Agricultural Innovation and Adaptation to Climate Change: Insights from Genetically Engineered Maize

Seungki Lee, Yongjie Ji, GianCarlo Moschini

Working Paper 21-WP 616
Updated
June 2022

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

Seungki Lee is Assistant Professor, Department of Agricultural, Environmental, and Developmental Economics, The Ohio State University, Columbus, OH 43210. E-mail: lee.10168@osu.edu

Yongjie Ji is Assistant Scientist, Center for Agricultural and Rural Development, Iowa State University, Ames, IA 50011. E-mail: yongjiej@iastate.edu

GianCarlo Moschini is Professor, Department of Economics, Iowa State University, Ames, IA 50011. E-mail: moschini@iastate.edu

For questions or comments about the contents of this paper, please contact GianCarlo Moschini, moschini@iastate.edu.

Agricultural Innovation and Adaptation to Climate Change:

Insights from Genetically Engineered Maize

Seungki Lee, Yongjie Ji, and GianCarlo Moschini *

Abstract. Climate change is one of the major threats to the global food supply, and adaptation by technological progress is believed to be essential. What is the scope of the required innovation tasks? To address this question, we estimate the yield gain in US maize production due to a major novel technology: genetically engineered (GE) varieties. Next, the yield model is used to extrapolate future expected yields given climate change projections from twenty large-scale models and two warming scenarios. We find that climate change entails significant yield shortfalls. The scale of these yield gaps, by the end of the century, ranges from about 2.8 to 6.3 times the total yield gains from GE varieties. These results suggest that the scope of adaptation is challenging. Ambitious and targeted R&D efforts, and innovation breakthroughs, may be required to close the yield gaps likely to arise from climate change.

Keywords: Adaptation, Agricultural Productivity, Climate Change, Genetically Engineered Varieties, Innovation, Maize.

JEL codes: Q16, O47, C23

* Seungki Lee is an Assistant Professor, Department of Agricultural, Environmental, and Development Economics, The Ohio State University. Yongjie Ji is an assistant scientist, and GianCarlo Moschini is Charles F. Curtiss Distinguished Professor and Pioneer Chair in Science and Technology Policy, both in the Center for Agricultural and Rural Development and the Department of Economics, Iowa State University.

This research was partially supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, grant No. 2018-67023-27682.

1. Introduction

Production agriculture is known to depend heavily on exogenous environmental conditions, including weather. As such, agriculture is acutely vulnerable to the deleterious long-run effects of climate change. Indeed, mounting evidence suggests large negative impacts (Mendelsohn, Nordhaus, and Shaw 1994; Fisher et al. 2012). Specifically, the yields of major staple crops will likely be adversely affected (Schlenker and Roberts 2009; Lobell, Schlenker, and Costa-Roberts 2011; Tack, Barkley, and Nalley 2015; Chen, Chen, and Xu 2016; Gammans, Merel, and Ortiz-Bobea 2017; Miller, Tack, and Bergtold 2021). The long-run health of the food supply may thus need deliberate mitigation and/or adaptation measures to deal with global warming. First-best mitigation strategies aimed at containing climate change, chiefly by reducing greenhouse gas (GHG) emissions, are proving problematic, due to the global and dynamic nature of the externalities involved with climate change (Carattini, Levin, and Tavoni 2019). Adaptation strategies, to blunt and counteract the damaging consequences of climate change, are perhaps more promising because they are less vulnerable such strategic problems—whereas GHG emission reduction is a global public good, investments in adaptation often have local payoffs and substantial private good aspects (Tol 2005; Hasson, Löfgren, and Visser 2010).

Whereas much hope must rest on adaptation, agriculture's ability to adapt to climate change includes unresolved questions. Successful efforts to cope with a hostile environment are a major component of the history of agriculture (Olmstead and Rhode 2011). Yet, some have suggested limited adaptation potential (Burke and Emerik 2016), although teasing out adaptability from observed yield responses to weather fluctuations involves subtle econometric considerations (Carter et al. 2018). In any event, adaptation depends largely on purposeful and targeted investments.[1] In particular, innovation has been recognized as a major element of adaptation to

[1] Perhaps with some risk of oversimplification, Tol (2018) puts it thus: "Adaptation to weather shocks is therefore limited to immediate responses: put up an umbrella when it rains, close the flood doors when it pours. In contrast, adaptation to climate change extends to changes in the capital stock: buy an umbrella, invest in flood doors."

climate change (Stern 2007), which highlights the critical role of research and development (R&D) investments (Lobell et al. 2008; Lybbert and Sumner 2012).

Technologies to foster agriculture's adaptation may include, inter alia, new crop varieties with traits enhancing their broad resistance to pest, disease, and environmental stress, particularly heat tolerance and resistance to drought and salinity (Lybbert and Sumner 2012). Varieties with shorter growing cycles and earlier maturation, precision agriculture technologies, and more efficient water management and expanded irrigation are also expected to be critical. Technological innovations beyond the farm level are also envisioned (Zilberman et al. 2018), including institutional innovation with a focus on adoption incentives and an appreciation for the role of learning, networks, and social capital (Zilberman, Zhao, and Heiman 2012). Ultimately, all this requires major R&D investments, from both the public and private sectors, to support enhanced innovation efforts in adaptation-enabling new technologies.

What is the scope of such an R&D challenge? Harnessing the potential of modern biotechnology, beyond traditional breeding, is obviously critical in this setting. Thus, in this paper, we propose to use the estimated impact of first-generation genetically engineered (GE) varieties in US maize production as a yardstick. Specifically, we estimate the contribution of GE traits separate from the long-run productivity improvements that have characterized maize yields with the diffusion of hybrid varieties. The estimated model, along with weather projections from mainstream climate change models, permit us to forecast expected yields at mid-century and end-of-century, and thus characterize the ceteris paribus "yield gaps" due to anticipated climate change. Comparison of such yield gaps with the one-time yield gains due to first-generation GE traits in maize provides a useful metric for the innovation challenge posed by climate change.

First commercialized in 1996, GE seeds rapidly replaced conventional varieties, and have exceeded 80% of plantings in maize since 2008 and in soybeans since 2003. In maize, commercially successful first-generation GE varieties have embedded agronomic traits—herbicide tolerance (HT), chiefly tolerance to glyphosate (aka Roundup), and insect resistance

(IR), specifically resistance to the European corn borer and corn rootworms. Farmers' keen interest in GE adoption attests to their perceived profitability, despite higher seed prices (Ciliberto et al., 2019). Whether and how GE traits affected yields is a somewhat distinct question. The National Academy of Sciences, Engineering, and Medicine opined that "the nation-wide data on maize, cotton, or soybean in the United States do not show a significant signature of genetic-engineering technology on the rate of yield increase" (NASEM 2016). But in fact, experimental evidence points to a significantly positive impact of GE on maize yields (Nolan and Santos 2012; Chavas, Shi, and Lauer 2014), and observational data on realized yields also shows a clear positive impact (Xu et al. 2013; Huffman, Jin, and Xu 2018; Lusk, Tack, and Hendricks 2019).

To characterize the impact of GE maize varieties on yield, as in related studies, we use a panel of county-level yield data. Our approach improves on previous work by a somewhat more nuanced measurement of maize GE adoption rates, and by an explicitly representation of the interaction effects between GE adoption and weather variables. Our GE adoption data exploit a large set of farm-level seed choices by US farmers assembled by Kynetec USA. Existing studies have been constrained by US Department of Agriculture (USDA) data on GE adoption rates, provided only for a limited number of states and, perhaps most importantly, only available starting in year 2000 (and thus missing the first four critical years of GE trait diffusion). Similar to previous work, our yield model is estimated conditional on historical weather metrics based on daily temperature and precipitation from the Parameter-elevation Regressions on Independent Slopes Model (PRISM). Our analysis covers 1,774 counties in 36 states over the 1981 to 2020 period, for a total of 60,400 panel observations.

Our results confirm the finding of some existing studies (Xu et al. 2013; Lusk, Tack, and Hendricks 2018; Ortiz-Bobea and Tack 2018) that GE varieties have led to significant productivity gains in maize production. Specifically, our base model (without interaction between GE and weather) suggests that full adoption of GE traits is associated with yield gains of about 16.6 bushels per acre. We find that the interaction effects between maize GE traits and weather variables are important, with the net interaction effects reducing the yield gain

attributable to GE varieties by about 2.4 bushels per acre. The yield gains from the underlying technical progress, separate from the adoption of GE traits, is estimated at about 0.95 bushel per acre per year (average across all US counties). Similar to other studies, we also find that accounting for weather conditions is essential in order to identify the role of technology in maize production—yields are significantly positively impacted by growing degree days, are negatively impacted by excess heat, and are sensitive to precipitation and water stress.

The estimated model is used for counterfactual simulations to determine the expected yield impacts of anticipated climate change. Our analysis relies on weather projections from all 20 global climate models (GCMs) from the Multivariate Adaptive Constructed Analogs (MACA) dataset (Abatzoglou and Brown 2012). Mid-century (2040–2059) and end-of-century (2080–2099) weather predictions and future climate conditions are used to compare scenarios under two GHG concentration pathways. We find sizeable yield shortfalls due to the changing weather predicted by these GCMs. Comparing such "yield gaps" with the model-estimate yield gains attributable to GE varieties provide a useful characterization of the innovation-adaptation challenge posed by climate change. Depending on GHG concentration pathways, average yield gaps due to climate change at mid-century range from 2.0 times to 2.7 times the entire yield gains made possible by the adoption of GE varieties. By the end of the century, the estimated average yield gaps range from 2.9 times to 6.3 times the GE yield gain.

Following Burke et al. (2015), we separate model uncertainty and climate uncertainty, and find that climate uncertainty accounts for virtually all of the variance of estimated yield shortfalls. We also find that the projections based on HadGEM2-ES and NorESM1-M, perhaps the two most widely used models to contrast the impact of future climate predictions on agricultural productivity, turn out to be quite pessimistic and not representative of the ensemble means of the full set of GCMs.

Our main conclusion is plainly stated. If future maize yield gains are to remain on the trajectory expected under historical climate conditions, considerable additional R&D investments into the development of suitable adapting technology are needed. To appreciate the comparison with

the historical impact of first-generation GE seeds, it is important to appreciate the distinctive attributes of that experience. The technology of GE crop varieties was developed in the 1980s from the application of revolutionary recombinant DNA techniques discovered in the 1970s (Moschini 2008; Bennet et al. 2013). The commercialization, and eventual widespread adoption of GE varieties in maize, soybeans, and cotton, was made possible by massive R&D investments, mostly by the private sector. In the process, the company leading this development, Monsanto, radically transformed itself from a chemical concern to the largest seed company in the world (Clancy and Moschini 2017). That the yield effects needed to offset the likely impacts of global climate change are, possibly, several times larger than what was accomplished with the development and diffusion of first-generation maize GE varieties underscores the scope of the inherent R&D challenge.

2. Data

The main variable to be explained is the US maize yield at the county level. Similar to other studies in this area, the focus is on rainfed agriculture. Specifically, as in Xu et al. (2013), counties are included in the sample if the fraction of harvested cropland that is irrigated is less than 10% (based on the USDA census of 2002). Furthermore, the analysis is limited to data since 1981, as in Ortiz-Bobea and Tack (2018), to ensure that daily precipitation and temperature data are available. **Figure X1** in the Supplementary Appendix illustrates the geographic area explored in our analysis, which includes 1,774 counites across 36 states.

2.1 Yields

The agricultural productivity metric of interest is the maize yield (bushel/acre). We draw on county-level maize yield data from the USDA-NASS for the period from 1981 to 2020. **Figure 1** illustrates the observed yield pattern with spatial and temporal variations. Note that NASS does not provide maize production and yield data for all counties over the entire period of interest—some marginal counties may not appear in some years. On average, 1,510 counties are observed per year, which is about 85% of the total number of counties we considered. Hence, the final data we end up using in estimation has the nature of an unbalanced panel.

2.2 GE adoption

In this paper we exploit a more refined measure of GE adoption rates than used by previous studies. The USDA survey-based state-level GE adoption data used by previous studies have two drawbacks: they are available only staring in year 2000, thereby missing the crucial early years of GE diffusion, and they cover only a limited number of large maize-growing states.[2] Our adoption data are largely constructed from an extensive set of farm-level observations of seed choices by US maize farmers assembled by Kynetec USA, a market research organization that collects agriculture-related survey data. These proprietary data are based on annual surveys of random, large samples of US farmers (approximately 4,700 maize farmers every year). The data include the quantity and attributes (including embedded GE traits) of farmers' seed purchases, along with the projected acres planted to each seed variety. These data are available to us from 1996 to 2016, thereby encompassing the entire period from the introduction of GE traits to their virtually complete adoption. Based on these data, we are able to construct reliable adoption rates for 21 states.[3] The pattern of spatial and temporal variation of these GE adoption rates are illustrated in **Figure 2**. We can see that adoption rates are spatially heterogeneous, especially in the first half of the adoption period, but this variability is reduced with adoption rates converging to more than 90% in the last few years.

2.3 Weather data

Daily temperature and precipitation are procured from the 4km by 4km grid cell PRISM dataset, and are used to calculate the model's weather variables.[4] The county-level temperature

[2] Specifically, 11 states starting in 2000 (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio, South Dakota, and Wisconsin), with two more states added starting in 2005 (North Dakota and Texas).

[3] Details on data handling are provided in the Supplementary Appendix. For the last four year of analysis, Kynetec data are supplemented with USDA data, without much loss of generality because adoption in the last few years has been stationary.

[4] PRISM Climate Group, historical daily data. Available online at http://prism.oregonstate.edu/.

and precipitation variables are obtained via an area-weighted scheme (Sacks et al. 2010). Specifically, we first intercept county polygon shape file with the PRISM grid cells in ArcGIS to get the overlapping areas, then we construct county-level weather variables by averaging these variables at the grid cells that are overlapped with a county. For example, if there are $i = 1,\ldots,I$ PRISM cells overlapped with county c, with overlapped area of A_{ic}, then the weather variable of interest for this particular county, W_c, is constructed as

$$W_c = \sum_{i=1}^{I} \frac{A_{ic}}{\sum_{r=1}^{I} A_{rc}} W_i$$

where W_i is the value of the weather variable for grid cell i. As explained in detail in what follows, the weather variables of interest are temperature and precipitation.

2.3.1 Heat: Growing degree days and excess heat degree days

Two standard metrics of weather heat used in modeling crop yields are growing degree days (GDD) and excess heat degree days (HDD) (Roberts, Schlenker and Eyer 2013). We adopt the specification of Xu et al. (2013), who construct GDDs and HDDs as follows. Given a county i and day d in year t, then

$$GDD_{i,t,d} \equiv \frac{\min[\max[T_{i,t,d}^{\max},10],30] + \min[\max[T_{i,t,d}^{\min},10],30]}{2} - 10 \quad (1)$$

$$HDD_{i,t,d} \equiv \frac{\max[T_{i,t,d}^{\max},32.22] + \max[T_{i,t,d}^{\min},32.22]}{2} - 32.22 \quad (2)$$

where $T_{i,t,d}^{\max}$ and $T_{i,t,d}^{\min}$ are, respectively, the maximum and minimum daily temperature of the observation unit (degree Celsius).[5] These daily GDDs and HDDs are aggregated over the growing season—defined as the months March to August, as in Schlenker and Roberts (2009) and Lusk, Tack, and Hendricks (2018)—to produce the annual measures of beneficial temperature and heat stress for the county in question (GDD_{it} and HDD_{it}).

[5] There are two widely accepted versions of GDD; the one in equation (1) is more commonly used for corn (McMaster and Wilhelm 1997).

2.3.2 Water stress: Precipitation and vapor pressure deficit

Schlenker and Roberts (2009) merely add total volume of precipitation and its quadratic in the growing season, March through August. However, Roberts, Schlenker, and Eyer (2013) find that the correlation between precipitation and yield is weak, and suggest that insufficient moisture can be better captured by additionally considering vapor pressure deficit (VPD). Hence, we use VPD along with cumulative precipitation to represent water stress. VPD relates to the difference between how much moisture the air can hold when saturated and the actual air moisture. It is related to relative humidity, but it additionally accounts for the effects of temperature on the water holding capacity of the air (Sinclair 2011). When the actual air moisture is not observed, it can be approximated by using the day minimum temperature in lieu of the dew point. Specifically, the VPD formula suggested by Roberts, Schlenker, and Eyer (2013) is used, where for county i and day d in year t we put:

$$VPD_{i,t,d} \equiv 0.6107 \times \left[\exp\left(\frac{17.269 \times T_{i,t,d}^{\max}}{237.3 + T_{i,t,d}^{\max}} \right) - \exp\left(\frac{17.269 \times T_{i,t,d}^{\min}}{237.3 + T_{i,t,d}^{\min}} \right) \right]. \tag{3}$$

Similar to GDD and HDD, daily VPDs are aggregated over the growing season to form annual county-level measures VPD_{it}. As in Roberts, Schlenker, and Eyer (2013), we utilize two VPD metrics, the VPD for the March–August growing season, and the VPD for the July–August months only. Cumulative precipitation with its quadratic are also included in the model.[6]

Figure 3 displays spatial and temporal variation of each weather variable. **Table 1** reports summary statistics for the main weather variables used in the analysis over the study period (1981–2020).

[6] An alternative approach to accounting for water stress, used by Xu et al. (2013), is to rely on the so-called Palmer Z index, a measure of dryness relative to local climatic norms. The advantage of instead using precipitations and VPD, as in this study, is that these variables can be easily obtained from future forecasts of climate change models, a crucial requirement for the counterfactual simulations of this paper. Hendricks (2018) develops his own metrics of water deficit and water surplus for use in his Ricardian rent functions.

2.4 Future weather projections

Because GCMs produce weather projections at coarse spatial cells, Auffhammer et al. (2013) highlighted the need for downscaling and bias correction. In this paper, we relied on a set of 20 downscaled and bias-corrected GCM projections available in the MACA dataset. For each climate model (or set of models), future weather data are obtained under two warming scenarios defined by the GHG representative concentration pathways (RCP), specifically RCP 4.5 and RCP 8.5. Forecasted temperature and precipitation for years 2040 to 2059 are used to generate weather variables at mid-century climatic conditions, and, correspondingly, forecasts for the years 2080 to 2099 are used to generate end-of-century weather data.

We also note that, as discussed in Burke at el. (2015), comparing future weather projected by climate models with historically *observed* weather can lead to bias in the estimates of weather change. Thus, when calculating the difference in weather variables between the historical period and future periods, we use model-implied temperature and precipitation for all periods. Specifically, the weather variables for the stationary climate scenarios are assumed to be the model-implied estimates over the period 1981–2005 (climate models provide current modeled temperature and precipitation only through 2005). **Table 2** presents the summary statistics of the predicted weather variables for the reference historical period and the two future periods (mid- and end-century), obtained from all the 20 GCMs considered.

3. Yield Response

The model to be estimated postulates that observed (realized, end-of-season) county maize yields (production per acre) are determined, inter alia, by the technology of production and realized climatic conditions (weather). We are particularly interested in separating the one-time impact of GE trait adoption from the underlying continuous technical progress due to all other improvement/breeding activities. The models we estimate can be written as:

$$(4) \qquad y_{it} = \alpha_i + \mathbf{X}_{it}\boldsymbol{\beta} + \gamma G_{st} + G_{st}\mathbf{X}_{it}\boldsymbol{\delta} + \tau_s T_t + \varepsilon_{it}$$

where i is county index; t indicates year; s indicates the state of county i; the conditioning (row) vector $\mathbf{X}_{it}$ includes all the weather variables of interest, discussed earlier; G_{st} is the GE adoption rate (measured, as noted, at the state level); and, $T_t = 1, 2, ..., 40$ is the linear trend variable. The parameter vector $\boldsymbol{\delta}$ captures the interaction effects between GE adoption and weather variables. Note that the trend coefficient τ_s, meant to capture the underlying technical change beyond that embedded in GE traits, is allowed to vary at the state level. Finally, the intercept α_i is county-specific, and captures heterogeneous factors impacting yield (e.g., soil quality) that are unobserved but largely time-invariant.

The parameterization of GE trait effect in (4) may deserve some additional discussion. Here we are maintaining what Xu et al. (2013) call the "adoption shift" model—full adoption of GE traits leads to a one-time shift in the yield trajectory.[7] Alternatively, one could postulate that GE trait adoption changes the trend slope, which is presumed to reflect the overall impact of technical change. Xu et al. (2013) note that such an "adoption slope" model is virtually indistinguishable, in sample, from the shift model. In fact, in the context of our data, this in-sample equivalence is documented in the Supplementary Appendix (Tables X4 and X5). Mid-century and end-century projections, however, would differ as the adoption slope parameterization likely produces larger effects the further away in the future. We believe that, in the context of our analysis, the adoption shift model is both more conservative and more appropriate. In particular, we again emphasize that we are not interest in projecting the unconditional impacts of current and future GE technologies. Rather, we merely wish to use yield gain attributable to GE traits as a yardstick to assess the extent of yield shortfalls under climate change scenarios. This yardstick is not meant to capture all future productive impact of GE technologies, but rather the realized yield effects of hitherto-adopted first-generation GE varieties embedding agronomic traits.

Two issues need to be addressed to make the model in equation (4) operational. One concerns the representation of GE adoption rates. As discussed earlier, both HT and IR traits have been introduced into maize varieties, alone or in combination, and at different times. It is likely that

[7] Insofar as adoption is gradual, however, such shift is also gradual over time.

HT and IR traits affect yield differently.[8] Whereas it would be desirable to reflect this differential in the model, the concern is that, apart from the early years, HT and IR traits have been jointly adopted via stacked traits (see Figure X2 in the supplementary Appendix), which makes it difficult to credibly identify the separate conribution of the two types of traits. Hence, similar to previous studies, we represent the impact of GE thorugh a single adoption variable.

The second issue to be resolved concerns the representation of the left-hand-side of equation (4). Many studies in this area have adopted a log transformation of yields, resulting in a "semi-log" functional form. That is, if Y_{it} denotes the actual yield (bushels per acre), then $y_{it} \equiv \ln Y_{it}$ (e.g., Schlenker and Roberts 2009; Roberts, Schlenker, and Eyer 2013; Burke and Emerick 2016; Ortiz-Bobea and Tack 2018; Malikov, Miao, and Zhang 2020). Alternatively, others (e.g., Nolan and Santos 2012; Xu et al. 2013; Lusk, Tack, and Hendricks 2019) have assumed a fully "linear" functional form; that is, $y_{it} \equiv Y_{it}$. In what follows, we present empirical evidence on the choice between the linear and semi-log models based on a general transformation of the LHS of equation (4).

Identification of the GE and trend impacts as modeled by equation (4) relies on the temporal and spatial variation in the adoption of GE varieties, a strategy also relied upon by Lusk, Tack, and Hendricks (2019). The crux of the argument is that the timing of commercialization of specific GE varieties is largely exogenous—see the related extensive discussion presented in Ciliberto, Moschini, and Perry (2019). Other sources of productivity growth, such as continuing

[8] This hypothesis can be motivated within the damage-control optics of Lichtenberg and Zilberman (1986). GE traits conferring insect resistance provided a novel technology to control infestations, such as those by the European corn borer, that had hitherto been only partially treated. By contrast, GT traits simply provided a new (cost effective) avenue to weed control that, however, had already been effectively managed with alternative herbicides. The large body of experimental evidence analyzed by Nolan and Santos (2012) indicates strong yield effects from IR traits, and essentially no impact of HT on maize yield.

germplasm improvement by conventional breeding,[9] is also traditionally taken as exogenous to farmers' decisions, and the model captures that by a linear trend. The fact that the GE adoption was technologically constrained to zero up to 1996 helps with the identification of GE effects, separate from other sources of yield growth that presumably operated throughout the sample period. Furthermore, we recognize that the latter may operate differently under dissimilar growing conditions, and the model permits trends to differ across states. Other confounding factors, beyond the weather effects that we explicitly model (such as pest pressure differing across regions), are assumed to be accounted for by county-specific fixed effects.

We note, at this juncture, that the objective of this paper is close in spirit to Ortiz-Bobea and Tack (2018). They, however, do not use information on the (gradual) adoption of GE varieties, and instead rely on the timing of GE crop introduction, identifying the GE effect on yield by the difference between the slopes from piecewise linear trend segments, before and after the initial 1996 commercialization of the GE technology. By contrast, we explicitly introduce the adoption rates in the model and represent this effect as an additive factor, again under the presumption that the gains from the full adoption of these first-generation agronomic traits is a one-time occurrence (notwithstanding the fact that genetic engineering, going forward, may be essential to sustain the trajectory of productivity gains captured by the underlying linear trend).

3.1. Linear or semi-log functional form

The choice of whether the linear or the semi-log model is more appropriate for our purposes can be cast in the context of transformation analyzed in the seminal paper by Box and Cox (1964). With this transformation, the LHS of model (4) is represented as:

$$(5) \qquad y_{it} = \frac{Y_{it}^{\theta} - 1}{\theta}$$

[9] A distinctive vehicle by which improvement has been achieved in maize has been the ability of modern hybrids to withstand crowding, thus allowing plant populations (number of plants per acre) to steadily increase over time (Duvick 2005; Perry, Hennessy, and Moschini 2022).

where θ is an unrestricted parameter. This transformation is appropriate for applications when the dependent variable is positive, which is the case for yields. It is apparent that this transformation nests both functional forms of interest to us: when $\theta = 1$ one obtains the linear model, whereas $\theta \to 0$ yields the semi-log model.

Table 3 reports the results of the estimated Box-Cox transformation. We present the results for two models that are fully analyzed below—model 1 refers to the specification in equation (4), but omitting the interaction effects between GE and weather variables (i.e., subject to $\boldsymbol{\delta} = \mathbf{0}$), whereas model 2 is the full model. For both models, the hypothesis $H_0 : \theta = 0$, corresponding to the semi-log functional form, is decisively rejected by the likelihood ratio statistics. The hypothesis $H_0 : \theta = 1$, corresponding to the linear functional form, is not rejected at any conventional significance level for the more general model 2 . Beyond these statistical tests, we note that the objective in this setting is mainly to settle on one of these two popular functional form representations of the yield function. It is clear from **Table 3** that the LR statistics overwhelmingly favor the linear model relative to the semi-log model. Based on these results, we focus on the linear formulation for the remainder of the paper. Corresponding results that use the semi-log model, however, are fully reported in the supplementary appendix.

3.2. Results

Results for the GE effect on maize yield from the historical data from 1981 to 2020 are reported in **Table 4**.[10] As noted in the foregoing, in this table model 1 and model 2 are defined as in equation (4) when, respectively, the interaction effects between GE adoption and weather variables are not and are included. It is apparent that conditioning yield response by realized weather variables is crucial. The *F* statistics of the null hypothesis that the coefficients for all weather variables are jointly equal to zero is, respectively, $\hat{F}(6, 58583) = 2,456.0$ for model 1 (no interaction effects) and $\hat{F}(12, 58577) = 1,349.5$ for model 2. It is apparent that the null of no weather impacts is conclusively rejected (p-values are less than 0.0001 in all cases). The GE-

[10] Estimation of the panel data model relies on the REGHDFE module in Stata (Correia 2019).

weather interaction effects included in model 2 are also significantly different from zero in their own: the correcting F statistics is $\hat{F}(6,58577) = 174.62$. Consistent with previous research results, for all models reported in **Table 4**, we find that the GDD variable has a positive and significant impact on yields, whereas heat stress, captured by the HDD variable, has a negative and significant impact.

We note at this juncture that our inferences, here and in what follows, are based on the so-called Huber-White covariance that is robust to heteroscedasticity (White, 1980). The question that naturally arises in this setting, noted by a reviewer, is whether concerns about the nature of the panel data at hand (arising, for example, from correlated weather across counties in the same state, or from the use of state-level adoption rates for GE traits) should suggest the use of clustered standard errors. The issue is moot. As discussed by Abadie et al. (2017), much of the conventional wisdom on this matter appears misplaced. The key is to distinguish motivations for clustering based on sampling-design reasons from experimental-design issues. In a context such as ours, where the universe of counties of interest is included in the sample, the former does not seem to apply. As for the latter, recognizing when clustering matters and when one ought to cluster are distinct matters.[11] Weather assignments, whereas not purely random, are not clustered either. As for measuring GE adoption at the state level, that is not equivalent to the GE treatment being clustered at the state level. Rather, as noted by Lusk, Tack, and Hendricks (2019), one has a measurement error instance that may not induce bias. In general, it seems that common clustering adjustments leads to standard errors that are unnecessarily too conservative. In any event, although the analysis in the text relies on Huber-White standard errors, we report the results of a few alternative clustering strategies in the Supplementary Appendix (**Table X6**).

[11] Abadie et al. (2017) conclude that correlations between residuals within clusters, and/or correlations between regressors within clusters, are neither necessary not sufficient to justify clustering.

Water stress matters as well, in a substantial way. Rainfall precipitation positively affects production, and the quadratic terms show that the yield response is (predictably) concave. Since the model includes season-total VPD and July–August VPD separately, the effect of VPD on yield from July to August can be measured by summing the two coefficients. The positive coefficient of the season-total VPD implies a positive effect of VPD on yield in the early-to-middle growing season. For July and August, however, the corresponding VPD coefficient is negative and much larger than that of the season-long variable, indicating an overall negative impact of VPD on yield. This is consistent with evidence that water stress is particularly detrimental to yield during certain stages of crop development (Ortiz-Bobea et al. 2019, Lobell et al. 2014)—for the case of maize, July and August are critical months for plant growth. Water stress coefficients are found to be highly statistically significant.

The effects of technology on maize output is large in magnitude and statistically significant. Model 2, the most general specification and our baseline parameterization, shows an average gain of 0.95 bushels per acre per year across all counties over the period 1981–2020 (this estimate is a simple average of estimated state-specific trend coefficients). State trend effects are quite heterogeneous—the *F*-test of the null hypothesis for the equality of state-specific trend coefficients gives a statistic of $\hat{F}(35, 58577) = 93.27$, clearly rejecting the null (p-value less than 0.0001). For illustration, we also report the state-specific trend coefficients for the three states with the largest contribution to US maize production—Iowa, Illinois, and Indiana. Clearly, in these three corn-belt states, the annual yield gain from technological improvements captured by the time trend is much higher than in the rest of the country.

The estimated coefficients of the GE variable permit a detailed assessment of the additional contribution of GE varieties. Model 1 suggests that the full adoption of GE traits, per se, contributes a one-time gain of about 16.6 bu./acre. This effect is large, equivalent to the gains of almost 20 years of underlying trend effects. Model 2, which accounts for interaction effects of GE with weather variables, provides a slightly attenuated estimate. Note that all weather variables were demeaned (using the overall average over the historical period) in order to ensure that the coefficients of the GE variable are directly comparable between model 1 and

model 2. From **Table 4** it is apparent that the overall impact of allowing interaction between weather and GE traits actually lowers the estimated GE effect, to 14.2 bu./acre (at full adoption of first-generation GE varieties). Thus, in this model, the GE effects is equivalent to about 16 years of the underlying trend effect.[12]

The estimated coefficients for the interaction GE-weather terms are suggestive of the potential effect of GE varieties under climate change. The interaction effect between GE and GDD is essentially nil. The interaction between GE adoption and the HDD variable suggests a moderate positive effect, with GE traits making maize yield more resilient to high temperatures. This estimated effects, however, is not statistically significant. A similar mitigating effect has been found by Wang et al. (2021) with field trial data from Wisconsin. What stands out are the interaction effects between GE adoption and VPD variables, both of which are negative, large, and statistically significant. Whereas the growing-season VPD has a positive coefficient absent GE, the net effect is negative after the full GE adoption. These findings appear consistent with a strand of literature documenting an increased yield vulnerability to weather stress after the introduction of GE varieties (e.g., Lobell et al., 2014). GE-precipitation interaction, on the other hand, seem to reduce yield sensitivity to higher rainfalls in the growing season.

Because model 2 is more general, by capturing the effects of GE traits and weather variables in a more nuanced fashion, and the addition of the interaction effects is supported by the F test discussed earlier, in what follows we rely on this model to analyze the impact of climate change on yields. For completeness, however, corresponding results obtained via model 1 are included in the supplementary appendix.

4. Yield Projections with Climate Change

Using the estimated yield models, we can assess the extent to which technology improvements, and climate change, are likely to affect future yields. For this purpose, future weather variables

[12] Our estimated GE effect is remarkably close to that inferred by Nolan and Santos (2012) who, with a very different methodology and data, estimated the yield gains for GE maize in the range of 13.3 and 14.8 bu./acre.

are used to forecast maize yields under alternative growth regimes and climate change scenarios. If we denote future variables by the superscript "*f*", and the estimated parameters of the model in equation (4) by a hat, the projections for future yields are expressed as:

$$\hat{y}_{it}^{f} = \hat{\alpha}_i + \mathbf{X}_{it}^{f}\left(\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\delta}} G_s^0\right) + \hat{\gamma} G_s^0 + \hat{\tau}_s T_t \quad \text{(forecast with projected future weather)} \tag{6}$$

$$\tilde{y}_{it}^{f} = \hat{\alpha}_i + \mathbf{X}_{i}^{h}\left(\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\delta}} G_s^0\right) + \hat{\gamma} G_s^0 + \hat{\tau}_s T_t \quad \text{(forecast with model-implied historical weather)} \tag{7}$$

where $t \in \{2040,\ldots,2059\}$ for mid-century projections and $t \in \{2080,\ldots,2099\}$ for end-of-century projections. In equation (7), $\mathbf{X}_i^h$ denotes the (row) vector of weather variables under the presumption of a stationary climate, proxied by their average value of weather variables predicted by the relevant climate model over the historical period 1981–2005. By contrast, $\mathbf{X}_{it}^f$ in equation (6) denotes the (row) vector of projected future weather variables according to the relevant climate change model and warming scenario. Note that, in all future periods, the adoption rates of GE varieties are held constant at G_s^0. Specifically, we fix this rate at the observed 2020 level.[13]

Although this may be obvious, we emphasize that the counterfactual projections based on equations (6) and (7) are not meant to be unconditional forecasts—among other things, we do not attempt at forecasting what future agronomic technologies may be. Rather, these projections are conditional forecasts about how anticipated future climatic conditions are likely to affect realized maize yields given the estimated response to weather variables over the historical period, and also the continuation of crop improvement as captured by the underlying trend.

Predicted counterfactual yields under a specific climate change scenario (using equation (6)) and under a stationary climate scenario (using equation (7)) are obtained at the county level. For our purposes, however, we wish to aggregate such conditional predictions to the national level. To that end, we aggregate county-level predictions as follows:

[13] The national average for this adoption rate is about 91%. We view this as a somewhat more conservative assumption than the alternative of full adoption (i.e., $G_s^0 = 1,\ \forall s$).

(8) $$\hat{y}_p = \frac{1}{T}\sum_{t=t_0}^{t_1}\sum_{i=1}^{N}\omega_i \hat{y}_{it}^f \qquad \text{and} \qquad \tilde{y}_p = \frac{1}{T}\sum_{t=t_0}^{t_1}\sum_{i=1}^{N}\omega_i \tilde{y}_{it}^f$$

where the subscript p denotes the period (either mid-century, 2040–2059, or end-of-century 2080-2099); t_0 and t_1 are the beginning and ending years of period p; and, T is the number of years in the future reference period (thus, $T = 20$ for both the mid-century and the end-of-century periods). Here, the county-level weighting constants ω_i are based on the acreage of harvested maize in the 2017 Census of Agriculture (by construction, $\sum_i \omega_i = 1$). Use of this particular Census provides weights as close as possible to the last year of the sample.

4.1. Yield gap due to climate change

We denote as the "yield gap" due to climate change, in this context, the difference between the expected yield with the anticipated climate change and the yield one would expect given stationary climate conditions (at recent historical levels)—in both cases, conditional on normal technological progress as captured by the underlying trend. Given the foregoing, this gap can be defined (as a positive number) as $gap \equiv \tilde{y}_p - \hat{y}_p$. Because of its ceteris paribus nature, we believe that this effect of climate change is of direct policy interest.

The estimated yield shortfalls due to anticipated climate change are reported in **Table 5**. It is apparent that climate change is predicted to have sizeable impacts on maize yield. Across all 20 GCMs, the ensemble mean at mid-century indicate a yield shortfall of 25.5 bu./acre for the RCP 4.5 scenario, and 37 bu./acre for the RPC8.5 scenario. The estimate yield gaps are larger at the end-century, with an average shortfall due to climate change of 34.7 bu./acre for RCP 4.5 and 82.4 bu./acre for RCP 8.5. It is also apparent that the various GCMs lead to considerable variability in predicted outcomes. For mid-century, predicted yield shortfalls range from a minimum of 3.8 bu./acre (model MRI-CGCM3 under RCP 4.5) to a maximum of 69.9 bu./acre (model MIROC-ESM-CHEM under RCP 8.5). Similarly, at end-century, predicted yield shortfalls range from a minimum of 7.3 bu./acre (again with model MRI-CGCM3 under RCP 4.5) to a maximum of 146.8 bu./acre (model HadGEM2-CC365 under RCP 8.5). Also worth

noting, climate forecasts from HadGEM2-ES and NorESM1-M, perhaps the two most widely used to contrast the impact of future climate predictions on agricultural productivity (e.g., Warszawski et al. 2014; Chen and Chen 2018; Malikov, Miao, and Zhang 2020; Ortiz-Bobea 2020), turn out to be quite pessimistic and not representative of the ensemble means of the full set of GCMs. In particular, HadGEM2-ES is the model that predicts the worst outcomes for future maize yields.

The estimated yield gaps attributable to climate change appear quite significant in magnitude. Again, these gaps are estimated conditional on the continuation of the gradual yield gains that have characterized maize production since the introduction of hybrid varieties, and conditional on the (already realized) yield gains from GE varieties. To appreciate the magnitude of the estimated yield gaps, we can use the estimated yield gains realized by the adoption of first-generation GE varieties as of 2020 as a benchmark. As reported in **Table 6**, the yield gain due to the (nearly complete) adoption of GE varieties, as implied by our estimated model (A-B), is 13.01 bu./acre. Thus, the estimated yield gaps due to climate change at mid-century range from approximately 2.0 times to 2.7 times the entire realized yield gains made possible by the development and widespread adoption of GE varieties. By the end of the century, the estimated yield gaps range from 2.8 times to 6.3 times the GE yield gain. Quite apparently, if the objective were to ensure that future maize yield gain remained on the trajectory expected under historical climate conditions, these results suggest that considerable R&D investments into the development of suitable adapting technology are needed.

The results discussed in the forgoing are illustrated in **Figure 4** (panel A pertains to the RCP 4.5 pathway, and panel B is for RCP 8.5). The effects of anticipated climate change on maize yields are quite dramatic. For the more pessimistic RCP 8.5 pathway, climate change totally offset the yield gains from the underlying technical progress (as captured by the trend in the estimated model), such that mid-century and end-century projected yields are actually lower than 2020 yields.

Figure 5 provides more details on the spatial distribution of estimated yield gaps, depending on the period and scenario of reference. It is apparent that there exists considerable spatial variation in the yield gaps attributable to climate change. Counties in the southern region turn out to be more sensitive and vulnerable to warming climate change compared to the northern regions. This result is consistent with the observation that southern counties are more likely to be exposed to climatic conditions exceeding a critical threshold (e.g., Schlenker and Roberts (2009) report 29 degrees Celsius as a crucial threshold for maize growth).

4.2. Heat stress and water stress

It may be of some interest to identify the primary weather factors that contribute to the estimated yield gaps. One way to do so is to categorize our six weather variables into two groups—heat stress and water stress. As in Hendricks (2018), we define heat stress by the change in both GDDs and HDDs. Holding other weather variables constant, we quantify the yield gap attributable to the change in these two weather variables. The predicted damage from water stress is similarly defined by the yield gap due to changes in VPD and precipitation. Because the estimated yield model is linear in the weather variables, an explicit decomposition is possible.

The county-level yield gap with all the six weather variables can be expressed as:

$$(9) \qquad gap_i \equiv \frac{1}{T}\sum_{t=t_0}^{t_1}\left(\hat{y}_{it}^{f}-\tilde{y}_{it}^{f}\right) = \frac{1}{T}\sum_{t=t_0}^{t_1}\sum_{k=1}^{6}\left(\hat{\beta}_k+\hat{\delta}_k G_s^0\right)\left(X_{ik}^{h}-X_{ikt}^{f}\right) = \sum_{k=1}^{6}\left(\hat{\beta}_k+\hat{\delta}_k G_s^0\right)\left(X_{ik}^{h}-\overline{X}_{ik}^{f}\right)$$

where $\overline{X}_{ik}^{f}$ is the average forecasted weather variable k over T years of the period of interest. At the national level, using the county-specific weights ω_i introduced earlier, the yield gap can be written as:

$$(10) \qquad gap \equiv \sum_{i=1}^{N}\omega_i\left(\sum_{k=1}^{6}\left(\hat{\beta}_k+\hat{\delta}_k G_s^0\right)\left(X_{ik}^{h}-\overline{X}_{ik}^{f}\right)\right) = \sum_{k=1}^{6}\left(\hat{\beta}_k+\hat{\delta}_k G_s^0\right)\Delta X_k$$

where $\Delta X_k = \sum_{i=1}^{N} \omega_i \left(X_{ik}^h - \bar{X}_{ik}^f \right)$ is the weighted average change of the weather variable *k* in the future period of interest, relative to historical levels.

Table 7 reports the decomposition of the estimated future yield gaps according to projected heat and water stress determinants (for the two RCP scenarios, and for both mid-century and end-of-century periods). In this table, the fraction of yield losses attributable to heat stress ranges from 34% to 47%, whereas the amount attributable to water stress ranges from 53% to 65%. Thus, perhaps somewhat surprisingly, it appears that water stress is actually quantitatively more important than heat stress. We note, however, that virtually all of the water stress component is attributable to the VPD weather variables. Whereas these variables are meant to capture the likely water stress that crops experience under rising temperatures, it is a fact that these variables are computed as a transformation of the underlying temperature metrics (just as GDD and HHD are). Hence, meaningful separation of heat and water stress in our model is inherently problematic (Ortiz-Bobea et al., 2019). In any event, it is apparent that adaptation strategies targeting both abiotic stresses is necessary for successful adaptation to climate change. The magnitude of the water stress impacts, if confirmed, would have a direct policy implication. Specifically, it suggests that already available and emerging technologies, related to irrigation and precision agriculture, could be deployed to offset a sizeable portion of the negative climate change yield impact.

5. Robustness

To investigate the robustness of the results discussed in the foregoing, in this section we briefly discuss the implications of changing some of the features of the baseline analysis.

5.1. Semi-log yield function

Earlier we presented evidence that, based on the Box-Cox framework, a linear functional form for the yield response is decisively preferred relative to the semi-log model. However, in the supplementary appendix we report the results using the semi-log model, as it is a widely used model. Specifically, **Table X7** reports the coefficients for the semi-log formulation of the two

model specifications we have investigated (with and without GE-weather interaction effects). This table is the analog of **Table 4** in the text. Whereas the non-linear transformation of the left-hand side in the semi-log model requires some caution when estimates are compared with those of the linear model, we conclude that the qualitative results from the semi-log model are similar to those of the linear model. The stand-alone effect of GE variety adoption is significant—for model 2, full adoption is estimated to increase yields by about 11.3%,[14] and it is equivalent to about 13 years of traditional yield improvement as captured by the underlying trend. Given that the weighted average of US county yield estimates over the entire period is 112.35 bu./acre (Table 1), this estimate implies an average GE yield bump of about 12.7 bu./acre, a bit lower than that of the linear model. Alternatively, using the 2020 benchmark of estimated yield absent GE traits, the estimated adoption rate coefficient implies an average GE yield bump from complete adoption of about 14.6 bu./acre, about the same as that implied by the linear model.

Similarly, the underlying crop improvement captured by the time trend is quite comparable across linear and semilog models. The estimated average time trend coefficient in Table X7 implies an annual gain of 0.84% which, when applied to the average sample yield of 112.35 bu./acre, implies average annual gains of 0.95 bu./acre per year, virtually identical to that of the linear model in Table 4. As for the effects of weather variables, their impacts in the semilog model are broadly consistent with those of the linear model. Again, we find that the GE-weather interaction effects are statistically significant. Also, permitting interaction effects (included in model 2 and excluded in model 1) leads to some reduction in the estimated stand-alone GE effect.

Whereas the in-sample qualitative and quantitative estimated mean effects are similar between the linear and the semi-log models, the implications of the two models for the yield gaps due to climate change at mid-century and end-of-century are quite different. This is documented in **Table X8** and **Figure X3** in the Supplementary Appendix, which show that the yield gap from the semi-log model are considerably higher than those predicted by the linear model: they

[14] The calculations here use the correction noted by Halvorsen and Palmquist (1980).

ranges from 3 to 4 times the entire GE yield gain at mid-century, and from 6.2 to 11.6 times the GE gain at the end of the century. All told, these results suggest that using the linear yield model provides more conservative inferences vis-à-vis the impact of climate change on maize yields.

5.2. Geographic scope of the analysis

The criterion used for sample inclusion noted in section 2 were simply that a county had less than 10% irrigated harvested cropland. Alternatively, Xu et al. (2013) and Lusk, Tack, and Hendricks (2019) included only counties for which yield data were available for at least two-thirds of the years in both the pre-GE and post-GE periods. Because those studies were specifically focused on identifying the GE traits effects on realized yields, the intent was to provide enough data variability to facilitate identification of the impacts of interest. As a robustness check, here we consider the implications, in our context, of such a sampling rule. Estimation results are reported in **Tables 10** of the supplementary appendix. It appears that the results reported in the text are quite robust to this alternative sampling rule. One minor difference is that, because this sampling rule ends up excluding marginal counties that are less productive, the estimated average trend effect is somewhat larger. The forecasted yield gap under climate change based on the estimated yield parameters with this alternative rule, reported in **Tables X11** and **Figure X4** of the Supplementary Appendix, are virtually undistinguishable from the baseline model results of the main text.

5.3 Uncertainty of Climate Change

Burke et al. (2015) emphasize the importance of considering the role of climate uncertainty when making inferences about economic outcomes of interest. In particular, they note the problem that may arise when relying on only one or a selective few climate change models. For example, while a substantial body of literature studying the impact of climate change relies heavily on the projections from Hadley models (such as the HadGEM2-ES model included in our baseline analysis), the models are sometimes quite different from the central tendency of the full ensemble of climate models. To get a better sense of the distribution of possible outcomes,

including the severity of worst-case scenarios, Burke et al. (2015) suggest the use of a model democracy approach, whereby all available climate models are used to investigate the possible impacts of changing weather patterns due to climate change. In the context of a model where projected weather variables from climate change scenarios impact the outcome variable of interest (maize yield in our case), Burke et al. (2015) also discuss the possible impact of a separate distinct source of uncertainty, which they call regression uncertainty (i.e., the fact that the parameters defining the mapping between climate variables and economic outcomes are not known with certainty but need to be estimated).

Our analysis is well positioned to gauge the impact of climate uncertainty, as we have relied on a broad set of 20 GCMs available to us in MACA. As for the impact of parameter uncertainty, we estimate the coefficients of the yield model 1,000 times using bootstrap samples. We then combine the resulting 1,000 sets of estimated yield model coefficients with the projected climate change input—the median change of each weather variable of interest—across all 20 climate change models in a given RCP. Thus, in total we have 20,000 projected outcomes, the distribution of which reflects both the degree of climate uncertainty, as arising from the heterogeneous weather implication from the 20 climate change models, and the regression uncertainty from the estimated yield models.

Results are reported in **Figure 6**, which show the range of climate impacts by the type of uncertainty for the two warming scenarios considered, RCP 4.5 (panel A) and RCP 8.5 (panel B). More precisely, in these figures the whiskers contain 90% of estimates by taking 5th and 95th percentiles of the bootstrap replications, and the boxes cover the interquartile range with the middle line standing for the median. These diagrams illustrate the fact that the uncertainty of our estimates arises mostly from climate uncertainty, whereas the role of regression uncertainty is minimal. Climate change uncertainty, as anticipated from the results reported in Table 5, has considerable impact on the estimated maize yield gap. When allowing for both regression and climate uncertainty, for the RCP 4.5 scenario the 90% confidence interval for the mid-century maize yield is [-42.1, -5.1] bu./acre and for end-of-century maize yield is [-67.3, -12.5] bu./acre,

whereas for RCP 8.5 the mid-century confidence interval is [-57.3, -9.1] bu./acre and the same for end-of-century is [-135.6, -31.7] bu./acre.

6. Discussion and Conclusion

Agriculture is at the forefront of anticipated impacts of climate change, and considerable evidence has accumulated to suggest that, without countervailing actions, large negative consequences are probable. Whereas a number of strategies might be helpful to blunt climate change's impacts on the food supply, there is a growing sense that major adaptation efforts will be necessary. Successful adaptation may require purposeful, directed investments in R&D to develop suitable new technologies. Just how large is the innovation effort required for successful adaptation in agriculture? To shed some light on this question, in this paper we focus on maize production in the United States. Maize is the most important field crop in the country, and one that has benefited greatly from major technological advances over the last few decades, including the development and widespread adoption of GE varieties. The latter constitutes the most prominent set of agricultural innovations since the green revolution. Because the nature and scale of this GE revolution is well understood, in this paper we propose to use it as a yardstick—that is, gauge the scope of the innovation task required for adaptation, to offset the impacts of anticipated climate change, in terms of multiples of what was achieved by the widespread adoption of first-generation GE traits in maize.

To be clear, our focus on first-generation GE traits in not meant to suggest that GE technologies have no further role to play going forward. Our point is that the GE productivity gains captured by our yield model relate to a clearly defined set of innovation—first generation GE varieties embedding agronomic traits—that were rapidly diffused (essentially to full adoption) over a relatively short time period. They are best viewed as a once-time bump in yields and, as such, provide an attractive yardstick to measure the extent of the innovation challenge posed by adaptation to climate change. Going forward, genetic engineering is expected to continue to play a key role in crop improvement. Indeed, promising new GE technologies such as CRISPER, which offer novel methods to control and improve crops' genome, are only beginning to be

deployed in this setting (Chen et al., 2017). For these new generations of GE technologies to affect yields in the face of changing climatic conditions, however, targeted new R&D investments will be needed—over and above those required to sustain the underlying trend of yield improvement estimated over the sample period (which we maintain in all counterfactuals).

Based on a dataset that permits a more nuanced representation of GE adoption rates, using county yield responses conditioned on an array of weather variables, we confirm the finding of previous studies that GE traits have contributed significantly to increase maize yields. The estimated parameters suggest that the full adoption of (first-generation) GE traits leads to yield improvement in the range of 14.2 to 16.6 bushels per acre. Next, we use the model to forecast the yield impact, at both mid-century and end-of-the-century, of weather patterns projected from a broad set of twenty climate models under two warming scenarios (RCP 4.5 and RCP 8.5). We find that the average yield shortfalls arising from adverse climate developments are large. The ensemble means across 20 GCMs indicate that climate change is expected to decrease maize yields in the range of 25.5 to 37 bu./acre at mid-century, and in the range of 34.7 to 82.4 by end century. These yields shortfalls correspond to a range of 2.0 to 2.7 times the yield gain from GE over the observed historical period at mid-century, and to a range of 2.8 to 6.3 times the GE yield gain at end-century. Finally, we establish that little uncertainty originates from the estimated yield regression model, and that virtually all of the estimated variability of predicted yield impacts of climate change is due to climate uncertainty.

Extrapolation of the yield gaps due to climate change that we have identified to general agricultural productivity is subject to some caveats, of course. In particular, when aggregating county-level results, including the counterfactuals under climate change conditions, we have relied on fixed county-specific weights. In other words, our results do not account for a margin of adjustment that has been recognized as very relevant in this context—the possibility that climate change may affect comparative advantage enough that the kind of crops grown, and their intensities, may spatially relocate (Costinot, Donaldson, and Smith 2016). Crop switching, when feasible, can of course reduce the overall impact of climate change on agricultural

productivity (Rising and Devineni 2020). Because our modeling framework is ill-suited to characterize this channel of possible adaptation, the results we present may be interpreted as an upper bound.

Notwithstanding the foregoing qualifications, the results we have presented imply that the scope of adaptation to climate change, vis-à-vis agricultural productivity, is very challenging. For the case of US maize, severe yield shortfalls are to be expected by the end of the century under a wide range of climate model projections, especially for the warming scenario RCP 8.5. In particular, we find that the estimated yield shortfalls are several times larger than the entire productivity gains due to the adoption of first-generation GE traits. This is particularly significant given that the development and diffusion of such GE traits were made possible by a unique confluence of propitious circumstances. By leveraging breakthrough advances in recombinant DNA techniques, unprecedented research efforts by agrochemical and seed companies led to the invention of insect resistant and herbicide tolerant traits that, once introduced into elite germplasm, were rapidly adopted by farmers. The underlying R&D investments were substantial—for example, private sector expenditure on crop seed and biotechnology R&D are estimated to have increased eight-fold, in real terms, between 1994 and 2010 (Heisey and Fuglie, 2011). Ortiz-Bobea and Tack (2019) ask whether another genetic revolution is needed to offset climate change impacts on US maize yields. This paper supports and extend their conclusions, as we find that that the magnitude of the impacts of the (first generation) GE revolution in crop improvement may need to be replicated several times to offset the damaging impact of climate change on maize yield. It is apparent that large, sustained, and targeted research efforts might be required to counter the negative implications of anticipated climate change.

References

Abadie, A., Athey, S., Imbens, G.W. and Wooldridge, J., 2017. When should you adjust standard errors for clustering? NBER Working Paper No. 24003, National Bureau of Economic Research.

Abatzoglou, J. T., and Brown, T. J. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, 32(5):772–780.

Auffhammer, M., Hsiang, S.M., Schlenker, W. and Sobel, A. (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*, 7(2):181–198.

Bennett, A.B., Chi-Ham, C., Barrows, G., Sexton, S., and Zilberman, D. (2013). Agricultural biotechnology: Economics, environment, ethics, and the future. *Annual Review of Environment and Resources*, 38: 19.1–19.31.

Box, G.E., and Cox, D.R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2):211–243.

Burke, M., Dykema, J., Lobell, D.B., Miguel, E., and Satyanath, S. (2015). Incorporating climate uncertainty into estimates of climate change impacts. *Review of Economics and Statistics*, *97*(2):461–471.

Burke, M., and Emerick, K. (2016). Adaptation to climate change: Evidence from US agriculture. *American Economic Journal: Economic Policy*, 8(3):106–40.

Carattini, S., Levin, S., and Tavoni, A. (2019). Cooperation in the climate commons. *Review of Environmental Economics and Policy*, 13(2):227–247.

Carter, C., Cui, X., Ghanem, D., and Mérel, P. (2018). Identifying the economic impacts of climate change on agriculture. *Annual Review of Resource Economics*, 10:361–380.

Chavas, J.P., Shi, G., and Lauer, J. (2014). The effects of GM technology on maize yield. *Crop Science*, *54*(4), 1331-1335.

Chen, K., Wang, Y., Zhang, R., Zhang, H. and Gao, C., 2019. CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annual Review of Plant Biology*, 70:667–697.

Chen, X., and Chen, S. (2018). China feels the heat: Negative impacts of high temperatures on China's rice sector. *Australian Journal of Agricultural and Resource Economics*, 62(4):576–588.

Chen, S., Chen, X., and Xu, J. (2016). Impacts of climate change on agriculture: Evidence from China. *Journal of Environmental Economics and Management*, 76:105–124.

Ciliberto, F., Moschini, G., and Perry, E.D. (2019). Valuing product innovation: Genetically engineered varieties in US corn and soybeans. *The RAND Journal of Economics*, 50(3):615–644.

Clancy, M.S., and Moschini, G. (2017). Intellectual property rights and the ascent of proprietary innovation in agriculture. *Annual Review of Resource Economics,* 9:53–74.

Correia, S. (2019). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects.

Costinot, A., Donaldson, D., and Smith, C. (2016). Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world. *Journal of Political Economy*, 124(1):205–248.

Duvick, D.N. (2005). The contribution of breeding to yield advances in maize (Zea mays L.). *Advances in Agronomy* 86: 83-145

Fisher, A.C., Hanemann, W.M., Roberts, M. J., and Schlenker, W. (2012). The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: comment. *American Economic Review*, 102(7):3749–60.

Gammans, M., Mérel, P., and Ortiz-Bobea, A. (2017). Negative impacts of climate change on cereal yields: Statistical evidence from France. *Environmental Research Letters*, 12(5):054007.

Hasson, R., Löfgren, Å., and Visser, M. (2010). Climate change in a public goods game: Investment decision in mitigation versus adaptation. *Ecological Economics*, 70(2):331–338.

Hendricks, N.P. (2018). Potential benefits from innovations to reduce heat and water stress in agriculture. *Journal of the Association of Environmental and Resource Economists*, 5(3):545–576.

Halvorsen, R. and Palmquist, R. (1980). The interpretation of dummy variables in semilogarithmic equations. ***American Economic Review***, 70(3), 474-475.

Heisey, P.W., and K.O. Fuglie (2011). Private Research and Development for Crop Genetic Improvement. Chapter 2 in: *Research Investments and Market Structure in the Food Processing, Agricultural Input, and Biofuel Industries Worldwide*. Economic Research Service Report No. 130.

Huffman, W.E., Jin, Y., and Xu, Z. (2018). The economic impacts of technology and climate change: New evidence from US corn yields. *Agricultural Economics*, 49(4):463–479.

Lichtenberg, E. and Zilberman, D. (1986). The econometrics of damage control: why specification matters. *American Journal of Agricultural Economics*, 68(2):261–273.

Lobell, D.B., Burke, M.B., Tebaldi, C., Mastrandrea, M.D., Falcon, W.P., & Naylor, R.L. (2008). Prioritizing climate change adaptation needs for food security in 2030. *Science*, 319(5863):607–610.

Lobell, D. B., Schlenker, W., and Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science*, 333(6042):616–620.

Lobell, D.B., Roberts, M.J., Schlenker, W., Braun, N., Little, B.B., Rejesus, R.M. and Hammer, G.L. (2014). Greater sensitivity to drought accompanies maize yield increase in the US Midwest. *Science*, 344(6183), 516-519.

Lusk, J.L., Tack, J., and Hendricks, N.P. (2019). Heterogeneous yield impacts from adoption of genetically engineered corn and the importance of controlling for weather. *Agricultural Productivity and Producer Behavior*, 11.

Lybbert, T. J., and Sumner, D.A. (2012). Agricultural technologies for climate change in developing countries: Policy options for innovation and technology diffusion. *Food Policy*, *37*(1): 114–123.

Malikov, E., Miao, R. and Zhang, J. (2020). Distributional and temporal heterogeneity in the climate change effects on U.S. Agriculture. *Journal of Environmental Economics and Management*, 2020.

McMaster, G.S., and Wilhelm, W.W. (1997). Growing degree-days: One equation, two interpretations. *Agricultural and Forest Meteorology*, *87*(4):291–300.

Mendelsohn, R., Nordhaus, W.D., and Shaw, D. (1994). The impact of global warming on agriculture: A Ricardian analysis. *American Economic Review*, 753–771.

Miller, N., Tack, J. and Bergtold, J. (2021). The impacts of warming temperatures on US sorghum yields and the potential for adaptation. *American Journal of Agricultural Economics*, 103(5), 1742-1758.

Moschini, G. (2008). Biotechnology and the development of food markets: Retrospect and prospects. *European Review of Agricultural Economics*, 35(3):331–355.

National Academies of Sciences, Engineering, and Medicine (NASEM) (2016). *Genetically engineered crops: Experiences and prospects*. National Academies Press.

Nolan, E., and Santos, P. (2012). The contribution of genetic modification to changes in corn yield in the United States. *American Journal of Agricultural Economics*, 94(5):1171–1188.

Olmstead, A.L., and Rhode, P.W. (2011). Adapting North American wheat production to climatic challenges, 1839–2009. *Proceedings of the National Academy of Sciences*, 108(2): 480–485.

Ortiz-Bobea, A. (2020). The role of nonfarm influences in Ricardian estimates of climate change impacts on US agriculture. *American Journal of Agricultural Economics*, 102(3):934–959.

Ortiz-Bobea, A., and Tack, J. (2018). Is another genetic revolution needed to offset climate change impacts for US maize yields? *Environmental Research Letters*, 13(12):124009.

Ortiz-Bobea, A., Wang, H., Carrillo, C.M. and Ault, T.R. (2019). Unpacking the climatic drivers of US agricultural yields. *Environmental Research Letters*, 14(6):064003.

Perry, E.D., Hennessy, D.A. and Moschini, G. (2022). Uncertainty and learning in a technologically dynamic industry: Seed density in US maize. *American Journal of Agricultural Economics*. https://doi.org/10.1111/ajae.12276

Rising, J., and Devineni, N. (2020). Crop switching reduces agricultural losses from climate change in the United States by half under RCP 8.5. *Nature Communications*, 11(1):1–7.

Roberts, M. J., Schlenker, W., and Eyer, J. (2013). Agronomic weather measures in econometric models of crop yield with implications for climate change. *American Journal of Agricultural Economics*, *95*(2):236–243.

Sacks, W.J., D. Deryng, J.A, Foley, and N. Ramankutty (2010). Crop Planting dates: An analysis of global patterns. *Global Ecology and Biogeography*, 19: 607–620

Schlenker, W., and Roberts, M.J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37):15594–15598.

Sinclair, T.R. (2011). Precipitation: The thousand-pound gorilla in crop response to climate change. In D. Hillel and C. Rosenzweig, eds., *Handbook of Climate Change and Agroecosystems: Impacts, Adaptation, and Mitigation*, chapter 9, pp. 179-190.

Stern, N. (2007). *The economics of climate change: The Stern review*. Cambridge, UK: Cambridge University Press.

Tack, J., Barkley, A., and Nalley, L.L. (2015). Effect of warming temperatures on US wheat yields. *Proceedings of the National Academy of Sciences*, 112(22):6931–6936.

Tol, R.S. (2005). Adaptation and mitigation: Trade-offs in substance and methods. *Environmental Science & Policy*, 8(6):572–578.

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

Wang, R., Rejesus, R.M., Tack, J.B. and Aglasan, S. (2021). Do higher temperatures influence how yields respond to increasing planting density? *Agricultural and Resource Economics Review*, 50(2), 273-295.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., and Schewe, J. (2014). The inter-sectoral impact model intercomparison project (ISI–MIP): Project framework. *Proceedings of the National Academy of Sciences*, 111(9):3228–3232.

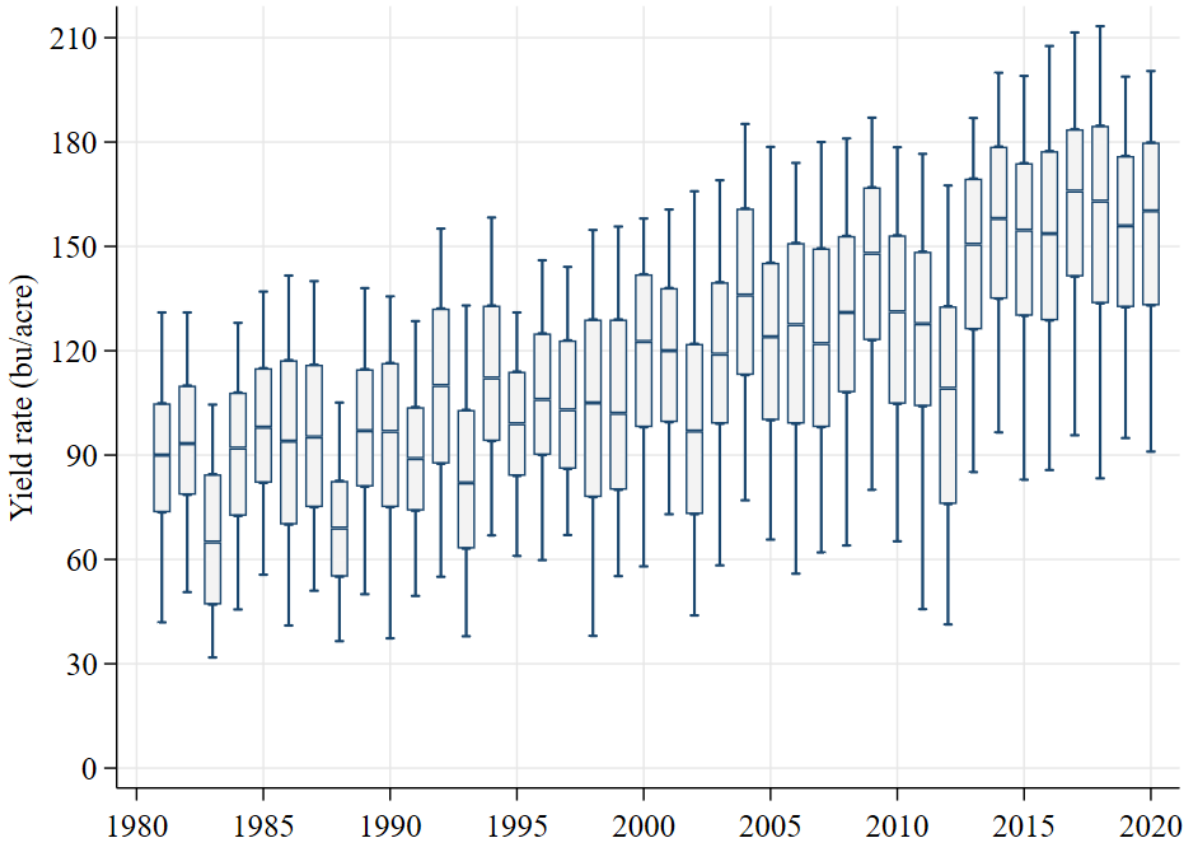
White, H. L., Jr. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48: 817–838.

Xu, Z., Hennessy, D.A., Sardana, K., and Moschini, G. (2013). The realized yield effect of genetically engineered crops: US maize and soybean. *Crop Science*, 53(3):735–745.

Zilberman, D., Zhao, J., and Heiman, A. (2012). Adoption versus adaptation, with emphasis on climate change. *Annual Review of Resource Economics* 4(1):27–53.

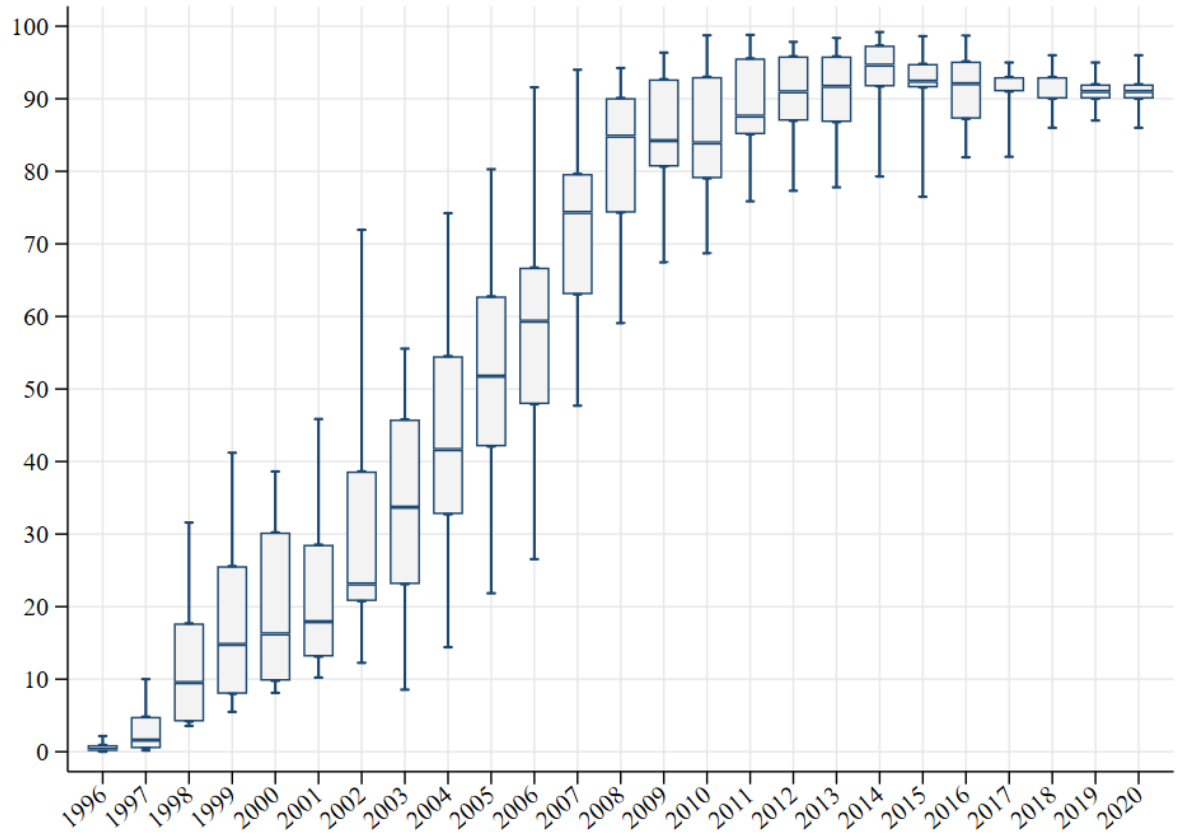
Zilberman, D., Lipper, L., McCarthy, N., and Gordon, B. (2018). Innovation in response to climate change. In Lipper et al. (Eds.) *Climate smart agriculture.* (pp. 49–74). Springer, FAO.

Figure 1. Observed county yields, 1981-2020



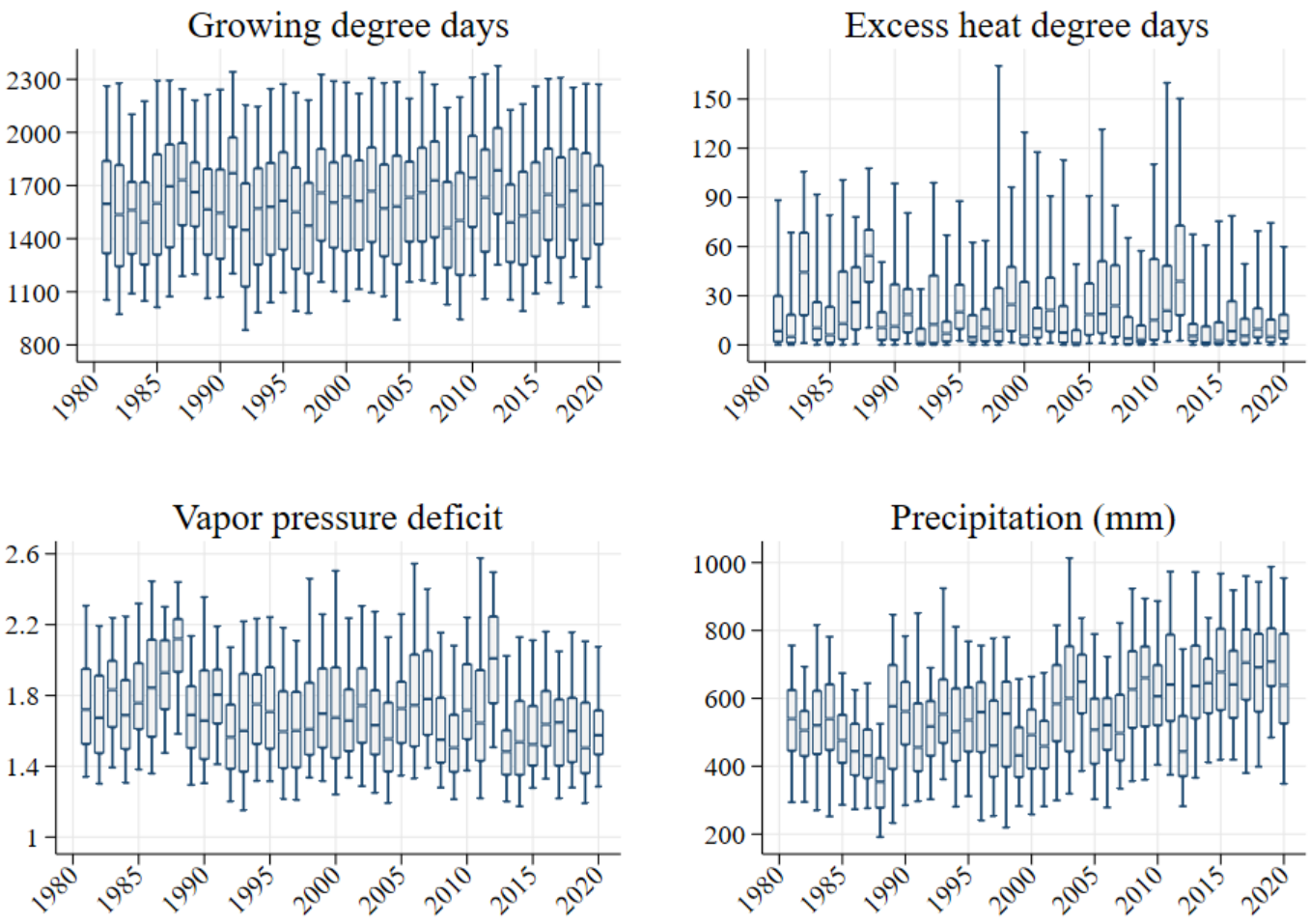
Note: The boxes are generated by the first and third quartile of yields for each year. Whiskers range from the 5th percentile to 95th percentile of yield observations.

Figure 2. Spatial and temporal variation of GE adoption rates, 1996-2020



Note: The y-axis measures GE adoption as a percentage. The boxes are generated by the first and third quartile of the adoption rates for all GE traits in the states of the counties included in our study. Whiskers range from the 5th percentile to 95th percentile of the adoption rates for each year. The middle line inside the box indicates the median adoption rate in each year.

Figure 3. Spatial and temporal variation of weather variables, 1981-2020



Note: The boxes are generated by the first and third quartile of four weather variables for each year (growing degree days in the top left chart, excessive heat degree days in the top right chart, vapor pressure deficit in the bottom left chart, and precipitation in the bottom right chart). Whiskers range from the 5th percentile to 95th percentile of the corresponding weather variable for each year. The middle line inside the box indicates the median of the corresponding weather variable in each year.

Figure 4. Forecasted yields under climate change, ensemble mean of 20 GCMs

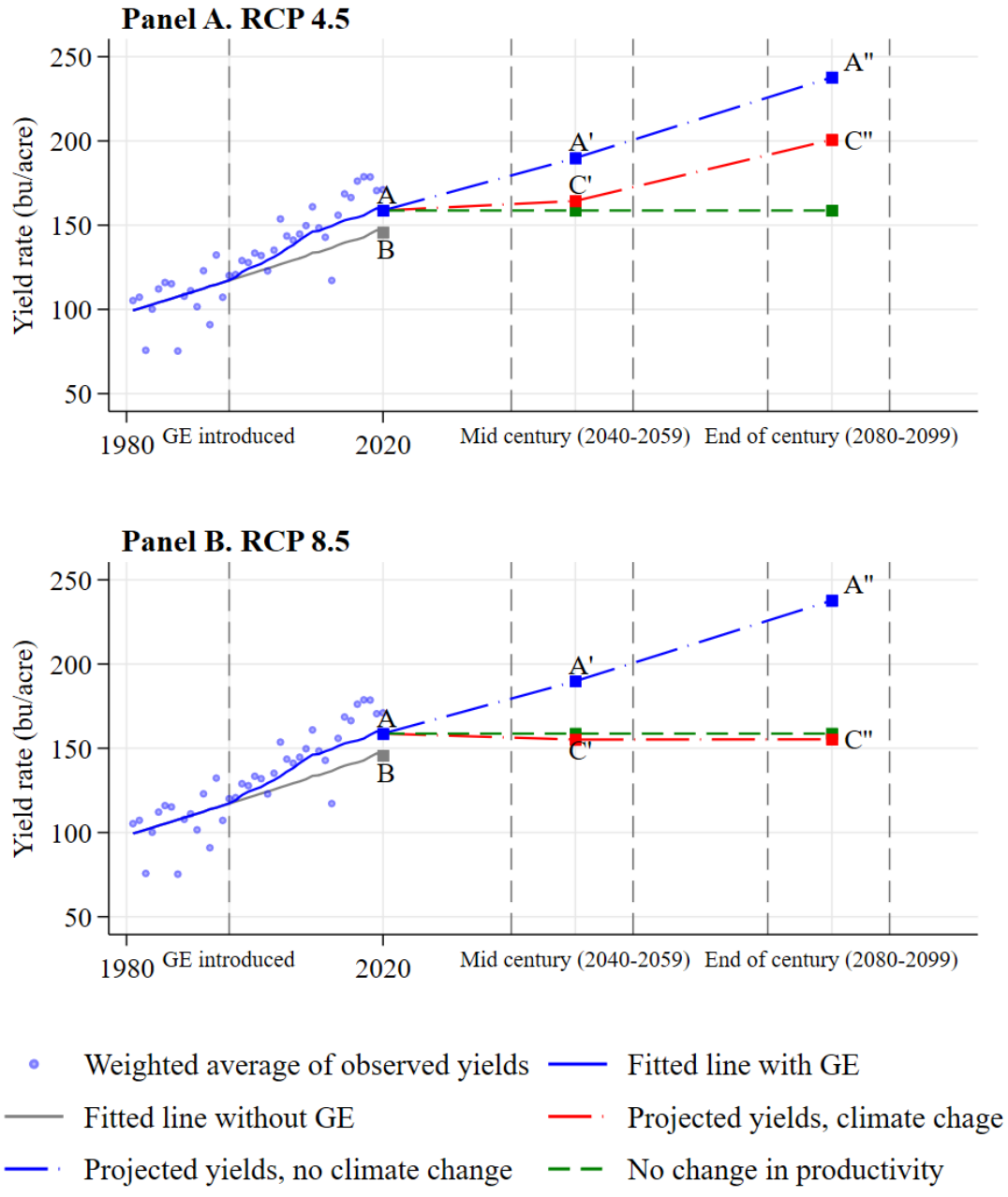
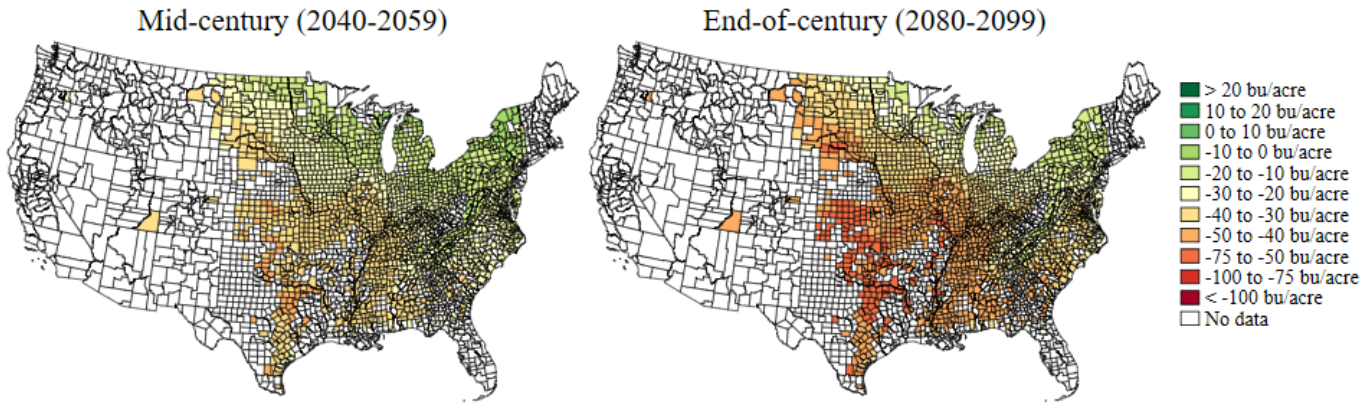


Figure 5. County-wise yield gap under climate change, ensemble mean of 20 GCMs

Panel A. RCP 4.5



Panel B. RCP 8.5

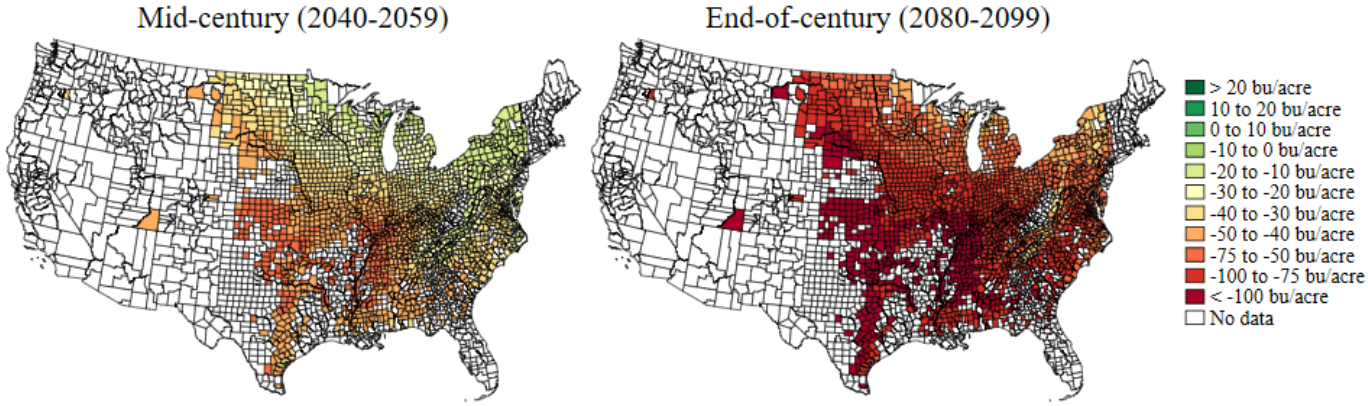
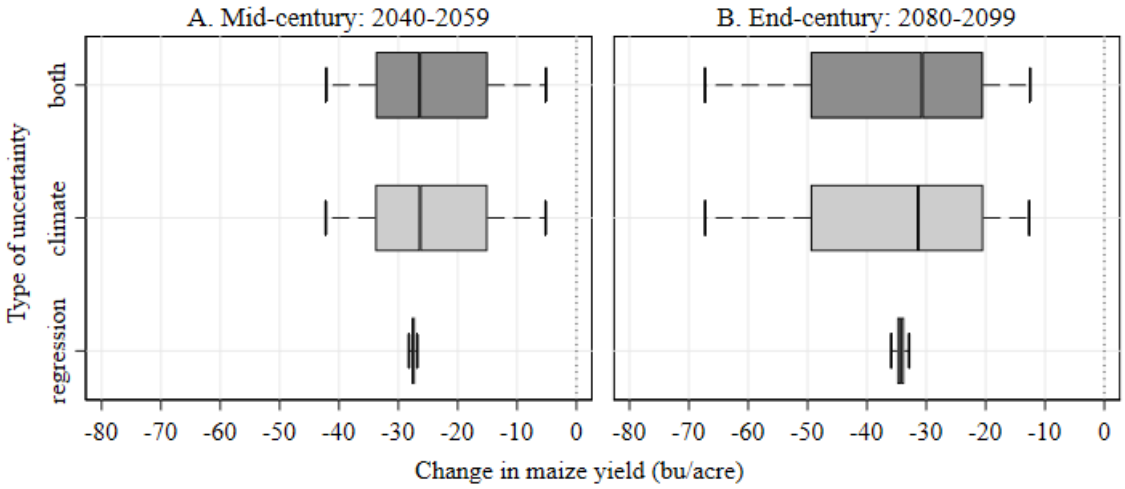
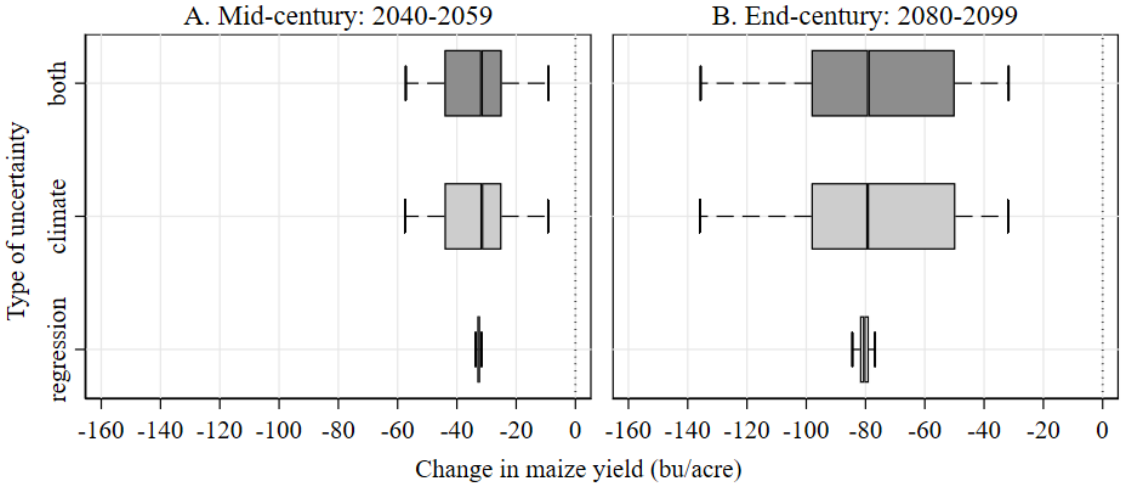


Figure 6. Uncertainty range from all 20 climate models

Panel A. RCP 4.5



Panel B. RCP 8.5



Note: For each type of uncertainty, the box shows the upper quartile and the lower quartile across bootstrap replications for yield gaps and the middle line indicates the median of the estimated yield gaps. Whiskers range from the 5th percentile to 95th percentile of the bootstrap replications.

Table 1. Descriptive statistics of yield model variables, 1981-2020

	Mean	S.D.	Min	Max
Yield rate (bu./acre)	112.35	38.88	4.50	246.70
Growing degree days (GDD)	1612.05	366.93	676.77	2928.50
Excess heat degree days (HDD)	24.60	33.48	0.00	332.18
Vapor pressure deficit (VPD)	1.72	0.32	0.91	3.10
VPD, July-August	2.22	0.42	1.09	4.28
Precipitation (mm)	549.65	171.13	38.50	1550.38

Note: Yield rates data are from the USDA-NASS. Growing degree days, Excess heat degree days, and Vapor pressure deficit data are generated by observed county-level daily temperatures during the corn growing season (March to August). VPD, July-August is obtained by using the daily temperatures only in July and August. Precipitation is calculated by summing county-level rainfalls during the growing season (March to August) of each year.

Table 2 – Summary statistics of predicted weather variables: 20 GCMs

	Historical (1981–2005)	*Mid-century (2040–2059)*		*End-of-century (2080–2099)*	
		RCP 4.5 Mean (S.D.)	RCP 8.5 Mean (S.D.)	RCP 4.5 Mean (S.D.)	RCP 8.5 Mean (S.D.)
Growing degree days (GDD)	1659.0	1878.9	1946.3	1955.9	2209.5
	(380.6)	(381.5)	(378.3)	(379.8)	(381.1)
Excess heat degree days (HDD)	34.68	86.65	108.7	110.4	222.2
	(43.50)	(77.28)	(87.75)	(88.96)	(135.1)
Vapor pressure deficit (VPD)	1.841	2.102	2.179	2.191	2.536
	(0.329)	(0.391)	(0.403)	(0.409)	(0.501)
VPD, July-August	2.396	2.769	2.896	2.900	3.393
	(0.431)	(0.539)	(0.570)	(0.587)	(0.750)
Precipitation (mm)	585.2	597.0	596.2	602.9	604.2
	(171.6)	(185.6)	(184.3)	(186.2)	(195.8)

Note: For any given period, daily temperature and precipitation are provided by each GCM, with a specific RCP, where the finer spatial observations (at around 6km by 6km spatial resolution) are aggregated into county-level daily levels. Growing degree days, Excess heat degree days, and Vapor pressure deficit are generated by the projected county-level daily temperatures during the corn growing season (March to August) each year. VPD, July-August is obtained by using the daily temperatures only in July and August months. Precipitation is also calculated by summing the county-level projected rainfalls during the growing season of each year.

Table 3 - Box-Cox regression

Functional form	Model 1		Model 2	
	Parameter θ	log likelihood	Parameter θ	log likelihood
Box-Cox	0.95403	-260,652.4	1.00165	-259,806.3
	(0.00758)		(0.00773)	
Linear	1	-260,670.7	1	-259,806.3
Semi-log	0	-269,913.0	0	-269,643.8

Note: Standard error for the estimated Box-Cox parameter is reported in parentheses.

Table 4. Estimated Yield model, 1981-2020 (36 States)

	Model 1			Model 2		
	coefficient	s.e.	p-value	coefficient	s.e.	p-value
GE	16.60	0.533	0.000	14.22	0.528	0.000
GDD	0.0120	0.0012	0.000	0.0144	0.00129	0.000
HDD	-0.316	0.009	0.000	-0.293	0.0102	0.000
VPD	3.396	1.338	0.011	10.47	1.399	0.000
VPD, July-Aug	-29.13	0.619	0.000	-27.22	0.704	0.000
PPT	0.0510	0.00327	0.000	0.0814	0.00442	0.000
PPTsq	-0.0000485	2.61E-06	0.000	-0.0000799	3.85E-06	0.000
GE×GDD				0.000647	0.00163	0.691
GE×HDD				0.0286	0.0195	0.143
GE×VPD				-13.75	2.602	0.000
GE×VPD, July-Aug				-12.93	1.597	0.000
GE×PPT				-0.0110	0.00837	0.190
GE×PPTsq				0.0000252	6.59E-06	0.000
Avg. time trend	0.826	0.105	0.000	0.945	0.102	0.000
State-specific trend:						
Illinois	1.320	0.0296	0.000	1.300	0.0292	0.000
Indiana	1.208	0.0287	0.000	1.157	0.0288	0.000
Iowa	1.520	0.0282	0.000	1.433	0.0283	0.000
Constant	86.31	0.205	0.000	85.56	0.206	0.000
Adj. R2	0.776			0.782		
N	60,400			60,400		

Note: All weather variables are demeaned so as to have mean zero over the estimation sample, such that the effect of the coefficient of the GE variable is directly comparable between Model 1 and Model 2 (which includes interaction effects).
Standard errors are robust …

Table 5. Maize yield shortfall due to anticipated climate change: 20 climate models

	Mid Century		End of Century	
Global climate model	RCP 8.5	RCP 8.5	RCP 8.5	RCP 8.5
HadGEM2-ES365	44.0	64.6	67.1	129.5
HadGEM2-CC365	37.7	68.7	44.8	146.8
MIROC-ESM	34.9	66.7	52.2	133.8
MIROC-ESM-CHEM	43.5	69.9	49.6	121
NorESM1-M	38.1	49.7	52.3	110.6
CCSM4	32.2	52.4	46.6	91.7
bcc-csm1-1-m	28.7	38.8	40.8	91.1
MIROC5	35.0	45.8	26.5	79.6
BNU-ESM	28.3	40.5	33.8	83.3
CSIRO-Mk3-6-0	25.9	35.5	36.2	80.4
bcc-csm1-1	31.5	29.7	28.4	84.7
IPSL-CM5A-MR	23.5	22.7	35.7	83.5
CanESM2	19.3	19.8	31.4	78.9
IPSL-CM5A-LR	19.0	23.3	31.5	64.7
GFDL-ESM2G	15.5	23.8	23.8	53
CNRM-CM5	11.8	25	25.4	50.6
inmcm4	15.7	18	27	47.3
GFDL-ESM2M	13.9	18.5	21.4	43
IPSL-CM5B-LR	7.4	18.2	12	50.5
MRI-CGCM3	3.8	8.2	7.3	23.9
Mean	25.5	37	34.7	82.4
Min	3.8	8.2	7.3	23.9
Max	44	69.9	67.1	146.8

Note: Models are listed in descending order based on the overall impact (average over mid-century and end-of-century, and over both RCP 4.5 and RCP 8.5).

Table 6. Summary of yield projections and technological needs under climate change

		RCP 4.5	RCP 8.5
Projected yields (bu/acre)			
Year 2020 with GE (A)	158.71		
Year 2020 w/o GE (B)	145.71		
Mid-century w/o climate change (A')	189.85		
End-century w/o climate change (A'')	237.73		
Mid-century with climate change (C')		164.37	155.15
End-century with climate change (C'')		200.73	155.33
Yield gaps (bu/acre)			
Mid-century (A' - C')		25.48	34.70
End-century (A'' - C'')		37.00	82.40
Innovation gap index			
Mid-century (A' - C')/(A - B)		1.96	2.67
End-century (A'' - C'')/(A - B)		2.84	6.34

Table 7. Decomposition of projected yield loss: heat stress and water stress

	Heat Stress	Water Stress		Total
	GDD, HDD	VPD, VPD, July-Aug	PPT, PPTsq	
Mid-century				
RCP 4.5	-8.79	-16.56	-0.12	-25.2
	34%	65%	0%	100%
RCP 8.5	-13.01	-21.53	-0.17	-34.7
	37%	62%	0%	100%
End-century				
RCP 4.5	-14.04	-22.81	-0.15	-37
	38%	62%	0%	100%
RCP 8.5	-38.41	-43.77	-0.22	-82.4
	47%	53%	0%	100%

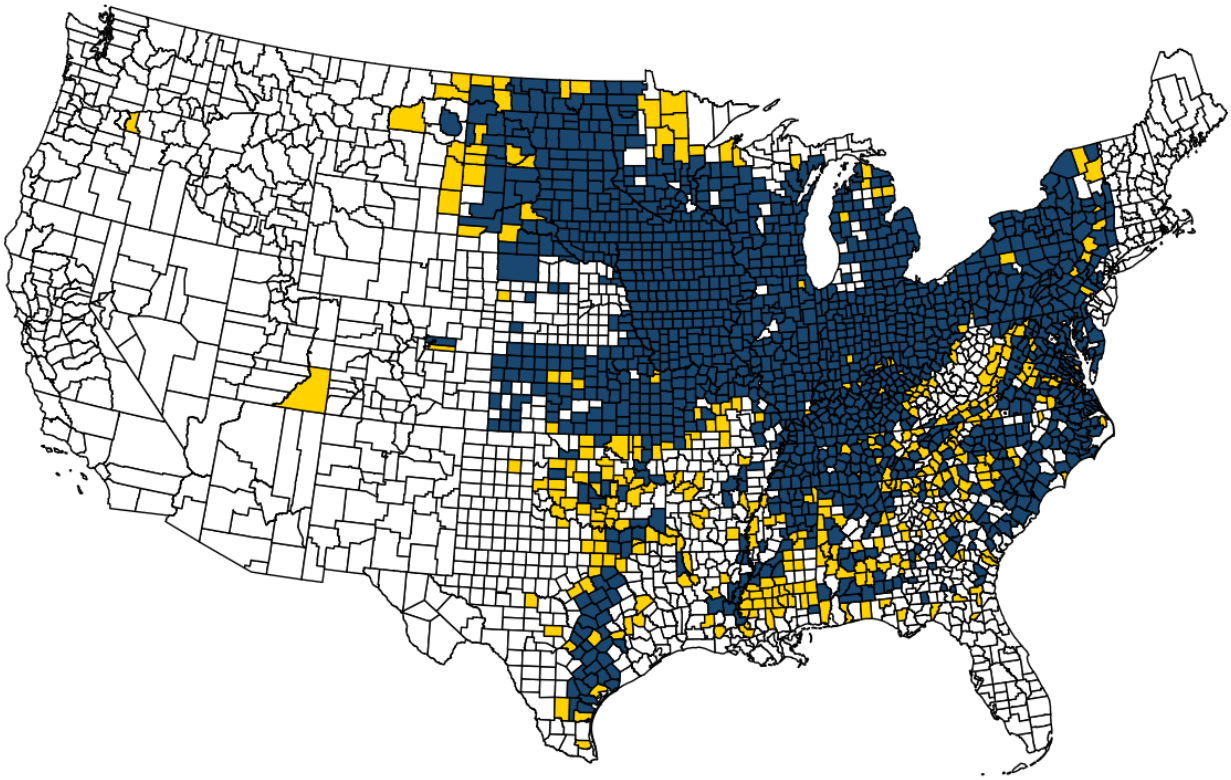
SUPPLEMENTARY APPENDIX

TABLE OF CONTENTS

Geographic scope of the analysis

The regions encompassed by the analysis are represented in Figure X1 below. Yellow counties are included in the analysis of the main text, but excluded in one of the cases considered in the Robustness section (when counties are required to have data for two thirds of the years in both the pre- and post-GE periods).

Figure X1. Counties in the analysis



GE adoption data

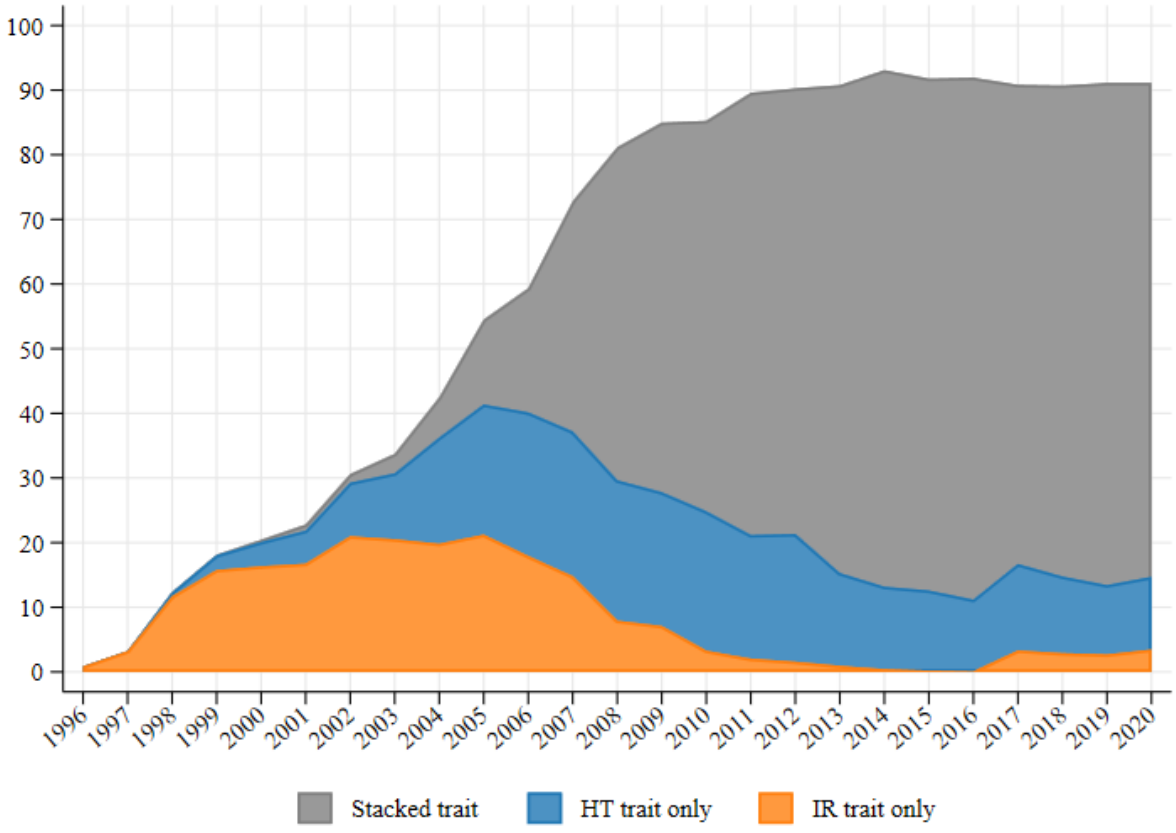
Previous work has relied on GE adoption data assembled by the National Agricultural Statistics Service (NASS) of the U.S. Department of Agriculture (USDA). These data are only available starting in year 2000, and they are available at most at the state level. Indeed, state-level adoption rates are consistently reported since 2000 only for 11 states (https://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-us/).

To get a more complete representation of the adoption of GE, we draw on an additional data source, Kynetec USA, Inc., a market research organization that collects agriculture-related survey data. These proprietary data are based on annual surveys of random, large samples of US farmers (approximately 4,700 maize farmers every year). For each surveyed farmer, Kynetec provides plot-level seed transaction data, including seed trait information and projected acres (multiple seed purchases for the same farmer are recorded if the farmers used more than one seed variety). In particular, Kynetec samples are structured to be representative at the crop reporting districts (CRD) level (CRDs are multi-county, sub-state regions identified by NASS).

Kynetec data are available to us for the 21-year period 1996-2016. Based on farm-level observations over this period, we can compute adoption rates for a broader set of states than those covered by the USDA data and, perhaps more importantly, we can do it starting with in 1996, the first year of commercialization of GE varieties. In terms of state representation, there is of course a tradeoff between precision and the extent of geographical coverage, because marginal states only contribute few farmers to the Kynetec sample. Thus, for individual state/years to be separately included in our data, we require that at least 75 distinct seed transactions per state-year be observed in the sample. Based on this condition, we generated a balanced panel of 21 state-level adoption rates. The specific states for which we compute state-level adoption rates are: Colorado, Illinois, Indiana, Iowa, Kansas, Kentucky, Maryland, Michigan, Minnesota, Missouri, Nebraska, New York, North Carolina, North Dakota, Ohio, Pennsylvania, South Dakota, Tennessee, Texas, Virginia, and Wisconsin. Additionally, similar to the USDA dataset, state-years failing the condition are pooled to provide a measure of the GE adoption rates for "other states." As noted, Kynetec data is available to us only up to 2016. For

the last four years (2017-2020), therefore, we rely on USDA's GE adoption data. This leads to some loss of detail, of course, but hopefully this is with little loss of generality because GE adoption rates have been quite stationary in recent years. Whereas both Kynetec and USDA data separately identify HT and IR traits, it seems that, at least since the mid-2000s, most commercialized maize hybrids have embedded stacked traits. This is illustrated in Figure X2. Hence, for the purpose of this paper, our focus has been on the combined adoption rate for all GE traits.

Figure X2. GE adoption rates, 1996-2020



Note: Adoption rates are calculated based on the weighted average of 1,774 counties analyzed in this study, using Kynetec data and USDA data, respectively, for the period of 1996-2016 and 2017-2020.

MACA data

Projected future weather variables are drawn from the Multivariate Adaptive Constructed Analogs (MACA) data. The MACA is a specific downscaling method, which is designed to obtain a higher spatial resolution from the native coarse resolution of global circulation models (GCMs). The specific downscaled data version is at a spatial resolution of around 6km by 6km grid cells with weather variables like daily temperature and precipitation. These sub-county grid cell weather variables are aggregated to the county level in the same way as the aggregation of PRISM data, as explained in section 2.

The MACA dataset is developed from twenty GCMs in the Coupled Model Intercomparison Project Phase 5 (CMIP5). To be specific, this dataset is generated under two representative concentration pathways (RCPs) – RCP 4.5 and RCP 8.5. In addition, each GCM provides simulated historical weather. **Table X1** shows summary statistics of 20 GCMs for the simulated historical weather (the other two models we consider, HadGEM2-ES and NorESM1-M, are extensively described in the main text). **Table X2**, and **Table X3** report summary statistics for the RCP 4.5 and RCP 8.5 scenario, respectively, for the weather variables obtained from these 20 GCMs. As shown in the tables, there exist a considerable variation in projected weather variables across the climate models.

Table X1. Summary statistics of MACA data (historical simulations)

Historical period (1981-2005)

	GDD		HDD		VPD		VPD-Jul,Aug		Precipitation	
Model name	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
BCC-CSM1-1	1660.6	(371.0)	36.32	(42.04)	1.852	(0.308)	2.440	(0.402)	587.5	(175.7)
BCC-CSM1-1-M	1647.8	(372.3)	34.29	(39.57)	1.835	(0.316)	2.410	(0.393)	595.9	(167.9)
BNU-ESM	1684.6	(390.2)	38.50	(47.09)	1.861	(0.345)	2.424	(0.424)	589.2	(166.6)
CanESM2	1665.3	(374.3)	35.72	(42.07)	1.846	(0.315)	2.420	(0.429)	591.5	(183.5)
CCSM4	1661.2	(378.8)	33.13	(42.18)	1.843	(0.323)	2.381	(0.440)	586.7	(157.5)
CNRM-CM5	1647.8	(389.1)	29.15	(40.26)	1.817	(0.336)	2.337	(0.425)	586.9	(177.1)
CSIRO-Mk3-6-0	1668.7	(391.0)	34.93	(41.25)	1.859	(0.335)	2.403	(0.417)	572.9	(168.5)
GFDL-ESM2G	1662.8	(386.0)	34.36	(46.94)	1.842	(0.343)	2.377	(0.443)	586.4	(178.2)
GFDL-ESM2M	1647.6	(388.8)	32.19	(45.46)	1.819	(0.351)	2.375	(0.445)	594.5	(171.4)
HadGEM2-CC	1639.3	(379.1)	35.63	(47.16)	1.830	(0.337)	2.395	(0.457)	578.0	(191.2)
HadGEM2-ES	1667.3	(384.5)	38.74	(52.14)	1.866	(0.351)	2.423	(0.472)	573.2	(190.7)
INMCM4	1657.4	(379.6)	36.50	(45.52)	1.843	(0.362)	2.418	(0.491)	583.1	(171.8)
IPSL-CM5A-LR	1678.6	(378.1)	37.68	(46.03)	1.856	(0.327)	2.403	(0.445)	584.5	(177.7)
IPSL-CM5A-MR	1670.3	(379.7)	34.69	(42.36)	1.846	(0.329)	2.387	(0.430)	580.4	(181.9)
IPSL-CM5B-LR	1662.9	(365.2)	36.76	(42.89)	1.840	(0.315)	2.413	(0.436)	586.6	(168.9)
MIROC5	1648.9	(375.1)	32.77	(38.78)	1.829	(0.310)	2.391	(0.394)	582.9	(155.3)
MIROC-ESM	1659.2	(377.3)	32.85	(41.05)	1.835	(0.318)	2.387	(0.415)	581.2	(162.6)
MIROC-ESM-CHEM	1645.9	(388.7)	35.03	(44.59)	1.832	(0.327)	2.383	(0.430)	584.1	(168.0)
MRI-CGCM3	1637.7	(375.5)	30.67	(38.09)	1.813	(0.308)	2.357	(0.385)	583.0	(154.1)
NorESM1-M	1667.2	(382.1)	33.63	(40.61)	1.850	(0.319)	2.395	(0.418)	595.8	(155.0)
Total	1659.0	(380.6)	34.68	(43.50)	1.841	(0.329)	2.396	(0.431)	585.2	(171.6)

Table X2. Summary statistics of MACA data (20 GCMs in RCP 4.5)

Mid century (2040-2059)										
	GDD		HDD		VPD		VPD-Jul,Aug		Precipitation	
Model name	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
BCC-CSM1-1	1822.6	(379.9)	94.09	(66.10)	2.149	(0.363)	2.860	(0.433)	563.3	(159.7)
BCC-CSM1-1-M	1885.3	(374.2)	102.4	(73.78)	2.185	(0.358)	2.960	(0.474)	550.3	(166.8)
BNU-ESM	1916.0	(358.0)	87.32	(70.07)	2.103	(0.328)	2.817	(0.477)	605.0	(193.9)
CanESM2	1931.7	(365.2)	86.24	(68.24)	2.125	(0.336)	2.712	(0.434)	616.1	(187.3)
CCSM4	1846.1	(370.8)	85.77	(67.43)	2.157	(0.358)	2.866	(0.527)	605.7	(171.9)
CNRM-CM5	1802.5	(368.8)	53.35	(57.06)	1.928	(0.336)	2.534	(0.452)	623.5	(192.1)
CSIRO-Mk3-6-0	1909.2	(384.6)	86.16	(68.58)	2.081	(0.361)	2.710	(0.488)	648.3	(209.6)
GFDL-ESM2G	1822.2	(411.2)	75.81	(86.69)	2.064	(0.452)	2.659	(0.600)	593.0	(180.1)
GFDL-ESM2M	1788.5	(429.9)	73.88	(95.74)	2.016	(0.502)	2.605	(0.626)	607.6	(204.4)
HadGEM2-CC	1926.3	(366.0)	113.2	(85.77)	2.197	(0.399)	2.968	(0.529)	584.4	(202.6)
HadGEM2-ES	1946.1	(388.3)	134.3	(90.20)	2.247	(0.406)	3.168	(0.588)	583.6	(188.3)
INMCM4	1743.1	(386.4)	59.29	(62.88)	1.919	(0.389)	2.568	(0.566)	594.5	(193.5)
IPSL-CM5A-LR	1919.2	(386.4)	82.48	(76.91)	2.097	(0.399)	2.690	(0.480)	572.3	(189.3)
IPSL-CM5A-MR	1939.0	(369.8)	88.77	(75.12)	2.113	(0.370)	2.704	(0.487)	564.9	(194.2)
IPSL-CM5B-LR	1848.5	(364.8)	60.35	(55.61)	2.007	(0.328)	2.581	(0.416)	621.1	(178.1)
MIROC5	1950.8	(355.5)	105.2	(84.89)	2.179	(0.369)	2.905	(0.566)	603.8	(168.9)
MIROC-ESM	1938.0	(348.8)	92.42	(70.33)	2.136	(0.334)	2.840	(0.461)	610.8	(168.1)
MIROC-ESM-CHEM	1986.2	(359.5)	104.6	(86.17)	2.276	(0.383)	2.896	(0.537)	580.3	(179.6)
MRI-CGCM3	1781.2	(374.3)	47.73	(51.99)	1.873	(0.321)	2.421	(0.423)	603.9	(159.9)
NorESM1-M	1874.9	(363.2)	99.55	(77.40)	2.197	(0.389)	2.919	(0.550)	608.0	(184.8)
End of century (2080-2099)										
BCC-CSM1-1	1930.4	(378.5)	116.1	(73.29)	2.266	(0.358)	2.949	(0.398)	531.2	(175.6)
BCC-CSM1-1-M	1880.3	(378.0)	99.45	(71.92)	2.180	(0.362)	2.996	(0.482)	598.6	(170.6)
BNU-ESM	2004.3	(350.4)	115.4	(74.80)	2.237	(0.321)	3.005	(0.452)	597.1	(162.6)
CanESM2	1990.2	(357.3)	91.95	(72.43)	2.150	(0.339)	2.654	(0.452)	655.0	(221.0)
CCSM4	1930.9	(367.2)	123.6	(71.70)	2.313	(0.347)	3.124	(0.535)	582.8	(171.1)
CNRM-CM5	1899.9	(390.3)	89.37	(76.38)	2.040	(0.400)	2.795	(0.539)	608.6	(174.4)
CSIRO-Mk3-6-0	2014.8	(378.1)	119.1	(85.18)	2.177	(0.377)	2.791	(0.509)	667.0	(214.4)
GFDL-ESM2G	1841.9	(417.9)	87.47	(92.90)	2.111	(0.474)	2.797	(0.642)	597.3	(176.5)
GFDL-ESM2M	1823.5	(409.3)	77.45	(99.13)	2.060	(0.486)	2.628	(0.673)	623.9	(201.5)
HadGEM2-CC	2076.5	(351.7)	174.0	(94.91)	2.406	(0.397)	3.313	(0.556)	571.3	(194.5)
HadGEM2-ES	2084.6	(355.7)	175.4	(91.03)	2.381	(0.370)	3.369	(0.566)	581.4	(184.9)
INMCM4	1811.0	(391.8)	69.10	(71.36)	1.938	(0.422)	2.596	(0.636)	605.2	(193.3)
IPSL-CM5A-LR	2019.5	(367.8)	98.95	(82.03)	2.167	(0.364)	2.754	(0.476)	586.5	(193.1)
IPSL-CM5A-MR	1973.5	(380.8)	99.74	(83.63)	2.128	(0.391)	2.728	(0.486)	601.6	(193.0)
IPSL-CM5B-LR	1915.2	(367.2)	78.31	(64.14)	2.080	(0.332)	2.721	(0.435)	623.2	(173.8)
MIROC5	2031.8	(358.2)	148.6	(109.2)	2.304	(0.423)	3.175	(0.648)	605.2	(166.3)
MIROC-ESM	2032.3	(340.3)	136.9	(93.65)	2.280	(0.378)	3.056	(0.561)	623.0	(192.0)
MIROC-ESM-CHEM	2045.6	(348.6)	123.7	(92.78)	2.348	(0.380)	3.000	(0.556)	592.4	(184.6)
MRI-CGCM3	1840.9	(363.0)	56.82	(55.58)	1.923	(0.322)	2.486	(0.407)	594.5	(156.5)
NorESM1-M	1971.7	(347.2)	125.9	(89.78)	2.336	(0.413)	3.067	(0.608)	612.6	(167.1)

Table X3. Summary statistics of MACA data (20 GCMs in RCP 8.5)

	GDD		HDD		VPD		VPD-Jul,Aug		Precipitation	
Model name	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
Mid century (2040-2059)										
BCC-CSM1-1	1906.9	(373.8)	116.8	(72.74)	2.230	(0.362)	2.985	(0.434)	577.3	(183.1)
BCC-CSM1-1-M	1946.3	(393.8)	108.5	(83.58)	2.221	(0.397)	2.885	(0.507)	595.9	(191.9)
BNU-ESM	1963.0	(353.6)	105.1	(73.41)	2.186	(0.339)	2.919	(0.446)	606.1	(181.7)
CanESM2	2004.6	(363.8)	116.5	(75.63)	2.235	(0.347)	2.859	(0.429)	618.8	(203.9)
CCSM4	1946.8	(365.2)	117.0	(80.18)	2.272	(0.369)	3.008	(0.521)	584.6	(173.9)
CNRM-CM5	1886.9	(365.3)	81.81	(67.11)	2.046	(0.357)	2.751	(0.456)	590.4	(177.4)
CSIRO-Mk3-6-0	1959.9	(363.3)	106.1	(70.18)	2.138	(0.334)	2.809	(0.441)	637.3	(199.1)
GFDL-ESM2G	1865.9	(412.8)	98.28	(100.8)	2.108	(0.481)	2.806	(0.640)	614.7	(179.3)
GFDL-ESM2M	1840.8	(385.2)	79.82	(82.10)	2.037	(0.404)	2.710	(0.532)	622.4	(171.5)
HadGEM2-CC	2001.3	(366.2)	138.3	(90.73)	2.272	(0.412)	3.087	(0.595)	581.4	(200.2)
HadGEM2-ES	2028.8	(388.0)	195.2	(119.1)	2.455	(0.486)	3.526	(0.720)	529.2	(174.5)
INMCM4	1820.7	(394.9)	77.91	(72.59)	1.993	(0.401)	2.750	(0.615)	588.4	(186.5)
IPSL-CM5A-LR	2016.0	(370.7)	110.5	(85.23)	2.188	(0.371)	2.829	(0.512)	557.7	(185.6)
IPSL-CM5A-MR	2001.2	(369.2)	113.7	(85.90)	2.174	(0.385)	2.819	(0.499)	554.4	(211.4)
IPSL-CM5B-LR	1915.4	(372.6)	73.86	(59.89)	2.060	(0.324)	2.650	(0.394)	620.0	(174.5)
MIROC5	2028.9	(358.7)	134.2	(101.1)	2.287	(0.416)	3.082	(0.602)	589.5	(150.7)
MIROC-ESM	2016.8	(344.5)	130.3	(89.28)	2.290	(0.372)	3.107	(0.578)	599.8	(170.3)
MIROC-ESM-CHEM	1994.9	(358.5)	94.22	(76.39)	2.221	(0.354)	2.826	(0.461)	617.5	(177.5)
MRI-CGCM3	1812.3	(370.8)	55.06	(57.75)	1.896	(0.331)	2.460	(0.431)	613.5	(159.1)
NorESM1-M	1968.0	(364.3)	119.9	(88.39)	2.272	(0.379)	3.058	(0.614)	625.4	(185.2)
End of century (2080-2099)										
BCC-CSM1-1	2125.4	(368.3)	221.1	(98.38)	2.595	(0.398)	3.520	(0.540)	568.8	(197.2)
BCC-CSM1-1-M	2194.2	(376.4)	236.8	(112.4)	2.651	(0.433)	3.528	(0.572)	566.4	(185.9)
BNU-ESM	2268.1	(351.2)	215.2	(102.6)	2.531	(0.369)	3.352	(0.535)	628.8	(200.9)
CanESM2	2287.2	(347.0)	249.8	(102.5)	2.632	(0.379)	3.184	(0.511)	610.3	(190.2)
CCSM4	2152.7	(374.2)	218.5	(105.0)	2.638	(0.411)	3.568	(0.645)	642.9	(198.5)
CNRM-CM5	2092.0	(370.3)	146.4	(92.74)	2.182	(0.385)	2.960	(0.498)	637.5	(188.4)
CSIRO-Mk3-6-0	2276.8	(353.7)	216.7	(106.5)	2.476	(0.389)	3.225	(0.563)	688.4	(220.6)
GFDL-ESM2G	2112.9	(392.2)	168.9	(135.5)	2.426	(0.519)	3.261	(0.740)	624.1	(195.2)
GFDL-ESM2M	2023.1	(392.9)	145.9	(123.2)	2.307	(0.486)	3.094	(0.673)	636.8	(186.7)
HadGEM2-CC	2364.5	(323.7)	365.9	(121.4)	2.964	(0.436)	4.164	(0.660)	509.5	(173.5)
HadGEM2-ES	2320.4	(366.9)	333.6	(123.1)	2.830	(0.458)	4.192	(0.720)	570.0	(184.2)
INMCM4	2000.2	(409.0)	132.3	(104.2)	2.108	(0.463)	2.890	(0.687)	594.4	(188.1)
IPSL-CM5A-LR	2274.7	(369.6)	215.9	(132.5)	2.445	(0.460)	3.118	(0.660)	581.8	(226.5)
IPSL-CM5A-MR	2317.3	(363.2)	252.9	(121.4)	2.554	(0.441)	3.299	(0.568)	528.9	(181.7)
IPSL-CM5B-LR	2145.7	(353.4)	168.8	(92.41)	2.346	(0.374)	3.099	(0.470)	618.3	(192.0)
MIROC5	2364.2	(353.3)	295.0	(153.1)	2.777	(0.530)	3.780	(0.774)	606.0	(166.2)
MIROC-ESM	2363.6	(344.8)	315.9	(158.6)	2.855	(0.552)	3.937	(0.825)	574.4	(161.9)
MIROC-ESM-CHEM	2304.2	(321.5)	212.1	(128.7)	2.655	(0.423)	3.360	(0.641)	608.6	(204.7)
MRI-CGCM3	2004.0	(370.8)	99.93	(87.64)	2.041	(0.368)	2.674	(0.510)	658.2	(179.8)
NorESM1-M	2198.2	(328.3)	232.9	(113.4)	2.702	(0.429)	3.653	(0.783)	630.8	(190.6)

Shift and Slope Models Compared

The analysis in the text relied on the so-called "adoption shift" specification of the yield response model, as articulated in equation (1). The alternative, sometime considered in related work, is the so-called "adoption slope" model, which can be written as

$$(A1) \qquad y_{it} = \alpha_i + \mathbf{X}_{it}\boldsymbol{\beta} + G_{st}\mathbf{X}_{it}\boldsymbol{\delta} + \gamma G_{st}T_t + \tau_s T_t + \varepsilon_{it}$$

Estimation results for this specification are presented in Table X9, side-to-side with the adoption shift model used in the main text. It is apparent that the in-sample fit of the two models is essentially identical.

To further compare the performances of these two specifications, we consider out-of-sample predictions within the estimation period 1981-2000. Specifically, we estimate the model over 10 subsamples, each time leaving out 4 years, and evaluate the out-of-sample forecasts for the 4 omitted years. (This procedure is similar to that used by D'Agostino and Schlenker, *Agricultural Economics* 2016). Table X10 summarized the root mean square errors of the out-of-sample forecasts for the two specifications. It is apparent that the performance of the two models is extremely close.

Table X4. Estimation results for "shift" and "slope" specifications (model 2)

	Adoption Shift Model		Adoption Slope Model	
	coefficient	s.e.	coefficient	s.e.
GE	14.22***	(0.528)		
GE×trend			0.460***	(0.0158)
Avg. trend	0.945***	(0.102)	0.912***	(0.102)
GE×GDD	0.000647	(0.00163)	-0.00162	(0.00161)
GE×HDD	0.0286	(0.0195)	0.0201	(0.0193)
GE×VPD	-13.75***	(2.602)	-11.44***	(2.603)
GE×VPD-JA	-12.93***	(1.597)	-12.41***	(1.591)
GE×PPT	-0.0110	(0.00837)	-0.0191**	(0.00834)
GE×PPTsq	0.0000252***	(0.00000659)	0.0000283***	(0.00000655)
GDD	0.0144***	(0.00129)	0.0160***	(0.00129)
HDD	-0.293***	(0.0102)	-0.285***	(0.0102)
VPD	10.47***	(1.399)	8.839***	(1.411)
VPD-JA	-27.22***	(0.704)	-27.52***	(0.706)
PPT	0.0814***	(0.00442)	0.0814***	(0.00444)
PPTsq	-0.0000799***	(0.00000385)	-0.0000796***	(0.00000388)
Constant	85.56***	(0.206)	86.37***	(0.213)
County FE	Y		Y	
N	60,400		60,400	
Adj. R2	0.782		0.783	

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table X5. Comparison of "shift" and "slope" models: Out-of-sample forecasts

Leave-out period	RMSE	
	Adoption Shift	Adoption slope
1981 - 1984	17.0	16.7
1985 - 1988	16.9	16.8
1989 - 1992	17.5	17.6
1993 - 1996	19.1	19.2
1997 - 1999	18.6	18.7
2001 - 2004	20.1	20.2
2005 - 2008	19.1	19.3
2009 - 2012	23.2	22.9
2013 - 2016	20.6	20.5
2017 - 2020	20.6	19.8
Avg.	19.3	19.2

Table X6. Clustered standard errors for model 2

	Baseline		Clustered Standard Errors, by:				
	Estimated parameter	Standard error	year	state	state & year	ge phase	year & ge phase
GE	14.22	0.528***	5.386**	1.699***	5.005***	11.69	11.7
GE×GDD	0.000647	0.00163	0.0106	0.00815	0.012	0.00245	0.00252
GE×HDD	0.0286	0.0195	0.116	0.0604	0.107	0.0703	0.0708
GE×VPD	-13.75	2.602***	14.95	11.47	15.92	9.797	9.832
GE×VPD-JA	-12.93	1.597***	9.094	9.411	11.3	3.041	3.132
GE×PPT	-0.0110	0.00837	0.0421	0.0258	0.0391	0.0213	0.0212
GE×PPTsq	0.000025	0.000007***	0.00003	0.00002	0.00003	0.00002	0.00002
GDD	0.0144	0.00129***	0.0131	0.00775*	0.0138	0.00405	0.00411
HDD	-0.293	0.0102***	0.0708***	0.0822***	0.0986***	0.0638***	0.064***
VPD	10.47	1.399***	12.83	10.16	15.01	6.941	6.979
VPD-JA	-27.22	0.704***	4.707***	6.551***	7.246***	0.474***	0.691***
PPT	0.0814	0.00442***	0.0245***	0.026***	0.0292***	0.0143***	0.0143***
PPTsq	-0.00008	0.000004***	0.00002	0.00002	0.00003	0.00001	0.00001
Avg. trend	0.945	0.102***	0.165***	0.0693***	0.139***	0.328***	0.329***
Constant	85.56	0.206***	1.924***	0.841***	1.83***	3.055***	3.056***
r2_a	0.782						
N	60400						

Note: The Baseline reports the estimates of the main text (with the EW robust standard errors). The last five columns report standard errors with alternative clustering strategies. Stars next to the standard errors flag te level of significance of the corresponding parameter estimate: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Robustness: Semi-log Model Results

Table X4 presents the yield estimation results for the semi-log specification of the three estimated models. **Table X5** reports the estimated yield gaps and technological needs under climate change, implied by the semi-log specification (Model 2). **Figure X3,** which is comparable to Figure 5 of the main text, visualizes the larger yield gaps implied by the use of the semi-log model,.

Table X7. Estimated semi-log model, 1981-2020

	Model 1		Model 2	
	coefficient	s.e.	coefficient	s.e.
GE	0.118***	(0.00609)	0.107***	(0.00601)
GE×GDD			-0.0000841***	(0.0000182)
GE×HDD			0.000713***	(0.000243)
GE×VPD			-0.187***	(0.0298)
GE×VPD-JA			0.0249	(0.0189)
GE×PPT			-0.000284***	(0.0000956)
GE×PPTsq			0.000000385***	(7.37e-08)
GDD	0.000257***	(0.0000136)	0.000292***	(0.0000152)
HDD	-0.00478***	(0.000118)	-0.00479***	(0.000146)
VPD	-0.0184	(0.0161)	0.0380**	(0.0177)
VPD-JA	-0.240***	(0.00798)	-0.250***	(0.00989)
PPT	0.000621***	(0.0000383)	0.000910***	(0.0000550)
PPTsq	-0.00000059***	(2.95e-08)	-0.00000089***	(4.65e-08)
Avg. time trend	0.0080***	(0.0011)	0.0084***	(0.0011)
Illinois	0.00849***	(0.000296)	0.00840***	(0.000293)
Indiana	0.00843***	(0.000266)	0.00816***	(0.000266)
Iowa	0.0102***	(0.000281)	0.00925***	(0.000283)
Constant	4.421***	(0.00247)	4.417***	(0.00249)
Adj. R2	0.731		0.733	
N	60,400		60,400	

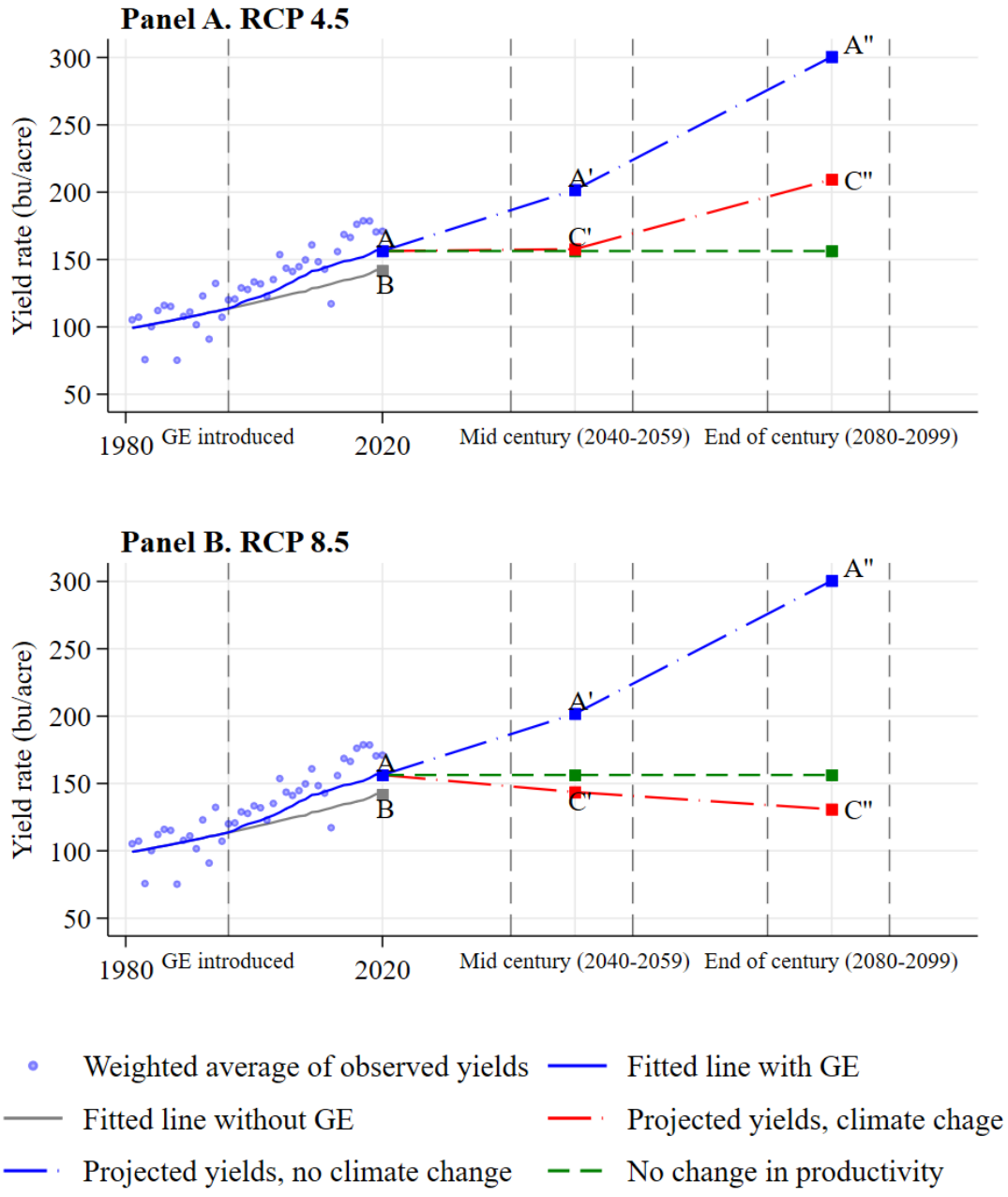
Note: All weather variables are demeaned so as to have mean zero over the estimation sample, such that the effect of the coefficient of the GE variable is directly comparable between Model 1 and Model 2 (which includes interaction effects). The *N*=60,400 observation encompass 1,774 counties in 36 states. Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. In addition to state-specific trends, all models include county-specific fixed effects.

Table X8. Summary of yield projections under climate change (Semi-log model)

		RCP 4.5	RCP 8.5
	Mean [range]	Mean [range]	Mean [range]
Projected yields (bu/acre)			
Year 2020 with GE (A)	156.33		
Year 2020 w/o GE (B)	141.71		
Mid-century w/o climate change (A')	201.79		
	[196.9, 206.7]		
End-century w/o climate change (A'')	300.54		
	[293.5, 307.9]		
Mid-century with climate change (C')		157.73	143.63
		[127.6, 198.1]	[102.9, 190.9]
End-century with climate change (C'')		209.48	130.83
		[144.0, 282.3]	[57.9, 238.7]
Yield gaps (bu/acre)			
Mid-century (A' - C')		44.05	58.16
		[7.88, 70.7]	[15.1, 94.0]
End-century (A'' - C'')		91.07	169.72
		[24.3, 154.6]	[67.9, 240.8]
Innovation gap index			
Mid-century (A' - C')/(A - B)		3.01	3.98
		[0.54,4.83]	[1.03,6.43]
End-century (A'' - C'')/(A - B)		6.23	11.60
		[1.66,10.57]	[4.64,16.47]

Note: The ln(yield) projections of the semi-log model are translated to yields (bushels per acre).

Figure X3. Forecasted yields under climate change (Semi-log model)



Robustness: Projections with Model 1

In the main text we presented the forecasted yield gaps based on Model 2, where the impact of GE adoption is captured by the interaction between GE adoption rates and 6 weather variables as well as the GE adoption itself. The alternative of using Model 1 (which omits GE-weather interaction effects) are presented in **Table X6.** It is apparent that the results reported in the corresponding **Table 6** in the main text are quite robust.

As for the geographical scope of the analysis, as noted in the text, we considered a more restrictive sample that requires included counties to have data for two thirds of the years in both the pre- and post-GE periods. The geographic coverage implications of this sampling rule are illustrated in **Figure X1**. With these samples, estimates for the yield models are reported in **Tables X7** and the corresponding projected yield gaps due to climate change are reported in **Table X8**. In addition, yield projections with the sample using the two thirds cutoff of yield observations, corresponding to Table X8, are illustrated in **Figure X4**. The results are similar to those reported in the main text, supporting robustness of our reported findings.

Table X9. Summary of yield projections under climate change (Model 1)

		RCP 4.5	RCP 8.5
	Mean [range]	Mean [range]	Mean [range]
Projected yields (bu/acre)			
Year 2020 with GE (A)	160.02		
Year 2020 w/o GE (B)	144.83		
Mid-century w/o climate change (A')	192.21		
	[190.4, 194.5]		
End-century w/o climate change (A'')	240.09		
	[238.3, 242.4]		
Mid-century with climate change (C')		169.23	160.45
		[150.3, 190.0]	[128.9, 186.5]
End-century with climate change (C'')		206.17	161.77
		[176.6, 234.1]	[100.0, 218.0]
Yield gaps (bu/acre)			
Mid-century (A' - C')		22.98	31.76
		[4.10, 40.03]	[7.59, 61.44]
End-century (A'' - C'')		33.92	78.32
		[7.85, 64.42]	[23.94, 139.78]
Innovation gap index			
Mid-century (A' - C')/(A - B)		1.51	2.09
		[0.27, 2.64]	[0.50, 4.05]
End-century (A'' - C'')/(A - B)		2.23	5.16
		[0.52, 4.24]	[1.58, 9.21]

Robustness: Geographic scope of the analysis

As noted in the text, here we considered a more restrictive sample that requires included counties to have data for two thirds of the years in both the pre- and post-GE periods. The geographic coverage implications of this sampling rule are illustrated in **Figure X1**. With these samples, estimates for the yield models are reported in **Tables X7** and the corresponding projected yield gaps due to climate change are reported in **Table X8**. In addition, yield projections with the sample using the two thirds cutoff of yield observations, corresponding to Table X8, are illustrated in **Figure X4**. The results are similar to those reported in the main text, supporting robustness of our reported findings.

Table X10. Estimated yield model with the "two thirds years" cutoff

	Model 1		Model 2	
	coefficient	s.e.	coefficient	s.e.
GE	16.82***	(0.557)	14.27***	(0.550)
GE×GDD			-0.000642	(0.00173)
GE×HDD			0.0703***	(0.0214)
GE×VPD			-19.61***	(2.703)
GE×VPD-JA			-10.16***	(1.667)
GE×PPT			-0.0240***	(0.00888)
GE×PPTsq			0.0000389***	(0.00000710)
GDD	0.0115***	(0.00126)	0.0140***	(0.00137)
HDD	-0.374***	(0.0102)	-0.356***	(0.0112)
VPD	8.386***	(1.426)	16.66***	(1.493)
VPD-JA	-30.54***	(0.663)	-29.06***	(0.752)
PPT	0.0574***	(0.00349)	0.0970***	(0.00495)
PPTsq	-0.0000545***	(0.00000281)	-0.0000955***	(0.00000438)
Avg. time trend	0.986***	(0.0301)	1.096***	(0.0300)
Illinois	1.318***	(0.0296)	1.331***	(0.0292)
Indiana	1.216***	(0.0289)	1.198***	(0.0290)
Iowa	1.510***	(0.0282)	1.446***	(0.0282)
Constant	88.15***	(0.219)	87.14***	(0.221)
Adj. R2	0.776		0.782	
N	53,765		53,765	

Note: All weather variables are demeaned so as to have mean zero over the estimation sample, such that the effect of the coefficient of the GE variable is directly comparable between Model 1 and Model 2 (which includes interaction effects). The *N*=53,765 observation encompass 1,431 counties in 32 states. Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. In addition to state-specific trends, all models include county-specific fixed effects.

Table X11. Summary of yield projections (Model 2 with the "two thirds years" cutoff)

		RCP 4.5	RCP 8.5
	Mean [range]	Mean [range]	Mean [range]
Projected yields (bu/acre)			
Year 2020 with GE (A)	160.01		
Year 2020 w/o GE (B)	146.96		
Mid-century w/o climate change (A')	190.60		
	[188.2, 193.3]		
End-century w/o climate change (A'')	239.38		
	[237.0, 242.1]		
Mid-century with climate change (C')		164.27	154.64
		[142.9, 188.7]	[119.0, 184.9]
End-century with climate change (C'')		201.01	153.39
		[167.7, 232.8]	[85.7, 216.3]
Yield gaps (bu/acre)			
Mid-century (A' - C')		26.34	35.96
		[4.1, 45.3]	[7.9, 69.2]
End-century (A'' - C'')		38.38	85.99
		[8.8, 72.5]	[25.3, 153.3]
Innovation gap index			
Mid-century (A' - C')/(A - B)		2.02	2.76
		[0.3, 3.5]	[0.6, 5.3]
End-century (A'' - C'')/(A - B)		2.94	6.59
		[0.7, 5.6]	[1.9, 11.7]

Figure X4. Forecasted yields under climate change ("two thirds yesr"cutoff)