_temp_+	-0.1174***	-0.1129***	-0.1149***	-0.1166***	-0.0541*	-0.0521
	(0.0225)	(0.0240)	(0.0215)	(0 .0289)	(0.0289)	(0.0336)
Sum_delta_temp_-	-0.1216	-0.2158*	-0.1408	-0.2228*	0.0990	0.0545
	(0.0908)	(0.1276)	(0.0864)	(0.1264)	(0.1156)	(0.1689)
Sum_delta_prec_+ (x 100)	-0.0004	0.0023				
	(0.0057)	(0.0077)				
Sum_delta_prec_- (x 100)	-0.0129	-0.0156				
	(0.0092)	(0.0095)				
Sum_delta_temp_+ x I(province i is poor)					-0.1156***	-0.1226***
					(0.0441)	(0.0459)
Sum_delta_temp_- x I(province i is poor)					-0.4294***	-0.5125**
					(0.1779)	(0.2583)
Sum_lagged_y	1.2630***	1.1820***	1.2567***	1.1744***	1.2478***	1.1663***
	(0.0376)	(0.0306)	(0.0371)	(0.0307)	(0.0376)	(0.0307)
No of Provinces (N)	110	110	110	110	110	110
T	31	31	31	31	31	31
No of Observations ($N\ x\ T$)	3410	3410	3410	3410	3410	3410

Notes: Specification 1 is given by $\Delta y_{it} = a_i + \sum_{l=1}^{p} \varphi_l \Delta y_{i,t-l} + \sum_{l=0}^{p} \beta_l \Delta \boldsymbol{x}_{i,t-l} + \varepsilon_{it}$, where y_{it} is the log of real GDP per capita of province i in year t, $\boldsymbol{x}_{i,t} = (delta\_temp_{it}^{+}, delta\_temp_{it}^{-}, delta\_prec_{it}^{+}, delta\_prec_{it}^{-})$ with $delta\_temp_{it}$ and $delta\_prec_{it}$ the average temperature and precipitations deviations (positive and negative) from the historical norms (moving averages of the past 30 years), and p=3 the number of lags. The long-run effect, Sum_i, are calculated from the OLS estimates of the short-run coefficients: $sum = \phi^{-1} \sum_{l=0}^{p} \beta_l$ with $\phi^{-1} = 1 - \sum_{l=1}^{p} \varphi_l$. Specification 2 drops precipitation variables from the model 1. Specification 3 interact the temperature variables with a dummy for “poor” provinces (see text). xtreg means fixed effects estimator; xtspj means half-panel jackknife fixed effects estimator of Chudik et al. (2018). Standard errors clustered by provinces in paratheses.

*, **, *** indicate significance at 90%, 95% and 99% confidence levels, respectively.

Table A5: Percent Loss in GDP per capita and GVA per-worker by 2050 and 2100 under the RCP4.5 and RCP8.5 Scenarios, based on Kahn et al. (2019)

	% loss in GDP per-capita		**% loss in GVA per-worker**	
	Year 2050 (h = 36)	Year 2100 (h = 86)	Year 2050 (h = 36)	Year 2100 (h = 86)
Italy				
RCP 4.5	-0.06%	-0.15%	-0.08%	0.02%
RCP 8.5	1.01%	2.72%	3.89%	10.44%
Rich provinces				
RCP 4.5	-0.19%	-0.51%	-0.16%	-0.40%
RCP 8.5	1.18%	3.43%	2.01%	5.16%
Poor provinces				
RCP 4.5	0.11%	0.44%	0.16%	0.27%
RCP 8.5	0.63%	1.85%	5.26%	13.87%

Notes: Positive (negative) numbers mean loss (gains) from global warming. Figures represent the impact of persistent increases in temperatures based on the RCP 4.5 and RCP 8.5 scenarios, on the GDP per-capita and GVA per-worker projections in 2050 and 2100. To obtain these estimates we used equations 22 in Kahn et al. (2019), and econometric estimated coefficients from XTSPJ specifications 2 (average effects for Italy) and XTSPJ specification 3 (average effects for rich and poor provinces) of Tables A.4 and A.5, respectively. Numbers are GDP (2014) weighted averages of the estimated impacts at provinces level. See Supporting Information, Section E, for details and additional results at province level.

Table A6: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Log (GDP per-capita)	3,850	2.981	0.316	1.792	3.869
Growth GDP per-capita	3,740	0.008	0.034	-0.210	0.308
Log (GVA per-worker)	3,850	3.370	0.601	0.956	5.209
Growth GVA per-worker	3,740	0.024	0.159	-0.855	0.803
Temperature (°C)	3,850	14.995	1.957	9.575	20.217
Temperature^2	3,850	228.669	59.788	91.681	408.714
Precipitation (100 mm)	3,850	8.194	3.072	1.981	27.120
Precipitation^2	3,850	76.578	64.329	3.924	735.494
Δ (Temperature)	3,850	0.052	0.572	-1.458	1.833
Δ (Temperature)^2	3,850	1.563	17.005	-43.953	47.147
Δ (Precipitation)	3,850	0.025	2.513	-16.503	13.835
Δ (Precipitation)^2	3,850	0.744	55.536	-622.774	559.003
Temp_dev_positive (m=30)	3,850	0.586	0.405	0	1.606
Temp_dev_negative (m=30)	3,850	0.035	0.115	0	0.856
Precip_dev_posisitive (m=30)	3,850	68.336	125.775	0	1050.865
Precip_dev_negative (m=30)	3,850	79.445	101.651	0	681.32
Poor dummy	3,850	0.50	0.50	0	1

Source: See text.