Estimating Continuous Treatment Effects in Panel Data using Machine Learning with an Agricultural Application

Max Vilgalys[1] and Sylvia Klosin[2]

4 October 2022

[1]MIT IDSS, Social and Engineering Systems [2]MIT Economics

Average temperatures 1°C warmer than pre-industrial levels

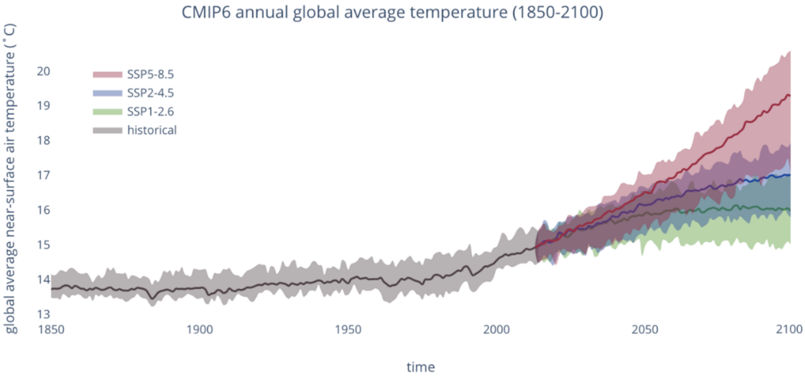


Figure 1: Historical and projected average temperature, under various climate scenarios. From Copernicus Climate Change Service, ECMWF. `https://climate.copernicus.eu/latest-projections-future-climate-now-available`

Climate change already has severe impacts

- Damages in agriculture, mortality, productivity, energy
- Driven by extreme weather and shifting averages
- **Today's talk: damaging heat in U.S. corn production**
- Methods relevant for other sectors

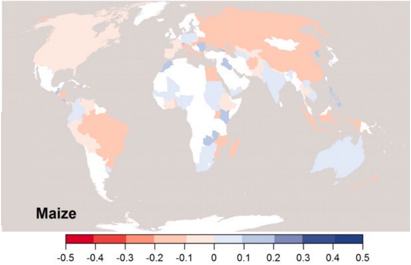


Figure 2: How much climate change impacted global corn yields from 1980-2008. Units are share of overall trend. Figure from Lobell et al. (2011).

Studying recent climate change can prepare for future impacts

- Can use data from recent change to:
  - Anticipate damages from projected climate change
  - Evaluate how much adaptation may offset damages
- Helps inform policy questions:
  - Quantify benefit of climate change mitigation
  - Identify sectors with limited adaptation
  - Useful for long-term budget forecasts, social cost of carbon
- Requires estimating economic impact of weather shocks

Can estimate impact of weather shocks from data

- Key measurement challenges:
  - Address fixed effects to isolate weather shock (Deschênes and Greenstone, 2007; Dell et al., 2012)
  - Nonlinear impacts of weather (Schlenker and Roberts, 2009; Deschênes and Greenstone, 2011)
  - Identifying adaptation from panel variation (Barreca et al., 2016; Burke and Emerick, 2016; Lemoine, 2018)
- Measurement targets:
  - Elasticity of weather feature (Schlenker and Roberts, 2009; Burke and Emerick, 2016; Downey et al., 2021)
  - Damage function (Hsiang et al., 2017; Rode et al., 2021)

Our application: extreme heat and crop yields

- Crop yields decline with high temperatures (Schlenker and Roberts, 2009)
- Major component of climate change impacts (Hsiang et al., 2017)
- Little evidence of adaptation (Burke and Emerick, 2016; Lemoine, 2018)

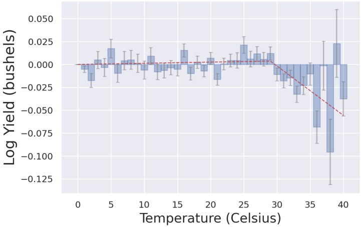


Figure 3: Change in corn yields with exposure to one additional day in each temperature bin. Estimates from OLS estimation with county and year fixed effects, from 1980-2019 U.S. corn production.

Measurement target: damages from extreme heat

- Elasticity of yield with respect to damaging heat
- Lets us estimate future damages:
  - How does extreme heat impact crop yields?
- Lets us measure adaptation:
  - Is elasticity changing over time?

We introduce new measurement approach

- Panel double machine learning (DML) estimator
- Addresses fixed effects without parametric restrictions
- Good for high-dimensional and/or nonlinear settings
- Learns function with less data than standard approaches (e.g. Dell et al. 2014; Hsiang 2016)
- Properties of estimator are not main focus today

With this estimator, we can:

- Consider flexible functional forms
- Include high-dimensional weather features
- Measure elasticity in short panels
- Preserve low standard errors

Today's talk: preview of results

- How does extreme heat impact crop yields?
    - *DML finds significantly greater damages than OLS*
- Is elasticity changing over time?
    - *No evidence that elasticity changes over time*

Related Literature

- Economics of climate change:
  - Developing a damage function: **Auffhammer et al. (2013)**, Dell et al. (2014), Hsiang (2016), Burke et al. (2015), Hsiang et al. (2017), Rode et al. (2021)
  - Evaluating adaptation: **Barreca et al. (2016),** Burke and Emerick (2016), Lemoine (2018), Mérel and Gammans (2021)
- Extreme heat in agriculture: **Schlenker and Roberts (2009), Burke and Emerick (2016),** Lobell et al. (2013), Butler and Huybers (2015), Liu et al. (2016)
- Machine learning in environmental economics: **Crane-Droesch (2018),** Knittel and Stolper (2019), Deryugina et al. (2019), Stetter et al. (2022)
- Double/debiased machine learning: **Belloni et al. (2016)**, Chernozhukov et al. (2018), Colangelo and Lee (2020), Klosin (2021), Rothenhäusler and Yu (2019), **Chernozhukov et al. (2022)**

Today's Talk

New estimator to measure climate change damages

- Develop estimator of continuous treatment effects in panel data
- *In paper: prove asymptotic normality, debiasedness*
- Show that estimator works via simulation exercise
- Measure impacts of extreme heat in U.S. agriculture
    - What is the extreme heat elasticity of crop yields?
    - Is this changing over time?

Why (double) machine learning

- Don't always have expert guidance to develop linear model
- ML: flexible modeling and low standard errors
- Requires less data than classical nonparametric methods
- DML: overcome regularization/overfitting bias from ML

Notation for estimating impact of extreme heat

- Schlenker and Roberts (2009) introduce a parsimonious linear model:

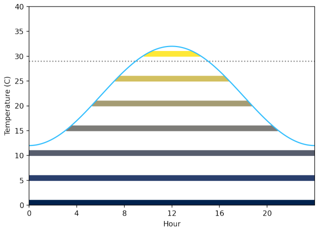
$$y_{it} = a_i + \beta_1 lower_{it} + \beta_2 higher_{it} + g(prec_{it}) + \varepsilon_{it} \tag{1}$$

- y_{it} is log corn yield, $lower_{it}$ is beneficial heat exposure, $higher_{it}$ is damaging heat exposure, $prec_{it}$ precipitation
- We allow general interactions, polynomial terms:

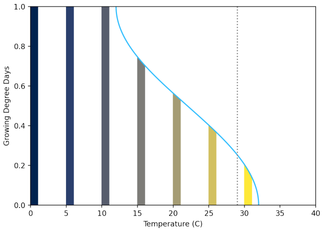
$$y_{it} = a_i + \gamma_0(D_{it}, X_{it}) + \varepsilon_{it} \tag{2}$$

- D_{it} treatment variable (damaging heat), X_{it} control variables (beneficial heat, precipitation)
- Elasticity is β_2 in 1, $\partial\gamma_0/\partial D_{it}$ in 2

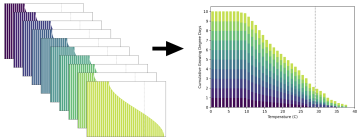
Clarification: Growing Degree Days/Extreme Heat



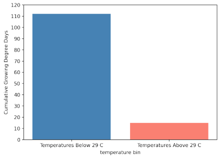
(a)



(b)



(c)



(d)

Figure 4: Transformations from daily minimum and maximum temperature records into the weather variables used in the analysis.

Today's Talk

Our panel double machine learning method

Our approach:

1. Take basis function transformations of D, X
2. Address fixed effects via first differences (Wooldridge, 2010)
3. Fit regression function $\hat{\gamma}$ via Lasso/other machine learning (ML) details
4. Take analytical derivative of $\hat{\gamma}$ details
5. Correct estimates using $\hat{\alpha}$, a second ML (Chernozhukov et al., 2022)

Introducing α, the Riesz Representer

- α is the *Riesz Representer* to the derivative operator

$$\mathbb{E}\left[\frac{\partial \Delta h(D_{i,t}, X_{i,t})}{\partial D_{i,t}}\right] = \mathbb{E}[\alpha(D_{i,t}, X_{i,t})\Delta h(D_{i,t}, X_{i,t}))] \qquad (3)$$

- Can find average derivative of an arbitrary function h using α instead of taking a derivative [1]
- *Example: average derivative of* $\gamma(X)$ *where* $X \sim N(0,1)$. *Via integration by parts after expanding the expectation,* $\mathbb{E}\left[\frac{\partial \gamma(X)}{\partial X}\right] = \mathbb{E}[X\gamma(X)]$.
- Estimate $\hat{\alpha}$ separately from estimating $\hat{\gamma}$, as in Chernozhukov et al. (2022) details

[1] Delta is the first difference operator; for a function: $\Delta h(D_{i,t}, X_{i,t}) := h(D_{i,t}, X_{i,t}) - h(D_{i,t-1}, X_{i,t-1})$

DML combines two ML estimates

$$\hat{\tau} = \mathbb{E}\left[\frac{\partial \Delta\hat{\gamma}_{i,t}}{\partial D_{i,t}} + \hat{\alpha}(D_{i,t}, X_{i,t})(\Delta Y_{i,t} - \Delta\hat{\gamma}_{i,t})\right] \tag{4}$$

- Conceptually: γ does regression, α takes a derivative
- Combining these, we correct biases from either alone (Chernozhukov et al., 2022)
- α is Riesz Representer to the derivative operator

Full estimator includes cross folds; asymptotic variance accounts for within-unit clustering details

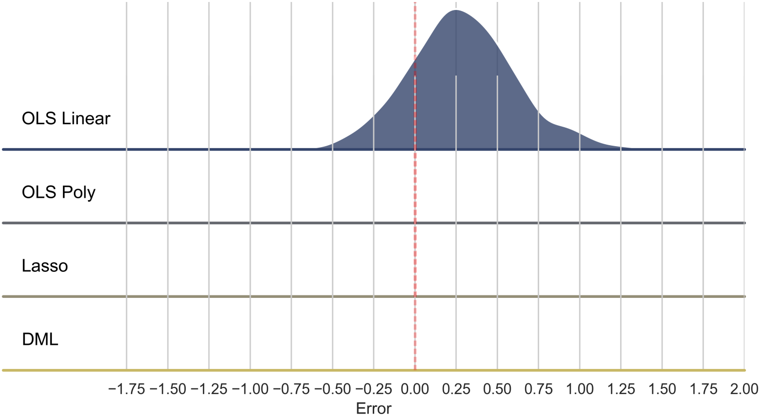
Today's Talk

Simulation of nonlinear panel setting

- Evaluate average derivative with OLS, Lasso, and DML
- Nonlinear function of treatment, including interactions
- Correlation between treatment, covariates, fixed effect
- Panel dataset: $N = 1000$, $T = 2$
- Treatment variable $D \in \mathbb{R}$; control variables $X \in \mathbb{R}^{20}$
- 1,000 Monte Carlo draws

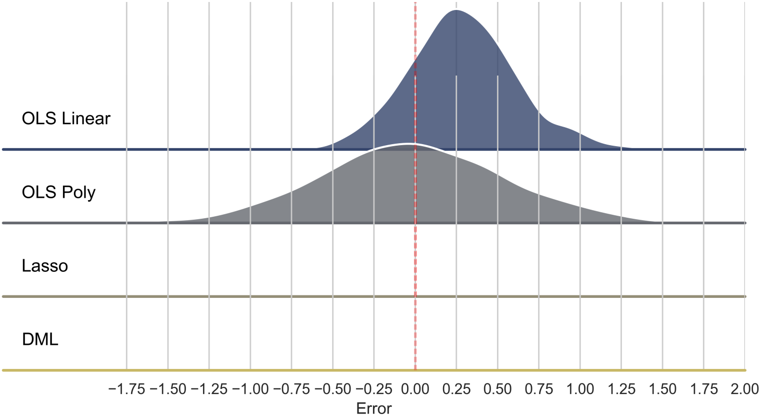
Details

Linear models can induce biases



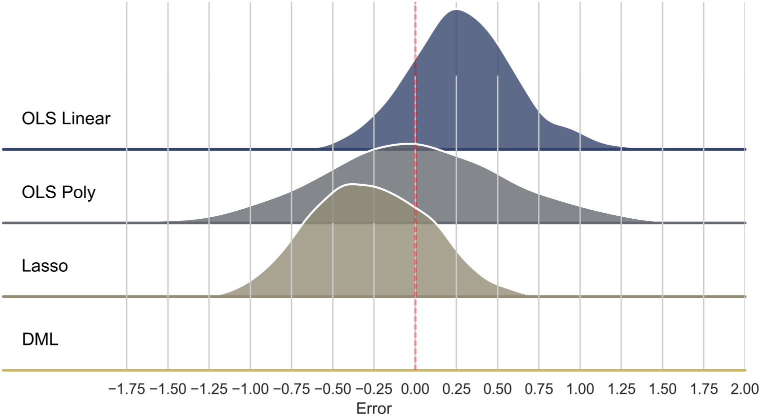
Density plot of errors from 1000 simulation trials. Zero error in red.

Flexible modeling with OLS has large standard errors



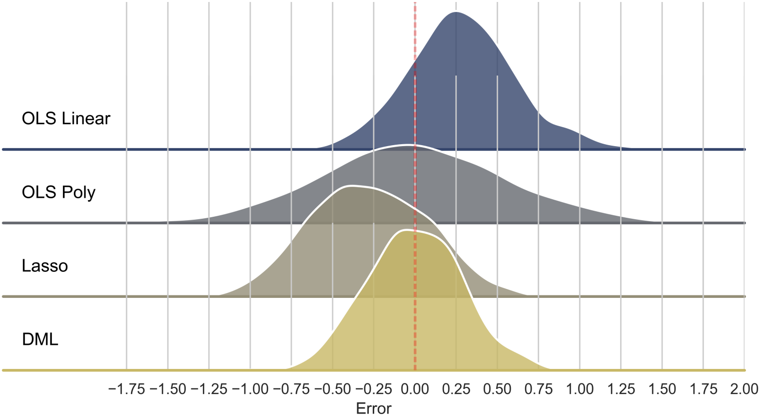
Density plot of errors from 1000 simulation trials. Zero error in red.

Regularization (Lasso) induces biases



Density plot of errors from 1000 simulation trials. Zero error in red.

Double machine learning corrects biases



Density plot of errors from 1000 simulation trials. Zero error in red.

Today's Talk

Measuring damages from extreme heat

- Measurement questions:
  - How does extreme heat impact crop yields?
  - Is elasticity changing over time?
- We estimate elasticity of U.S. corn yield with respect to extreme heat (GDD above 29°C)
- One large channel where climate change impacts crops (Schlenker and Roberts, 2009)
- With this new method, we can:
  - Include polynomials and interactions
  - Preserve low standard errors
  - Estimate elasticity flexibly in short panels

Data

- U.S. crop yields and weather from 1980-2019
- Counties east of 100°W
- Corn yield and area from USDA Survey of Agriculture
- Weather from Abatzoglou (2013), averaged from daily gridded values to county level
- Relevant weather features: precipitation, Growing Degree Days
- Additional exercise includes 7 additional weather features

Estimation: comparing 4 approaches

1. Schlenker and Roberts (2009) functional form
2. OLS with flexible set of basis functions
3. ML: Lasso without bias correction
4. DML: Lasso with our bias correction

- All address fixed effects via first differences
- All observations weighted by acres of corn per county
- ML/DML approaches use 5-folds cross validation

Summary of results: better model fit, larger damages

	OLS Linear	OLS Poly	Lasso	DML
Average Derivative	-0.005193	-0.005657	-0.005821	-0.005823
	(0.000099)	(0.000135)	(0.000011)	(0.000073)
MSE In Sample	0.080929	0.077958	0.077878	0.077878
MSE Cross Folds	0.080975	0.078353	0.078079	0.078079
Observations	63662	63662	63662	63662
Covariates	3	36	36	36

Table 1: Estimates of elasticity of corn yields with respect to increase in growing season exposure to extreme heat. Standard errors clustered at county level.

Difference is statistically, economically significant

- Differences are statistically significant:
  - OLS Linear and DML: p value < 0.001
  - OLS Linear and OLS Poly: p value 0.006
- Following Auffhammer et al. (2013), translate to economic damage from climate scenarios
- Damage estimates under median climate scenario:
  - OLS Linear: Central est. $16.2, 95% conf int $[15.2, 17.1]$
  - DML: Central est. $17.7, 95% conf int $[17.0, 18.4]$

Translating coefficient estimate into damages

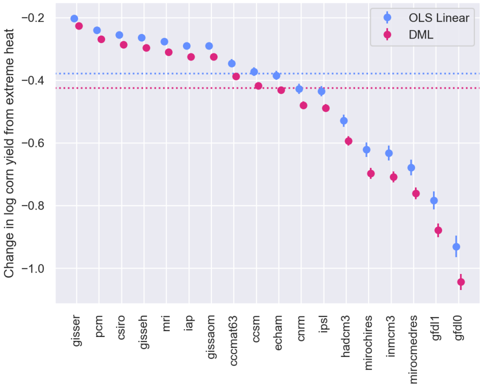


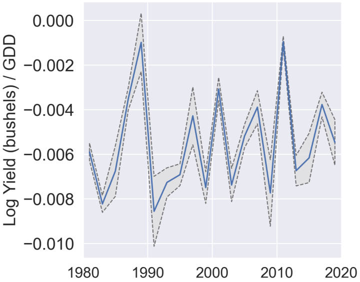
Figure 5: Declines in corn yields by 2050 due to increased exposures to temperatures above 29°C, under a range of climate scenarios. County-level processed data of climate projections generously provided by Burke and Emerick (2016).

Evaluating adaptation: Is elasticity changing over time?

- Damage declining over time could be evidence of adaptation
- Adaptation in the sense studied by Barreca et al. (2016) [2]
- Captures intensive margin, system-level adaptation
- Estimate in 2-year panels from 1980-2019

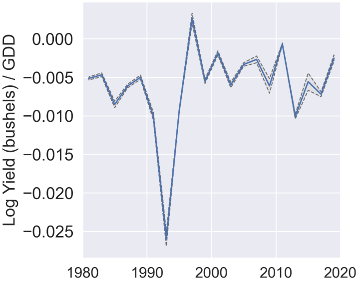
[2] Distinct from Burke and Emerick (2016), which asks if elasticity varies with exposure to climate change

Cannot reject that elasticities are constant over time



OLS Linear

Trend 3.43e-05, p-val 0.433



DML

Trend 9.52e-05, p-val 0.387

Estimating elasticity of corn yield over time, in 2-year panels. Line shows central estimate, and grey band shows 95% confidence interval.

Why are these estimates so variable?

- Some significant weather events (e.g. 1993 flood) more
- Other sources of variability:
  - Extreme heat more damaging during certain phases of plant growth (Ortiz-Bobea, 2013)
  - Regional differences (Butler and Huybers, 2015; Ortiz-Bobea et al., 2019)

Conclusions

- Panel DML model works well in simulations
- In empirical application, similar results to Schlenker and Roberts (2009)
- DML finds significantly greater damages
- No evidence that damages are decreasing over time
- Analysis with more weather variables finds similar results

Abatzoglou, J. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology 33*(1), 121–131.

Auffhammer, M., S. Hsiang, W. Schlenker, and A. Sobel (2013). Using Weather Data and Climate Model Output in Economic Analyses of Climate Change. *Review of Environmental Economics and Policy 7*(DOE/ER64640-2).

Barreca, A., K. Clay, O. Deschenes, M. Greenstone, and J. S. Shapiro (2016). Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the Twentieth Century. *Journal of Political Economy 124*(1), 105–159.

Belloni, A., V. Chernozhukov, C. Hansen, and D. Kozbur (2016). Inference in high-dimensional panel models with an application to gun control. *Journal of Business & Economic Statistics 34*(4), 590–605.

Bibliography ii

Burke, M. and K. Emerick (2016). Adaptation to climate change: Evidence from US agriculture. *American Economic Journal: Economic Policy 8*(3), 106–140.

Burke, M., S. Hsiang, and E. Miguel (2015). Global non-linear effect of temperature on economic production. *Nature 527*(7577), 235–239.

Butler, E. E. and P. Huybers (2015). Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase. *Environmental Research Letters 10*(3), 034009.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. K. Newey, and J. Robins (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal 21*(1), C1–C68.

Chernozhukov, V., W. K. Newey, and R. Singh (2022). Automatic debiased machine learning of causal and structural effects. *Econometrica 90*(3), 967–1027.

Colangelo, K. and Y.-Y. Lee (2020). Double debiased machine learning nonparametric inference with continuous treatments. *arXiv preprint arXiv:2004.03036*.

Crane-Droesch, A. (2018). Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. *Environ. Res. Lett 13*, 114003.

Dell, M., B. F. Jones, and B. A. Olken (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics 4*(3), 66–95.

Dell, M., B. F. Jones, and B. A. Olken (2014). What Do We Learn from the Weather? The New Climate-Economy Literature. *Journal of Economic Literature 52*(3), 740–798.

Deryugina, T., G. Heutel, N. H. Miller, D. Molitor, and J. Reif (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review 109*(12), 4178–4219.

Deschênes, O. and M. Greenstone (2007). The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *American Economic Review 97*(1), 354–385.

Deschênes, O. and M. Greenstone (2011). Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. *American Economic Journal: Applied Economics 3*(4), 152–185.

Downey, M., N. Lind, and J. G. Shrader (2021). Adjusting to rain before it falls. Technical report, Working Paper.

Hsiang, S. (2016). Climate econometrics. *Annual Review of Resource Economics 8*(1), 43–75.

Hsiang, S., R. Kopp, A. Jina, J. Rising, M. Delgado, S. Mohan, D. J. Rasmussen, R. Muir-Wood, P. Wilson, M. Oppenheimer, K. Larsen, and T. Houser (2017). Estimating economic damage from climate change in the United States. *Science 356*(6345), 1362–1369.

Klosin, S. (2021). Automatic Double Machine Learning for Continuous Treatment Effects. *arXiv preprint arXiv:2104.10334*.

Knittel, C. R. and S. Stolper (2019). Using machine learning to target treatment: The case of household energy use. Technical report, National Bureau of Economic Research.

Lemoine, D. (2018). Estimating the Consequences of Climate Change from Variation in Weather. *Manuscript. National Bureau of Economic Research* (w25008).

Liu, B., S. Asseng, C. Müller, F. Ewert, J. Elliott, D. B. Lobell, P. Martre, A. C. Ruane, D. Wallach, J. W. Jones, C. Rosenzweig, P. K. Aggarwal, P. D. Alderman, J. Anothai, B. Basso, C. Biernath, D. Cammarano, A. Challinor, D. Deryng, G. De Sanctis, J. Doltra, E. Fereres, C. Folberth, M. Garcia-Vila, S. Gayler, G. Hoogenboom, L. A. Hunt, R. C. Izaurralde, M. Jabloun, C. D. Jones, K. C. Kersebaum, B. A. Kimball, A. K. Koehler, S. N. Kumar, C. Nendel, G. J. O'Leary, J. E. Olesen, M. J. Ottman, T. Palosuo, P. V. Prasad, E. Priesack, T. A. Pugh, M. Reynolds, E. E. Rezaei, R. P. Rötter, E. Schmid, M. A. Semenov, I. Shcherbak, E. Stehfest, C. O. Stöckle, P. Stratonovitch, T. Streck, I. Supit, F. Tao, P. Thorburn, K. Waha, G. W. Wall, E. Wang, J. W. White, J. Wolf, Z. Zhao, and Y. Zhu (2016). Similar estimates of temperature impacts on global wheat yield by three independent methods. *Nature Climate Change 6*(12), 1130–1136.

Lobell, D. B., G. L. Hammer, G. McLean, C. Messina, M. J. Roberts, and W. Schlenker (2013). The critical role of extreme heat for maize production in the United States. *Nature Climate Change 3*(5), 497–501.

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

Mérel, P. and M. Gammans (2021). Climate Econometrics: Can the Panel Approach Account for Long-Run Adaptation? *American Journal of Agricultural Economics 103*(4), 1207–1238.

Ortiz-Bobea, A. (2013). Is Weather Really Additive in Agricultural Production? Implications for Climate Change Impacts. *SSRN Electronic Journal 1*, 3–41.

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

Rode, A., T. Carleton, M. Delgado, M. Greenstone, T. Houser, S. Hsiang, A. Hultgren, A. Jina, R. E. Kopp, K. E. McCusker, et al. (2021). Estimating a social cost of carbon for global energy consumption. *Nature 598*(7880), 308–314.

Rothenhäusler, D. and B. Yu (2019). Incremental causal effects. *arXiv preprint arXiv:1907.13258*.

Schlenker, W. and M. J. Roberts (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences of the United States of America 106*(37), 15594–8.

Bibliography ix

Stetter, C., P. Mennig, and J. Sauer (2022). Using Machine Learning to Identify Heterogeneous Impacts of Agri-Environment Schemes in the EU: A Case Study. *European Review of Agricultural Economics*.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Detail on Estimating Lasso

1. Construct the dictionary of basis functions b, a $p \times 1$ dictionary. We do so by using polynomial basis functions of terms and interactions, although other approaches like kernel functions or splines could be used as long as the derivatives are bounded.
2. Set each function in the dictionary b to have mean 0 and variance 1
3. Find a vector of coefficients $\hat{\beta}$ for our dictionary such that $\Delta\hat{\gamma}(D_{i,t}, X_{i,t}) := \Delta b(D_{i,t}, X_{i,t})'\hat{\beta}$ is a sparse linear approximation of $\Delta\gamma_0(D_{i,t}, X_{i,t})$. We do so by solving the following Lasso problem:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} (\Delta Y_{i,t} - \Delta b(D_{i,t}, X_{i,t})'\beta)^2 + r_L|\beta|_1 \right\} \tag{5}$$

This procedure depends on the regularization weight r_L, which we determine by finding values that minimize test-set error in a cross-folds procedure.

return

Detail on Estimating Derivatives

1. Construct the dictionary b_D, a $p \times 1$ dictionary of derivatives of each basis function in b. For each basis function b^j for $j = 1, \ldots, p$ in our dictionary of basis functions, define its derivative as follows:

$$b_D^j(D, X) = \frac{\partial b^j(D, X)}{\partial D} \tag{6}$$

2. Estimate the average derivative as:

$$\mathbb{E}\left[\frac{\partial \Delta\hat{\gamma}(D_{i,t}, X_{i,t})}{\partial D_{i,t}}\right] = \mathbb{E}[b_D(D_{i,t}, X_{i,t})'\hat{\beta}] \tag{7}$$

return

Simple example: how we estimate derivatives

Consider a simple setting where $X_{i,t} \in \mathbb{R}$, and where $\gamma_0(D_{i,t}, X_{i,t}) = D_{i,t}^2 X_{i,t}$. Our basis function dictionary is $b(D_{i,t}, X_{i,t}) = \{D_{i,t}, X_{i,t}, D_{i,t}^2 X_{i,t}\}$. In our linear representation, $\beta_0 = \{0, 0, 1\}$.

In step 1, we obtain an estimate $\hat{\beta}$ using Lasso. In step 2, we first define the derivative of the basis functions. Here, $b_D(D_{i,t}, X_{i,t}) = \{1, 0, 2D_{i,t} X_{i,t}\}$. The estimated average derivative is then: $\hat{\beta}_1 + \mathbb{E}[2D_{i,t} X_{i,t}]\hat{\beta}_3$, where $\hat{\beta}_j$ is the j^{th} component of $\hat{\beta}$.

return

Riesz Representer estimation details

We assume that α_0 has a sparse linear form:
$\alpha_0(D_{i,t}, X_{i,t}, D_{i,t-1}, X_{i,t-1}) = \Delta b(D_{i,t}, X_{i,t})'\rho_0.$, and estimate $\hat{\rho}$

$$\hat{\rho} = \underset{\rho}{\operatorname{argmin}} \left\{ \mathbb{E}[(\alpha_0(D_{i,t}, X_{i,t}, D_{i,t-1}, X_{i,t-1}) - \Delta b(D_{i,t}, X_{i,t})'\rho)^2] + r_\alpha |\rho|_1 \right\} \tag{8}$$

This equality holds regardless of the function b, so we estimate α from data independently of estimating the function γ. We follow Chernozhukov et al. (2022) and find $\hat{\rho}$ to minimize the squared loss between α_0 and $\hat{\alpha}$:

We introduce a novel optimization based approach; we show that it results in lower bias, variance, and MSE than an iterative approach in simulations

return

Riesz Representer estimation details

Find the estimator $\hat{\alpha}(W_{i,t})$ that minimizes the mean squared error (MSE) between $\hat{\alpha}$ and α_0:

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}}\, \mathbb{E}[(\alpha_0(W_{i,t}) - \alpha(W_{i,t}))^2]$$

With assumption that $\alpha_0 = \rho_0 b(W_{i,t})$, and that ρ_0 is sparse:

$$\begin{aligned}
\hat{\rho} &= \underset{\rho}{\operatorname{argmin}}\, \mathbb{E}\left[(\alpha_0(W_{i,t}) - \Delta b(W_{i,t})\rho)^2\right] + \lambda|\rho|_1 \\
&= \underset{\rho}{\operatorname{argmin}} -2\mathbb{E}[b_D(W_{i,t})]\rho + \rho' \mathbb{E}[\Delta b(W_{i,t})' \Delta b(W_{i,t})]\rho + \lambda|\rho|_1 \\
&= \underset{\rho}{\operatorname{argmin}} -2\hat{M}\rho + \rho' \hat{Q}\rho + \lambda|\rho|_1
\end{aligned}$$

2nd equality: from Riesz representation theorem,
$\mathbb{E}[\alpha_0(W_{i,t})h(W_{i,t})] = \mathbb{E}[\partial h(W_{i,t})/\partial D]$
where $\hat{M} := \mathbb{E}[b_D(W_{i,t})]$ and $\hat{Q} := \mathbb{E}[\Delta b(W_{i,t})' \Delta b(W_{i,t})]$

return

Estimator with cross folds

$$\hat{\tau} = \frac{1}{n(T-1)} \sum_{\ell=1}^{L} \sum_{i \in \ell} \sum_{t=2}^{T} \hat{\tau}_{\ell;i,t}$$

$$\hat{\tau}_{\ell;i,t} = \frac{\partial \Delta \hat{\gamma}_\ell(D_{i,t}, X_{i,t})}{\partial D_{i,t}} + \hat{\alpha}_\ell(D_{i,t}, X_{i,t}, D_{i,t-1}, X_{i,t-1})(\Delta Y_{i,t} - \Delta \hat{\gamma}_\ell(D_{i,t}, X_{i,t})) \tag{9}$$

We assume that errors have a constant correlation within a panel unit but are uncorrelated between panel units. Let $\hat{\tau}_{\ell;i} = 1/(T-1)\sum_{t=2}^{T} \hat{\tau}_{\ell;i,t}$. Then the asymptotic variance is:

$$\hat{V} = \frac{1}{n(T-1)} \sum_{\ell=1}^{L} \sum_{i \in \ell} \left\{ \sum_{t=2}^{T} (\hat{\tau}_{\ell;i,t} - \hat{\tau})^2 + 2 \sum_{t=2}^{T-1} \sum_{t'=t+1}^{T} (\hat{\tau}_{\ell;i,t} - \hat{\tau}_{\ell;i})(\hat{\tau}_{\ell;it'} - \hat{\tau}_{\ell;i}) \right\} \tag{10}$$

return

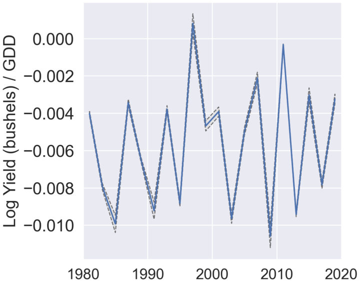
Details of simulated dataset

$$Y_{i,t} = a_i + D_{i,t} + D_{i,t}^2 + D_{i,t}^3 + D_{i,t}X_{i,t}^{(1)} + .1\theta \mathbf{X}_{i,t} + \epsilon_{i,t} \tag{11}$$

- $\theta^{(j)} = 1/j^2$
- Fixed effects a_i, covariates $\mathbf{X}_{i,t}$, and random noise $\epsilon_{i,t}$ are Gaussian R.V.
- Treatment correlated with $X_{i,t}$, with noise from Beta distribution $D_{i,t} \sim .1\theta \mathbf{X}_{i,t} + Beta(1,7)$

Return

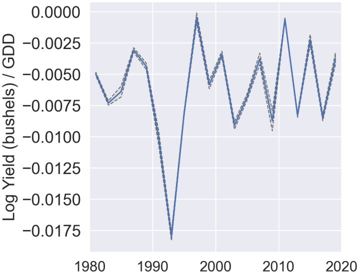
Elasticities over time, removing power and interactions



DML

No interactions

Trend 3.7e-05, p-val 0.579



DML

No polynomials

Trend 3.43e-05, p-val 0.433

Estimating elasticity of corn yield over time, in 2-year panels. Line shows central estimate, and grey band shows 95% confidence interval. Some variability is driven by interactions, especially from 1993 drought. return