Quantifying Climate Damages When Regions Trade: A Structural Gravity Approach

Jeanne Astier,* Geoffrey Barrows,† Raphael Calel,‡ Hélène Ollivier§

Feb 28, 2025

Preliminary and Incomplete.
Please do not cite and do not distribute.

Abstract

This paper presents a method for estimating treatment effects of local cost shocks when regions trade with each other. Because of spillovers induced by trade flows, comparing the evolution of outcomes between pre-shock and post-shock periods in regions exposed versus unexposed to local shocks leads to a biased estimate of treatment effect. We model these across-region dependencies using standard assumptions from international trade theory. We use our model-consistent estimation strategy to revisit the literature on the evaluation of impacts from climate change onto country-level gross output using year-to-year variation in temperature and precipitation.

Keywords: climate change , spillovers, trade, gravity
JEL Classification: Q48, L1, L5

*Center for Research in Economics and Statistics (CREST), Institut Polytechnique de Paris. Email: jeanne.astier@ensae.fr.

†Center for Research in Economics and Statistics (CREST), CNRS, Ecole polytechnique, GENES, ENSAE Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France. Email: geoffrey-masters.barrows@polytechnique.edu.

‡McCourt School of Public Policy, Georgetown University. Email: raphael.calel@georgetown.edu

§Paris School of Economics - CNRS, Address: Paris School of Economics, 48 Boulevard Jourdan 74014 Paris, France. Email: helene.ollivier@psemail.eu.

1 Introduction

The world has warmed considerably over the last 50 years, and is expected to continue to do so for the foreseeable future.[1] Quantifying the economic effects of this warming is essential for setting optimal policy (Nordhaus, 1992), and for understanding the divergent economic paths of nations (Acemoglu et al., 2001).

Estimates of the economic effects of climate are mostly based on historical correlations between weather and outcomes (Auffhammer et al, 2013). Researchers argue that identification in these models follows from the conditional mean independence of year-to-year variation in weather with respect to year-to-year variation in unobserved shocks to supply and demand (Deschênes & Greenstone, 2007; Dell et al., 2012; Deryugina & Hsiang, 2014; Burke et al., 2015). However, when regions trade with each other, weather-induced shocks to supply or demand in one region can affect the economic outcomes of other regions through input-output linkages, competition, and income effects (Eaton & Kortum, 2002; Hsieh & Ossa, 2016). In this case, the stable unit treatment value assumption (SUTVA) likely fails, making it harder to argue credibly that reduced-form models recover unbiased estimates of target parameters (Donaldson, 2015).

In this paper, we propose a structural approach to estimating the effects of climate change when regions trade with each other. We start by characterizing the outputs of popular reduced-form estimation strategies when the true data generating process features heterogeneous effects and spillovers across units. We pay particular attention to the two-way fixed effect (TWFE) estimator because it is the most commonly used empirical model, but we also study the performance of the heterogeneous treatment robust estimator from de Chaisemartin et al. (2024b), the estimator from Das et al. (2022); Feng et al. (2023); Zappalà (2024) that controls for weighted average shocks upstream and downstream, and the global time-series estimator of Bilal & Känzig (2024). In a general setting, there are more parameters to estimate than data points, so restrictions on the statistical relationships between inputs and outputs are required to derive any of these reduced-form estimation equations. By contrast, we leverage quantitative trade theory to model spillovers, allowing

[1] As of 2024, the world has warmed 1.55°C relative to pre-industrial temperature. Even if greenhouse gas emissions were cut to zero in 2025, the world would likely still reach 2°C warming relative to pre-industrial temperature (IPCC, 2021).

for flexible across-unit dependencies.

Analytically, we find in a general setting that the TWFE estimator recovers unbiased estimates of what we call the *average slope of expected treatment effects (ASETE)*—an estimand closely related to the average slope of switchers' potential outcomes function discussed in de Chaisemartin et al. (2022)—if and only if (i) the expected treatment effect of the entire vector of warming is proportional to own-region warming, and (ii) either higher-order effects are negligible and unobserved factor growth is mean zero, or the partial effect of outcomes to unobserved factors is constant across regions. Proportionality to own-region warming is unlikely in a trade equilibrium, so it is unlikely that the TWFE recovers the ASETE. But even if the first condition fails, the TWFE estimator may still recover an unbiased estimate of the best linear approximation to the conditional expectation function, as long as the second condition holds. In Monte Carlo simulations, we find that the TWFE estimator indeed generates an estimate that is sometimes close to the slope of this function. This means that in trade settings, the TWFE estimator can provide useful information about *relative* effects, but not the overall *level* of effects.[2]

To make progress on estimating the *level* of treatment effects, we adopt a standard multi-sector trade framework, drawing from Costinot & Rodríguez-Clare (2014). In the model, climate affects productivity, and productivity shocks propagate through the system via input-output linkages, output market competition, and income effects. Leveraging the structure of the model, we estimate sector-specific elasticities of productivity to climate from sector-level trade and production data, and compute counterfactuals using standard exact hat algebra techniques (Dekle et al., 2007).

In Monte Carlo experiments, we find that the TWFE tends to understate the damages from climate change at the country level when regions trade. This is because spillover effects are mostly negative, meaning that adverse climate shocks in any given region tend to reduce gross output in other regions. These spillover effects generate a negative intercept in the best linear approximation to the conditional expectation function. As the TWFE tends to recover the *slope* of this function, but not the intercept, estimated effects are shifted up relative to the true effects. Dissaggregating to the sector level, we find that

[2]The failure of reduced-form estimators to recover level effects is sometimes referred to as the "missing intercept" problem (Adao et al., 2020).

the TWFE can overstate or understate the damages for a given sector, depending on the relative strength of the elasticity of sector-specific productivity to climate. By contrast, our structural approach recovers the full vector of treatment effects under a range of scenarios.

Extensions to the TWFE estimator perform only marginally better. The heterogeneous robust estimator from de Chaisemartin et al. (2024b) performs qualitatively similar to the TWFE estimator. This is because neither the TWFE nor the heterogeneous robust estimator explicitly account for spillover effects. The upstream/downstream estimator from Das et al. (2022); Feng et al. (2023); Zappalà (2024) *does* explicitly account spillover effects, but we find in the simulations that this estimator tends to understate country-level damages and the damages in manufacturing, just like the TWFE and the heterogeneous robust estimator. This result indicates that the *ad hoc* controls for upstream/downstream shocks are not flexible enough to capture heterogeneous spillover effects which arise naturally in structural trade models. Lastly, we find that the global time-series estimator from Bilal & Känzig (2024) tends to overstate the global damages from climate change.

We use our structural approach to provide new estimates of the effect of country-level temperature on country-level gross output and welfare. Using annual bilateral trade and production data in agriculture and manufacturing (Mayer et al., 2023; Fontagné et al., 2023), country-level labor and material shares in production (Aguiar et al., 2016), country-sector labor allocations from International Labor Organization (ILO), and country-level environmental variables from the ERA-5 dataset (Hersbach et al., 2023), we quantify the effect of the observed change in temperature between 1991 and 2019 on country-sector gross output and country-level real wage.

We find that an additional day in the year with maximum temperature above 30°C lowers annual productivity in agriculture by about 0.35%, with slightly stronger effects for days with maximum temperature above 35°C, and no corresponding effect in manufacturing. These productivity effects translate into modest annual gross output and real wage losses computed for the last year in the sample, 2019, on the order of roughly 1% aggregated across the whole world. These aggregate effects mask substantial heterogeneity though, with some countries losing as much as 8% of aggregate gross output, and as much as 25% of gross agricultural output. We also find evidence of trade-induced spillovers, as countries that hardly warmed at all also suffered as a result of climate change. We also find evidence of substantial reallocation of labor across sectors. We find that climate change *increased* gross output in manufacturing for countries that warmed substantially between 1991 and 2019. This is because agricultural productivity suffered so much in these countries that the relative labor demand in manufacturing increased, drawing labor into the sector, and thus increasing gross output. The TWFE overstates the effect on agricultural output for most countries, but understates the effect on total gross output, as does the heterogeneous robust estimator, and the upstream/downstream estimator.

Our estimation procedure builds on a large body of work in quantitative trade (Eaton & Kortum, 2002; Hsieh & Ossa, 2016; Caliendo & Parro, 2015; Redding & Venables, 2004; Donaldson & Hornbeck, 2016; Shapiro & Walker, 2018; Bartelme, 2018; Anderson et al., 2020; Adao et al., 2020; Bartelme et al., 2020). Most of this work aims to quantify the effects of changes to trade costs or changes to overall productivity levels. We show how the same tools can be used to quantify the effects of observable determinants of productivity, a pursuit of interest even outside the field of International Trade.

One notable comparison from the quantitative trade literature that *does* study the effects of an observed determinant of productivity is Cruz & Rossi-Hansberg (2024). Similar to our approach, Cruz & Rossi-Hansberg (2024) infers productivity from a general equilibrium model and then projects onto observable factors. Our model is more general in that it incorporates multiple sectors, input-output linkages, different market structures, and asymmetric trade costs. However, it is less general in other respects, as it remains static and abstracts from migration, among other dynamics. The two approaches are complementary, as they rely on different assumptions and data requirements.

Our paper also builds on previous work that analyzes the output of TWFE models under general conditions. De Chaisemartin & d'Haultfoeuille (2020) and related work demonstrate that the TWFE may yield biased estimates of average treatment effects when the assumption of treatment effect homogeneity fails. There has been much less attention paid to SUTVA violations. Two notable exceptions are Borusyak et al. (2022) and Alves et al. (2024), who analyze the performance of the TWFE estimator in the context of migration models. As in our multi-sector trade model, local shocks affect equilibrium outcomes in these models, undermining SUTVA. Borusyak et al. (2022) and Alves et al. (2024) both find that the TWFE estimator recovers *relative* treatment effects, assuming a first-order approximation to the equilibrium conditions. In our setting, we find that there are conditions under which even this limited interpretation of the TWFE estimator is invalid.[3]

Finally, our paper contributes to the environmental literature devoted to quantifying the impact of weather shocks on aggregate economic outcomes. While most existing work relies on TWFE estimators, a few previous papers account for across-unit dependencies with trade models. Dingel et al. (2019) and Costinot et al. (2016) study climate effects on agricultural output, for which price and productivity data are available. When productivity measures are observed, the elasticity of productivity to environmental factors can be directly assessed with econometric techniques, under some assumptions.[4] We show how to estimate these key structural parameters even for sectors for which productivity measures are not observed. Nath (2020) estimates productivity effects of climate variables in multiple industries and locations from microdata, and then aggregates using a trade model similar to the one we use. This "bottom up" approach is complementary to the method we propose, though we note that we require only aggregate data, whereas Nath (2020) requires detailed input and output data in all countries and sectors studied. Rudik et al. (2022) and Osberghaus & Schenker (2022) also use structural gravity frameworks to estimate the effects of weather shocks, but both papers exploit restrictive assumptions with respect to the correlation between multilateral resistance terms and unobserved productivity shocks. Our estimation strategy relaxes these assumptions.

2 General Model, Estimands, and Standard Estimators

In this section, we formalize a general model of how observable and unobservable factors influence endogenous outcomes, allowing for heterogeneous effects and spillovers across

[3]There is also a broader literature in Urban Economics and Network Econometrics that studies the identification of treatment effects in settings in which SUTVA fails. Canonical works include Manski (1993), Sobel (2006), Hudgens & Halloran (2008). More recent efforts include Butts (2021), Leung (2020), and Vazquez-Bare (2023). In most settings, researchers exploit sparsity of networks to identify spillover effects. In quantitative trade models, all units can potentially affect all other units, so this strategy is not available to us.

[4]Even yield measures might be contaminated with equilibrium responses, since the choice of crops to grow on a parcel as well as the rotations of crops within a year would respond to equilibrium conditions, in particular market prices.

units. We then study conditions under which standard estimation procedures recover unbiased estimates of target values. Finally, we characterize the expected values of the TWFE estimator when these conditions fail.

2.1 General Model

We consider an economy in which a finite number of units $i = 1, \ldots, N$ are endowed each period t with a realization of an observable factor z and an unobservable (to the researcher) factor ε. Units could represent geographical regions or countries, as well as firms or workers, depending on the application. Observable factors could be policies or operating conditions such as local taxes, infrastructure, climate conditions, or natural disasters, the effects of which we would like to quantify. Unobservable factors could represent local productivity levels, amenities, or worker skills. For simplicity, we consider a single observable factor, a single unobservable factor, and two periods, though allow for multiple observables, unobservables, and periods in the empirical application.

Each period, units interact to determine the vector of endogenous outcomes $y_t \equiv \{y_{1t}, y_{2t}, \ldots y_{Nt}\}$. The endogenous outcome could be gross output, welfare, revenues, or wages depending on the application. To fix ideas, in the empirical application below, units are countries, z is a climate variable, ε is unobserved productivity, and y is gross output.

We assume a static equilibrium each period and write the mapping between inputs and outputs as

$$\ln y_t = f(z_t, \varepsilon_t), \tag{1}$$

where $\ln y_t \equiv \{\ln y_{1t}, \ln y_{2t}, \ldots \ln y_{Nt}\}$ denotes the vector of endogenous outcomes, $z_t \equiv \{z_{1t}, z_{2t}, \ldots z_{Nt}\}$ denotes the vector of observed factors, and $\varepsilon_t \equiv \{\varepsilon_{1t}, \varepsilon_{2t}, \ldots \varepsilon_{Nt}\}$ denotes the vector of unobserved factors.[5] The supports of z and ε are bounded subsets of $\Re$. We assume $f(\cdot)$ is continuously differentiable for all orders. The mapping $f(\cdot)$ nests many economic environments and data generating processes, all of which admit heterogeneous effects of z_{it} and ε_{it} on $\ln y_{it}$, as well as spillover effects of z_{jt} and ε_{jt} on $\ln y_{it}$, for all $i \neq j$.[6]

[5] We take the natural log of the outcome variable because in our application we study growth rates, but the model could just as easily be expressed in levels.

[6] Adão et al. (2024) adopt the same general structure to study equilibrium effects of observable factors, as

We put no restrictions on the joint distribution of z_t and ε_t in levels, but impose that the change in the unobservable factor is mean independent from the change in the observable factor:

Assumption 1. *(Conditional Mean Independence)* $E[\Delta\varepsilon|\Delta z] = \kappa$,

where $\Delta v_i \equiv v_{i1} - v_{i0}$ for an arbitrary variable v and $\kappa \in \Re$. Allowing for non-zero κ rationalizes general technological progress or trends in preferences and skills. This assumption is necessary for identification in virtually any estimation strategy.

2.2 Estimands

Our primary object of interest is the causal effect of the change in observables from z_0 to z_1. In terms of the model, this effect can be expressed as

$$\Delta \ln y^{\dagger}(z_1, z_0; \varepsilon_1) \equiv f(z_1, \varepsilon_1) - f(z_0, \varepsilon_1). \tag{2}$$

We refer to (2) as the *realized treatment effect* (RTE), and its expectation as the *expected treatment effect* (ETE), where expectations are taken over $\Delta\varepsilon$, given $\varepsilon_1 = \varepsilon_0 + \Delta\varepsilon$. The RTE, as its name implies, depends on the *realized* vector $\Delta\varepsilon$. By contrast, the ETE is independent of the influence of any given realization of $\Delta\varepsilon$, and thus represents a suitable target value for estimators that impose separability between $\Delta\varepsilon$ and Δz. Both the RTE and the ETE may vary by unit, in general, and allow for spillover effects from any unit j to any other unit i.

While researchers would ideally like to recover the entire vector of treatment effects, statistical power often precludes the estimation of such rich heterogeneity. To benchmark estimators that instead target average treatment effects, we define at the economy level the

do many papers in the network econometric literature (Hudgens & Halloran, 2008; Leung, 2020). In settings with spillovers, researchers often assume a network structure that generates restrictions on the matrix of partial effects from any given unit onto any other unit. In our structural estimator below, we allow that all units can affect the outcomes in all other units, or what is sometimes referred to as a "complete network."

average slope of expected treatment effects as

$$ASETE \equiv \frac{1}{N}\sum_i \frac{E\left[\Delta \ln y_i^{\dagger}\left(z_1, z_0; \varepsilon_0 + \Delta\varepsilon\right)\right]}{\Delta z_i}. \tag{3}$$

This estimand is very similar to the average slope of switchers' potential outcomes function presented in de Chaisemartin et al. (2022). The difference is that the expected treatment effect for a given unit in our framework depends not only on its own treatment status z_{it}, as in de Chaisemartin et al. (2022), but also on the full vector of z in both periods, as well as ε_0.

Finally, we define the *marginal effect* to unit i as the effect of a one-unit increase in z_{it} to $\ln y_{it}$

$$ME_i(z_t, \varepsilon_t) \equiv f_i(z_{t(i)}, \varepsilon_t) - f_i(z_t, \varepsilon_t) \tag{4}$$

where $z_{t(i)}$ corresponds to the vector z_t but with the ith element replaced with $z_{it}+1$. The $ME_i(z_t, \varepsilon_t)$ is in fact a general equilibrium effect, because movements from z_t to $z_{t(i)}$ could influence outcomes throughout the system, but $ME_i(z_t, \varepsilon_t)$ gives the effect to just unit i's outcome. Evaluating at $t = 0$ and taking averages over all units, we obtain the *average marginal effect*:

$$AME(z_0, \varepsilon_0) \equiv \frac{1}{N}\sum_i ME_i(z_0, \varepsilon_0). \tag{5}$$

The AME will also be useful for benchmarking the performance of existing estimators.

2.3 Some Existing Estimators

Empirical research often aims to recover treatment effects like the RTE or the ETE, or at least the $ASETE$, from data on $\Delta \ln y$ and Δz, without imposing any structural assumptions on demand, supply, or industrial organization. But even without structural assumptions on economic fundamentals, restrictions on the statistical relationships between inputs and outputs are needed for identification.

To illustrate the identification problem, we take a Taylor series expansion of (1) around

(z_0, ε_0), assuming $f(\cdot)$ is smooth and differentiable to the required order, to find:

$$\Delta \ln y = \frac{\partial f(z_0, \varepsilon_0)}{\partial z_0}\Delta z + \frac{\partial f(z_0, \varepsilon_0)}{\partial \varepsilon_0}\Delta \varepsilon + \sum_{k=2}^{\infty} \frac{1}{k!} D^k f(z_0, \varepsilon_0) \left[\Delta z, \Delta \varepsilon\right]^k \tag{6}$$

where $D^k f(z_0, \varepsilon_0)$ is the k-th derivative tensor and $[\Delta z, \Delta \varepsilon]^k$ is the tensor product of Δz and $\Delta \varepsilon$, taken k times. It will be convenient to write this expansion in matrix form, denoting the higher-order terms with the vector c:

$$\begin{pmatrix} \Delta \ln y_1 \\ \Delta \ln y_2 \\ \vdots \\ \Delta \ln y_N \end{pmatrix} = \underbrace{\begin{pmatrix} b^z_{11} & b^z_{12} & \cdots & b^z_{1n} \\ b^z_{21} & b^z_{22} & \cdots & b^z_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b^z_{N1} & b^z_{N2} & \cdots & b^z_{NN} \end{pmatrix}}_{\equiv B^z} \begin{pmatrix} \Delta z_1 \\ \Delta z_2 \\ \vdots \\ \Delta z_N \end{pmatrix} + \underbrace{\begin{pmatrix} b^\varepsilon_{11} & b^\varepsilon_{12} & \cdots & b^\varepsilon_{1n} \\ b^\varepsilon_{21} & b^\varepsilon_{22} & \cdots & b^\varepsilon_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b^\varepsilon_{N1} & b^\varepsilon_{N2} & \cdots & b^\varepsilon_{NN} \end{pmatrix}}_{\equiv B^\varepsilon} \begin{pmatrix} \Delta \varepsilon_1 \\ \Delta \varepsilon_2 \\ \vdots \\ \Delta \varepsilon_N \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \tag{7}$$

with the elements of the B^z matrix defined as $b^z_{ij} \equiv \frac{\partial f_i(z_0, \varepsilon_0)}{\partial z_{0j}}$ and the elements of the B^ε matrix defined as $b^\varepsilon_{ij} \equiv \frac{\partial f_i(z_0, \varepsilon_0)}{\partial \varepsilon_{0j}}$ for all i, j.

As it appears clearly from this matrix expression, the dimensionality of the problem makes purely reduced-form estimation of the causal effect of the observable factors on the endogenous outcomes very difficult. Indeed, even ignoring higher order terms, there are N^2 elements in the B^z matrix, whereas researchers only observe data for N units. As a result, estimating each b^z_{ij} elements is not feasible without imposing *some* restrictions.[7]

Two-way Fixed Effect Estimator One of the most commonly-used estimators in empirical economic research is the TWFE estimator (De Chaisemartin & d'Haultfoeuille, 2023). Expressed in first differences, this econometric model can be written as

$$\Delta \ln y_i = \beta^{FE} \Delta z_i + \alpha^{FE} + \Delta \xi_i, \tag{8}$$

[7] Extending the sample to multiple periods does not necessarily help unless one assumes the b^z_{ij} coefficients are constant over time, which they need not be. Even in this case, higher order terms in the Taylor series expansion make purely reduced-form estimation very challenging. Additionally, one would need a very long time series to be able to estimate the b^z_{ij} coefficients nonparametrically, a point raised by Adão et al. (2024).

where α^{FE} denotes a constant that absorbs common trends and ξ_i denotes the error term.[8] The estimator is computed as

$$\check{\beta}^{FE} = \sum_i \frac{\left(\Delta \ln y_i - \overline{\Delta \ln y}\right)\left(\Delta z_i - \overline{\Delta z}\right)}{\left(\Delta z_i - \overline{\Delta z}\right)\left(\Delta z_i - \overline{\Delta z}\right)} \tag{9}$$

where $\overline{\Delta \ln y}$ and $\overline{\Delta z}$ indicate the average value of $\Delta \ln y_k$ and Δz_k in the economy.

This estimation strategy is often motivated by assumption (1), i.e., conditional mean independence. But given (7), we see that restrictions beyond (1) are required to derive (8). First, one must impose treatment effect homogeneity, i.e., $b^z_{ii} = \beta^{FE}$ for all i. Second, one must impose no spillover effects (hence, SUTVA is assumed to hold) so that $b^z_{ij} = 0$ for all $j \neq i$. Third, one must impose no higher-order terms, hence $c_i = 0$ for all i. If these restrictions hold, we have $\Delta \ln y^{\dagger}_i (z_1, z_0; \varepsilon_0 + \Delta\varepsilon) = \beta^{FE} \Delta z_i = E\left[\check{\beta}^{FE}\right] \Delta z_i$, and $ASETE = \beta^{FE} = E\left[\check{\beta}^{FE}\right]$; but clearly, these three restrictions are unlikely to hold in any case in which units can influence each others' outcomes.

Heterogeneous-Robust Estimators A recent literature surveyed by De Chaisemartin & d'Haultfoeuille (2023) develops estimators that recover unbiased estimates of target values when treatment effects are heterogeneous. The majority of these estimators are only unbiased when treatment is discrete or continuously distributed in the post-shock periods. The only estimator that is designed for a shock that is continuously distributed in every period with no "stayer" observations—which is the setting of our empirical application—is de Chaisemartin et al. (2024b).

de Chaisemartin et al. (2024b) identify a weighted average of the slopes of units' potential-outcome function by comparing the observed changes in outcomes to estimated changes in outcomes under the counterfactual of no treatment. In our notation, de Chaise-

[8] We adopt a first-differenced econometric model following (8), commonly described as a "level-effects" model, where the observable factor influences only contemporaneous outcomes. This contrasts with "growth-effects" models, where the observed factor also affects future outcomes. Newell et al. (2021) show that this distinction is inconsequential in our context, as both models yield nearly identical prediction errors (as long as we do not include country-specific parametric trends). Additionally, Nath et al. (2024) find that while climate variables may have persistent effects on growth, they are not permanent.

martin et al. (2024b) first estimate $E[f(z_0,\varepsilon_1)-f(z_0,\varepsilon_0)|z_0]$, and then compute

$$\check{\beta}^{HR}=\frac{1}{N}\sum_i \frac{sign(\Delta z_i)\times\left(\Delta\ln y_i-\check{E}[f(z_0,\varepsilon_1)-f(z_0,\varepsilon_0)|z_0]\right)}{|\Delta z_i|}. \tag{10}$$

When $b^z_{ij}=0$ for all $j\neq i$ and $c_i=0$ for all i, this estimator is an unbiased estimator of the $ASETE$, even if b^z_{ii} varies across i. Hence, de Chaisemartin et al. (2024b) allow for heterogeneous effects of Δz_i, but still impose no spillover effects.

Estimator with Upstream and Downstream Spillovers One approach to incorporating spillover effects in a parsimonious way is to include "upstream" and "downstream" exposure measures to shocks in foreign units (Das et al., 2022; Feng et al., 2023; Zappalà, 2024). A typical estimation equation takes the form

$$\Delta\ln y_i=\beta^{Own}\Delta z_i+\beta^{Upstream}\sum_{j\neq i}\pi_{ij,0}\Delta z_j+\beta^{Downstream}\sum_{k\neq i}\gamma_{ki,0}\Delta z_k+\alpha^{UD}+\xi_i, \tag{11}$$

where $\pi_{ij,0}$ denotes the share of expenditures that i imports from j in the pre-period, and $\gamma_{ki,0}$ denotes the share of i-production that i sends to k in the pre-period. This econometric model decomposes the B_z matrix into three components: first, a diagonal matrix capturing the homogeneous effect of own-unit treatment; second, an upstream shock term, using shocks in upstream countries weighted by import trade shares, and third, a downstream shock term, using shocks in downstream countries weighted by export trade shares. The restrictions are: i/ $b^z_{ii}=\beta^{Own}$ for all i, ii/ $b^z_{ij}=\beta^{Upstream}\pi_{ij,0}$ for all $j\neq i$, iii/ $b^z_{ji}=\beta^{Downstream}\gamma_{ji,0}$ for all $j\neq i$, and iv/ $c_i=0$ for all i.

This strategy is less restrictive than the TWFE model, as it allows for non-zero off-diagonal elements b^z_{ij}, but it assumes these elements are functions of first-degree trade shares links. While first-degree trade shares likely matter, factors such as preferences, competition, and sectoral allocation also shape these spillovers, suggesting that modeling spillovers as linear in weighted average foreign shocks may be overly restrictive.

Global Spillovers Another approach to incorporating spillover effects is to aggregate to a broader geographic level, as in Bilal & Känzig (2024), who analyze data at the world level. In our notation, their econometric model is:

$$y_{t+h} - y_t = \alpha_h + \beta_h z_t^{shock} + \sum_{l=1}^{L} \gamma_{h,l} X_{t-l} + \xi_{t+h} \tag{12}$$

where $y_{t+h} - y_t$ is world-level outcome growth h periods ahead of t, z_t^{shock} is a one-year-ahead forecast error of world-average z, and X_{t-l} is a vector of controls up to l lags. Model (12) is estimated by OLS using purely time series data.

While world-level aggregation makes the estimate β_h to some extent inclusive of spillover effects, it does not fully resolve the identification problem. Deriving (12) from (1) requires assuming $c_i = 0$ for all i and that all columns of B^z sum to β_h/N. This implies that a shock to any unit affects the world-level outcome equally, regardless of location—an assumption that is unlikely to hold in the context of international trade.

Although one could specify alternative models that populate the B^z matrix with off-diagonal elements under other restrictions than the ones presented above, researchers face a challenge with no ideal solution. Any *ad hoc* restriction on the off-diagonal elements may lead to misspecification in linkages across units so that we do not know whether these alternative models represent an improvement over the standard TWFE model that ignores spillovers or not. The comparison between these models should therefore be in terms of their ability to recover the true treatment effects. This is what we test below using Monte Carlo simulations.

2.4 Two-way Fixed Effect Estimator in the full model

Though the conditions required to derive (8) from (1) are unlikely to hold, the TWFE may still be informative. In this subsection, we discuss what can be learned from the TWFE model when the restrictions presented above fail.[9]

In full generality, we can decompose the observed change in outcomes into the ETE

[9] We test the performance of other existing estimators in Monte Carlo experiments.

and components that depend on the unobserved factors:

$$\begin{aligned}\Delta \ln y &= f(z_1,\varepsilon_1) - f(z_0,\varepsilon_1) + f(z_0,\varepsilon_1) - f(z_0,\varepsilon_0) \\ &= E\left[\Delta \ln y^{\dagger}\left(z_1,z_0;\varepsilon_0+\Delta\varepsilon\right)\right] + a + f(z_0,\varepsilon_1) - f(z_0,\varepsilon_0) \end{aligned} \tag{13}$$

where $a \equiv \Delta \ln y^{\dagger}\left(z_1,z_0;\varepsilon_0+\Delta\varepsilon\right) - E\left[\Delta \ln y^{\dagger}\left(z_1,z_0;\varepsilon_0+\Delta\varepsilon\right)\right]$, i.e., the deviation of the realized general equilibrium effect from the expected general equilibrium effect.

Substituting this expression into (9) and taking expectations over $\Delta\varepsilon$ yields

$$E\left[\check{\beta}^{FE}\right] = \sum_i \zeta_i \left[\frac{E\left[\Delta \ln y_i^{\dagger}\left(z_1,z_0;\varepsilon_0+\Delta\varepsilon\right)\right]}{\Delta z_i} + \frac{\sum_{k=1}^{\infty}\frac{1}{k!}\frac{\partial^k f_i(z_0,\varepsilon_0)}{\partial \varepsilon_0^k} m(k)}{\Delta z_i}\right] \tag{14}$$

$$E\left[\check{\alpha}^{FE}\right] = \frac{1}{N}\sum_i \left[E\left[\Delta \ln y_i^{\dagger}\left(z_1,z_0;\varepsilon_0+\Delta\varepsilon\right)\right] + \sum_{k=1}^{\infty}\frac{1}{k!}\frac{\partial^k f_i\left(z_0,\varepsilon_0\right)}{\partial \varepsilon_0^k} m(k)\right] - E\left[\check{\beta}^{FE}\right]\overline{\Delta z}, \tag{15}$$

with $\zeta_i \equiv \frac{\Delta z_i\left(\Delta z_i - \overline{\Delta z}\right)}{\sum_k \Delta z_k\left(\Delta z_k - \overline{\Delta z}\right)}$, $\sum_i \zeta_i = 1$, and $m(k) \equiv E\left[(\Delta\varepsilon)^k\right]$, the k-th moment of the distribution of $\Delta\varepsilon$. The second term inside the brackets results from taking a Taylor series approximation to $f(z_0,\varepsilon_1) - f(z_0,\varepsilon_0)$ around the point (z_0,ε_0).

We obtain the following proposition:

Proposition 1. *For a DGP described in* (1) *and under Assumption 1, we obtain* $E\left[\check{\beta}^{FE}\right] = ASETE$ *iff the following two conditions hold:*
1) $E\left[\Delta \ln y_i^{\dagger}\left(z_1,z_0;\varepsilon_0+\Delta\varepsilon\right)\right] = \nu \Delta z_i$ *for all i, and*
2) either all second-order and above terms in the Taylor series approximation to $f(z_0,\varepsilon_1)$ *around* (z_0,ε_0) *are zero and* $E[\Delta\varepsilon|\Delta Z] = 0$, *or* $\frac{\partial^k f_i(z_0,\varepsilon_0)}{\partial \varepsilon_0^k}$ *is constant across units for all k.*

Hence, the TWFE can recover the $ASETE$, even if there are spillover effects, under conditions (1) and (2).[10] For example, we obtain condition (1) if $f(z_t,\varepsilon_t) = \alpha + B_t^z z_t + B_t^{\varepsilon}\varepsilon_t$, with $B_t^z = \nu\mathbb{I}$ for all t, where $\mathbb{I}$ represents the identity matrix. This imposes a constant intercept, separability between the effects of z and ε on the endogenous outcome, and

[10] This positive result, although restrictive, is reminiscent of the result obtained by de Chaisemartin et al. (2024a) in the context of heterogeneous adoption design without stayers. They show that under parallel trends and if treatment effects are mean-independent of the treatment variable, the TWFE estimates a well-defined effect and inference based on $\check{\beta}^{FE}$ can at worst be conservative.

linearity of treatment effects with a constant slope across units and over time. Spillovers are not necessarily incompatible with condition 1), but the combination of direct and indirect effects from the observed factor needs to be the same across units.

Condition 2) is also restrictive, as general productivity growth would rule out $E[\Delta\varepsilon|\Delta Z] = 0$, even though second-order and above terms would cancel out by linear approximation. Alternatively, the condition requires a constant responsiveness of gross output to unobserved factors across units. As a result, an unobserved productivity shock in the US could not have a different effect on the gross output of Canada as compared to India, for instance. Hence, in general, $E\left[\check{\beta}^{FE}\right]$ could be larger or smaller in magnitude than the $ASETE$, and even take the opposite sign.

An alternative interpretation of $E[\check{\beta}^{FE}]$ is that it represents the average effect on $\Delta \ln y_i$ of an extra unit of Δz_i, holding all $\Delta z_j = 0$, $j \neq i$ (Borusyak et al., 2022). This is what we define as $AME(z_0, \varepsilon_0)$. Does $E[\check{\beta}^{FE}] = AME(z_0, \varepsilon_0)$? To assess this possibility, we take a Taylor series expansion of $AME(z_0, \varepsilon_0)$ around (z_0, ε_0) and average over units to find

$$AME(z_0, \varepsilon_0) \quad = \quad \frac{1}{N}\sum_i \sum_{k=1}^{\infty} \frac{1}{k!} \frac{\partial^k f_i(z_0, \varepsilon_0)}{\partial z_{0i}^k}$$

where z_{0i} corresponds to the i-th component of the vector z. We compare it with:

$$E\left[\check{\beta}^{FE}\right] \quad = \quad \sum_i \frac{\zeta_i}{\Delta z_i}\left[\sum_{k=1}^{\infty} \frac{1}{k!} D^k f(z_0, \varepsilon_0) \sum_{r=0}^{\infty} (\Delta z)^r m(k-r) + \sum_{k=1}^{\infty} \frac{1}{k!} \frac{\partial^k f_i(z_0, \varepsilon_0)}{\partial \varepsilon_0^k} m(k)\right]$$

Two main differences emerge. First, spillover effects intervene in $E[\check{\beta}^{FE}]$, while they do not in $AME(z_0, \varepsilon_0)$. Second, $E[\check{\beta}^{FE}]$ represents a weighted average of individual effects, while AME is a simple average.

Thus, the TWFE estimator recovers neither the $ASETE$ nor the AME, in general. However, linear regression models are often justified not on the basis that they recover unbiased estimates of the average affects, but rather because they represent the best linear approximation to unknown and potentially non-linear conditional expectation function (Angrist & Pischke, 2009). Does the linear approximation theorem apply in this case?

To check this possibility, we compute the slope and intercept of the best linear approximation to the conditional expectation function relating ETE to own-unit treatment:

$$\beta^{LA} = \sum_i \zeta_i \left[\frac{E\left[\Delta \ln y_i^{\dagger}(z_1, z_0; \varepsilon_0 + \Delta\varepsilon)\right]}{\Delta z_i} \right] \tag{16}$$

$$\alpha^{LA} = \frac{1}{N}\sum_i E\left[\Delta \ln y_i^{\dagger}(z_1, z_0; \varepsilon_0 + \Delta\varepsilon)\right] - \sum_i \zeta_i \left[\frac{E\left[\Delta \ln y_i^{\dagger}(z_1, z_0; \varepsilon_0 + \Delta\varepsilon)\right]}{\Delta z_i} \right] \overline{\Delta z} \tag{17}$$

and compare to (14) and (15). In general, we find

Proposition 2. *For a DGP described in* (1) *and under Assumption 1, we obtain*
1) $E\left[\check{\beta}^{FE}\right] = \beta^{LA}$ *iff either all second-order and above terms in the Taylor series approximation to* $f(z_0, \varepsilon_1)$ *around* (z_0, ε_0) *are zero and* $E[\Delta\varepsilon|\Delta Z] = 0$*, or* $\frac{\partial^k f_i(z_0,\varepsilon_0)}{\partial \varepsilon_0^k}$ *is constant across units for all* k*;*
2) $E\left[\check{\alpha}^{FE}\right] = \alpha^{LA}$ *iff* $E\left[\check{\beta}^{FE}\right] = \beta^{LA}$ *and* $m(k) = 0$ *for all* k*.*

The first condition in Proposition 2—which is the same as the second condition of Proposition 1—is sufficient and necessary for the slope of the TWFE regression to recover the slope of the best linear approximation to the expected treatment effects, in expectation. Proposition 2 yields a striking result: even without explicitly modeling spillovers, the TWFE can still capture meaningful comparisons of treatment effects across units, *inclusive of spillovers*. To illustrate, consider two units ℓ and m for which $\Delta z_\ell = \Delta z_m + 1$. Assuming the first condition of Proposition 2 holds, the approximated difference in treatment effects between ℓ and m from the entire vector of exogenous changes, Δz, is $\beta^{LA} = E\left[\check{\beta}^{FE}\right]$. This result generalizes the findings of Borusyak et al. (2022), who finds that the linear approximation theorem holds given a specific model of migration.[11]

If the second condition of Proposition 2 holds, the TWFE estimator recovers both the slope and intercept of the best linear approximation to expected treatment effects. Researchers typically include a constant in the TWFE model to account for nonzero average growth in the unobserved factor. However, when spillover effects are present, the intercept

[11] Borusyak et al. (2022) shows that the linear approximation theorem always holds in their migration model where they consider a linear approximation to the equilibrium conditions, and hence assume away higher order terms.

also absorbs average spillovers across units. Since average spillovers and unobserved factor growth are not separately identified, $\breve{\alpha}^{FE}$ is usually ignored when computing treatment effects. But if unobserved factor growth is zero, then $E\left[\breve{\alpha}^{FE}\right]$ solely captures average spillovers, meaning $E\left[\breve{\alpha}^{FE}\right] = \alpha^{LA}$.

Proposition 2 establishes conditions under which the linear approximation theorem applies. As always, its usefulness depends on how well the linear approximation captures the true ETE. If the relationship between ETE and own-unit treatment Δz_i is highly nonlinear, the TWFE estimates may deviate significantly from the truth. Nevertheless, to the extent the conditions hold, Proposition 2 indicates that the TWFE estimator can be justified on the same grounds as in standard i.i.d. settings, even with heterogeneous spillover effects.

In summary, reduced-from estimation strategies impose restrictions on the statistical relationships between inputs and outputs to circumvent an identification problem. Only under highly restrictive conditions do reduced-form estimation strategies recover unbiased estimates of the $ASETE$. Under slightly less restrictive conditions, the TWFE estimator may approximate *relative* effects, though these conditions are unlikely to hold in a trade context. To make progress on estimating treatment effects, we propose a theory-based solution to the identification problem.

3 A Structural Approach Based on Quantitative Trade

In this section, we present a quantitative trade model in which productivity, and hence output, depends on observable factors. Our main objective is to generate constraints on the relationship between inputs and outputs that are stringent enough to ensure identification while remaining sufficiently flexible to accommodate heterogeneous cross-unit spillover effects. Additionally, we use our standard trade model to assess whether the reduced-from restrictions discussed in the previous section hold. The model nests microfoundations for both perfect competition and imperfect competition within a general framework, as in Costinot & Rodríguez-Clare (2014). Most of the elements of the model are quite standard, so we leave the majority of the derivations for the appendix. At the end of the section, we present a multi-step procedure to estimate structural parameters of the model and compute counterfactuals.

3.1 Model

The world is divided into N distinct regions, which we refer to as "countries", and S different sectors of the economy. In each period t, each country n is endowed with a fixed number of worker-consumers L_{nt} that endogenously sort into one of the S sectors. All worker-consumers receive the country-time-specific marginal product of their labor, w_{nt}, for supplying their one unit of labor inelastically, as well as average investment income $\bar{r}_{nt}$. The worker-consumers are assumed to be immobile between countries but perfectly mobile between sectors, such that the wage is different across countries but the same for all sectors in each country.

Worker-consumers from each country n have Cobb-Douglas preferences over the different sectors $s \in \{1,...,S\}$, spending a share α^C_{ns} of their income on goods from sector s, with $\sum_s \alpha^C_{ns} = 1$. Aggregate expenditures on final good consumption on goods from sector s is thus $X^F_{nst} \equiv \alpha^C_{ns} L_{nt} (w_{nt} + \bar{r}_{nt})$. Within each sector s, worker-consumers in all countries have identical preferences over a continuum of varieties $j \in \Lambda_{nst}$ with a constant elasticity of substitution $\sigma_s > 0$, where Λ_{nst} denotes the set of varieties from sector s consumed in country n at time t. The resulting Cobb-Douglas price index of country n at time t is: $p_{nt} = \prod_{s=1}^{S} \left(\frac{p_{nst}}{\alpha^C_{ns}} \right)^{\alpha^C_{ns}}$, where p_{nst} is the CES price index of sector s in country n at date t. All worker-consumers in country n at time t receive the same real wage $RW_{nt} \equiv (w_{nt} + \bar{r}_{nt})/p_{nt}$, which we treat as the metric of welfare.

Producers in each sector s and each country i at date t produce output Q_{ist} using a Cobb-Douglas constant returns to scale technology, requiring labor in proportion $\eta_{is} \in [0,1]$ and intermediate inputs in proportion $1 - \eta_{is}$. These intermediate inputs combine output from each sector in a Cobb-Douglas fashion, whereby each sector h output is used in proportion α^M_{ish} for the intermediate input of sector s, with $\sum_h \alpha^M_{ish} = 1$. Total expenditure on industry s varieties in country i, X_{ist}, is the sum of final good expenditures and purchases from other firms. Productivity in country i sector s and time t depends on a structural parameter A_{ist}, which depends on both observable factors and unobservable factors.

As in Costinot & Rodríguez-Clare (2014), we allow for different market structures that all generate structural gravity (Head & Mayer, 2014). In particular, we consider two standard microfoundations: (1) if the market structure is perfect competition as in Eaton & Kortum (2002), then our model collapses to Caliendo & Parro (2015); (2) if the market structure is monopolistic competition as in Melitz (2003), then our model collapses to Hsieh & Ossa (2016) or Gouel & Jean (2023).[12] We use the parameter $\chi \in \{0,1\}$ to indicate microfoundations, where $\chi = 0$ indicates perfect competition and $\chi = 1$ monopolistic competition.

Under either assumption on market structure, bilateral export flows X_{nist} from country i to country n in sector s at date t take the standard structural gravity form

$$X_{nist} = \frac{Y_{ist}}{\Omega_{ist}} \frac{X_{nst}}{\Phi_{nst}} \phi_{nist} \tag{18}$$

with

$$\Phi_{nst} = \sum_k \frac{Y_{kst}}{\Omega_{kst}} \phi_{nkst} \;\;,\;\; \Omega_{nst} = \sum_k \frac{X_{kst}}{\Phi_{kst}} \phi_{knst} \tag{19}$$

where Y_{ist} denotes country i's income from selling goods of sector s at date t and Ω_{ist} denotes outward multilateral resistance for exporting country i in sector s at time t—a measure of market access for country i—, Φ_{nst} denotes inward multilateral resistance term for country n in sector s at time t—a measure of accessibility of n to producers—, and ϕ_{nist} represents bilateral accessibility (the inverse of trade frictions). As seen in (19), the multilateral resistance terms depend on endogenous outcomes in all markets, as well as trade costs. These terms will channel spillover effects. Under both assumptions on market structure, the parameter θ_s represents the elasticity of trade flows to trade costs, which are embedded in ϕ_{nist}.

To connect observable factors to equilibrium outcomes, we parameterize the country-sector-time-specific productivity parameter, which partly determines the exporter i's capabilities as a supplier, as a function of observable factors z^v, $v \in \{1,...,\mathscr{V}\}$, as well as

[12] Hsieh & Ossa (2016)'s model shares most of the features of Gouel & Jean (2023). The key difference is that Hsieh & Ossa (2016) assumes firms pay the fixed costs of exporting in the destination, while Gouel & Jean (2023) assumes that firms pay the fixed costs of exporting in the origin, as we do.

unobserved (by the researcher) factors of productivity ψ_{is}, ι_{st} and ω_{ist}:

$$A_{ist} \propto \exp\left(\sum_{v=1}^{\mathscr{V}} \mu_s^v z_{it}^v\right) \exp\left(\psi_{is} + \iota_{st} + \omega_{ist}\right) \tag{20}$$

where ψ_{is} reflects time-invariant sector-country-specific base productivity, which is influenced by factors such as longitude, latitude, infrastructure and institutions, ι_{st} reflects time-varying productivity shocks that are common across all countries for a particular sector, and ω_{ist} captures country-sector-time unobserved factors. Parameters μ_s^v govern how observed factors affect local productivity for a given sector.

Given the distribution of labor across countries $L_t \equiv \{L_{it}\}_{i=1,\ldots N}$, country-sector-specific productivities $A_t \equiv \{A_{ist}\}_{s=1,\ldots S, i=1,\ldots N}$, and bilateral sectoral trade costs, we obtain a static equilibrium for each period t that satisfies the following equations, in addition to (19):

$$X_{ist} = \sum_{h=1}^{S} \left(\alpha_{is}^C \eta_{ih} + (1-\eta_{ih})\, \alpha_{ihs}^M\right) Y_{iht} \tag{21}$$

$$L_{ist}/L_{it} = \eta_{is} Y_{ist} / \left(\sum_h \eta_{ih} Y_{iht}\right), \tag{22}$$

and

$$Y_{ist} = \exp\left(\sum_{v=1}^{\mathscr{V}} \beta_{z,v}^{i,s} z_{it}^v\right) L_{ist}^{\beta_L^{i,s}} \Omega_{ist}^{\beta_\Omega^{i,s}} \left(\prod_{h=1}^{S} \Phi_{iht}^{\beta_{\Phi_h}^{i,s}}\right) \left(\prod_{h=1}^{S} X_{iht}^{\beta_{X_h}^{i,s}}\right)^{\chi} \exp\left(\delta_{is} + \delta_{st} + \varepsilon_{ist}\right) \tag{23}$$

where δ_{is} and δ_{st} are country-sector and sector-time fixed effects, respectively, ε_{ist} is a country-time-specific unobserved factor that combines ω_{ist} and structural parameters. The reduced-form parameters $\{\beta_{z,v}^{i,s}\}_{v=1}^{\mathscr{V}}$, $\beta_L^{i,s}$, $\beta_\Omega^{i,s}$, $\{\beta_{\Phi_h}^{i,s}\}_{h=1}^{H}$, $\{\beta_{X_h}^{i,s}\}_{h=1}^{H}$ are functions of structural parameters (θ_s, σ_s, μ_s^v, χ, η_{is}, α_{ish}^M, a_{is}^C), their exact expressions depending on the microfoundations considered.

From equation (23), we see that gross output in a given country-sector depends directly on observed determinants of productivity in the country, z_{it}^v, and indirectly on both observable and unobservable determinants of productivity in all country-sectors through labor allocations and multilateral resistance terms (and potentially expenditures, in the case of monopolistic competition). These across-unit dependencies invalidate the SUTVA, as in the general framework from Section 2.

3.2 Evaluating the conditions for identification

With our structural model, we can derive analytical expressions for the elements in (7) and check if the conditions for identification of the reduced-form approaches identified in section 2 hold.

To compare with (7), we consider a version of our structural model with a single sector, a single observable factor, two periods, and perfect competition. Using this model, we derive a first-order approximation of $\Delta \ln Y$ (see Appendix A.4 for details):

$$\Delta \ln Y = B\left[\operatorname{diag}(\beta_z)\Delta z + \operatorname{diag}(\beta_L)\Delta L + \Delta\xi\right], \tag{24}$$

where

$$\begin{aligned} B &= \left[I - \left(\operatorname{diag}(\beta_\Phi)\Pi + (\operatorname{diag}(\beta_\Omega) - I)(\operatorname{diag}(\beta_\Phi) - I)\Pi(I - \operatorname{diag}(\beta_\Omega)\Gamma)^{-1}\Gamma\right)\right]^{-1} \\ &\times \left[\Pi + (\operatorname{diag}(\beta_\Omega) - I)\Pi(I - \operatorname{diag}(\beta_\Omega)\Gamma)^{-1}\Gamma\right] \end{aligned} \tag{25}$$

where $\operatorname{diag}(\beta_a)$ denotes the N-by-N matrix with diagonal elements β_a^i for all i and off-diagonal elements 0, and where Π and Γ are the N-by-N matrices of import shares and export shares in period 0, respectively. The matrix B^z from (7) corresponds to $B\operatorname{diag}(\beta_z)$ in our model, while B^ε corresponds to B.

Evaluating individual elements of the matrix B yields, for $i = m_0$,

$$\begin{aligned} b_{ij} &= \sum_{K=0}^{\infty} \sum_{m_1,\dots,m_K} \prod_{\ell=1}^{K} \Bigg(\beta_\Phi^{m_{\ell-1}} \pi_{m_{\ell-1},m_\ell} + (\beta_\Omega^{m_{\ell-1}} - 1)(\beta_\Phi^{m_{\ell-1}} - 1) \\ &\quad \times \sum_r \pi_{m_{\ell-1},r} \sum_{d=0}^{\infty} \sum_{s_1,\dots,s_d} (\beta_\Omega^{m_{\ell-1}})^d \gamma_{rs_1}\gamma_{s_1s_2}\dots\gamma_{s_d,m_\ell} \Bigg) \\ &\quad \times \left(\pi_{m_K,j} + (\beta_\Omega^{m_K} - 1) \sum_r \pi_{m_K,r} \sum_{d=0}^{\infty} \sum_{s_1,\dots,s_d} (\beta_\Omega^{m_K})^d \gamma_{rs_1}\gamma_{s_1s_2}\dots\gamma_{s_d,j} \right). \end{aligned} \tag{26}$$

Recall from section 2 that the TWFE and heterogeneous-robust estimators are correctly specified only if $b^z_{ij} = 0$ for all $i \neq j$, or equivalently $b_{ij}\beta^j_z = 0$ in the notation of our structural model. Furthermore, the TWFE yields the slope of the best fit line if, given Assumption 1 with $\kappa \neq 0$, the elements of B^ε are constant across units (proposition 2, condition 1). Similarly, the upstream/downstream estimator is correctly specified if $b^z_{ij} = \beta^{Upstream}\pi_{ij} + \beta^{Downstream}\gamma_{ij}$ for $i \neq j$. Lastly, the global time-series estimator is correctly specified if the sum of the elements in each column of matrix B^z is equal to a constant. Investigating (26) reveals that none of these conditions are met if countries trade with each other. The b_{ij} terms are not zero (unless all trade shares are zero), nor are they linear functions of π_{ij} and γ_{ij}. Additionally, it is very unlikely that $\sum_j b_{ij}\beta^j_z$ are constant across units i.

It is instructive to assess the conditions under which $b_{ij}\beta^j_z$ *are* constant across units. We show in Appendix A.4 that, if we impose constant trade shares $\pi_{ij} = \gamma_{ij} = \frac{1}{N}$ and constant labor shares $\eta_i = \eta$ across all units, $b_{ij}\beta^j_z$ simplifies to a constant. In this case, the TWFE estimator recovers the slope of the best fit line of expected treatment effects as functions of own-unit treatment. The conditions for identification of the global time-series estimator would also be met. But of course, these conditions are not likely to hold in a world with heterogeneous trade costs.

3.3 Estimation

In this subsection, we outline our procedure for estimating structural parameters and computing counterfactuals. We proceed in five steps. First, we estimate gravity regressions to recover estimates of the trade elasticity and bilateral trade costs. Second, conditional on estimated trade costs, we solve the non-linear system for Φ_{ist} and Ω_{ist}. Third, for a given assumption of market structure (either perfect or imperfect competition), we substitute into (23), and solve for A_{ist}. Fourth, we project $\breve{A}_{ist}$ on all observable determinants z^v to estimate μ^v_s. Finally, we compute counterfactuals via exact hat algebra (Dekle et al., 2007).

Gravity and Multilateral Resistance Terms The structural gravity equation (18) relates bilateral trade flows of sector s goods from country i to country n, X_{nist}, to bilateral trade frictions, along with endogenous terms. Collecting multilateral resistance terms and aggregate income and expenditures into importer-sector-year and exporter-sector-year fixed effects, we have

$$\begin{aligned} X_{nist} &= e_{ist} \times m_{nst} \times \exp\Bigg(-\theta_s \nu_s^{dist} \ln dist_{ni} - \theta_s \nu_s^{Contig} Contig_{ni} - \theta_s \nu_s^{Border} Border_{ni} \\ &- \theta_s \nu_s^{ComLang} ComLang_{ni} - \theta_s \nu_s^{Colonial} Colonial_{ni} - \theta_s \ln\left(ShippingCost_{nist}\right)\Bigg) \rho_{nist} \quad (27) \end{aligned}$$

where we adopt the standard parametric representation: $\phi_{nist} = \exp\left(-\theta_s \sum_g \nu_s^g g_{ni}\right)$ with $g \in \{\ln dist, Contig, Border, ComLang, Colonial, ShippingCost\}$ identifying the usual gravity variables, and ρ_{nist} is a statistical error term. Equation (27) is estimated using Pseudo-Poisson Maximum Likelihood (PPML) as Silva & Tenreyro (2006).

Examining (27), we see that for the first 5 gravity variables—log distance, contiguous neighbors, common border, common official language, and common colonial connection—PPML identifies the product of θ_s and parameters ν_s^g, the latter of which represent the elasticity of trade costs to trade variable g.[13] But the parameter on shipping costs directly identifies θ_s. This is because θ_s represents the elasticity of trade flows to trade costs, and shipping costs are denominated in monetary terms. Without data on shipping costs, we take estimates of θ_s identified from the usual gravity regression from the literature.[14] With estimates of ν_s^g and θ_s, we compute estimated trade frictions $\check{\phi}_{nist}$.

Next, with $\check{\phi}_{nist}$ in hand, we solve the system

$$\widetilde{m}_{nst} = \sum_k \frac{Y_{kst}}{\widetilde{e}_{kst}} \check{\phi}_{nkst} \quad , \quad \widetilde{e}_{ist} = \sum_k \frac{X_{kst}}{\widetilde{m}_{kst}} \check{\phi}_{kist} \tag{28}$$

[13] The microfoundations indicate that $\phi_{nist} = -\theta_s \tau_{nist}$, where τ_{nist} represent iceberg trade costs. This is why all the gravity variables are multiplied by $-\theta_s$ as well as parameters ν_s^g.

[14] In the empirical application, we do not observe shipping costs for all bilateral partners, so we take θ_s from Shapiro (2016), who compiled detailed sector-specific shipping data for two importing countries—the US and Australia—and estimated θ_s via regressions like (27). We can also use the estimates from Caliendo & Parro (2015), which are derived from variation in tariffs. In either case, we still estimate (27) with the data at hand to compute $\check{\phi}_{nist}$.

for $\widetilde{m}_{nst}$ and $\widetilde{e}_{ist}$. Normalizing $\widetilde{m}_{1st} = \widetilde{e}_{1st} = 1$, and denoting the normalized values of $\widetilde{m}_{nst}$ and $\widetilde{e}_{ist}$ by $\check{\Phi}_{nst}$ and $\check{\Omega}_{ist}$, respectively, we have $\check{\Phi}_{ist} = \frac{\Phi_{nst}}{\Phi_{1st}}$ and $\check{\Omega}_{ist} = \frac{\Omega_{ist}}{\Omega_{1st}}$. Hence, multilateral resistance terms are only identified up to a normalization.[15]

Inverting the model and Estimating μ With estimates of $\check{\Phi}_{ist}$, $\check{\Omega}_{ist}$, and θ_s, we have almost all the elements needed to invert (23). The only variables that are missing are the sector-country labor production shares η_{is}, intermediate production shares α^M_{ish}, and consumption shares α^C_{is}. We assume all these series are available from input-output tables.

Substituting these parameters into all the reduced-form parameters (the βs), we can solve for $\check{A}_{ist}$. We then estimate TWFE regressions in first-difference sector-by-sector

$$\Delta\check{A}_{ist} = \delta_t + \sum_{v=1}^{\mathscr{V}} \delta^v_s \Delta z^v_{it} + \Delta\omega_{ist} \tag{29}$$

and recover $\check{\mu}^v_s$ from regression parameters $\check{\delta}^v_s$. While $\check{\delta}^v_s$s are estimated via TWFE, and are thus potentially vulnerable to the same critique we make of β^{FE} in model (8), most models assume weather shocks in one region do not affect *productivity* elsewhere, indicating that SUTVA might plausibly hold in this case. The assumption is less defensible in (8), wherein the outcome variable is an equilibrium outcome like gross output.

Finally, we note that the exclusion restrictions we exploit in estimating (29) are different from those exploited by Rudik et al. (2022) to estimate the elasticity of productivity to weather shocks. Specifically, the exclusion restriction in Rudik et al. (2022) is that the difference in unobserved components of productivity between countries i and n and the difference in material input prices between countries i and n are orthogonal to the difference in weather shocks and the difference in wages between countries i and n.[16] However,

[15] This normalization does not influence the computation of counterfactuals, as we show below.

[16] To illustrate, in the case of perfect competition, we divide bilateral trade flows by self-trade flows to derive

$$\ln\left(\frac{X_{nist}}{X_{nnst}}\right) = \sum_{v=1}^{\mathscr{V}} \mu^v_s\left(z^v_{it} - z^v_{nt}\right) + \left(\psi_{is} - \psi_{ns}\right) + \left(\omega_{ist} - \omega_{nst}\right) - \theta_s \ln\left(\frac{w_{it}^{\eta_{is}}}{w_{nt}^{\eta_{ns}}}\right) - \theta_s \ln\left(\frac{\sum_{h=1}^{S} p_{iht}^{\alpha^M_{ish}(1-\eta_{is})}}{\sum_{h=1}^{S} p_{nht}^{\alpha^M_{nsh}(1-\eta_{ns})}}\right) + \phi_{nist}.$$

Rudik et al. (2022) parameterize ϕ_{nist} in the usual way and then estimate μ^v_ss directly from the equation above via PPML, controlling for the relative wage rates. Lacking data on material input prices, these variables are given our model, wages are jointly determined with material input prices, both of which depend on unobserved components of productivity, so it is unlikely that this strategy identifies μ_s^vs. For this reason, we develop a strategy based on estimating (29) instead.

Computing Counterfactuals The last step is to compute counterfactual outcomes under counterfactual vector z_t'. In particular, we can solve for the counterfactual outcome x_t', for a generic variable x_t, assuming the vector of observables had been the same as the base year values, $z_t' = z_0$. Subtracting this quantity from the observed outcomes x_t yields the RTE.

Armed with estimates of multilateral resistance terms, trade frictions, and structural parameters, we solve the system (denoting $\widehat{x_t} \equiv \frac{x_t'}{x_t}$):

$$\widehat{\Phi}_{nst} = \frac{\sum_k \frac{\widehat{Y}_{kst} Y_{kst}}{\widehat{\Omega}_{kst} \check{\Omega}_{kst}} \check{\phi}_{knst}}{\sum_k \frac{Y_{kst}}{\check{\Omega}_{kst}} \check{\phi}_{nkst}} \quad , \quad \widehat{\Omega}_{nst} = \frac{\sum_k \frac{\widehat{X}_{kst} X_{kst}}{\widehat{\Phi}_{kst} \check{\Phi}_{kst}} \check{\phi}_{knst}}{\sum_k \frac{X_{kst}}{\check{\Phi}_{kst}} \check{\phi}_{knst}} \tag{30}$$

$$\widehat{Y}_{ist} = \exp\left(\sum_{v=1}^{\mathscr{V}} \beta_{z,v}^{i,s} \left((z_{it}^v)' - z_{it}^v \right) \right) \widehat{L}_{ist}^{\beta_L^{i,s}} \widehat{\Omega}_{ist}^{\beta_\Omega^{i,s}} \left(\prod_{h=1}^{S} \widehat{\Phi}_{iht}^{\beta_{\Phi_h}^{i,s}} \right) \left(\prod_{h=1}^{S} \widehat{X}_{iht}^{\beta_{X_h}^{i,s}} \right) \tag{31}$$

$$\widehat{X}_{ist} X_{ist} = \sum_h (\alpha_{is}^C \eta_{ih} + (1 - \eta_{ih}) \alpha_{ihs}^M) \widehat{Y}_{iht} Y_{iht} \tag{32}$$

$$\widehat{L}_{ist} L_{ist} = L_{it} \frac{\eta_{is} \widehat{Y}_{ist} Y_{ist}}{\sum_h \eta_{ih} \widehat{Y}_{iht} Y_{iht}} \tag{33}$$

for $\widehat{\Phi}_{nst}$, $\widehat{\Omega}_{nst}$, $\widehat{Y}_{nst}$, , $\widehat{L}_{nst}$ and $\widehat{X}_{nst}$. This system is solved by fixed point iteration, conditional on an assumption of market structure.

With the solution to this system in hand, the proportional change in the real wage is

included in the error term.

computed, without loss of generality, as[17]

$$\widehat{RW}_{it} = \frac{\widehat{w}_{it}}{\widehat{p}_{it}} = \frac{\widehat{Y}_{i1t}/\widehat{L}_{i1t}}{\left(\prod_{h=1}^{S}\widehat{\Phi}_{iht}^{-\frac{\alpha_{ih}^{C}}{\theta_h}}\right)\left(\prod_{h=1}^{S}\widehat{X}_{iht}^{\frac{\alpha_{ih}^{C}}{\theta_h}+\frac{\alpha_{ih}^{C}}{\sigma_h-1}}\right)^{\chi}}. \tag{34}$$

4 Data

In this section, we present data that will be used in the empirical application and the Monte Carlo experiments that follow.

Trade and Production Data We model gross output and bilateral trade flows in the agriculture and manufacturing sectors at the country level using data from Mayer et al. (2023) and Fontagné et al. (2023) over the periods 1991 - 2019.[18] Mayer et al. (2023) and Fontagné et al. (2023) combine data on annual cross-border international bilateral flows (UN Commodity Trade Statistics Database (COMTRADE) for manufacturing and Food and Agriculture Organization of the United Nations Statistics Division (FAOSTAT) for agriculture) with production data (the UNIDO Industrial Statistics database (INDSTAT) for manufacturing and again FAOSTAT for agriculture) to compile a square matrix of bilateral trade flows for each sector, including self-trade. It is important to have measures of self-trade in order to compute counterfactuals (Head & Mayer, 2014; Yotov, 2021).[19] The underlying data is disaggregated at the 6-digit level for manufacturing and the 4-digit level for agriculture. We aggregate the 9 industries into one manufacturing sector.

Weather Data To measure the observable determinants z_{it}^{v}, we use temperature and precipitation data from the global reanalysis ERA-5 dataset compiled by the European Centre for Medium-Range Weather Forecasts (Hersbach et al., 2023). We start from daily data

[17] Since the wage is equalized across sectors, we have $\widehat{w}_{it} = \frac{\widehat{Y}_{ist}}{\widehat{L}_{ist}}$ for any s.

[18] That data run through 2022, but we stop the panel in 2019, the year before the COVID 19 shock.

[19] Domestic sales are calculated as the difference between the value of production and total exports. When the sum of total exports exceeds the value of production, Mayer et al. (2023) and Fontagné et al. (2023) set the value to missing and extrapolate. This affects less than 1% of observations.

resolved at the 0.25° x 0.25° resolution grid (corresponding to cells of 30km x 30km at most, at the Equator) from 1991 to 2019. We then aggregate over time and space to get yearly information at the country level.[20]

Our aggregation method relies on temperature and precipitation bins to capture the potential nonlinear effects of weather on economic variables, while imposing minimal restrictions on the response functional form (Schlenker & Roberts, 2006; Deschênes & Greenstone, 2011). For temperature, we follow Somanathan et al. (2021) and consider six temperature bins (expressed in celsius): $(-\infty; 0℃]$, $(0℃; 20℃]$, $(20℃; 25℃]$, $(25℃; 30℃]$, $(30℃; 35℃]$, $(35℃; +\infty)$. For each grid cell, we compute the number of days in the year where maximum daily temperature falls within a given temperature bin.[21] We build similar precipitation variables, following Akyapi et al. (2022). We consider four precipitation bins (expressed in milliliters): $[0; 1]$, $(1; 10]$, $(10; 20]$, $(20; +\infty)$, and compute the number of days in the year where total daily precipitation falls within a given precipitation bin.[22] We then derive country-level variables by aggregating grid-level information using population-weighted averages, to take into account the uneven distribution of people and economic activities across space (Dell et al., 2012). We use population weights of year 2000 from the Socioeconomic Data and Application Center's UN WPP-Adjusted Gridded Population of the World dataset (CIESIN , 2018).

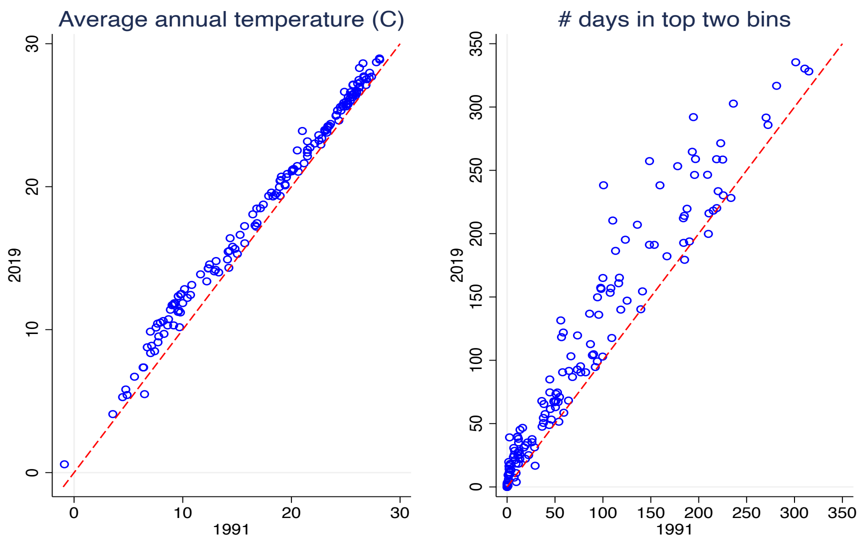
Figure 1 shows average annual temperature in ℃ as well as the number of days with maximum temperature above 30℃ (the top two bins) at the beginning and at the end of the sample. We find that almost all countries warmed between 1991 and 2019. The average growth in annual temperature was 1.2℃, and the average increase in the number of days in the top two bins was 22—a substantial rightward shift in the climate distribution over the period. In the left panel, we see that the largest increases in number of days with maximum temperature above 30℃ occurred in countries that had medium to large counts for these bins in 1991.

[20] Following Hsiang (2016), we compute (nonlinear) temporal aggregation at the grid cell level before aggregating values across space.

[21] Formally, for a bin $(x_1; x_2]$, we build the yearly variable $\sum_j \mathbb{1}_{\{x_1 < T_j^{max} \leq x_2\}}$, where T_j^{max} is the maximum temperature of day j.

[22] In robustness checks, we also consider an alternative temporal aggregation method by computing simple averages of the daily mean temperature and total precipitation over the year.

Figure 1: Warming between 1991-2019



Notes: Figure plots average temperature degrees C (left) and number of days with maximum temperature over 30 degrees C (right) in the year 2019 (y-axis) against the year 1991 (x-axis). Each dot correpsonds to a country.

Table 1 shows significant variation in weather variables year-to-year over the period. The extreme temperature bins seem to be the most volatile: 61% of the observations deviate more than 75% from the country mean for the $(-\infty℃, 0℃]$ bin, and more than a third of the observations for the $(35℃, +\infty℃)$ bin. Conversely, the middle bins ($(20℃, 25℃]$, $(25℃, 30℃]$) are more stable over time for a given country. As is standard in this literature, we exploit these deviations from country averages over time to identify effects (see for example Deryugina & Hsiang 2014).[23]

[23] Our variables describe annual (short-term) *weather* variations rather than climate conditions, the latter corresponding to the long term distribution of weather variables. Hsiang (2016) shows, using the envelope theorem, that for an optimized variable, variation in weather is isomorphic to variation in climate. Yet, there is a caveat that arises from adaptation behaviors (Kolstad & Moore, 2020). Some methods have been proposed to include these adaptative responses in the estimation of weather response functions, see for instance Hultgren et al. (2022).

Table 1: Observed variation in annual temperature and precipitation measured with daily bins (1991-2019)

	Proportion of country-year observations with number of days in each weather bin [...]% above/below country mean				
	1%	10%	25%	50%	75%
temperature bins					
inf-0°C	0.98	0.89	0.76	0.66	0.61
0-20°C	0.93	9.49	0.31	0.22	0.18
20-25°C	0.94	0.49	0.19	0.08	0.05
25-30°C	0.91	0.41	0.13	0.04	0.03
30-35°C	0.95	0.57	0.31	0.18	0.12
35-inf°C	0.98	0.80	0.63	0.46	0.35
precipitation bins					
0-1 mm	0.84	0.14	0.02	0.00	0.00
1-10 mm	0.90	0.28	0.04	0.01	0.00
10-20 mm	0.95	0.56	0.20	0.04	0.02
20-inf mm	0.97	0.71	0.39	0.15	0.07

Lecture: for the $(25°C, 30°C]$ variable, 91% of observations deviate more than 1% from the country mean while only 3% of observations deviate more than 75% from the country mean.

Gravity Data The gravity variables required to estimate equation (27) come from the Gravity database developed by CEPII (Conte et al., 2022).[24] Each observation corresponds to a combination of an exporter country, an importing country and a year. Data spans from 1948 to 2019, and includes 252 countries (with a history of past territorial configurations of countries). We consider bilateral characteristics r_{ni} such as the geographical distance between two countries i and n, whether they share a common language, whether they are contiguous, whether they were in a colonial relationship or shared a common colonizer, whether they have a regional trade agreement.

GTAP Data We compute production and consumption shares from the Global Trade Analysis Project (GTAP) Data Base version 9 (Aguiar et al., 2016). We consider domestic and import expenditures at purchaser's price for the reference year 2011. The database contains information on 57 commodities for 116 countries and 24 aggregate regions. We aggregate these commodities into three broader sectors (mapping reported in Table C.1):

[24] We downloaded version V202211 of the database.

agriculture, manufacturing, other. We attribute the values of the aggregate regions to their individual constitutive countries (see Table C.2).

We compute α^C_{is} by dividing country i's household consumption (domestic and imported) in each sector s by total consumption. We compute α^M_{ish} by dividing input expenditures in each sector h goods (domestic and imported) coming from sector s in country i by total expenditures in all sectors. The parameter η_{is} is computed as the ratio of labor expenses by firms in country i and sector s over the sum of their total variable input expenditures (domestic and imported).

Labor Data Country-sector-year employment levels correspond to International Labor Organization modeled estimates (ILOEST).[25] We aggregate multiple industries into three broader sectors (see Table C.3): agriculture, manufacturing, other.

Final Dataset The merged dataset covers 132 countries from 1991 to 2019 with information on trade, production, weather, labor allocation, and consumption for both the agricultural and manufacturing sectors.

5 Monte Carlo Experiments

In this section, we evaluate the performance of the TWFE estimator, extensions thereof, and our structural procedure in Monte Carlo experiments. We use the model from section 3 as the data generating process, and tailor the experiments to the climate application that follows, using data from section 4.

5.1 Set Up

We simulate fictitious world economies as follows. In each experiment, we treat the 132 countries from the merged dataset in section 4 as the "units", each endowed with their observed characteristics from section 4 (geography, population, production parameters, etc.).

[25] Starting from labor force survey or household survey data, the ILO relies on econometric models using economic and demographic variables to fill in missing observations.

For simplicity, we consider a single continuous treatment variable—the number of days in the year with maximum temperature above 30°C (the top two bins). We first estimate trade costs for all importer-exporter pairs using bilateral trade flows and gravity variables. We then construct country-sector-year productivity A_{ist} using (20), choosing a value for $\mu^{30°C}$. We set ψ_{is} proportional to baseline temperature and construct ω_{ist} following a random walk process, with $\Delta\omega_{ist} \sim N(0, 0.01)$. We also draw stochastic error term $\rho_{nist} \sim N(0, 0.01)$ to build τ_{nist}. Then, for an assumed market structure and for assumed values of σ_s and θ_s, we solve the system of equations (19) and (21), (22), and (23) in levels period by period for $\{X_{st}\}$,$\{Y_{st}\}$,$\{\Phi_{st}\}$,$\{\Omega_{st}\}$, $\{L_{st}\}$, by fixed point iteration for sectors $s \in (1, ...S)$. This process yields many simulated datasets, or "replications", each of which approximating the true data.[26]

We simulate equilibrium levels each year from 1991 until 2019 using observed temperature realizations, and then simulate the counterfactual equilibrium in the last year of the sample imposing treatment from the first year of the sample, $z'_{i,2019} = z_{i,1991}$. Taking the difference between these two equilibria yields the RTEs for each replication. Taking averages of the RTEs across replications yields the ETE.[27]

For each replication, we then implement the TWFE estimator, the heterogeneous-robust estimator, the upstream/downstream estimator, the global time-series estimator, and our structural estimator, and compare estimated effects to the "true" ETEs.

[26] These simulations are similar to those in Head & Mayer (2014) and Baier & Bergstrand (2009), who also simulate unobserved components of trade costs and then test the performance of gravity estimation models. Key differences here are that we also simulate an unobserved component of productivity, and that the gravity estimation is just one step in our procedure to recover estimates of the effects of observed determinants of productivity.

[27] In the data, expenditures are not equal to income, i.e. countries operate either in surplus or in deficit in a given year. We abstract from this feature in the simulations. In the empirical application, we first compute counterfactual outcomes in the final year assuming weather corresponds to the observed 2019 weather, but deficits go to zero. We then compute counterfactuals setting $z'_{i,2019} = z_{i,1991}$ also assuming zero deficits. We compute the treatment effect of the change in z as the difference between equilibrium levels in the two counterfactuals. Researchers often use this strategy to address deficits (Ossa, 2016).

5.2 Single Sector Simulation Results

In our first Monte Carlo experiments, we simulate equilibria with $S = 1$. We set $\mu^{30℃} = -.003$, which means that each day with maximum temperature above 30℃ decreases productivity by 0.3%.[28] We set $\sigma = 3.6$, the average price elasticity of substitution from Broda & Weinstein (2006), and set $\theta = 3.3$, the average trade elasticity estimated by Shapiro (2016) for the agricultural sector.

Figure 2 presents results from experiments simulated under the assumption of perfect competition. The black dashed line plots the ETE on gross output in 2019 across 100 replications. The x-axis represents the change in the number of days with maximum temperature above 30℃ between 1991 and 2019. The blue solid line presents the median estimate from the indicated estimator, and the blue shaded region depicts the interquartile range of the estimates across the 100 replications.

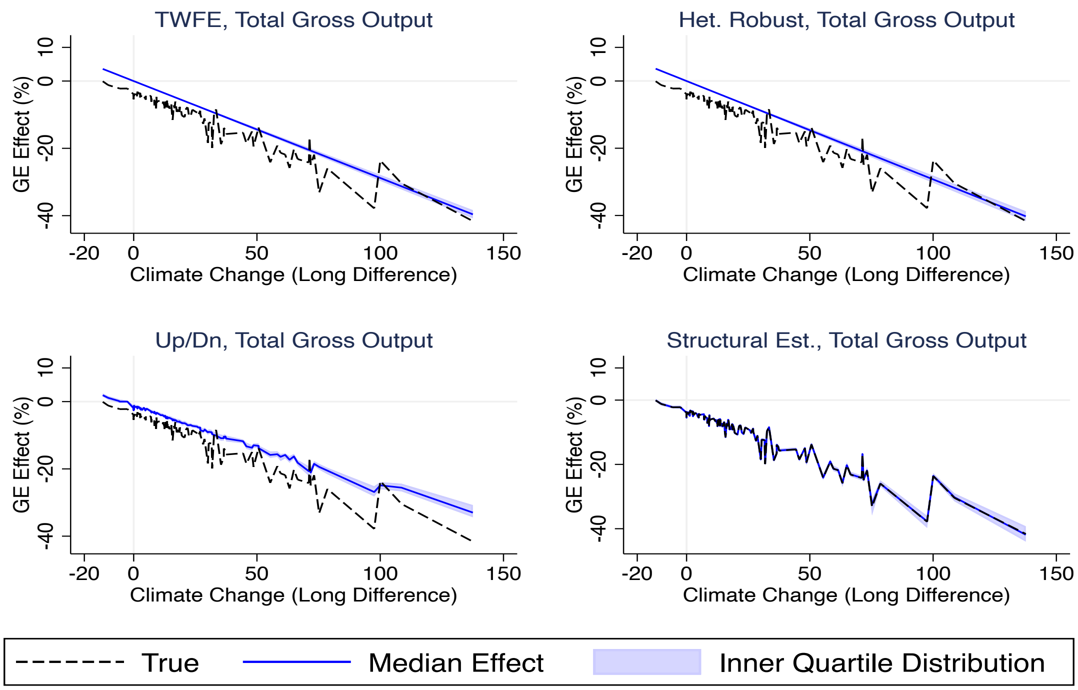
In Figure 2, we see that gross output falls in all countries as a result of the change in the treatment vector from base-year weather to end-of-sample weather. The ETE varies from 0% to -42% across countries, and tends to increase in magnitude with treatment, but there is substantial variation in ETE effects across countries with very similar warming. This variation arises because of variation in network position (τ_t), size (L_t), production structure (η_t) and preferences (α_t^C). The fact that all countries suffer in the simulations—even those with *negative* treatment (i.e., countries that cooled slightly)—indicates that there are spillover effects. Moreover, this fact suggests that spatial linkages amplify treatment effects, as in Adao et al. (2020). Hence, when productivity falls in any one country because of a temperature increase, output in other countries falls as well.[29]

In the top left panel of Figure 2, we see that the TWFE estimator *understates* the damages of warming, for most countries. Proposition 2 provides an explanation since, under its conditions, the TWFE recovers the slope of the best fit line through the ETEs,

[28] This number is equivalent to the correlation estimated by Somanathan et al. (2021) between firm-level annual output and the number of day with maximum temperature above 30℃.

[29] It is not obvious that spatial linkages should amplify treatment effects. In general equilibrium, a negative productivity shock in one country could lead to increased sales in another country, as competing firms in other countries capture market share. In this case, one would imagine that spatial linkages should dampen treatment effects. We find that, in our setting, reduced demand and higher input costs from negative productivity shocks dominate the substitution effect, leading to lower gross output in other countries.

Figure 2: Estimated vs True Treatment Effects, Single Sector



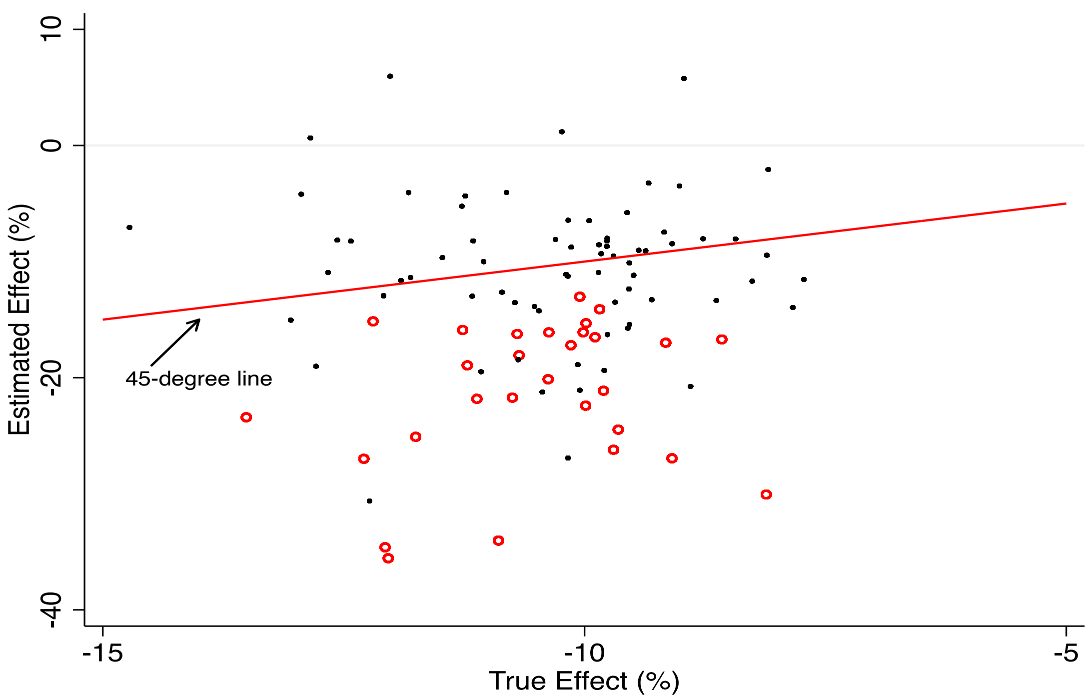
Notes: Figure plots the treatment effect in % on the y-axis against the change in the # of days with maximum temperature above 30℃ between 1991 and 2019. The black dashed line plots the ETE on gross output in 2019 across 100 replications, i.e. the average true effect computed across replications. The blue line presents the median estimate from the indicated estimator, and the blue shaded region depicts the interquartile range of the estimates across 100 replications. Simulations include 132 countries and 29 time periods. Data is aggregated to a single sector. Market structure is perfect competition. Parameters are set at $\theta = 3.3$, $\sigma = 3.6$, $\mu = -.003$.

but not the intercept. While the conditions of Proposition 2 are not met in our simulations, the TWFE still recovers a value close to the slope of the best fit line through the ETEs. By construction, the TWFE imposes an intercept of 0 (zero own-country treatment implies zero treatment effect), hence the TWFE estimates are shifted up relative to the best fit line through the true expected treatment effects.[30]

[30] Borusyak et al. (2022) find the same direction of bias in the TWFE estimates in the context of a small open economy with intra-region migration flows.

In the top right (bottom left) panel of Figure 2, we find that the the heterogeneous robust estimator from de Chaisemartin et al. (2024b) and the upstream/downstream estimator generate estimates that are qualitatively similar to the results from the TWFE: both estimators mostly understate the true ETE. The upstream/downstream estimator is more flexible than the TWFE estimator, as it explicitly models spillover effects; but the constraints imposed on these spillover effects appear not to match the data generating process from this standard quantitative trade model, as the estimated treatment effects are still quite far from the true ETEs.

Figure 3: Estimated vs True Treatment Effects, Global



Notes: Figure plots the estimates treatment effect on global gross output in % on the y-axis against the true treatment effect on global gross output in % on the x-axis, where the estimates are computed via the local projection approach from Bilal & Känzig (2024). The 45-degree line is plotted in red. Red dots indicate replications for which the point estimate on contemporaneous temperature shocks are statistically significant at the 5% level. Simulations include 132 countries and 29 time periods. Data is aggregated to a single sector. Market structure is perfect competition. Parameters are set at $\theta = 3.3$, $\sigma = 3.6$, $\mu = -.003$.

Finally, in the bottom right panel of Figure 2, we see that the structural estimator recovers very well the ETEs on gross output. In appendix Figure B.6, we find that the structural estimator also recovers the ETEs on real wage.[31] The median estimate across replications aligns quite well with the ETE, and the interquartile distribution across replications is roughly centered on the truth. Since the structural estimator is designed for the data generating process, it is not surprising that the estimator performs well. Nevertheless, it is not obvious that the estimator should be successful in recovering unbiased estimates in finite samples. These Monte Carlo experiments indicate that the structural estimator can succeed in finite samples of the size we encounter in the data.

Figure 3 presents results from the global time-series estimator of Bilal & Känzig (2024). The Figure plots the estimated treatment effect on global gross output against the true treatment effect on global gross output, by replication. Since our model is static, we abstract from the dynamic considerations in Bilal & Känzig (2024). To implement the estimator, we simply aggregate global gross output and temperature and then estimate model (12) on the simulated data. Figure 3 shows that the global time series estimator yields biased estimates, and that the bias is away from zero. The mean (median) estimated effect is -13.0% (-13.7%), whereas the true mean (median) effect is -10.4% (-10.1%). Hence, the global time-series estimator tends to overstate the true global treatment effect by about 30%. If we condition on statistically significant estimates (red dots), the mean (median) estimated effect increases in magnitude to -21.4% (-20.1%). These estimates suggest that the treatment effect estimated in Bilal & Känzig (2024) could overstate the economic damages from climate change.[32]

We repeat the single-sector experiments imposing a higher value for the trade elasticity ($\theta = 8.3$), and assuming imperfect competition for the data generating process, and present results in the Appendix (Figures B.2 and B.3). Results are qualitatively the same. The TWFE, heterogeneous robust estimator, and the upstream downstream estimators understate the loss to gross output from warming, while the median estimate from the struc-

[31] We do not estimate effects on real wage with the reduced-form estimators because we assume real wage is not observed. It is only in virtue of equation (34) that we can estimate effect on the real wage with the structural estimator.

[32] These results do not appear to be driven by small sample bias, as we find approximately the same results when we double the length of the panel (see Figure B.1).

tural estimator aligns well with the ETE, with errors roughly symmetrically distributed around 0.[33]

5.3 Two Sector Simulation Results

We next simulate world economies assuming two sectors in each country, thus allowing for across-sector spillover effects. We use agricultural consumption and production structures as well as trade costs for the first sector, and manufacturing consumption and production structures as well as trade costs for the second sector. We set $\sigma = 3.6$ in both agricultural and manufacturing from Broda & Weinstein (2006),[34] and set trade elasticities $\theta_1 = 8.3$ in sector 1 and $\theta_2 = 8.5$ in sector 2, the estimated trade elasticities for agriculture and manufacturing, respectively, from Caliendo & Parro (2015).[35]

We still consider just a single continuous treatment variable, but impose heterogeneous parameters across sectors. We impose $\mu_1^{30℃} = -.008$ for sector 1 and $\mu_2^{30℃} = 0$ for sector 2, in the baseline case. Hence, we assume *no productivity effect* in manufacturing. This parametrization highlights the role of general equilibrium spillovers across sectors, as effects on gross output in sector 2 cannot arise from productivity changes.[36]

Figure 4 presents results from experiments simulated under the assumption of perfect competition. As before, the black dashed line plots the ETEs on gross output in 2019 across 100 replications, and the x-axis represents the change in the number of days with maximum temperature above 30℃ between 1991 and 2019. The dark blue solid line presents the median estimate from the indicated reduced-from estimator considering each sector independently (first two columns), or considering aggregate output (last column). Blue shaded regions depict the interquartile range of the distribution of estimates.

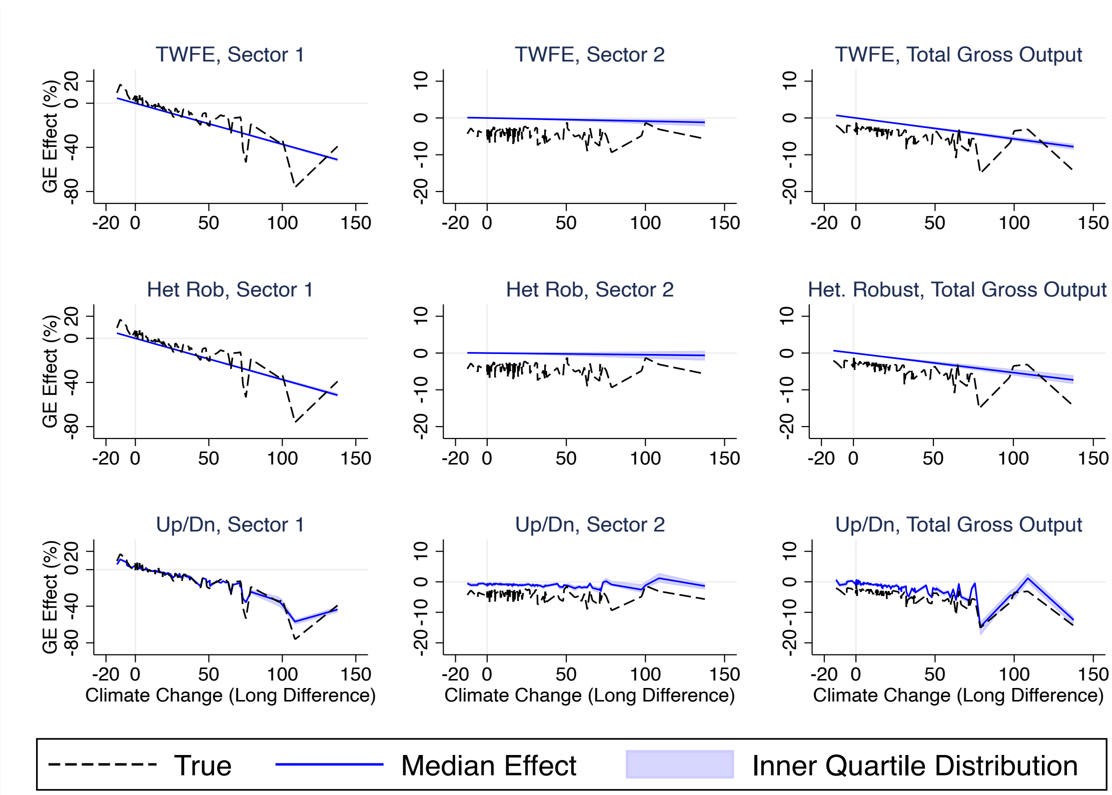
[33]When the true underlying data generating process is monopolistic competition, the global time series estimator becomes wildly imprecise (Figure B.3) with estimated errors ranging from -550% to +300%, and with hardly any point estimates distinguishable from zero.

[34]The average σ across 3-digit industries in the manufacturing sector (SITC Rev3 3-digit codes 231 - 971) computed by Broda & Weinstein (2006) is 3.59. The average σ across 3-digit industries in the agricultural sector (SITC Rev3 3-digit codes 001 - 223) is 3.62, excluding one outlier industry (SITC 017, MEAT AND EDIBLE MEAT OFFAL, PREPARED OR PRESERVED N.E.S.).

[35]Alternatively, we use the trade elasticity from Shapiro (2016) for the agricultural sector.

[36]In the appendix, we repeat the exercise allowing for moderate productivity effects in manufacturing, ($\mu = -.003$), and find qualitatively similar results.

Figure 4: Estimated vs True Treatment Effects, Reduced-Form Estimators, Two Sectors



Notes: Figure plots the treatment effect in % on the y-axis against the change in the # of days with maximum temperature above 30°C between 1991 and 2019. The black dashed line plots the ETEs on gross output in 2019 across 100 replications, i.e. the average true effects computed across replications. The blue solid line presents the median estimate from the indicated estimator, and the blue shaded region depicts the interquartile range of the distribution of these estimates. Simulations include 132 countries and 29 time periods and two sectors. Market structure is perfect competition. Parameters are set at $\theta_1 = 8.3$, $\theta_2 = 8.5$, $\sigma_1 = \sigma_2 = 3.6$, $\mu_1 = -.008$, $\mu_2 = 0$.

As in the one-sector simulations, we find that the TWFE estimator *understates* the loss of aggregate gross output from warming (top right panel). However, broken out by sector, we find different results. In sector 1, for which we impose a large productivity effect, we find that the TWFE overstates the loss in gross output (top left panel). In sector 2, for which we impose no productivity effect, true treatment effects are mostly negative, but practically flat in temperature (middle panels, dashed black line). The TWFE still picks up the slope of this line, which is roughly zero. So using the TWFE estimator, one might conclude that warming has no effect on output in sector 2, even though the warming decreased output in sector 2 for almost all countries (top row, middle panel).

These results can be explained as follows. In a two-sector model, spillovers dampen the loss in gross output in sector 1 (as revealed by the positive intercept), but exacerbate the loss in sector 2 (negative intercept). In countries that warm substantially, labor reallocates towards manufacturing, buoying losses in the sector. By contrast, countries with low increase in temperature gain market shares in agriculture, as their relative productivity in this sector increases. The TWFE estimator recovers the slope of the general equilibrium effects quite well, but misses the intercept, thereby overstating the effects for agriculture and understating the effects for manufacturing.

In the second and third rows of Figure 4, we find that the heterogeneous-robust estimator and upstream/downstream estimator generate results that are qualitatively similar to the TWFE estimates, though the upstream/downstream estimator seems to recover the effects on agriculture quite well. By contrast, Figure 5 shows that the structural estimator recovers very well the ETEs on gross output for each sector independently and for the aggregate, as well as on real wage.

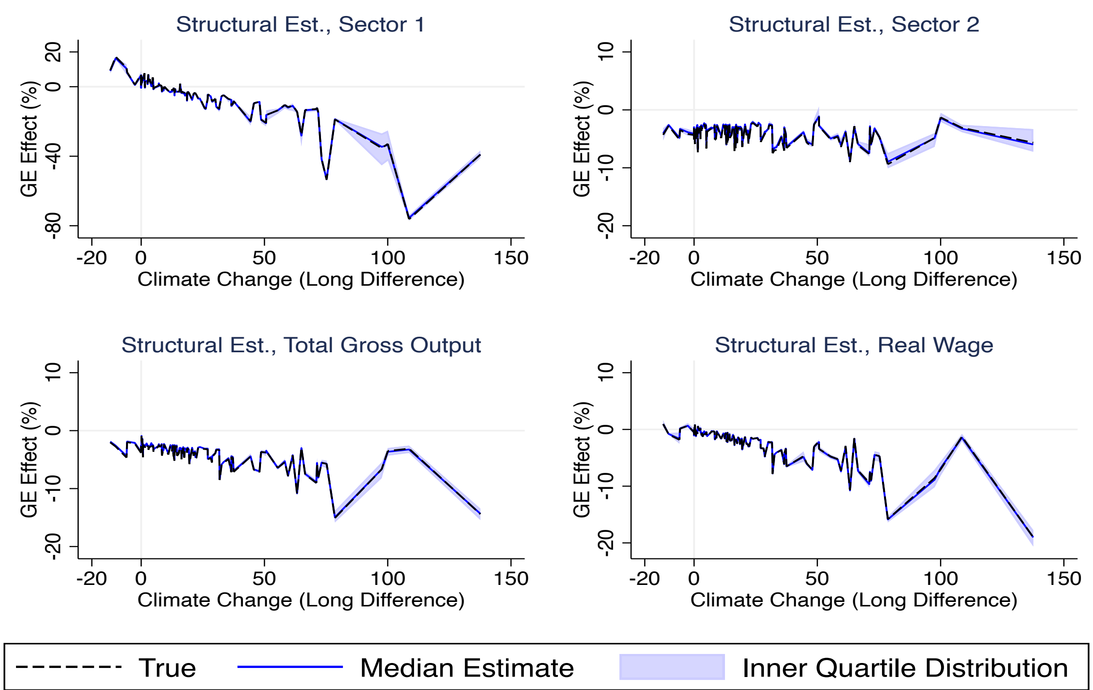
6 Empirical Application

In this section, we estimate the effect of warming between 1991 and 2019 on agricultural and manufacturing gross output, as well as total gross output and real wages, using data from section 4.

6.1 TWFE Estimates

We start by implementing the TWFE regression model (8), regressing year-on-year growth in agricultural and manufacturing gross outputs, and the sum of both, on the change in different temperature bins, taking $[20℃, 25℃]$ as the omitted category, and controlling for flexible continent-by-year fixed effects, as well as year-on-year growth in total labor and

Figure 5: Estimated vs True Treatment Effects, Structural Estimator, Two Sectors



Notes: Figure plots the treatment effect in % on the y-axis against the change in the # of days with maximum temperature above 30°C between 1991 and 2019. The black dashed line plots the ETEs on gross output in 2019 across 100 replications. The blue solid line presents the median estimate from the structural estimator, and the blue shaded region depicts the interquartile range of the distribution of these estimates. Simulations include 132 countries and 29 time periods and two sectors. Market structure is perfect competition. Parameters are set at $\theta_1 = 8.3$, $\theta_2 = 8.5$, $\sigma_1 = \sigma_2 = 3.6$, $\mu_1 = -.008$, $\mu_2 = 0$.

precipitation. We plot point estimates and 95% confidence intervals by outcome in Figure 6.

In left-most panel of Figure 6, we find that the number of days with extreme maximum temperature correlates negatively with gross agricultural output, as in many previous studies (Burke et al., 2015; Deryugina & Hsiang, 2014). We find that an additional day with maximum temperature above 30°C (35°C) is correlated with approximately 0.15% (0.20%) lower annual agricultural output, which is quite close to the values estimated in Deryugina & Hsiang (2014) using US county-level data. These point estimates are statistically significant at the 5% level. The effect from an additional day with negative maximum temperature is also negative, yielding the familiar inverted-U shape in temperature, though the correlation with cold days is not statistically significant.[37]

Figure 6 shows that the effect of the different temperature bins on manufacturing gross output, as well as on the aggregate of both sectors, is indistinguishable from zero, for all bins. Yet, there is evidence that, even though the estimates of β_v^{FE} for manufacturing are close to zero for all temperature bins v, the estimate for the number of days when the maximum temperature reaches at least 35℃ is slightly positive. This positive effect is reminiscent of our findings from the Monte Carlo simulations and suggests the existence of cross-sectoral spillovers. When we combine the two sectors together, the negative effect of temperature on agriculture gross output washes out when combined with the null effect on a larger sector, manufacturing, making it harder to detect any effect in the aggregate (far right panel).

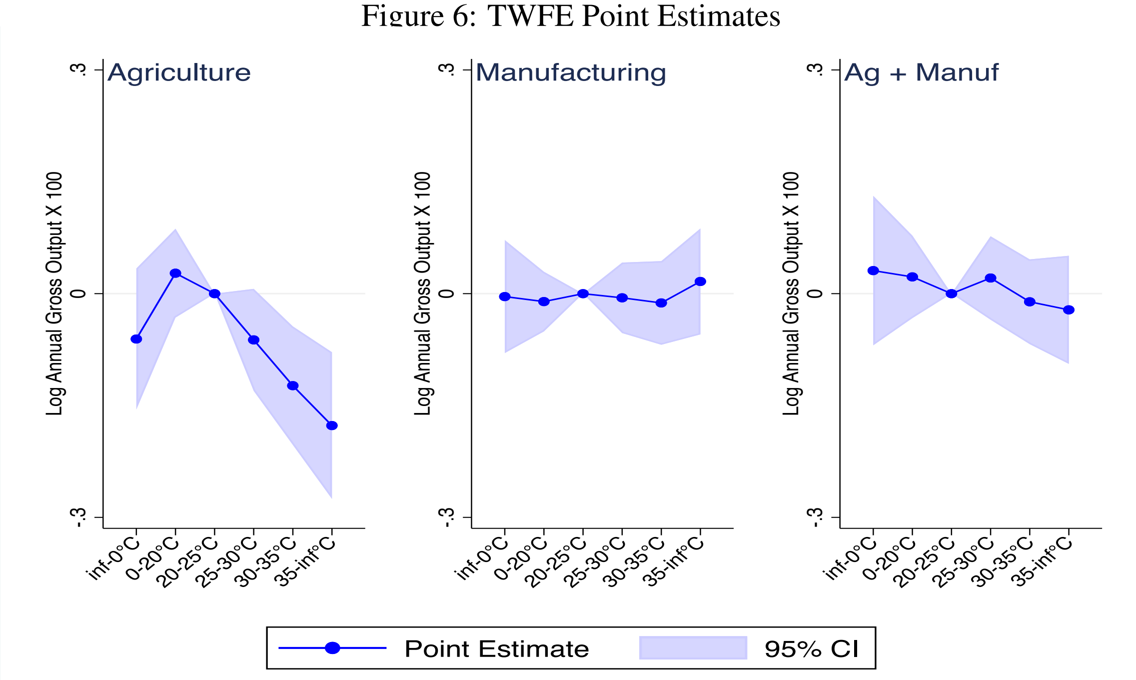
6.2 Structural Estimates

Next, we use the structural gravity model to estimate effects of warming and compare to the TWFE estimator, the heterogeneous robust estimator, and the upstream/downstream estimator.

In Figure 7, we plot point estimates and 95% confidence intervals resulting from estimating regression model (29) for agriculture (left panel) and manufacturing (right panel), where the dependent variable $\Delta\check{A}_{ist}$ is computed by inverting (23) and taking first differences for each sector. To compute $\Delta\check{A}_{ist}$, we impose perfect competition, and set $\theta^{Ag} = 8.3$ and $\theta^{Manuf} = 8.5$, the trade elasticities estimated by Caliendo & Parro (2015) for agriculture and manufacturing, respectively. In Figure 7, we find that each additional day with maximum temperature above 30℃ (35℃) reduces agricultural productivity by approximately 0.35% (0.40%). Both point estimates are statistically distinguishable from zero at the 5% level. We also find that cool days also seem to lower agricultural productivity, but

[37]The point estimates on precipitation bins in these regressions are statistically indistinguishable from zero, as they usually are in the literature (Schlenker & Roberts, 2009), and so we do not report them.

Figure 6: TWFE Point Estimates

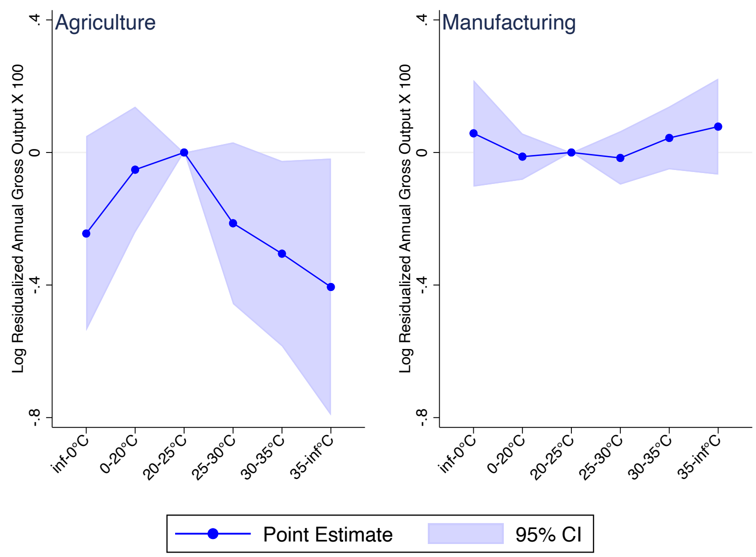


Notes: Figure plots the point estimates and 95% confidence intervals resulting from estimating regression model (8) sector by sector (left two panels), and aggregated together (right panel). Number of days in the year with maximum temperature between 20°C and 25°C serves as the omitted category. 132 countries included, spanning the period 1991 - 2019. Top and bottom 1% of observations in terms of year-over-year growth rates were omitted from the regression. Regression controls for flexible continent-by-year effects, number of days in the year with precipitation falling within given ranges, and total labor. Standard errors are clustered on the country.

these point estimates are not statistically distinguishable from zero. The effects of temperature on manufacturing are small and statistically indistinguishable from zero. Hence, we find largely the same pattern of correlations in Figure 7 as in Figure 6, but the estimates in Figure 7 are net of equilibrium effects like labor and trade reallocation.

Using the structural elasticities from Figure 7, we then solve for counterfactual outcomes assuming 2019 weather realizations had coincided with the 1991 realizations instead, and compute RTEs by country and sector. For these computations, we continue to impose perfect competition and set $\theta^{Ag} = 8.3$ and $\theta^{Manuf} = 8.5$. In our preferred specification, we only include effects on productivity for bins for which point estimates are statistically significant in Figure 7, i.e., the last two temperature bins for agriculture, though

Figure 7: Structural Estimates of the elasticity of productivity to temperature



Notes: Figure plots the point estimates and 95% confidence intervals resulting from estimating regression model (29) for agriculture (left panel) and manufacturing (right panel). Dependent variable $\Delta\check{A}_{ist}$ is computed by inverting (23) and taking first differences, imposing perfect competition and two sectors. We set $\theta^{Ag} = 8.3$, $\theta^{Manuf} = 8.5$, the trade elasticities estimated by Caliendo & Parro (2015) for Agriculture and Manufacturing, respectively. Number of days in the year with maximum temperature between $20°C$ and $25°C$ serves as the omitted category. Period spans 1991 - 2019. Top and bottom 1% of observations in terms of residualized year-over-year growth rates were omitted from the regression. Regression controls for flexible continent-by-year effects, number of days in the year with precipitation falling within given ranges, and total labor. Standard errors are clustered on the country.

estimated effects do not change much if we include the entire vector of point estimates to construct counterfactual productivity.[38]

Figure **??** reports the percentage change in 2019 aggregate gross output (panel a), real wage (panel b), agricultural gross output (panel c), and manufacturing gross output (panel d) resulting from the change in temperature observed between 1991 and 2019 by color on the world map. Warmer colors indicate losses in the observed equilibrium relative to the

[38] To compute these counterfactuals, we proceed in two steps, as discussed above. We first compute counterfactual outcomes in the final year assuming weather corresponds to the observed 2019 weather, but deficits go to zero. We then compute counterfactuals setting $z'_{i,2019} = z_{i,1991}$ also assuming zero deficits. We compute the treatment effect of the change in z as the difference between equilibrium levels in the two counterfactuals, as discussed in Ossa (2016).

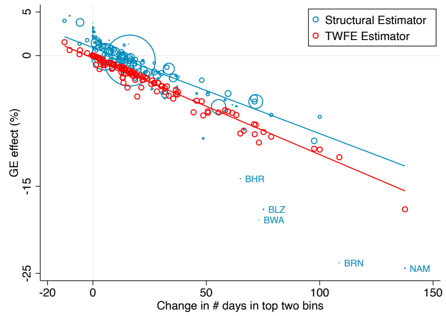
counterfactual no-warming scenario, and cooler colors indicate gains in the observed data relative to the counterfactual.

In panel a, we find that aggregate gross output fell in almost all countries as a result of climate change. The mean (median) loss in aggregate gross output was 0.9% (0.5%), with the largest losses sustained in Belize (8.3%), Laos (5.4%), and Namibia (3.3%). The three outlier countries in which aggregate gross output increased slightly (less than 0.5%) were Canada (0.08%), Ecuador (0.13%), and Eritrea (0.30%). All countries outside of these 6 sustained losses between 0-3%, with slightly larger damages observed in the Tropics and the Southern Hemisphere, where the warming was the most severe. The magnitude and distribution of damages in real wage (panel b) are roughly the same. For comparison, Costinot et al. (2016) compute an aggregate loss of 0.26% in global real wage resulting from expected warming over the 21st century (roughly 2°C). Other studies find slightly larger magnitudes, on the order of 1-3% damages from 1°C warming (Nath, 2020; Dell et al., 2012). So our aggregate estimate is of similar magnitude to previous work, though our modeling assumptions are different.

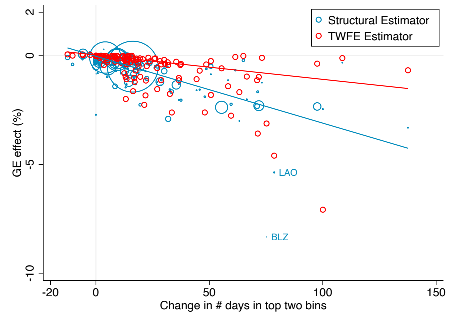
However, when we break down impacts by sector, we observe more heterogeneity. Panel c shows that agricultural gross output *increased* in roughly a third of the countries, while other countries lost up to 15-20% of agricultural gross output (e.g., Namibia and Bahrain). Agricultural losses correlated with warming, for the most part, but many Northern countries that warmed slightly actually increased agricultural gross output. In panel d, we find that most countries lost gross output in manufacturing as a result of the warming, but that countries that warmed a lot (e.g, Namibia, Bahrain, Belize, Congo) experienced both large agricultural losses and modest gains in manufacturing. This pattern indicates that climate-induced losses in agricultural productivity triggered labor reallocation away from agriculture in countries that warmed substantially, thereby raising manufacturing output in these countries, and generating agriculture gains in countries that warmed less, exactly as we found in the simulations. With the effects on agriculture and manufacturing often running in opposite directions, the aggregate effect of warming at the country level is more evenly distributed.

Finally, we compare these structural estimates to effects resulting from reduced-form estimators. In Figure 8, we report estimated effects for agriculture (left column) and

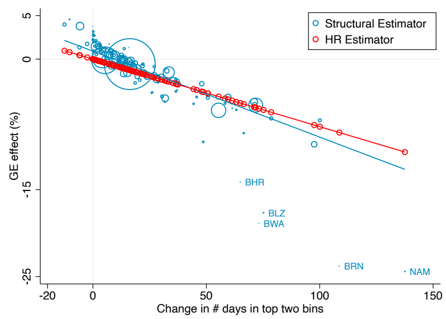
Figure 8: Structural Estimates vs TWFE Estimates



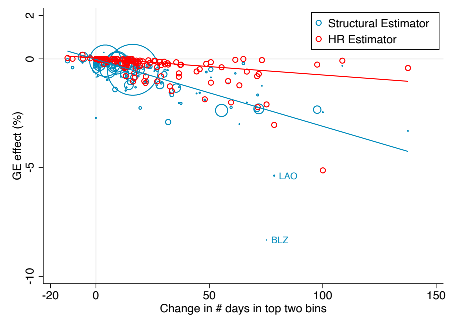
(a) STR vs TWFE, Agriculture



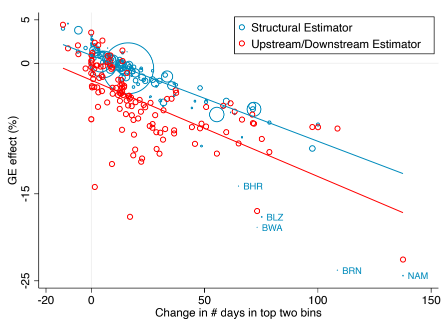
(b) STR vs TWFE, Gross Output



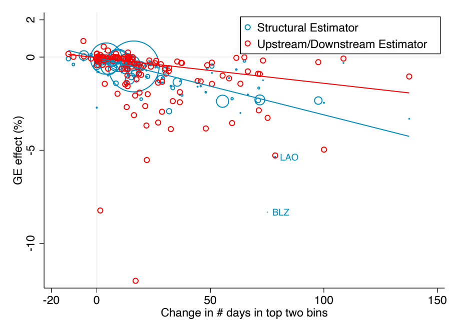
(c) STR vs Het Rob, Agriculture



(d) STR vs Het Rob, Gross Output



(e) STR vs Up/Down, Agriculture



(f) STR vs Up/Down, Gross Output

Notes: Figures plot the % change in 2019 gross output for agricultural (left column), and aggregate gross output (right column) resulting from changes from 1991 weather to 2019 weather (y-axis) against the change in the number of days with maximum temperature above $30°C$ (the top 2 bins) (x-axis) computed using the structural estimator (blue) and the specified reduced-form estimator (red). For the structural estimator, we impose perfect competition, assume a two-sector model, and set $\theta_1 = 8.3$, $\theta_2 = 8.5$. Marker sizes indicate the share of output for each country in total 2019 output. ISO country label reported for select countries.

aggregate gross output (right column) for the structural gravity estimator (blue) and the TWFE estimator (first row), the heterogeneous-robust estimator (second row), and the upstream/downstream estimator (third row), all in red. Estimated treatment effects (RTEs) are plotted on the y-axis in % against the change in the number of days in the year with maximum temperature above 30°C.For each outcome and estimator, we also plot the linear fit between the RTE and the change in the annual number of days with maximum temperature above 30°C. Marker sizes indicate the share of output for each country in total 2019 output. We report ISO country labels for select countries.

Since the estimator from de Chaisemartin et al. (2024b) is only designed for a single continuous treatment, we impose a single treatment variable—the annual number of days with maximum temperature above 30°C—for the heterogeneous robust estimator. As before, we only include estimated effects that are statistically significant. Hence, we compute treatment effects on aggregate gross output by adding counterfactual output in agriculture to observed output in manufacturing and then taking the difference with the observed gross output for the TWFE and the heterogeneous-robust estimators.[39]

In the top row, we find that the TWFE estimator tends to overstate the treatment effects on agricultural output and understate the effects on total gross output, relative to the structural estimator, just as we found in the Monte Carlo simulations. Here, the TWFE actually picks up the slope of the best-fit line quite well in panel a (though not in panel b), meaning that the TWFE recovers relative effects in agriculture, but not the level of effects.[40] For agriculture, there is a slightly positive intercept in the best-fit line for the structural estimator, indicating positive spillover effects from warming, as in the simulations. The population-weighted average estimated RTE computed via the TWFE is -0.44%, 50.1% smaller than the population-weighted average estimate computed using the structural estimator.

In the second row, we find that the heterogeneous robust estimator delivers smaller effects on agriculture relative to the TWFE estimator, which brings the best-fit line closer

[39]Point estimates from the TWFE and the heterogeneous-robust estimators are not statistically significant when estimated for gross output. Hence, if we only consider statistically significant point estimates, the effects on gross output are zero by construction.

[40]Estimated treatment effects for the TWFE do not lie precisely on the best-fit line because we use the both point estimates from Figure 7 to compute effects.

to the best-fit line for the structural estimator. But this implies even smaller estimated effects on gross output relative to the TWFE, resulting in even greater negative bias of the treatment effects on gross output for the heterogeneous robust estimator relative to the TWFE. If the structural model is correctly specified, this result indicates that allowing for heterogeneous effects of own-country warming is not sufficient to recover the full general equilibrium effects of warming. Spillover effects across countries need also be accounted for.

In the last row of Figure 8, we report comparisons of the structural estimator with the upstream/downstream estimator. The upstream/downstream estimator accounts for spillover effects across countries and sectors, but imposes restrictive forms on the spillovers that are not in fact consistent with structural gravity. In the last row of Figure 8, we see that imposing these restrictions generates more heterogeneity in treatment effects relative to the other reduced-form estimators, conditional on warming. But these estimates do not look any closer to the model-consistent estimates in blue. The upstream/downstream estimator appears to overstate the damages on agricultural output, but understate the damages on aggregate output, relative to the structural estimator.[41]

7 Conclusion

In this paper, we show that, for observed shocks that continuously affect all units in the economy, researchers cannot assume away spillover effects as long as regions trade with each other. To estimate the effects of these shocks on endogenous outcomes, assumptions must be made because otherwise the dimensionality of the problem precludes any progress. The canonical TWFE estimator, while a priori relying on SUTVA, thereby ruling out spillovers, actually may recover an approximation of the average slope of the general equilibrium effects, under some conditions.

To make progress in recovering the general equilibrium effects from observed changes

[41] We omit comparisons of the structural estimator to the reduced-form estimator for manufacturing because point estimates from the reduced-form estimators are never distinguishable from zero for the manufacturing sector. Estimated effects are zero by construction, since we only consider statistically significant coefficients.

in temperature between 1991 and 2019, for instance, we use a structural framework relying on quantitative trade theory and gravity estimation to discipline the spillover effects coming from input-output linkages, output market competition, and income effects. In a two-sector model, we find that countries that hardly warmed experienced an increase in their agricultural gross output, a decrease in their manufacturing gross output, and a slight welfare loss in aggregate. By contrast, countries that warmed the most over the period experienced effects in opposite directions, reflecting strong sectoral reallocations. We find large heterogeneities in the general equilibrium effects across countries, even for similar temperature changes. We also find that the TWFE estimates recover neither the relative effects nor the level of the general equilibrium effects, and tend to underestimate these effects.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American economic review*, 91(5), 1369–1401.

Adao, R., Arkolakis, C., & Esposito, F. (2020). General equilibrium effects in space: Theory and measurement. *National Bureau of Economic Research Working Paper*.

Adão, R., Costinot, A., & Donaldson, D. (2024). Putting quantitative models to the test: An application to the us-china trade war. *The Quarterly Journal of Economics*, (pp. qjae041).

Aguiar, A., Narayanan, B., & McDougall, R. (2016). An overview of the gtap 9 data base. *Journal of Global Economic Analysis*, 1(1), 181–208.

Akyapi, B., Bellon, M. M., & Massetti, E. (2022). *Estimating macro-fiscal effects of climate shocks from billions of geospatial weather observations*. Number 2022-2156. International Monetary Fund.

Alves, G., Burton, W. H., & Fleitas, S. (2024). Difference-in-differences in equilibrium: Evidence from place-based policies.

Anderson, J. E., Larch, M., & Yotov, Y. V. (2020). Transitional growth and trade with frictions: A structural estimation framework. *The Economic Journal*, 130(630), 1583–1607.

Angrist, J. D. & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

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

Baier, S. L. & Bergstrand, J. H. (2009). Bonus vetus ols: A simple method for approximating international trade-cost effects using the gravity equation. *Journal of International Economics*, 77(1), 77–85.

Bartelme, D. (2018). Trade costs and economic geography: Evidence from the u.s. *Mimeo*.

Bartelme, D., Lan, T., & Levchenko, A. A. (2020). Specialization, market access and real income. *National Bureau of Economic Research*.

Bilal, A. & Känzig, D. R. (2024). *The Macroeconomic Impact of Climate Change: Global vs. Local Temperature*. Working Paper 32450, National Bureau of Economic Research.

Borusyak, K., Dix-Carneiro, R., & Kovak, B. (2022). Understanding migration responses to local shocks. *Mimeo*.

Broda, C. & Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly journal of economics*, 121(2), 541–585.

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

Butts, K. (2021). Difference-in-differences estimation with spatial spillovers. *arXiv preprint arXiv:2105.03737*.

Caliendo, L. & Parro, F. (2015). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies*, 82(1), 1–44.

CIESIN (2018). Gridded population of the world, version 4.11 (gpwv4): Population count. Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

Conte, M., Cotterlaz, P., Mayer, T., et al. (2022). The cepii gravity database.

Costinot, A., Donaldson, D., & 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.

Costinot, A. & Rodríguez-Clare, A. (2014). Trade theory with numbers: Quantifying the consequences of globalization. In *Handbook of international economics*, volume 4 (pp. 197–261). Elsevier.

Cruz, J.-L. & Rossi-Hansberg, E. (2024). The economic geography of global warming. *Review of Economic Studies*, 91(2), 899–939.

Das, S., Magistretti, G., Pugacheva, E., & Wingender, P. (2022). Sectoral spillovers across space and time. *Journal of Macroeconomics*, 72, 103422.

de Chaisemartin, C., Ciccia, D., D'Haultfœuille, X., & Knau, F. (2024a). Two-way fixed effects and differences-in-differences in heterogeneous adoption designs without stayers. *Available at SSRN 4284811*.

de Chaisemartin, C., D'Haultfœuille, X., Pasquier, F., & Vazquez-Bare, G. (2022). Difference-in-differences for continuous treatments and instruments with stayers. *Available at SSRN 4011782*.

de Chaisemartin, C., D'Haultfœuille, X., & Vazquez-Bare, G. (2024b). Difference-in-difference estimators with continuous treatments and no stayers. 114, 610–613.

De Chaisemartin, C. & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American economic review*, 110(9), 2964–2996.

De Chaisemartin, C. & d'Haultfoeuille, X. (2023). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. *The Econometrics Journal*, 26(3), C1–C30.

Dekle, R., Eaton, J., & Kortum, S. (2007). Unbalanced trade. *American Economic Review*, 97(2), 351–355.

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

Deryugina, T. & Hsiang, S. M. (2014). *Does the environment still matter? Daily temperature and income in the United States*. Technical report, National Bureau of Economic Research.

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

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

Dingel, J. I., Meng, K. C., & Hsiang, S. M. (2019). *Spatial correlation, trade, and inequality: Evidence from the global climate*. Technical report, National Bureau of Economic Research.

Donaldson, D. (2015). The gains from market integration. *Annual Review of Economics*, 7(1), 619–647.

Donaldson, D. & Hornbeck, R. (2016). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2), 799–858.

Eaton, J. & Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5), 1741–1779.

Feng, A., Li, H., & Wang, Y. (2023). *We Are All in the Same Boat: Cross-Border Spillovers of Climate Shocks through International Trade and Supply Chain*. CESifo Working Paper 10402, CESifo.

Fontagné, L., Lebrand, M. S. M., Murray, S., Santoni, G., & Ruta, M. (2023). Trade and infrastructure integration in africa. *Available at SSRN 4672520*.

Gouel, C. & Jean, S. (2023). Love of variety and gains from trade. *European Economic Review*, 158, 104558.

Head, K. & Mayer, T. (2014). Gravity equations: Workhorse,toolkit, and cookbook. In *Handbook of International Economics, Volume 4*.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Sabater, J. M., Nicolas, J., Peubey, C., Radu, R., Rozum, I., et al. (2023). ERA5 hourly data on single levels from 1940 to present.

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

Hsieh, C.-T. & Ossa, R. (2016). A global view of productivity growth in china. *Journal of international Economics*, 102, 209–224.

Hudgens, M. G. & Halloran, M. E. (2008). Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482), 832–842.

Hultgren, A., Carleton, T., Delgado, M., Gergel, D. R., Greenstone, M., Houser, T., Hsiang, S., Jina, A., Kopp, R. E., Malevich, S. B., et al. (2022). Estimating global impacts to agriculture from climate change accounting for adaptation. *Available at SSRN*.

IPCC (2021). *Climate Change 2021: The Physical Science Basis*. Cambridge University Press.

Kolstad, C. D. & Moore, F. C. (2020). Estimating the economic impacts of climate change using weather observations. *Review of Environmental Economics and Policy*.

Leung, M. P. (2020). Treatment and spillover effects under network interference. *Review of Economics and Statistics*, 102(2), 368–380.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3), 531–542.

Mayer, T., Santoni, G., & Vicard, V. (2023). *The CEPII trade and production database*. CEPII.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6), 1695–1725.

Nath, I. B. (2020). *The food problem and the aggregate productivity consequences of climate change*. Technical report, National Bureau of Economic Research.

Nath, I. B., Ramey, V. A., & Klenow, P. J. (2024). How much will global warming cool global growth? *NBER Working Paper No. 32761*.

Newell, R. G., Prest, B. C., & Sexton, S. E. (2021). The gdp-temperature relationship: Implications for climate change damages. *Journal of Environmental Economics and Management*, 108, 102445.

Nordhaus, W. D. (1992). An optimal transition path for controlling greenhouse gases. *The Energy Journal*, 13(1), 1–21.

Osberghaus, D. & Schenker, O. (2022). International trade and the transmission of temperature shocks. *ZEW-Centre for European Economic Research Discussion Paper*.

Ossa, R. (2016). Quantitative models of commercial policy. In *Handbook of commercial policy*, volume 1 (pp. 207–259). Elsevier.

Redding, S. & Venables, A. J. (2004). Economic geography and international inequality. *Journal of international Economics*, 62(1), 53–82.

Rudik, I., Lyn, G., Tan, W., & Ortiz-Bobea, A. (2022). The economic effects of climate change in dynamic spatial equilibrium ivan rudik, gary lyn, weiliang tan, and ariel ortiz-bobea.

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

Schlenker, W. & Roberts, M. J. (2006). Nonlinear effects of weather on corn yields. *Review of agricultural economics*, 28(3), 391–398.

Shapiro, J. & Walker, R. (2018). Why is pollution from us manufacturing declining? the roles of environmental regulation, productivity, and trade. *American Economic Review*, 108(12), 3814–3854.

Shapiro, J. S. (2016). Trade costs, co2, and the environment. *American Economic Journal: Economic Policy*, 8(4), 220–254.

Silva, J. S. & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics*, (pp. 641–658).

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate? causal inference in the face of interference. *Journal of the American Statistical Association*, 101(476), 1398–1407.

Somanathan, E., Somanathan, R., Sudarshan, A., & Tewari, M. (2021). The impact of temperature on productivity and labor supply: Evidence from indian manufacturing. *Journal of Political Economy*, 129(6), 1797–1827.

Vazquez-Bare, G. (2023). Identification and estimation of spillover effects in randomized experiments. *Journal of Econometrics*, 237(1), 105237.

Yotov, Y. (2021). *The variation of gravity within countries*. Technical report, LeBow College of Business, Drexel University.

Zappalà, G. (2024). Estimating sectoral climate impacts in a global production network. *Mimeo*.

Appendix

A Microfoundations

A.1 Details of the General Model

This section presents in further detail the general structural model. The world is divided into N countries, and there are S different sectors of the economy. In each period t, each country n is endowed with a fixed number of worker-consumers L_{nt} that endogenously sort into one of the S sectors. All worker-consumers receive the country-time-specific marginal product of their labor, w_{nt}, for supplying their one unit of labor inelastically, as well as average investment income $\overline{\pi}_{nt}$. The worker-consumers are assumed to be immobile between countries but perfectly mobile between sectors.

Worker-consumers from each country n have Cobb-Douglas preferences over the different sectors $s \in \{1,...,S\}$, spending a share α_{ns}^{C} of their income on goods from sector s, with $\sum_s \alpha_{ns}^{C} = 1$. Aggregate expenditures on final good consumption on goods from sector s is thus $X_{nst}^{F} \equiv \alpha_{ns}^{C} L_{nt} \left(w_{nt} + \overline{\pi}_{nt}\right)$.

Within each sector s, worker-consumers in all countries have identical preferences over a continuum of goods varieties $j \in \Lambda_{nst}$ with a constant elasticity of substitution $\sigma_s > 0$. They purchase goods in amounts $q_{nst}(j)$ to maximize their utility, given by

$$U_{nt} = \prod_{s=1}^{S} \left[\left(\int_{\Lambda_{nst}} q_{nst}(j)^{\frac{\sigma_s - 1}{\sigma_s}} \, dj \right)^{\frac{\sigma_s}{\sigma_s - 1}} \right]^{\alpha_{ns}^{C}} \tag{A.1}$$

The Cobb-Douglas price index of country n at time t is: $p_{nt} = \prod_{s=1}^{S} \left(\frac{p_{nst}}{\alpha_{ns}^{C}} \right)^{\alpha_{ns}^{C}}$, where p_{nst} is the CES price index of sector s in country n at date t $p_{nst} = \left(\int_{\Lambda_{nst}} p_{nst}(j)^{1-\sigma_s} \, dj \right)^{\frac{1}{1-\sigma_s}}$. All worker-consumers in country n at time t receive the same real wage

$$RW_{nt} \equiv \frac{w_{nt} + \overline{\pi}_{nt}}{p_{nt}}, \tag{A.2}$$

which we treat as the metric of welfare.

Producers in each sector s and each country i at date t produce output Q_{ist} using a Cobb-Douglas constant returns to scale technology, requiring labor (hired at price w_{it}) in proportion $\eta_{is} \in [0,1]$ and intermediate inputs in proportion $1-\eta_{is}$. These intermediate inputs combine output from each sector in a Cobb-Douglas fashion (each sector h output being used in proportion α^M_{ish} for the intermediate input of sector s with $\sum_h \alpha^M_{ish} = 1$). We define the aggregate input specific to industry s of country i as:

$$I_{ist} = \left(\frac{L_{ist}}{\eta_{is}}\right)^{\eta_{is}} \left[\prod_{h=1}^{S}\left(\frac{Q^M_{isht}}{\alpha^M_{ish}(1-\eta_{is})}\right)^{\alpha^M_{ish}}\right]^{(1-\eta_{is})} \tag{A.3}$$

where Q^M_{isht} is the quantity of inputs hired by manufacturing firms in sector s from sector h.

Cost minimization implies that the unit costs of the aggregate input can be written as a Cobb–Douglas aggregate of wages and industry price indices, $c_{ist} = w_{it}^{\eta_{is}} \prod_{h=1}^{S} (p_{iht})^{\alpha^M_{ish}(1-\eta_{is})}$, and that expenditures on inputs can be written $c_{ist} I_{ist} = \frac{1}{\eta_{is}} w_{it} L_{ist} = \frac{1}{1-\eta_{is}} X^M_{ist}$, where X^M_{ist} is the expenditure on intermediate goods. Total expenditure on industry s varieties in country i is the sum of final good expenditures and purchases from other firms: $X_{ist} = X^F_{ist} + \sum_{h=1}^{S} \alpha^M_{ihs} X^M_{iht}$.

As in Costinot & Rodríguez-Clare (2014), we allow for different market structures that all generate structural gravity (Head & Mayer, 2014). In particular, we consider two market structures: (1) perfect competition, and (2) monopolistic competition. Denoting X_{nist} as the bilateral export flows of sector s goods from country i to country n at date t, the share of country n's expenditures on sector s spent on goods sourced from country i at date t, π_{nist}, can be written as:

$$\pi_{nist} \equiv \frac{X_{nist}}{X_{nst}} = \frac{S_{ist}\phi_{nist}}{\Phi_{nst}} \quad , \quad \Phi_{nst} = \sum_k S_{kst}\phi_{nkst} \tag{A.4}$$

where S_{ist} is the exporter i's characteristic in sector s representing its capabilities as a supplier to all potential destinations, ϕ_{nist} captures the bilateral accessibility of sector s of country n to exporter i at date t (which depends on directional bilateral trade costs), and Φ_{nst} is the inward multilateral resistance term. The latter term reflects how accessible sector s of country n is for all other exporting countries.

Market clearing then yields the structural gravity equations of Head & Mayer (2014):

$$X_{nist} = \frac{Y_{ist}}{\Omega_{ist}} \frac{X_{nst}}{\Phi_{nst}} \phi_{nist} \tag{A.5}$$

with

$$\Phi_{nst} = \sum_k \frac{Y_{kst}}{\Omega_{kst}} \phi_{nkst} \quad , \quad \Omega_{nst} = \sum_k \frac{X_{kst}}{\Phi_{kst}} \phi_{knst} \tag{A.6}$$

where Y_{ist} denotes country i's income from selling goods of sector s at date t and Ω_{kst} denotes outward multilateral resistance.

Nesting both microfoundations, we can write the importer-sector-specific factor S_{ist}:

$$S_{ist} = A_{ist} w_{it}^{-\theta_s \eta_{is}} \left(\prod_{h=1}^{S} p_{iht}^{\alpha_{ish}^M} \right)^{-\theta_s(1-\eta_{is})} \left(a_0^s A_{ist}^{\theta_s - 1} Y_{ist} w_{it}^{\frac{\eta_{is}\theta_s}{1-\sigma_s}} \left(\prod_{h=1}^{S} p_{iht}^{\alpha_{ish}^M} \right)^{\frac{(1-\eta_{is})\theta_s}{1-\sigma_s}} \right)^{\chi} \tag{A.7}$$

where A_{ist} is a country-sector-time-specific productivity parameter, θ_s is the elasticity of trade flows to trade costs in sector s, and a_0^s is a functions of structural parameters or constants, defined differently depending on the micro-foundation. The parameter $\chi \in \{0, 1\}$ indicates micro-foundations. When $\chi = 0$, the market structure is perfect competition and the model collapses to Caliendo & Parro (2015). When $\chi = 1$, market structure is monopolistic competition and the model collapses to Hsieh & Ossa (2016) and Gouel & Jean (2023).

The different micro-foundations also yield that country i's price index in sector s is related to the country and sector's inward multilateral resistance term Φ_{ist} through the following relationship:

$$p_{ist} = a_1^s \Phi_{ist}^{-\frac{1}{\theta_s}} X_{ist}^{\chi\left(\frac{1}{\theta_s} - \frac{1}{\sigma_s - 1}\right)} \tag{A.8}$$

where a_1^s is a combinations of sector-specific structural parameters, defined differently depending on the micro-foundation considered. In the case of monopolistic competition, imposing zero expected profit yields

$$X_{ist} = \sum_{h=1}^{S} \left(\alpha_{is}^{C}\eta_{ih} + (1-\eta_{ih})\,\alpha_{ihs}^{M}\right) Y_{iht} \tag{A.9}$$

and

$$\frac{L_{ist}}{L_{it}} = \frac{\eta_{is}Y_{ist}}{\sum_h \eta_{ih}Y_{iht}}. \tag{A.10}$$

Finally, we parameterize technology A_{ist} as a function of $\mathscr{V}$ observable factors, as well as unobserved (by the researcher) factors of productivity ψ_{is}, ι_{st} and ω_{ist}:

$$A_{ist} = \left(\frac{\theta_s - 1}{\theta_s}\right)^{\chi} \exp\left(\sum_{v=1}^{\mathscr{V}} \mu_s^v z_{it}^v\right) \exp\left(\psi_{is} + \iota_{st} + \omega_{ist}\right) \tag{A.11}$$

where ψ_{is} reflects time-invariant sector-country-specific base productivity, which is influenced by factors such as longitude, latitude, infrastructure and institutions, ι_{st} reflects time-varying productivity shocks that are common across all countries for a particular sector, and ω_{ist} captures country-sector-time unobserved factors. Parameters μ_s^v govern how observed factors affect local productivity for a given sector.

Finally, we can relate Y_{ist} to observed and unobserved factors, as well as multi-lateral resistance terms:

$$Y_{ist} = \exp\left(\sum_{v=1}^{\mathscr{V}} \beta_{z,v}^{i,s} z_{it}^{v}\right) L_{ist}^{\beta_L^{i,s}} \Omega_{ist}^{\beta_\Omega^{i,s}} \left(\prod_{h=1}^{S} \Phi_{iht}^{\beta_{\Phi_h}^{i,s}}\right) \left(\prod_{h=1}^{S} X_{iht}^{\beta_{X_h}^{i,s}}\right) \exp\left(\delta_{is} + \delta_{st} + \varepsilon_{ist}\right) \tag{A.12}$$

where δ_{is} and δ_{st} are country-sector and sector-time fixed effects, respectively, ε_{ist} is a country-time-specific unobserved factor that combines ω_{ist} and structural parameters. The reduced-form parameters $\{\beta_{z,v}^{i,s}\}_{v=1}^{\mathscr{V}}$, β_L^s, $\beta_\Omega^{i,s}$, $\{\beta_{\Phi_h}^{i,s}\}_{h=1}^{H}$, $\{\beta_{X_h}^{i,s}\}_{h=1}^{H}$ are functions of structural parameters (θ_s, σ_s, μ_s^v, χ, η_{is}, α_{ish}^M, a_{is}^C), their exact expressions depending on the microfoundations considered.

Given the distribution of labor across countries $L_t \equiv \{L_{it}\}_{i=1,\ldots N}$, location-sector-specific productivities $A_t \equiv \{A_{ist}\}_{s=1,\ldots S, i=1,\ldots N}$, and bilateral sectoral trade costs, we define a static equilibrium for each period t as a vector of wages $w_t \equiv \{w_{it}\}_{i=1,\ldots N}$ satisfying equilibrium equations (A.6), (A.9), (A.10), and (A.12).

A.2 Ricardian comparative advantage (Eaton-Kortum)

Within each sector s, there is a set of varieties $j \in \Lambda_s$. Consumers in each country i in year t source variety j from the minimum cost supplier. Firms engage in perfect competition. Hence, there are no profits, so there is no investment income. In this case $X_{ist}^F = \alpha_s^C w_{it} L_{it}$.

With constant returns to scale, the cost of producing a unit of good j in sector s of country i at date t is $\frac{c_{ist}}{\upsilon_{ist}(j)}$ where $\upsilon_{ist}(j)$ denotes country i's efficiency in producing good j of sector s at date t and c_{ist} is given by the Cobb-Douglas combination $w_{it}^{\eta_{is}} \prod_{h=1}^{S} (p_{iht})^{\alpha_{ish}^M (1-\eta_{is})}$. Firms from sector s of country i selling in country n at date t also face iceberg trade costs τ_{nist}. Hence delivering a unit of good j from sector s produced in country i to country n at date t costs

$$p_{nist}(j) = \frac{w_{it}^{\eta_{is}} \left(\prod_{h=1}^{S} p_{iht}^{\alpha_{ish}^M} \right)^{1-\eta_{is}}}{\upsilon_{ist}(j)} \tau_{nist} \tag{A.13}$$

where $\upsilon_{ist}(j)$ is a country-sector-specific productivity, realization of a random variable (drawn independently for each i,s) from country-specific probability Fréchet distribution, with cdf $F_{ist}(\upsilon) = e^{-A_{ist}\mathbf{z}^{-\theta_s}}$. The (sector-country-specific) parameter $A_{ist} > 0$ governs the location of the distribution. The (sector-specific but common to all countries) parameter $\theta_s > 1$ reflects the amount of variation within the distribution.

Assuming perfect competition, $p_{nist}(j)$ is what buyers in country n would pay if they chose to buy good j from sector s of country i at date t. But the price buyers in country n actually pay for a good is the lowest across all sources i. Using the Fréchet distribution, this yields a distribution of prices in sector s of country n with cdf

$$G_{nst}(p) = 1 - e^{-\Phi_{nst} p^{\theta_s}} \tag{A.14}$$

where

$$\Phi_{nst} = \sum_{k=1}^{N} A_{kst} \left(w_{kt}^{\eta_{ks}} \left(\prod_{h=1}^{S} p_{iht}^{\alpha_{ksh}^M} \right)^{(1-\eta_{ks})} \tau_{nkst} \right)^{-\theta_s} \tag{A.15}$$

Using this distribution and substituting (A.13) allows to re-write the CES price index as:

$$p_{nst} = \gamma_s \Phi_{nst}^{-\frac{1}{\theta_s}} \tag{A.16}$$

where $\gamma_s = [\Gamma(\frac{\theta_s+1-\sigma_s}{\theta_s})]^{\frac{1}{1-\sigma_s}}$. We recognize the general model formula $p_{ist} = a_1^s \Phi_{ist}^{-\frac{1}{\theta_s}} X_{ist}^{\chi\left(\frac{1}{\theta_s}-\frac{1}{\sigma_s-1}\right)}$ with $a_1^s = \gamma_s$ and $\chi = 0$.

Moreover, the probability that country i provides a good of sector s at the lowest price in country n can be written as:

$$\pi_{nist} = \frac{A_{ist} w_{it}^{-\theta_s \eta_{is}} \left(\prod_{h=1}^{S} p_{iht}^{\alpha_{ish}^M}\right)^{-\theta_s(1-\eta_{is})} \tau_{nist}^{-\theta_s}}{\Phi_{nst}} \tag{A.17}$$

This probability is also the fraction of sector s goods that country n buys from country i at date t, i.e. $\pi_{nist} = \frac{X_{nist}}{X_{nst}}$. We recognize the gravity equation from the general model (18), where $S_{ist} = w_{it}^{-\theta_s \eta_{is}} \left(\prod_{h=1}^{S} p_{iht}^{\alpha_{ish}^M}\right)^{-\theta_s(1-\eta_{is})} A_{ist} \left(a_0^s A_{ist}^{\theta_s-1} Y_{ist} w_{it}^{\eta_{is}} \left(\prod_{h=1}^{S} p_{iht}^{\alpha_{ish}^M}\right)^{(1-\eta_{is})}\right)^{\chi}$ with $\chi = 0$, and $\phi_{nist} = \tau_{nist}^{-\theta_s}$.

Perfect competition implies that $c_{ist} I_{ist} = Y_{ist} = \frac{1}{\eta_{is}} w_{it} L_{ist} = \frac{1}{1-\eta_{is}} X_{ist}^M$. This allows us to write expenditures :

$$\begin{aligned} X_{ist} &= X_{ist}^F + \sum_{h=1}^{S} \alpha_{ihs}^M X_{iht}^M \\ &= \alpha_{is}^C \sum_{h=1}^{S} \eta_{ih} Y_{iht} + \sum_{h=1}^{S} \alpha_{ihs}^M (1-\eta_{ih}) Y_{iht} \end{aligned} \tag{A.18}$$

Combining the Euler equation from the different sectors, we have

$$\alpha_{ist}^L \equiv \frac{L_{ist}}{L_{it}} = \frac{\eta_{is} Y_{ist}}{\sum_{h=1}^{S} \eta_{ih} Y_{iht}} = \left(1 + \sum_{h \neq s}^{S} \frac{\eta_{ih} Y_{iht}}{\eta_{is} Y_{ist}}\right)^{-1} \tag{A.19}$$

This expression will be used for simulations and counterfactuals.

Combining market clearing with (A.17) and (A.16), we have:

$$Y_{ist} = \eta_{is}^{\frac{-\theta_s \eta_{is}}{1+\theta_s \eta_{is}}} \left(\prod_{h=1}^{S} \gamma_h^{\alpha_{ish}^M} \right)^{\frac{-\theta_s(1-\eta_{is})}{1+\theta_s \eta_{is}}} (A_{ist})^{\frac{1}{1+\theta_s \eta_{is}}} (L_{ist})^{\frac{\theta_s \eta_{is}}{1+\theta_s \eta_{is}}} \Omega_{ist}^{\frac{1}{1+\theta_s \eta_{is}}} \prod_{h=1}^{S} \Phi_{iht}^{\alpha_{ish}^M \frac{\theta_s}{\theta_h} \frac{(1-\eta_{is})}{1+\theta_s \eta_{is}}} \tag{A.20}$$

Substituting (20) (and assuming just one observable factor z^v to keep the equations a bit shorter) into (A.20) leads to to following equation:

$$Y_{ist} = \exp\left(\underbrace{\frac{\mu_s^v}{1+\theta_s \eta_{is}}}_{\equiv \beta_{z,v}^{i,s}} z_{it}^v \right) \left(\prod_{h=1}^{S} \Phi_{iht}^{\overbrace{\alpha_{ish}^M \frac{\theta_s}{\theta_h} \frac{(1-\eta_{is})}{1+\theta_s \eta_{is}}}^{\equiv \beta_{\Phi_h}^{i,s}}} \right) L_{ist}^{\overbrace{\frac{\theta_s \eta_{is}}{1+\theta_s \eta_{is}}}^{\equiv \beta_L^{i,s}}} \Omega_{ist}^{\overbrace{\frac{1}{1+\theta_s \eta_{is}}}^{\equiv \beta_\Omega^{i,s}}}$$

$$\exp\left(\underbrace{\frac{1}{1+\theta_s \eta_{is}} \left(-\theta_s(1-\eta_{is}) \sum_{h=1}^{S} \alpha_{ish}^M \ln \gamma_h - \theta_s \eta_{is} \ln \eta_{is} + \psi_{is} \right)}_{\equiv \delta_{is}} + \underbrace{\frac{1}{1+\theta_s \eta_{is}} \iota_{st}}_{\equiv \delta_{st}} + \underbrace{\frac{1}{1+\theta_s \eta_{is}} \omega_{ist}}_{\equiv \varepsilon_{ist}} \right) \tag{A.21}$$

A.3 Multi-sector model following Hsieh & Ossa (2016)

Worker-Consumers again have Cobb–Douglas preferences across industries and CES preferences across varieties within industries. There are $M_{ni,st}$ firms in industry s from country i serving market n. The consumers' utility function can be written as:

$$U_{nt} = \prod_{s=1}^{S} \left[\sum_{i=1}^{N} \left(\int_0^{M_{ni,st}} q_{nist}(j)^{\frac{\sigma_s - 1}{\sigma_s}} \, dj \right)^{\frac{\sigma_s}{\sigma_s - 1}} \right]^{\alpha_{is}^C} \tag{A.22}$$

Firms produce using inputs as in (**??**). Firm heterogeneity is captured by the following production process: entrants into industry s of country i have to hire f_{ist}^e units of I_{ist} to draw their productivities φ from a Pareto distribution

$$G_{ist}(\varphi) = 1 - \left(\frac{A_{ist}}{\varphi} \right)^{\theta_s}$$

where f_{ist}^e is a fixed cost of entry, A_{is} is the Pareto location parameter, and θ_s is the Pareto shape parameter. Entrants into industry s of country i wishing to sell to country n further need to hire $\frac{x_{nist}\tau_{nist}}{\varphi}$ units of I_{ist} and f_{nist} units of I_{ist} to deliver x_{nist} units of output to country n, where f_{nist} is a fixed cost of serving market n.

All labor and profit income is distributed to households. As a result, households in country i spend

$$X_{ist}^F = \alpha_{is}^C \left(\sum_{h=1}^{S} w_i L_{iht} + M_{iht}^e \bar{\pi}_{iht} \right)$$

on industry s varieties, where $\bar{\pi}_{iht}$ is are the expected profits of the M_{iht}^e entrant firms into industry h of country i. Total expenditure on industry s varieties in country i is defined can be written as

$$X_{ist} = \alpha_{is}^C \left(\sum_{h=1}^{S} w_{it} L_{iht} + M_{iht}^e \bar{\pi}_{iht} \right) + \sum_{h=1}^{S} \alpha_{ihs}^M \frac{1 - \eta_{ih}}{\eta_{ih}} w_{it} L_{iht} \tag{A.23}$$

Profit maximization requires that industry s firms from country i which serve market n

charges

$$p_{nist} = \frac{\sigma_s}{\sigma_s - 1} \frac{c_{ist}\tau_{nist}}{\varphi}.$$

where c_{ist} is given by

$$c_{ist} = w_{it}^{\eta_{is}} \prod_{h=1}^{S} (p_{iht})^{\alpha_{ish}^{M}(1-\eta_{is})} \tag{A.24}$$

However, the fixed market access costs imply that only sufficiently productive firms choose to serve market n. Given that the associated revenues are $r_{nist} = \left(\frac{\sigma_s}{\sigma_s-1}\frac{c_{ist}\tau_{nist}}{\varphi P_{nst}}\right)^{1-\sigma_s} X_{nst}$, the associated variable profits are $\pi_{nist}^{v} = \frac{1}{\sigma_s}\left(\frac{\sigma_s}{\sigma_s-1}\frac{c_{ist}\tau_{nist}}{\varphi P_{nst}}\right)^{1-\sigma_s} X_{nst}$ which only exceed the fixed market access costs $c_{ist}f_{nist}$ if $\varphi > \varphi_{nis}^{*}$, where

$$\varphi_{nis}^{*} = \frac{\sigma_s}{\sigma_s - 1} \frac{c_{ist}\tau_{ist}}{P_{nst}} \left(\frac{\sigma_s c_{ist} f_{nist}}{X_{nst}}\right)^{\frac{1}{\sigma_s - 1}}.$$

The ideal price indices are given by $P_{nst} = \left(\sum_{i=1}^{N} M_{nist}\, p_{nist}(\tilde{\varphi}_{nist})^{1-\sigma_s}\right)^{\frac{1}{1-\sigma_s}}$, where

$$\tilde{\varphi}_{nist} = \left(\int_{\varphi_{nist}^{*}}^{+\infty} \varphi^{\sigma_s - 1} dG_{ist}(\varphi|\varphi > \varphi_{nist}^{*})\right)^{\frac{1}{\sigma_s - 1}}$$

is an average productivity measure. Using the Pareto distribution, we get that

$$\tilde{\varphi}_{nist} = \left(\frac{\theta_s}{\theta_s - \sigma_s + 1}\right)^{\frac{1}{\sigma_s - 1}} \varphi_{nist}^{*}$$

The Pareto assumption also implies that the probability of drawing a productivity above the cutoff is given by $Prob(\varphi > \varphi_{nist}^{*}) = \left(\frac{A_{ist}}{\varphi_{nist}^{*}}\right)^{\theta_s}$ so that the relationship between the eventual number of firms and the initial number of entrants is simply $M_{nist} = \left(\frac{A_{ist}}{\varphi_{nist}^{*}}\right)^{\theta_s} M_{ist}^{e}$. This relationship can be used together with the pricing formula and the definitions of $\tilde{\varphi}_{nist}$ and φ_{nist}^{*} to rewrite P_{nst} as:

$$
\begin{aligned}
P_{nst} &= \left(\sum_{i=1}^{N} \frac{\theta_s}{\theta_s-\sigma_s+1} M_{ist}^{e} \left(\frac{\sigma_s}{\sigma_s-1}\frac{c_{ist}\tau_{nist}}{A_{ist}}\right)^{-\theta_s} \left(\frac{\sigma_s c_{ist} f_{nist}}{X_{nst}}\right)^{\frac{\sigma_s-\theta_s-1}{\sigma_s-1}}\right)^{-\frac{1}{\theta_s}} \\
&= \underbrace{\left(\frac{\theta_s}{\theta_s-\sigma_s+1}\right)^{-\frac{1}{\theta_s}} \left(\frac{\sigma_s}{\sigma_s-1}\right) \sigma_s^{\frac{1}{\sigma_s-1}-\frac{1}{\theta_s}}}_{\equiv a_1^s} \left(\underbrace{\sum_{i=1}^{N} M_{ist}^{e} \left(\frac{c_{ist}\tau_{nist}}{A_{ist}}\right)^{-\theta_s} \left(c_{ist} f_{nist}\right)^{\frac{\sigma_s-\theta_s-1}{\sigma_s-1}}}_{\equiv \Phi_{nst}}\right)^{-\frac{1}{\theta_s}} \\
&\times \quad \left(X_{nst}\right)^{\frac{1}{\theta_s}-\frac{1}{\sigma_s-1}} \qquad \text{(A.25)}
\end{aligned}
$$

The expected profits of an entrant into industry s of country i are

$$
\bar{\pi}_{ist} = \sum_{n=1}^{N} Prob(\varphi > \varphi_{nist}^{*}) \left[E(\pi_{nist}^{v} | \varphi > \varphi_{nist}^{*}) - c_{ist} f_{nist}\right] - c_{ist} f_{ist}^{e}.
$$

From the definition of π_{nist}^{v}, we can derive that $E(\pi_{nist}^{v}|\varphi > \varphi_{nist}^{*}) = \frac{1}{\sigma_s}\left(\frac{\sigma_s}{\sigma_s-1}\frac{c_{ist}\tau_{nist}}{\tilde{\varphi}_{nist}P_{nst}}\right)^{1-\sigma_s} X_{nst}$. Combining with the expression of $Prob(\varphi > \varphi_{nist}^{*})$ and the definitions of $\tilde{\varphi}_{nist}$ and φ_{nist}^{*}, we get:

$$
\bar{\pi}_{ist} = \sum_{n=1}^{N} \frac{\sigma_s - 1}{\sigma_s \theta_s} \frac{\left(c_{ist} f_{nist}\right)^{\frac{\sigma_s-\theta_s-1}{\sigma_s-1}} \left(\frac{c_{ist}\tau_{nist}}{A_{ist}}\right)^{-\theta_s}}{\sum_{k=1}^{N} M_{kst}^{e} \left(c_{kst} f_{nkst}\right)^{\frac{\sigma_s-\theta_s-1}{\sigma_s-1}} \left(\frac{c_{kst}\tau_{nkst}}{A_{kst}}\right)^{-\theta_s}} X_{nst} - c_{ist} f_{ist}^{e} \qquad \text{(A.26)}
$$

Input market clearing requires that

$$
c_{ist} I_{ist} = \underbrace{M_{ist}^{e} c_{ist} f_{ist}^{e}}_{\text{entry costs}} + \underbrace{M_{ist}^{e} c_{ist} E(i_{ist}^{v})}_{\text{production costs}} + \underbrace{\sum_{j=1}^{N} M_{ijst} c_{jst} f_{ijst}}_{\text{market access costs}},
$$

where $E(i_{ist}^{v})$ denotes the expected demand for inputs used directly in production. Using the fact that $c_{ist} E(i_{ist}^{v}) = \theta_s(\bar{\pi}_{ist} + c_{ist} f_{ist}^{e})$ and $\sum_{j=1}^{N} M_{ijst} c_{jst} f_{ijst} = \frac{\theta_s-\sigma_s+1}{\theta_s\sigma_s} X_{ist}$, we can substitute these terms into the input market clearing condition and solve for M^e_{ist}:

$$M^e_{ist} = \frac{\frac{1}{\eta_{ist}} w_{it} L_{ist} - \frac{\theta_s - \sigma_s + 1}{\theta_s \sigma_s} X_{ist}}{\theta_s \bar{\pi}_{ist} + c_{ist}(\theta_s + 1) f^e_{ist}} \tag{A.27}$$

Adding finally a labor market clearing condition :

$$L_{it} = \sum_{s=1}^{S} L_{ist} \tag{A.28}$$

We follow the long-term setting of Hsieh & Ossa (2016) where profits are null ($\bar{\pi}_{ist} = 0$) and M^e_{ist} is endogenous (free entry). We then have (A.23), (A.24), (A.25), (A.26), (A.27), (A.28) representing a system of $5NS + N$ equations of $5NS + N$ unknowns $\{X_{ist}, c_{ist}, P_{ist}, L_{ist}, M^e_{ist}, w_{it}\}$.

From (A.25) we see

$$P_{nst} = a^s_1 \Phi_{nst}^{-\frac{1}{\theta_s}} (X_{nst})^{\frac{1}{\theta_s} - \frac{1}{\sigma_s - 1}} \tag{A.29}$$

and

$$\Phi_{nst} = \sum_{i=1}^{N} M^e_{ist} \left(\frac{c_{ist} \tau_{nist}}{A_{ist}} \right)^{-\theta_s} (c_{ist} f_{nist})^{\frac{\sigma_s - \theta_s - 1}{\sigma_s - 1}} \tag{A.30}$$

$$\Phi_{nst} = \sum_{i=1}^{N} \frac{\sigma_s - 1}{\theta_s \sigma_s f^e_{is}} Y_{ist} A^{\theta_s}_{ist} c_{ist}^{-\theta_s + \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} - 1} \overbrace{\tau_{nist}^{-\theta_s} f_{nist}^{\frac{\sigma_s - \theta_s - 1}{\sigma_s - 1}}}^{\phi_{nist}} \tag{A.31}$$

We can solve for sector-specific labor

$$\frac{L_{ist}}{L_{it}} = \left(1 + \frac{\sum_{h \neq s} \frac{\eta_{ih}}{\eta_{is}} Y_{iht}}{Y_{ist}} \right)^{-1} \tag{A.32}$$

This equation will be used for simulations and conterfactuals.

The expenditure of country n on sector s goods sourced from country i is $X_{nist} = M_{nist} P_{nist} x_{nist}$. Using the expression of revenues r_{nis}, the definitions of $\tilde{\varphi}_{nist}$ and φ^*_{nist} and

(A.30), we get the following expression for trade shares:

$$\pi_{nist} = \frac{X_{nist}}{X_{nst}} = \frac{M^e_{ist} A^{\theta_s}_{ist} c_{ist}^{-\theta_s + \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1}} \overbrace{\tau_{nist}^{-\theta_s} f_{nist}^{\frac{\sigma_s - \theta_s - 1}{\sigma_s - 1}}}^{\phi_{nist}}}{\Phi_{nst}} \tag{A.33}$$

Combining this with the identity $Y_{ist} = \sum_{n=1}^{N} X_{nits} = \sum_{n=1}^{N} \pi_{nist} X_{nst}$ and (A.33), we get:

$$Y_{ist} = M^e_{ist} A^{\theta_s}_{ist} w_{it}^{\left(-\theta_s + \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1}\right)\eta_{is}} \left(\prod_h p_{iht}^{\alpha^M_{ish}}\right)^{\left(-\theta_s + \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1}\right)(1-\eta_{is})} \underbrace{\sum_n \frac{\phi_{nist}}{\Phi_{nst}} X_{nst}}_{\Omega_{ist}} \tag{A.34}$$

We want to get rid of M^e_{ist}. From (A.27), with zero profits, we get

$$M^e_{ist} = \frac{\frac{1}{\eta_{is}} w_{it} L_{ist} - \frac{\theta_s - \sigma_s + 1}{\theta_s \sigma_s} X_{ist}}{c_{ist}(\theta_s + 1) f^e_{ist}} \tag{A.35}$$

And

$$c_{ist} I_{ist} = Y_{ist} \tag{A.36}$$

which together yields

$$M^e_{ist} = \frac{\sigma_s - 1}{\theta_s \sigma_s} \frac{Y_{ist}}{c_{ist} f^e_{ist}} \tag{A.37}$$

Substituting (A.37), we have

$$1 = \frac{\sigma_s - 1}{\sigma_s \theta_s f^e_{is}} A^{\theta_s}_{ist} w_{it}^{\left(-\theta_s + \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} - 1\right)\eta_{is}} \left(\prod_h p_{iht}^{\alpha^M_{ish}}\right)^{\left(-\theta_s + \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} - 1\right)(1-\eta_{is})} \Omega_{ist} \tag{A.38}$$

Moving wages to the LHS

$$w_{it} = \left[\frac{\sigma_s - 1}{\sigma_s \theta_s f^e_{is}}\right]^{\frac{1}{\left(\theta_s - \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} + 1\right)\eta_{is}}} A_{ist}^{\frac{\theta_s}{\left(\theta_s - \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} + 1\right)\eta_{is}}} \left(\prod_h p_{iht}^{\alpha^M_{ish}}\right)^{\frac{-(1-\eta_{is})}{\eta_{is}}} \Omega_{ist}^{\frac{1}{\left(\theta_s - \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} + 1\right)\eta_{is}}} \tag{A.39}$$

now using cost minimization

$$Y_{ist} = \frac{1}{\eta_{is}} L_{ist} \left[\frac{\sigma_s - 1}{\sigma_s \theta_s f_{is}^e}\right]^{\frac{1}{\left(\theta_s - \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} + 1\right)\eta_{is}}} A_{ist}^{\frac{\theta_s}{\left(\theta_s - \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} + 1\right)\eta_{is}}} \left(\prod_h p_{iht}^{\alpha_{ish}^M}\right)^{\frac{-(1-\eta_{is})}{\eta_{is}}} \Omega_{ist}^{\frac{1}{\left(\theta_s - \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} + 1\right)\eta_{is}}} \tag{A.40}$$

Substituting with Φs and Productivity

$$Y_{ist} = Const_1 \times L_{ist} A_{ist}^{\frac{\theta_s}{\left(\theta_s - \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} + 1\right)\eta_{is}}} \left(\prod_h \Phi_{iht}^{\frac{\alpha_{ish}^M}{\theta_h}}\right)^{\frac{(1-\eta_{is})}{\eta_{is}}} \left(\prod_h X_{iht}^{\alpha_{ish}^M\left(\frac{1}{\theta_h} - \frac{1}{\sigma_h - 1}\right)}\right)^{\frac{-(1-\eta_{is})}{\eta_{is}}} \Omega_{ist}^{\frac{1}{\left(\theta_s - \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} + 1\right)\eta_{is}}} \tag{A.41}$$

where

$$Const_1 = \frac{1}{\eta_{is}} \left[\frac{(\sigma_s - 1)}{\theta_s \sigma_s f_{is}^e}\right]^{\frac{1}{\left(\theta_s - \frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} + 1\right)\eta_{is}}} \prod_h (a_1^h)^{-\alpha_{ish}^M \frac{1-\eta_{is}}{\eta_{is}}}$$

Let us consider for simplicity that the sector-country-specific productivity A_{ist} is a linear function of one observable z_{it}^v:

$$A_{ist} \equiv \frac{\theta_s - 1}{\theta_s} \exp(\mu_s^v z_{it}^v) \exp(\psi_{is} + \iota_{st} + \omega_{ist}) \tag{A.42}$$

Then, substituting in (A.41), we get (gathering also the fixed effects on the constant term):

$$Y_{ist} = Const_2 \times L_{ist} \exp\left(\frac{(\sigma_s - 1)\,\mu_s^v}{\sigma_s \eta_{is}} z_{it}^v\right) \left(\prod_h \Phi_{iht}^{\frac{\alpha_{ish}^M}{\theta_h}}\right)^{\frac{(1-\eta_{is})}{\eta_{is}}} \left(\prod_h X_{iht}^{\alpha_{ish}^M\left(\frac{1}{\theta_h} - \frac{1}{\sigma_h - 1}\right)}\right)^{\frac{-(1-\eta_{is})}{\eta_{is}}} \tag{A.43}$$

$$\times \quad \Omega_{ist}^{\frac{\sigma_s - 1}{\theta_s \sigma_s \eta_{is}}} \exp\left(\left(\frac{\sigma_s - 1}{\sigma_s \eta_{is}}\right)(\psi_{is} + \iota_{st} + \omega_{ist})\right) \tag{A.44}$$

where

$$Const_2 = \left(\frac{\theta_s - 1}{\theta_s}\right)^{\frac{\sigma_s - 1}{\sigma_s \eta_{is}}} \frac{1}{\eta_{is}} \Big[\underbrace{\frac{(\sigma_s - 1)}{\theta_s \sigma_s f_{is}^e}}_{\equiv a_0^s} \Big]^{\frac{\sigma_s - 1}{\theta_s \sigma_s \eta_{is}}} \prod_h (a_1^h)^{-\alpha_{ish}^M \frac{1-\eta_{is}}{\eta_{is}}}$$

A.4 Proof for the conditions of identification of the reduced-form approaches

Assuming a single sector, a single observable factor, two periods, perfect competition, and no deficits, a first-order approximation of the equilibrium conditions of the model can be written:

$$\begin{aligned} \Delta \ln Y &= \text{diag}(\beta_z)\Delta z + \text{diag}(\beta_L)\Delta L + \text{diag}(\beta_\Phi)\Delta \ln \Phi + \text{diag}(\beta_\Omega)\Delta \ln \Omega + \Delta\xi && \text{(A.45)} \\ \Delta \ln \Phi &= \Pi\left(\Delta \ln Y + \text{diag}(\beta_\Phi)\Delta \ln \Phi + (\text{diag}(\beta_\Omega) - I)\Delta \ln \Omega\right) && \text{(A.46)} \\ \Delta \ln \Omega &= \Gamma\left(\Delta \ln Y + (\text{diag}(\beta_\Phi) - I)\Delta \ln \Phi + \text{diag}(\beta_\Omega)\Delta \ln \Omega\right) && \text{(A.47)} \end{aligned}$$

where $\text{diag}(\beta_a)$ denotes the N-by-N matrix with diagonal elements β_a^i for all i and off-diagonal elements 0, and where Π and Γ are the N-by-N matrices of import shares and export shares in period 0, respectively.

Solving the system for $\Delta \ln Y$, we have

$$\Delta \ln Y = B\left(\text{diag}(\beta_z)\Delta z + \text{diag}(\beta_L)\Delta L + \Delta\xi\right) \tag{A.48}$$

where

$$\begin{aligned} B &= \left[I - \left(\text{diag}(\beta_\Phi)\Pi + (\text{diag}(\beta_\Omega) - I)(\text{diag}(\beta_\Phi) - I)\Pi(I - \text{diag}(\beta_\Omega)\Gamma)^{-1}\Gamma\right)\right]^{-1} \\ &\times \left[\Pi + (\text{diag}(\beta_\Omega) - I)\Pi(I - \text{diag}(\beta_\Omega)\Gamma)^{-1}\Gamma\right] && \text{(A.49)} \end{aligned}$$

with individual elements, for $i = m_0$,

$$\begin{aligned} b_{ij} \;\; = \;\; & \sum_{K=0}^{\infty} \sum_{m_1,\ldots,m_K} \prod_{\ell=1}^{K} \Bigg(\beta_{\Phi}^{m_{\ell-1}} \pi_{m_{\ell-1},m_\ell} + (\beta_{\Omega}^{m_{\ell-1}} - 1)(\beta_{\Phi}^{m_{\ell-1}} - 1) \\ & \times \sum_r \pi_{m_{\ell-1},r} \sum_{d=0}^{\infty} \sum_{s_1,\ldots,s_d} (\beta_{\Omega}^{m_{\ell-1}})^d \gamma_{rs_1}\gamma_{s_1 s_2}\ldots\gamma_{s_d,m_\ell} \Bigg) \\ & \times \left(\pi_{m_K,j} + (\beta_{\Omega}^{m_K} - 1) \sum_r \pi_{m_K,r} \sum_{d=0}^{\infty} \sum_{s_1,\ldots,s_d} (\beta_{\Omega}^{m_K})^d \gamma_{rs_1}\gamma_{s_1 s_2}\ldots\gamma_{s_d,j} \right). \end{aligned} \tag{A.50}$$

If we assume constant trade shares $\pi_{ij} = \gamma_{ij} = \frac{1}{N}$ and constant labor shares across units, where the latter restriction implies $\beta_{\Phi}^{i} = \frac{1-\eta}{1+\theta\eta}$ and $\beta_{\Omega}^{i} = \frac{1}{1+\theta\eta}$, the expression b_{ij} simplifies to

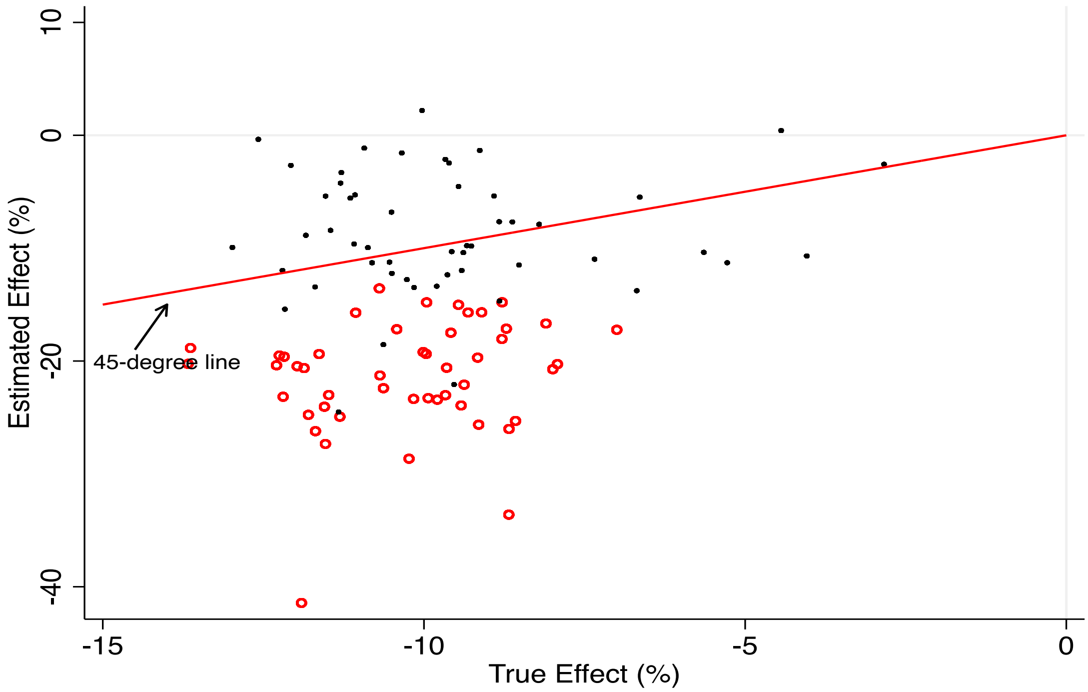
$$b_{ij} = \sum_{K=0}^{\infty} \sum_{m_1,\ldots,m_K} \prod_{\ell=1}^{K} \left(\frac{1-\eta}{1+\theta\eta} \cdot \frac{1}{N} + \frac{\theta(1+\theta)\eta^2}{(1+\theta\eta)\left((1+\theta\eta)N-1\right)} \right)$$

$$\times \left(\frac{1}{N} - \frac{\theta\eta}{1+\theta\eta} \cdot \frac{(1+\theta\eta)}{(1+\theta\eta)N-1} \right).$$

which is a constant. Furthermore, imposing constant labor shares across units also implies that $\beta_z^i = \mu/(1+\theta\eta)$, a constant. As a result, $b_{ij}\beta_z^j$ is identical for all i and j.

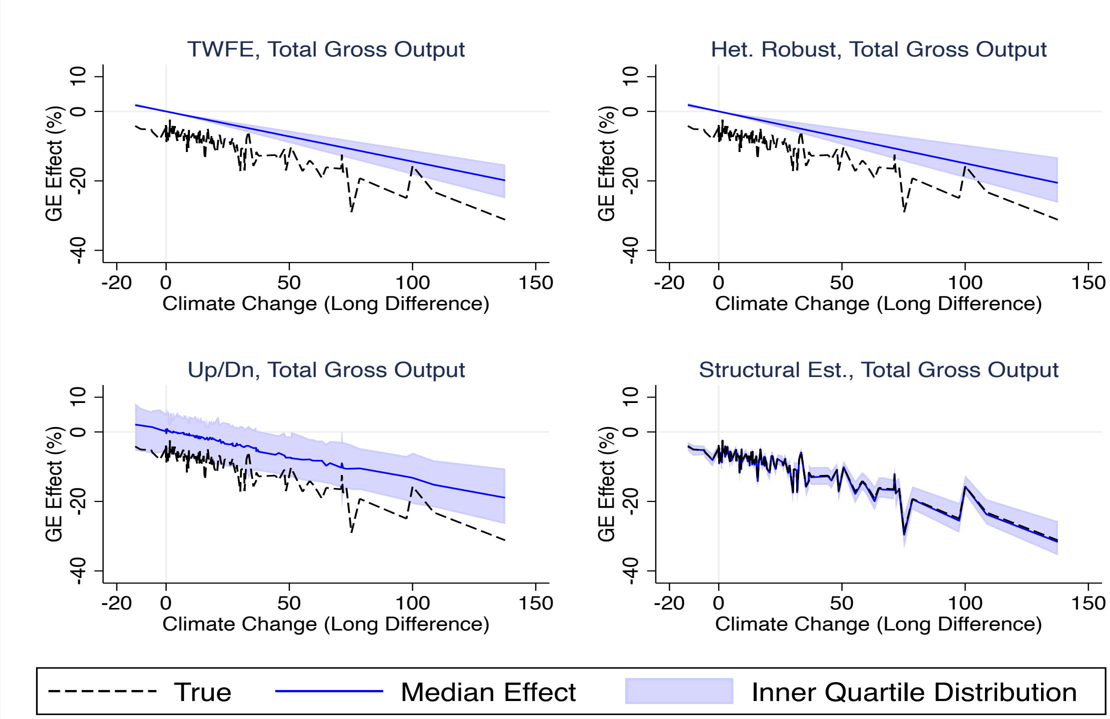
B Further Results

Figure B.1: Estimated vs True Treatment Effects, Global Long Panel



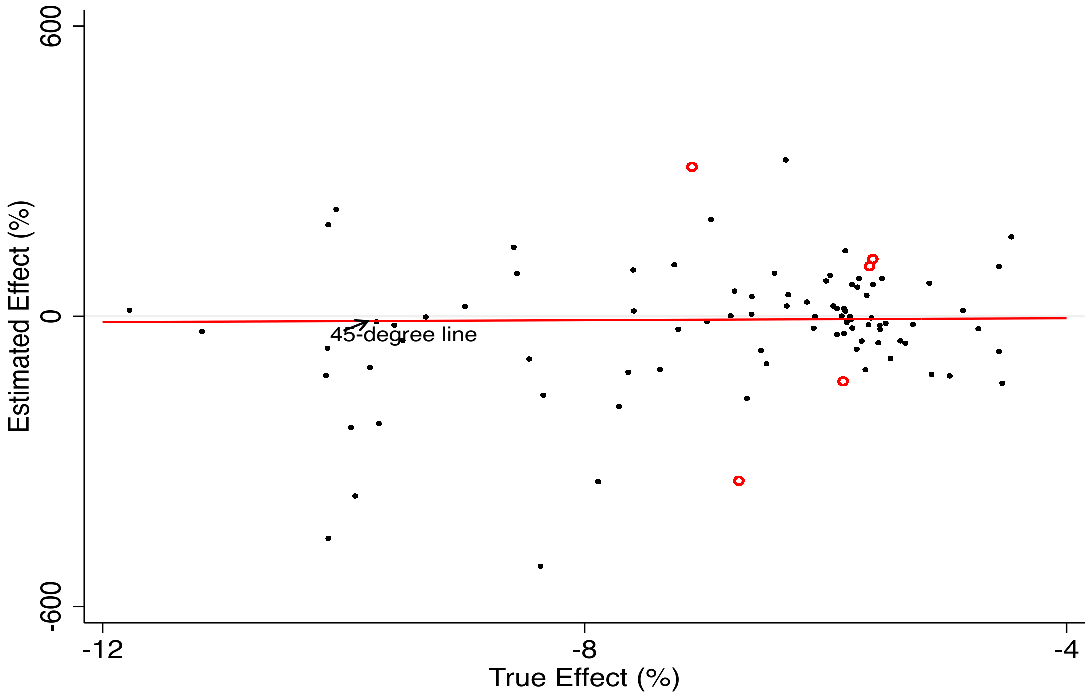
Notes: Figure plots the estimates treatment effect on global gross output in % on the y-axis against the true treatment effect on global gross output in % on the x-axis, where the estimates are computed via the local projection approach from Bilal & Känzig (2024). The 45-degree line is plotted in red. Red dots indicate replications for which the point estimate on contemporaneous temperature shocks are statistically significant at the 5% level. Simulations include 132 countries and 58 time periods. To increase the length of the panel from 29 to 58 periods, we split each year into two periods, each with the same weather realizations, but with different productivity draws. Data is aggregated to a single sector. Market structure is perfect competition. Parameters are set at $\theta = 3.3$, $\sigma = 3.6$, $\mu = -.003$.

Figure B.2: Estimated vs True Treatment Effects, Single Sector, Monopolistic Competition



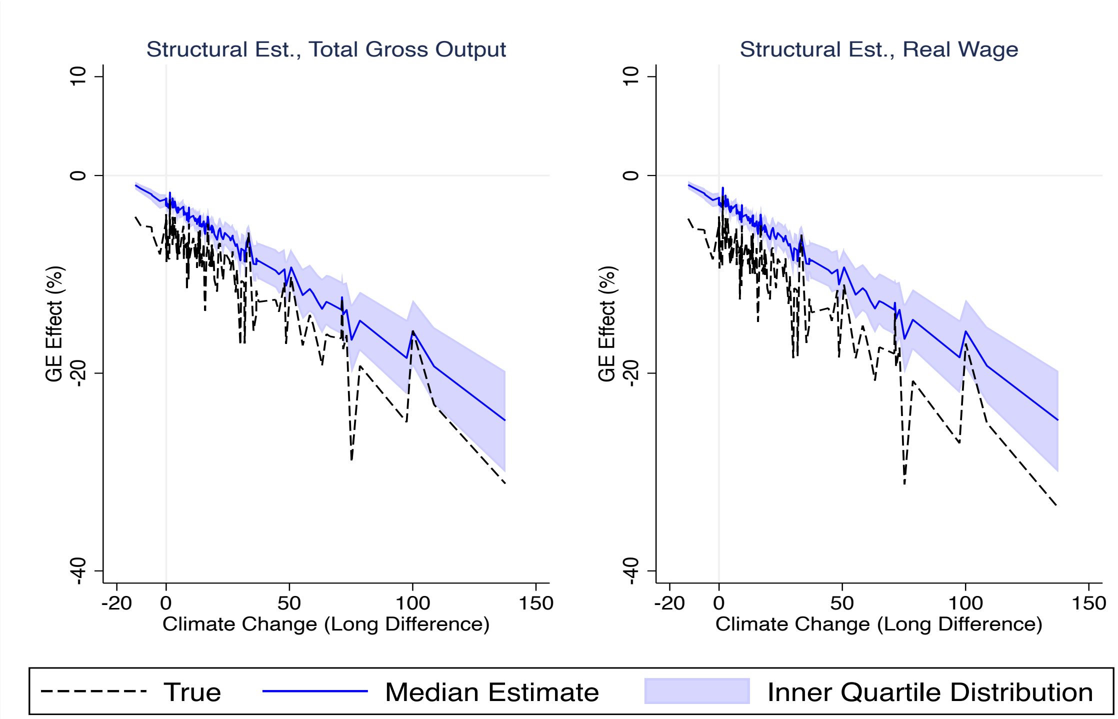
Notes: Figure plots the treatment effect in % on the y-axis against the change in the # of days with maximum temperature above 30 degrees between 1991 and 2019. The black dashed line plots the ETE effect on gross output in 2019 across 100 replications, i.e. the average true effect computed across replications. The blue line presents the median estimate from the indicated estimator, and the blue shaded region depicts the interquartile range of the estimates across 100 replications. Simulations include 132 countries and 29 time periods. Data is aggregated to a single sector. Market structure is monopolistic competition. Parameters are set at $\theta = 3.3$, $\sigma = 3.6$, $\mu = -.003$.

Figure B.3: Estimated vs True Treatment Effects, Global, Monopolistic Competition



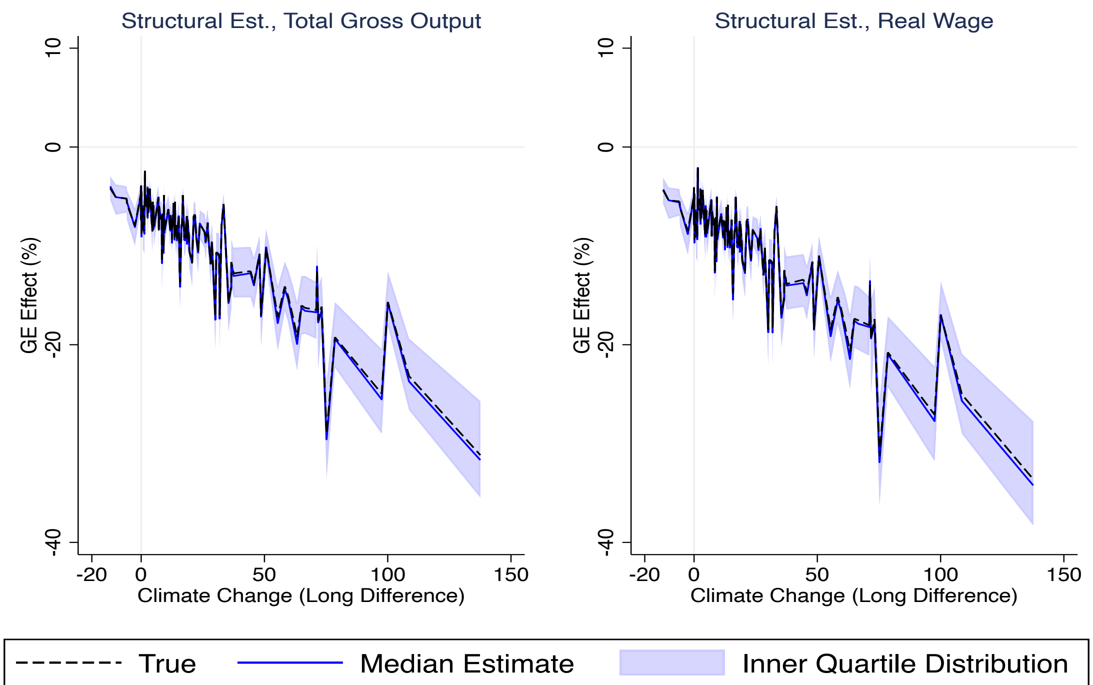
Notes: Figure plots the estimates treatment effect on global gross output in % on the y-axis against the true treatment effect on global gross output in % on the x-axis, where the estimates are computed via the local projection approach from Bilal & Känzig (2024). The 45-degree line is plotted in red. Red dots indicate replications for which the point estimate on contemporaneous temperature shocks are statistically significant at the 5% level. Simulations include 132 countries and 29 time periods. Data is aggregated to a single sector. Market structure is monopolistic competition. Parameters are set at $\theta = 3.3$, $\sigma = 3.6$, $\mu = -.003$.

Figure B.4: Estimated vs True Treatment Effects, Structural Estimator, Single Sector, Monopolistic Competition



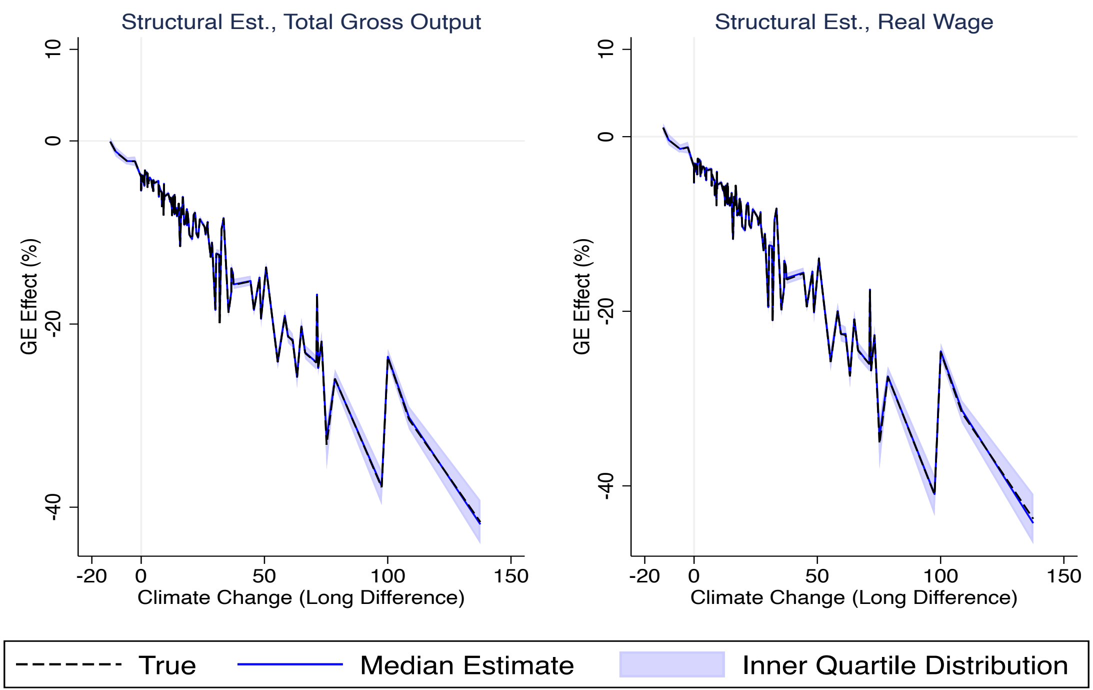
Notes: Figure plots the treatment effect in % on the y-axis against the change in the # of days with maximum temperature above 35 degrees between 1991 and 2019. The black dashed line plots the ETE effect on gross output in 2019 across 100 replications. The blue solid line presents the median estimate from the structural estimator, imposing perfect competition, and the blue shaded region depicts the interquartile range of the distribution of these estimates. Simulations include 132 countries and 29 time periods. Data is aggregated to a single sector. Market structure is monopolistic competition. Parameters are set at $\theta = 3.3$, $\sigma = 3.6$, $\mu = -.003$.

Figure B.5: Estimated vs True Treatment Effects, Structural Estimator, Single Sector, Monopolistic Competition



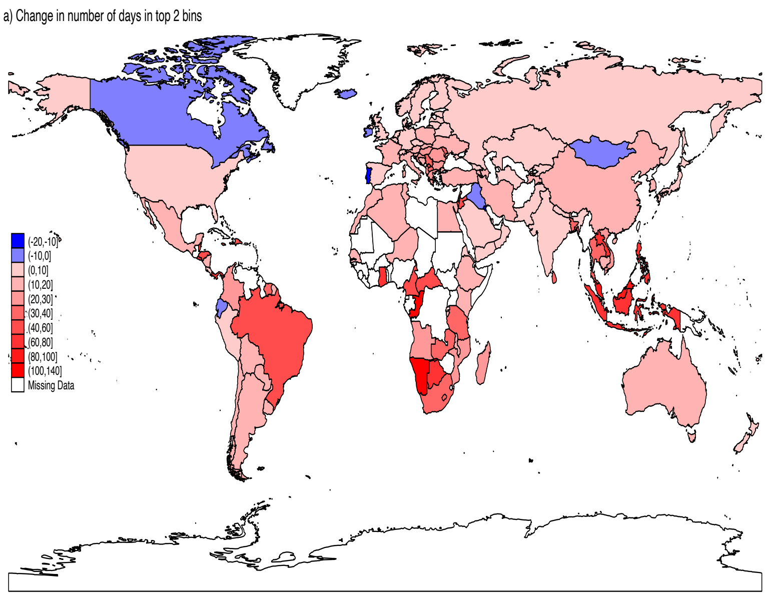
Notes: Figure plots the treatment effect in % on the y-axis against the change in the # of days with maximum temperature above 35 degrees between 1991 and 2019. The black dashed line plots the ETE effect on gross output in 2019 across 100 replications. The blue solid line presents the median estimate from the structural estimator, imposing monopolistic competition, and the blue shaded region depicts the interquartile range of the distribution of these estimates. Simulations include 132 countries and 29 time periods. Data is aggregated to a single sector. Market structure is monopolistic competition. Parameters are set at $\theta = 3.3$, $\sigma = 3.6$, $\mu = -.003$.

Figure B.6: Estimated vs True Treatment Effects, Structural Estimator, Single Sector



Notes: Figure plots the treatment effect in % on the y-axis against the change in the # of days with maximum temperature above 35 degrees between 1991 and 2019. The black dashed line plots the ETE effect on gross output in 2019 across 100 replications. The blue solid line presents the median estimate from the structural estimator, and the blue shaded region depicts the interquartile range of the distribution of these estimates. Simulations include 132 countries and 29 time periods. Data is aggregated to a single sector. Market structure is perfect competition. Parameters are set at $\theta = 3.3$, $\sigma = 3.6$, $\mu = -.003$.

Figure B.7: Estimated vs True Treatment Effects



Notes:

C Data

Table C.1: Aggregation of GTAP commodities into sectors

GTAP Sector	Aggregated Sector
Paddy rice, Wheat, Cereal grains nec	Agriculture
Vegetables, fruit, nuts	Agriculture
Oil seeds	Agriculture
Sugar cane, sugar beet	Agriculture
Plant-based fibers	Agriculture
Crops nec	Agriculture
Bovine cattle, sheep, goats, horses	Agriculture
Animal products nec	Agriculture
Raw milk	Agriculture
Wool, silk-worm cocoons	Agriculture
Forestry, Fishing	Agriculture
Coal, Oil, Gas, Minerals nec	Manufacturing
Meat: cattle, sheep, goats, horse	Manufacturing
Meat products nec	Manufacturing
Vegetable oils and fats	Manufacturing
Dairy products, Processed rice, Sugar	Manufacturing
Food products nec	Manufacturing
Beverages and tobacco products	Manufacturing
Textiles, Wearing apparel, Leather products	Manufacturing
Wood products	Manufacturing
Paper products, publishing	Manufacturing
Petroleum, coal products	Manufacturing
Chemical, rubber, plastic products	Manufacturing
Mineral products nec	Manufacturing
Ferrous metals, Metals nec, Metal products	Manufacturing
Motor vehicles and parts, Transport equipment nec	Manufacturing
Electronic equipment, Machinery and equipment nec	Manufacturing
Manufactures nec	Manufacturing
Electricity, Gas manufacture, distribution, Water	Other
Construction, Trade	Other
Transport nec, Sea transport, Air transport	Other
Communication, Financial services nec	Other
Insurance, Business services nec	Other
Recreation and other services	Other
Public Administration, Defense, Education, Health	Other
Dwellings	Other

Table C.2: Attribution of aggregate regions values to missing countries in GTAP

individual country	GTAP 9 Composite Region
Barbados	Rest of Caribbean
Central African Republic	Rest of Central Africa
Republic of the Congo	Rest of Central Africa
Belize	Rest of Central America
Burundi	Rest of Eastern Africa
Eritrea	Rest of Eastern Africa
Moldova	Rest of Eastern Europe
Bosnia and Herzegovina	Rest of Europe
Iceland	Rest of Europe
North Macedonia	Rest of Europe
Serbia	Rest of Europe
Tajikistan	Rest of Former Soviet Union
Turkmenistan	Rest of Former Soviet Union
Iraq	Rest of Western Asia
Lebanon	Rest of Western Asia
Palestine	Rest of Western Asia
Yemen	Rest of Western Asia
Algeria	Rest of North Africa
Fiji	Rest of Oceania
Suriname	Rest of South America
Maldives	Rest of South Asia
Angola	South Central Africa
Cape Verde	Rest of Western Africa
Gambia	Rest of Western Africa
Niger	Rest of Western Africa

Table C.3: Aggregation of ILO industries into sectors

ILO industry	sector
Agriculture	Agriculture
Manufacturing	Manufacturing
Mining and quarrying	Manufacturing
Electricity, gas and water supply	Manufacturing
Construction	Other
Trade, Transportation, Accommodation and Food	Other
Business and Administrative Services	Other
Public Administration, Community, Social	Other
Other Services and Activities	Other

Table C.4: Summary statistics on trade flows, production and labor (1991-2019)

		N	mean	sd
annual total production	Agriculture	4,495	9,658,591	4.58E+07
(by country)	Manufacturing	4,495	1.75E+08	8.30E+08
annual cross-border total exports	Agriculture	4,495	948,572.8	3.31E+06
(by country)	Manufacturing	4,495	4.65E+07	1.47E+08
annual self-trade flows	Agriculture	4,495	8,710,019	4.48E+07
(by country)	Manufacturing	4,495	1.28E+08	7.02E+08
annual cross-border bilateral trade flows	Agriculture	692,230	6,150	115,730
(by country-pair)	Manufacturing	692,230	301,821	3,563,093
annual employment	Agriculture	4,495	5,949	29,747
(by country)	Manufacturing	4,495	2,802	12,949

Note: the trade flows and production levels are expressed in millions of current US dollars, the employment in thousands of people.

Table C.5: Summary statistics on weather measures built from ERA5 data (1991-2019)

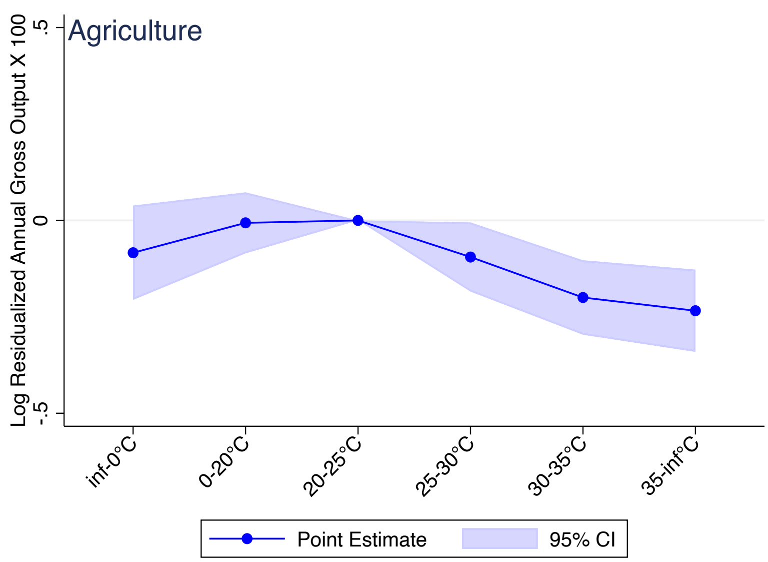
	N	mean	sd	min	max
average temperature (°C)	4,495	18.35	7.09	-1.44	29.37
average precipitation (mm)	4,495	3.1	2.1	0.02	14.01
temperture bins (# days)					
inf-0°C	4,495	11.04	22.84	0.00	149.99
0-20°C	4,495	104.53	101.11	0.00	359.53
20-25°C	4,495	58.67	43.37	0.00	269.52
25-30°C	4,495	100.83	76.22	0.00	365.00
30-35°C	4,495	67.05	64.72	0.00	285.21
35-inf°C	4,495	22.89	42.03	0.00	230.50
precipitation bins (# days)					
0-1 mm	4,495	220.27	70.96	26.77	362.81
1-10 mm	4,495	110.66	50.67	1.71	251.64
10-20 mm	4,495	24.32	18.47	0.00	111.05
20-inf mm	4,495	9.75	10.30	0.00	76.49

C.1 One Sector Model

We implement our structural estimation of the nonlinear effect of temperature on the gross output of one sector first. As in the Monte Carlo simulations, we use our estimation procedure assuming that the only sector in the economy is agriculture and firms operate under perfect competition. We set the trade elasticity to $\theta = 8.3$, following the estimate from Caliendo & Parro (2015) for agriculture.

Figure C.8 reveals that the negative impact of temperature on agricultural gross output is almost linearly increasing starting at 25℃, with statistically significant effect for all three bins above 25℃. The magnitude of the coefficients for these three bins is larger than the TWFE estimates, but not significantly different.

Figure C.8: Point Estimates, One Sector Model



Notes: Figure plots the point estimates and 95% confidence intervals resulting from estimating regression model (29). Dependent variable $\Delta\check{A}_{ist}$ is computed by inverting (23) and taking first differences, imposing perfect competition and a single sector. We set $\theta = 8.3$, the trade elasticity estimated by Caliendo & Parro (2015) for Agriculture. Number of days in the year with maximum temperature between $20°C$ and $25°C$ serves as the omitted category. Period spans 1991 - 2019. Top and bottom 1% of observations in terms of residualized year-over-year growth rates were omitted from the regression. Regression controls for flexible continent-by-year effects. Standard errors are clustered on the country.

Figure C.9 plots on a world map the percentage change in 2019 gross output (panel

a) and real wage (panel b) that results from a change in the treatment vectors from base-year weather to end-of-sample weather, focusing on the last two bins $[30°\text{C}, 35°\text{C}]$ and $[35°\text{C}, +\infty)$. The RTE effects on gross output varies from almost -40% to +2-3% across countries. The magnitude of the RTE effects on real wage are very similar, and the two pictures almost line up perfectly. Overall, most countries experienced a decrease in gross output and real wage due to the larger number of days with high temperatures. The countries that suffered the most from global warming are located in Eastern Europe, Africa and Central America. By contrast, countries like Canada, Portugal or Irak benefited from a decrease in the number of days in the top bins, which resulted in an slight increase in gross output and real wage.

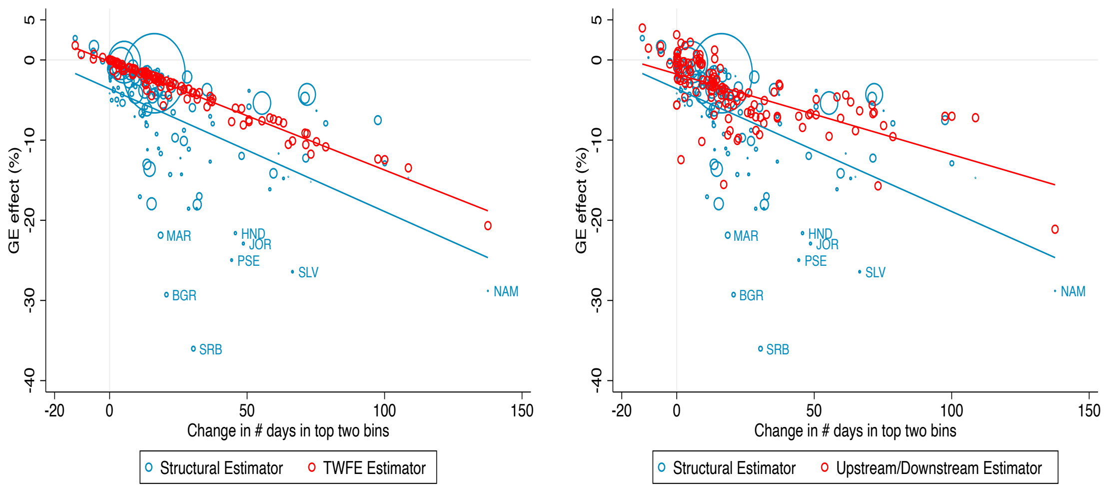
Figure C.10 presents the estimated effects on gross output from the change in the top two temperature bins observed from 1991 to 2019. We plot the percentage change in 2019 gross output against the change in the number of days in the top two bins, using our structural estimator (in blue) or the TWFE estimator (in red, panel a) or the Upstream/Downstream estimator (in red, panel b). Each circle indicates a country and we allow blue marker size to vary with the share of agricultural output for each country in 2019 total agricultural output.

The TWFE estimates in Figure C.10 reflect an almost linear relationship between percentage change in gross output and the change in the number of days in the top two bins. It is not fully linear because we combine $\check{\beta}^{FE}_{[30°\text{C},35°\text{C}]}$ and $\check{\beta}^{FE}_{[35°\text{C},+\infty]}$ with their respective change in the number of days in these bins for each country. Yet, the relationship clearly indicates that larger shocks entail larger negative effects (up to -20%) and rules out spillovers (zero intercept).

Our structural estimates differ from the TWFE estimates in three important ways. First, our estimates highlight the heterogeneity in RTE effects, even for similar degrees of warming. Second, a line going through these *RGE* effects entails a negative intercept (see blue line), which reflects the presence of negative spillovers from the vector of weather shocks in the trade equilibrium. Third, and consequently, the TWFE estimates tend to *understate* the RTE effects on agriculture and ignore the dramatic negative effects on some small developing countries like Bulgaria, Serbia, El Savaldor or Namibia (with effects up to -35%). However, the slope of the red and blue lines are relatively similar, which implies that the

In panel b), we compare our estimates with the upstream-downstream estimator. Even though the latter estimator is able to capture a negative intercept and some heterogeneous effects, it still *understates* the RTE effects relative to our structural estimator and understates the degree of heterogeneity, and it provides less accurate comparisons across countries (since the slope of the red line differs from the slope of the blue line).

Figure C.10: Structural Estimates vs TWFE or Upstream/Downstream Estimates, One Sector



(a) Structural vs TWFE (b) Structural vs Upstream/Downstream

Notes: Figures plot the % change in 2019 gross output resulting from change from 1991 weather to 2019 weather (y-axis) against the change in the number of days with maximum temperature above 30°C (the top 2 bins) (x-axis) computed using the structural estimator (blue) and the TWFE estimator (red in panel a) or the Upstream/Downstream estimator (red in panel b). For all estimators, we use just agricultural data and assume this sector comprises the entire economy. For the structural estimator, we impose perfect competition, and set $\theta = 8.3$. Marker sizes indicate the share of agricultural output for each country in total 2019 agricultural output. For countries with greater than 20% losses in gross output, we report the ISO country label.