Quantifying Aggregate Impacts in the Presence of Spillovers

Dave Donaldson
MIT

Federico Huneeus
Duke University & Central Bank of Chile

Vincent Rollet
MIT

March 2026

Abstract

A widespread threat to the validity of standard policy evaluation tools is the presence of spillovers between treated and untreated groups such as spatial regions. Economic interactions across units of analysis—due to the flow of goods, factors, and payments to and from the government, for instance—result in bias in standard estimates of objects of interest such as the average treatment effect or the total effect of a program. In this paper, we develop a suite of approaches that can enable researchers to use economic theory and data about economic flows and distortions in order to overcome this bias. We apply this methodology to estimate the aggregate economic impact of a large earthquake that struck Chile in 2010.

Author contacts: ddonald@mit.edu, federico.huneeus@duke.edu, and vrollet@mit.edu. We are grateful to Noah Siderhurst for outstanding research assistance, and to Rodrigo Adao, Keelan Beirne, Kirill Borusyak, Arnaud Costinot, and Tomás Domínguez-Iino for very helpful comments. The views expressed are those of the authors and do not necessarily reflect the views of the Central Bank of Chile or its board members. This study was developed within the scope of the research agenda conducted by the Central Bank of Chile (CBC) in economic and financial affairs of its competence. The CBC has access to anonymized information from various public and private entities, by virtue of collaboration agreements signed with these institutions.

1 Introduction

Economic research increasingly exploits experimental and quasi-experimental variation when estimating the impact of policy and non-policy shocks. When such "treatments" are as good as randomly assigned to units of analysis, the aggregate impact of the assignments that occurred is typically estimated by comparing treated units to untreated ones. However, estimates based on these comparisons are generally biased in the presence of treatment spillovers, which happen whenever treating one unit has an effect on any other. Such spillovers are not a mere nuisance—they are at the heart of our field. Whether the unit of analysis is a neighborhood, a region, a firm, or a household, economists expect that outcomes of interest will be co-determined in equilibrium, implying that economic units cannot be studied in isolation. And even when spillovers appear small at the level of any given unit, economists have long cautioned that such a finding should be expected precisely whenever spillovers affect many units, implying that their aggregate impact can still be substantial (e.g. Bradford, 1978).

While bias due to treatment spillovers is well understood, it is less clear what applied researchers can do to fix it. A predominant approach requires the analyst to know which untreated units are "pure controls"—ones that are completely unaffected by the treatment. Given this knowledge, standard program-evaluation methods can be used to estimate the total impact of the program by comparing pure control units to all other units. These methods have the advantage of being robust to unobserved treatment effect heterogeneity, as well as offering regression implementations that make inference and diagnostic tests (such as for lack of pre-trends) straightforward. However, a common critique of this approach is that the choice of pure control units may appear ad-hoc. For example, when they are selected on the basis of a metric of distance from treated units, what metric should be used?

Clearly, economic theory can be used to help researchers to de-bias their estimates of a treatment's effect on the directly treated units and evaluate ripple effects on others to measure aggregate effects. What is less clear is how to do so robustly, with help from information about the economic flows across units causing spillovers, and without jettisoning the attractive features of standard program evaluation methods.

In this paper we develop a set of tools that allow applied researchers to do exactly this. We draw on a suite of simple economic models that enable applied researchers to focus on their ideal estimand, articulate the types of spillovers that may occur in their setting, craft a strategy for de-biasing estimates that depends on the types of outcomes that are observable, and implement this strategy as an augmented version of an otherwise standard regression. We then illustrate this approach with an application evaluating the reduction in Chile's aggregate real GDP that was caused by the historic earthquake that occurred there in February 2010.

We begin in Section 2 by describing the general setup that we consider. A researcher observes that a treatment of interest (e.g., a credit intervention or a natural disaster) was assigned to a subset of units (e.g., firms or residential neighborhoods) in the economy. The treatment is presumed to have affected these units by potentially shifting the units' fundamental characteristics (e.g., their productivity levels or residential amenities). In line with the standard potential outcomes approach to program evaluation, we allow the treatment to have arbitrarily heterogeneous treatment effects on units' fundamentals.

In this environment, we consider a researcher whose goal is to estimate the effect of the entire set of treatment assignments that occurred on an aggregate quantity of interest (e.g., real GDP or a measure of aggregate inequality such as the Gini index). This aggregate impact can be approximated, to first order, as a weighted sum of the effects of the treatments on underlying fundamentals. Here, the weights characterize the responsiveness of the aggregate outcome to changes in each unit's fundamental characteristic, so we refer to them as *responsiveness weights*. One special case in which these weights are a simple function of pre-determined data is that corresponding to Hulten (1978): in a closed and undistorted economy, if the goal is to estimate the effect of treatment on real GDP, then responsiveness weights are simply each unit's share of sales in GDP.

While fundamentals may be unobserved, they are nevertheless linked to observable outcomes (e.g., firms' sales or workers' commuting choices across neighborhoods) available for the units of interest. Importantly, because of spillovers, changes in the fundamental characteristics of any unit j may affect the observed outcome of any other unit i (e.g., firms compete for the same customers or commuters change labor supply across neighborhoods in response to wage differences). We refer to the matrix that summarizes these spillover effects as the *exposure matrix*.

Section 3 describes an unbiased and consistent estimator of the aggregate impact of the treatment in the case where the responsiveness weights and exposure matrix are known to the researcher. Under the standard assumption that the treatments are as good as randomly assigned (perhaps conditional on predetermined covariates), we show that the aggregate impact is simply the slope coefficient from a regression (potentially augmented to include covariates) of units' observed outcomes on a measure of their treatment exposure. This novel measure of treatment exposure combines the exposure matrix and the responsiveness weights into one appropriately weighted measure of units' exposure to the latent fundamentals that treatments affect. Even though the underlying effects of treatment on fundamentals are arbitrarily heterogeneous, and latent fundamentals are unobserved, the proposed regression uses observed outcomes to reveal the weighted average of such underlying heterogeneous effects that matters for the goal at hand.

This result can be applied in any context where the researcher knows the responsiveness weights and exposure matrix that apply to their setting and to the aggregate impact that they aim to estimate. However, we believe there is value in delineating the steps that researchers can follow to arrive at these two ingredients from more primitive beliefs and data. Section 4 aims to make this straightforward for practitioners. It draws on the setup of Baqaee and Farhi (2019) to describe a flexible model economy that nests a wide range of applied problems in fields where economic spillovers take place. In particular, we consider an economy with: *(i)* a set of production units referred to as firms; *(ii)* a set of households who own firms, consume final goods, and supply factors for production; *(iii)* a government sector that levies taxes on certain transactions and makes lump-sum transfers to certain households; and *(iv)* a set of arbitrary distortions in production and consumption.

Given data on the economic linkages between units in the economy (the flows of goods from firms to households and other firms, the flows of factors from households to firms, and the flows of government taxes/transfers from and to firms and households), as well as the extent to which firms and households can substitute between the different goods and factors they consume, we derive how any set of shocks to fundamentals would affect any other endogenous outcome of interest as a function of baseline flow data and elasticities. We show how to map such effects into the exposure matrix and responsiveness weights that are required for our regression method to be applied.

Section 5 presents an application of our approach. We consider the case of the earthquake that hit Chile in February 2010—one of the most violent in recorded history—and seek to quantify the effect of this calamity on aggregate Chile's real GDP. Existing approaches to questions such as this one typically compare the path of output (or value-added) in regions of the country that were hit by a natural disaster to those that weren't. However, such an approach is only unbiased if these regions are autarkic in terms of trade, factor flows, and government taxes and transfers—a presumption that is clearly rejected by all available data. We find that in this context the estimated effects of the earthquake on real GDP are approximately 2 percent per year for at least five years after the event.

Related literature. The econometric issues raised by treatment spillovers (also known as treatment interference) have been widely documented. See Cox, 1958, Rubin, 1980, and Manski, 1993 for early discussions. In particular, Rubin's Stable Unit Treatment Value Assumption (SUTVA)—which posits that each unit's outcome depends only on its own treatment status and not on the treatments assigned to other units—is widely considered to be necessary for comparisons between treated and untreated units to deliver unbiased estimates of typical estimands of interest.

A leading strategy for addressing spillover bias relies on the researcher's *a priori* knowledge of a "treatment intensity" function that provides a complete mapping of any unit's outcome into the entire set of possible treatment assignments, both to that unit and all other units. An important special case of such a function arises when some units are known to be completely unaffected by the treatment assignments of other units, so they can be considered to be "pure controls". For example, in Hudgens and Halloran (2008) and Zigler and Papadogeorgou (2020) spillovers are assumed to take place only within disjoint clusters of units, as is also the case in the analysis of cluster-randomized trials that assign treatments to spatial units (such as schools in Miguel and Kremer, 2004, cities in Crépon et al., 2013, or urban neighborhoods in Franklin et al., 2024) that are believed to be large enough that all spillovers take place within them.[1] Another commonly used type of treatment intensity function allows for overlapping clusters but relies on knowledge of an explicit metric of distance between treated and untreated units.[2] For example, Aronow and Samii (2017) define treatment intensity via the existence of a treated neighbor, Hornbeck and Moretti (2018) do so via distance-weighted averages of other units' treatments (with distance measured using bilateral economic flow data such as that on trade and migration), and Leung (2020) does so via the number of treated neighbors.

As discussed above, one challenge facing any use of a treatment intensity function is

[1] A related (implicit) use of a treatment intensity function occurs in the widely-used strategy of excluding all units that lie within a distance "buffer"—according to a given metric—of the directly treated units. However, this strategy is ill-suited for estimating the aggregate effect of a program of interest because it discards units that are most likely to experience indirect impacts, thereby omitting a potentially important component of the treatment's aggregate effect.

[2] A relaxed version of this requirement (e.g. in Sävje et al., 2021 and Faridani and Niehaus, 2024) relies on the assumption that treatment spillovers decay with distance faster than a specified rate, rather than a complete specification of this distance dependence. Similarly, Munro et al. (2025) and Arkhangelsky and Rutgers (2025) describe methods for estimating aggregate impacts when treatment spillovers work via a given set of observable market prices and via centralized assignment processes, respectively.

the need for that function to be specified by the analyst *a priori*. It is therefore natural for economists to use economic models explicitly when doing so. This approach appears, for example, in "market access" studies approach to trade spillovers (e.g. Redding and Venables, 2004; Donaldson and Hornbeck, 2016; Egger et al., 2022), by Monte et al. (2018) in their study of labor market spillovers through commuting, by Borusyak et al. (2022) in their study of equilibrium migration behavior, and by Adao et al. (2019) in their study of the "China shock". The methodology developed in this paper builds on such approaches in two ways. First, it allows the analyst to develop a measure of treatment intensity tailored to their setting—oriented around the aggregate impact they aim to estimate, chosen from a menu of economic models featuring the types of spillovers they believe are important, and designed to match data on the flows that govern those spillovers.[3] Second, despite relying on a fully specified economic model, our methodology allows for arbitrarily heterogeneous direct effects of the shocks of interest, as in the canonical potential outcomes framework of Rubin (1974). This combination of explicit modeling, as in fully "structural" approaches, and unobserved treatment effect heterogeneity, as in conventional program evaluation techniques, is possible because many questions hinge only on identifying a particular weighted average of treatment effects, rather than each individual effect.

Finally, we contribute to a large literature on the economic costs of natural disasters. Barrot and Sauvagnat (2016), Boehm, Flaaen and Pandalai-Nayar (2019), and Carvalho, Nirei, Saito and Tahbaz-Salehi (2021) show, in the context of natural disasters in the U.S. and the 2011 Tōhoku earthquake in Japan, that disruptions caused by these shocks propagated through trade links, with clearly visible effects both upstream and downstream in the supply chains of affected firms. Boustan, Kahn, Rhode and Yanguas (2020) show that natural disasters in the U.S. tend to be followed by out-migration, and Strobl (2011) shows that such out-migration can account for a large share of the lower growth of U.S. counties after a hurricane hits them. Deryugina (2017) finds, again in the context of hurricanes in the U.S., that these catastrophes lead to substantial government transfers that largely offset their damages. Together, these findings highlight that spillovers severely impact the estimation of environmental disasters' economic effects. Several approaches have been explored to estimate the economic effects of natural

[3] This component draws on the literature describing how economic shocks propagate through networks, and in particular on the important theoretical advances of Long and Plosser (1983), Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2012) and Baqaee and Farhi (2019, 2020), as well as empirical investigations such as Di Giovanni, Levchenko and Mejean (2014), Barrot and Sauvagnat (2016), Carvalho et al. (2021), and Korovkin and Makarin (2022).

disasters despite this challenge. Some studies (e.g., Cavallo et al., 2013; Hsiang and Jina, 2014) have focused on cross-country analyses that are less subject to spillovers. Others (e.g., Felbermayr, Gröschl, Sanders, Schippers and Steinwachs, 2018; Lima and Barbosa, 2019) have relied on within-country analyses paired with the buffer approach to estimate spillover effects. Finally, a large literature (summarized in Botzen, Deschenes and Sanders, 2019) has leveraged computable general equilibrium models to estimate the costs of environmental catastrophes. As described above, our approach to the study of economic impacts of natural disasters illustrates how one can combine the complementary strengths of these theoretical and empirical methods.

2 Aggregate effects of shocks

2.1 Setting

We consider a setting in which the goal is to evaluate the effect of a vector of treatments T on some aggregate outcome denoted by W. The treatment shifts a latent, fundamental variable A, with potentially heterogeneous treatment effects β_i:

$$\Delta \log A_i = \beta_i T_i + \varepsilon_i^a. \tag{1}$$

While A is unobserved, it affects another variable, y, that is observed. We assume that we know the exposure matrix E, where $E_{ij} \equiv \frac{\mathrm{d}\log y_i}{\mathrm{d}\log A_j}$ measures the exposure of unit i to latent variable shocks affecting unit j.

To first order, shifts in observables are given by

$$\Delta \log y_i = \sum_{j=1}^{N} \underbrace{\frac{\mathrm{d}\log y_i}{\mathrm{d}\log A_j}}_{E_{ij}} \Delta \log A_j + \sum_{k=1}^{p} \gamma_k X_{ik} + \varepsilon_i^y, \tag{2}$$

where the X_{ik} capture p observable characteristics of firm i. Denoting by $\mathcal{T} \equiv \{i \mid T_i = 1\}$ the set of treated units, by $N_T = |\mathcal{T}|$ the number of treated units, by $E_T \in \mathbb{R}^{N \times N_T}$ the submatrix of E that keeps the columns indexed by $\mathcal{T}$, and by $\beta_T \equiv (\beta_j)_{j \in \mathcal{T}} \in \mathbb{R}^{N_T}$ the vector of size N_T that collects treatment effects for treated observations, the expression above can be simplified in vector notation as

$$y = E_T \beta_T + X\gamma + \varepsilon, \tag{3}$$

where $\varepsilon = \varepsilon^y + E\varepsilon^a$.

Our goal is to measure the aggregate impact of this set of treatments on W. To first order, this impact is given by:

$$\theta \equiv \Delta \log W = \sum_{i \in \mathcal{T}} \beta_i \frac{\mathrm{d} \log W}{\mathrm{d} \log A_i} = \sum_{i \in \mathcal{T}} \beta_i \kappa_i = \kappa_T^\top \beta_T, \tag{4}$$

where $\kappa_T \equiv (\kappa_j)_{j \in \mathcal{T}} \in \mathbb{R}^{N_T}$ is the vector collecting responsiveness weights for treated observations.

2.2 Examples

Three examples from recent work illustrate the notation introduced above.

Example 1: Natural disaster

- **Setting**: A natural disaster T takes place in treated regions of the country. The researcher's goal is to measure the effect of the disaster (i.e. the "treatments") on national real GDP W. This is the setting of Carvalho et al. (2021).
- **Unit of analysis** i: Firms.
- **Latent variable** A: Firm-specific capital. An unobserved share of the working capital of firms directly hit by the shock, due to their region of location, is destroyed. Non-treated firms' capital is assumed to be unaffected by the disaster.
- **Responsiveness weights** κ: Firms' capital expenditure shares times their sales share in the economy. Intuitively, capital losses will have a larger effect on GDP if they affect firms that heavily rely on capital and/or are larger in terms of sales.
- **Observed variable** y: Sales of each firm.
- **Source of spillovers** E: Input-output linkages across firms and space. When a supplying firm's capital is damaged this limits its ability to sell to downstream buying firms, so those buyers shrink due to limited access to inputs. Similarly, when a buying firm's capital is destroyed it demands less inputs, so its upstream suppliers shrink. Finally, demand spillovers (as in Example 1) are present for every buyer (firms and the final consumer).

Example 2: Public works program

- **Setting**: A public works program T provides attractive employment options to the residents of treated neighborhoods in a city. The researcher's goal is to measure the effect of the program on the total (and hence also per-person) income earned by less-educated workers in the city W. This is similar to the setting of Franklin et al. (2024).
- **Unit of analysis** i: Neighborhoods.
- **Latent variable** A: Residual labor supply to the private sector in each neighborhood.
- **Responsiveness weights** κ: Change in the local wage (of less-educated workers) in each neighborhood caused by the change in local residual labor supply, weighted by the initial labor supply in the neighborhood. Intuitively, treatments will shift the average wage of less-educated workers more if the treated areas are home to more of these workers and if wages are more sensitive to the labor supply there.
- **Source of spillovers** E: Commuting flows across neighborhoods. Reducing the residual labor supply of any given treated neighborhood's residents cuts the labor available to firms in other areas (both treated and untreated) that relied on its outbound commuters. Further, wage increases in any treated neighborhood, due to the increased labor scarcity there, draws commuters from other (treated and untreated) neighborhoods which, in turn, affects wages there.
- **Observed variable** y: Wages of less-educated workers in each neighborhood.

Example 3: Loan program

- **Setting**: A loan program T is offered to treated firms within an industry. The researcher's goal is to measure the effect of the program treatments on the total consumer surplus created by this industry W. This is the setting of Cai and Szeidl (2024), and is related to that in Rotemberg (2019).
- **Unit of analysis** i: Firms.
- **Latent variable** A: Firm productivity. Access to credit potentially increases (a notion of) the treated firms' productivity by lowering their marginal costs of production and/or by increasing the quality or variety of goods they can offer. Non-treated firms' productivity levels are assumed to be unaffected by the program.
- **Responsiveness weights** κ: Changes in prices caused by productivity shifts weighted by firms' market shares.
- **Observed variable** y: Sales of each firm.
- **Source of spillovers** E: Demand spillovers. Treated firms' productivity growth

causes them to expand by offering lower prices to customers, but part of this growth comes at the expense of other firms (treated and non-treated). These effects are strongest among local competitors and weaker for more distant rivals.

Next steps. Estimating θ is associated with two challenges. First, even when the responsiveness weights κ and the exposure matrix E are known, estimating θ is complicated by the fact that treatment shifts the latent variable A, and hence the treatment effects β_i in equation (1) cannot be directly estimated. Second, knowledge of κ and E is far from guaranteed. To make progress, we begin in Section 3 by describing how θ can be estimated when κ and E are known. Then, in Section 4, we show how κ and E can be computed in a broad class of models.

3 Estimating Aggregate Impacts when Exposure and Responsiveness Are Known

Equation (3) gives us a linear mapping between shocks to latent variables and observable outcomes. If treatment is randomly assigned, this linear mapping allows us to recover the desired weighted sum of treatment effects θ without having to observe or infer the latent variable (under typical rank and regularity assumptions). As we describe here, this can be done via a simple regression based on a modified version of exposure to treatment. This modified regressor depends on a mix of the responsiveness weights κ and exposure matrix E.

Assumption 1 (Exogeneity)**.** $\mathbb{E}[\varepsilon \mid X, T] = \mathbb{E}[\varepsilon \mid X] \equiv m(X)$, *where* $m(X)$ *is linear in* X.

Assumption 2 (Rank)**.** *The matrix* $[X\ E_T]$ *has full column rank.*

The first equality in Assumption 1 is a standard exogeneity assumption—namely, that treatments T are as good as randomly assigned, conditional on the covariates X; the second part of this assumption goes further and requires the researcher to have a correctly specified model of covariates that linearly predict the outcomes in the absence of treatment.[4] Assumption 2 avoids cases of perfect multicollinearity. It implies that both $X^\top X$ and that the Gram matrix of residualized treated exposures $G \equiv E_T^\top M_X E_T \in \mathbb{R}^{N_T \times N_T}$ (with $M_X \equiv I - X(X^\top X)^{-1} X^\top$) are invertible.

[4] In some applications, the researcher may be willing to dispense with this latter assumption and instead apply design-based knowledge (of the data generating process of T), following Borusyak and Hull (2020).

One situation in which Assumption 2 fails is when the covariate set includes an intercept and, for some treated unit j, the corresponding exposure column $E_{\cdot j}$ is proportional to a vector of ones. In this case, the effect of treating unit j, β_j, cannot be distinguished from shocks that affect all units equally. For example, suppose the outcome of interest is firm sales y_i, and that an increase in the productivity of a particular firm j—say, the economy's internet provider—has an identical proportional effect on every firm, so that $\mathrm{d}\log y_i/\mathrm{d}\log A_j = c_j$ for all i. In such a setting, productivity shocks to the internet provider are observationally equivalent to aggregate shocks influencing all firms, making β_j unidentified.

Assumption 3 (Regularity)**.** *Define $\Sigma_a \equiv \mathrm{Var}(\varepsilon^a \mid X, E)$ and $\Sigma_y \equiv \mathrm{Var}(\varepsilon^y \mid X, E)$. We assume:*

(i) Information growth: $\kappa_T^\top G^{-1}\kappa_T \to_p 0$.

(ii) Well-behaved errors and exposure matrix: $\|\Sigma_a\|_{\mathrm{op}} \leq C_a < \infty$, $\|\Sigma_y\|_{\mathrm{op}} \leq C_y < \infty$, $\|E\|_{\mathrm{op}} \leq C_E < \infty$, and $\varepsilon^a \perp \varepsilon^y$ conditional on (X, E).[5]

Here, limits are taken as $N \to \infty$ while the number of observable covariates p stays constant.

This assumption is sufficient for the consistency of our estimator. Part (i) rules out cases where the responsiveness weights κ_T are excessively concentrated on a small number of observations. While our definition of an aggregate W has so far been general, this assumption requires W to place weights that are sufficiently spread out relative to both the economy as a whole and the treatments that happened.

Proposition 1 (Estimation of weighted sums of treatment effects with spillovers)**.** *Under Assumptions 1 and 2, define weights w and the weighted exposure to treatment z as*

$$w \equiv \frac{G^{-1}\kappa_T}{\kappa_T^\top G^{-1}\kappa_T}, \qquad z \equiv E_T w. \tag{5}$$

Consider the OLS regression

$$y = \theta z + X\gamma + \varepsilon, \tag{6}$$

and let $\widehat{\theta}$ be the OLS coefficient on z. $\widehat{\theta}$ is a finite-sample unbiased estimator of $\theta = \sum_{i\in\mathcal{T}} \beta_i \kappa_i$. Under Assumption 3, this estimator is consistent.

We prove this result in Appendix A. The appendix further shows that the estimator of equation (6) is BLUE under homoskedasticity, and shows how Proposition 1 can be extended to GLS.

[5] Here, $\|\cdot\|_{\mathrm{op}}$ denotes the operator norm, defined as $\|A\|_{\mathrm{op}} = \sup_{v\neq 0} \frac{\|Av\|}{\|v\|}$.

Special case of homogeneous treatment effects. If treatment effects are homogeneous ($\beta_i = \beta$ for all $i \in \mathcal{T}$), then the target $\theta = \kappa_T^\top \beta_T$ reduces to $\left(\sum_{i\in\mathcal{T}} \kappa_i\right)\beta$. In this case, there is no need to compute w or invert G. To estimate θ, it suffices to compute the exposure $\tilde{T} = ET$ of each node to treatment, and regress changes in Domar weights on exposure to treatment and controls:

$$y = \beta\tilde{T} + X\gamma + \varepsilon.$$

The target can then be recovered as $\hat{\theta} = \left(\sum_{i\in\mathcal{T}} \kappa_i\right)\hat{\beta}$.

Average treatment effect on the treated. Proposition 1 is valid for any vector of responsiveness weights κ as long as Assumptions 1–3 hold. One particular aggregate W that may be of interest is the average treatment effect on the treated, which can be recovered using as responsiveness weights κ_T the uniform vector $\frac{1}{N_T}\mathbf{1}_{N_T}$.

4 Computing Responsiveness Weights and the Exposure Matrix

Implementing the estimation strategy of Proposition 1 requires knowledge of the exposure matrix $E_{ij} = \frac{\mathrm{d}\log y_i}{\mathrm{d}\log A_j}$ of nodes to latent shocks, as well as the responsiveness weights $\kappa_i = \frac{\mathrm{d}\log W}{\mathrm{d}\log A_i}$. These quantities typically cannot be directly observed or estimated, but they can be recovered given assumptions on the economic channels through which spillovers propagate (e.g., input-output linkages, income shocks, etc.) In this section, we draw on the results of Baqaee and Farhi (2019) to show how to compute exposure measures in a broad range of settings.

4.1 Setup

We consider a flexible model of an economy featuring F factors indexed by f, C consumers indexed by c, and N firms indexed by i. Goods in the economy are produced by firms through an aggregation of factors and goods, while factors are supplied ex nihilo.

Consumers. Each consumer c supplies a quantity L_{cf} of each factor f (at a price w_f) and consumes a composite final good $Y_c = \mathcal{D}_c(c_{c1}, \ldots, c_{cN})$, where c_{ci} is c's consumption of good i (exchanged at a price p_i). $\mathcal{D}_c$ is assumed to be homothetic, and consumers solve

$$\begin{aligned} \max \quad & U_c(\mathcal{D}_c(c_{c1}, \ldots, c_{cN}), L_{c1}, \ldots, L_{cF}) \\ \text{s.t.} \quad & \sum_{i=1}^{N} p_i c_{ci} = \sum_{f=1}^{F} w_f L_{cf} + \Pi_c, \end{aligned}$$

where Π_c denotes the income that c derives from government transfers or the profits of the firms they own. The total supply of factor f is given by $L_f = \sum_{c=1}^{C} L_{cf}$.

Firms. The production function of firm i is given by $y_i = A_i F_i(L_{i1}, \ldots, L_{iF}, x_{i1}, \ldots, x_{iN})$, where A_i is a Hicks-neutral productivity shifter, L_{ij} is firm i's usage of factor j, and x_{ij} is its consumption of intermediate input j. Firms sell the good they produce at a markup μ_i over their production cost.

The economy's production and consumption network can be summarized by a (revenue-based) input-output matrix Ω of size $(C + N + F) \times (C + N + F)$, where the first C rows and columns correspond to consumers, the following N rows and columns to firms, and the final F rows and columns to factors.[6] The entry Ω_{ij} of the matrix corresponds to node i's spending on inputs from j as a share of i's sales, $\Omega_{ij} \equiv \frac{p_j x_{ij}}{p_i y_i}$. Its Leontief inverse is then defined by $\Psi \equiv (I - \Omega)^{-1} = I + \Omega + \Omega^2 + \Omega^3 + \cdots$. While Ω measures the direct reliance of each node on every other node for inputs into production, Ψ captures both the direct and indirect input reliance of any node on every other node.

GDP and Domar weights. Nominal GDP corresponds to the value of final consumption: $\text{GDP} \equiv \sum_{i=1}^{N} \sum_{c=1}^{C} p_i c_{ci}$. The importance of a node in the economy is captured by its revenue-based Domar weight λ_i, equal to its sales as a fraction of GDP: $\lambda_i \equiv (p_i y_i)/\text{GDP}$. Because only a fraction of production is typically used in final consumption, the sum of the λ_i is typically larger than one. For convenience, we denote by Λ_f the Domar weight of factor f.

The importance of a good in final consumption is captured by the vector b of final demand expenditures as a share of GDP, $b_i \equiv \left(\sum_{c=1}^{C} p_i c_{ci} \right) / \text{GDP}$. Final demand is linked to producers' Domar weights by the Leontief inverse, via $\lambda^\top = b^\top \Psi = b^\top (I + \Omega + \Omega^2 + \Omega^3 + \cdots)$. Indeed, Domar weights equal final demand shares propagated through the

[6]The input-output matrix collects flows of consumption goods, intermediates, and factors. It is therefore convenient to use the slight abuse of notation of Baqaee and Farhi (2019) and use interchangeably w_f and p_{N+f} for factor prices; L_{if} and $x_{i,N+f}$ for factor usage; L_f and y_f for factor supply; and c_{ci} and x_{ci} for final consumption.

whole input–output network. This measure of final demand also allows us to measure changes in real GDP (denoted by Y), with $\mathrm{d}\log Y = \sum_{i=1}^{N} b_i \, \mathrm{d}\log c_i$.

On top of the revenue-based input-output matrix, we can define a cost-based input-output matrix $\tilde{\Omega}$, whose entry $\tilde{\Omega}_{ij}$ corresponds to the cost share of input j in the production of good i: $\tilde{\Omega}_{ij} = \mu_i \Omega_{ij}$. This allows us to define, by analogy, the cost-based Leontief inverse $\tilde{\Psi} = (I - \tilde{\Omega})^{-1}$, and the cost-based Domar weights $\tilde{\lambda}_i = b^\top \tilde{\Psi}$. For factors, we use the notation $\tilde{\Lambda}_f$. In the absence of distortions, cost-based and revenue-based objects are equal.

4.2 Computing responsiveness weights

Real GDP as the outcome of interest. A central case of interest is when W is real GDP. In this case, Baqaee and Farhi (2020) have shown that the elasticity of real GDP to productivity shocks is given by

$$\kappa_i^Y \equiv \frac{\mathrm{d}\log Y}{\mathrm{d}\log A_i} = \tilde{\lambda}_i - \sum_{f=1}^{F} \tilde{\Lambda}_f \frac{\mathrm{d}\log \Lambda_f}{\mathrm{d}\log A_i}. \tag{7}$$

The first term of the expression is a technology effect. In the case of an efficient economy, it is the only channel at work, and equation (7) simplifies to $\kappa_i^Y = \lambda_i$, which is Hulten's theorem.[7] The second term captures an allocative efficiency effect: it captures the change in aggregate output arising from the reallocation of factors across producers. When this reallocation move resources toward more distorted sectors, allocative efficiency improves since those sectors were too small to begin with; when they move toward less distorted ones, efficiency worsens.

When the economy is efficient, κ_i^Y can typically be observed in the data. When the economy is inefficient, computing κ_i^Y requires measures of the exposure of factor incomes to productivity shocks, $\mathrm{d}\log \Lambda_f / \mathrm{d}\log A_i$. We describe in Section 4.3 a method for computing such auxiliary terms.

Prices (and functions of them) as the outcome of interest. Other aggregate outcomes of interest are (potentially nonlinear) functions of prices in the economy. For instance, one may wish to evaluate the effect of a treatment on the cost of living P_c of consumer

[7] This is because an undistorted economy has $\mu_i = 1$ for all i, and hence $\tilde{\lambda}_i = \lambda_i$ for all i and $\tilde{\Lambda}_f = \Lambda_f$ for all f. Hence, the second term in (7) becomes $\sum_{f=1}^{F} \mathrm{d}\Lambda_f / \mathrm{d}\log A_i$, which is zero due to the identity $\sum_{f=1}^{F} \Lambda_f = 1$.

c, with $P_c(p) \equiv \min_{c \geq 0} \{\sum_i p_i c_i \mid D_c(c_{c1}, \ldots, c_{cN}) \geq 1\}$. By Shephard's lemma, the first-order effect of a productivity shock on P_c is given by

$$\kappa_i^{P_c} \equiv \frac{\mathrm{d} \log P_c}{\mathrm{d} \log A_i} = \sum_k b_{ck} \frac{\mathrm{d} \log p_k}{\mathrm{d} \log A_i},$$

where $b_{ci} \equiv \frac{p_i c_{ci}}{\sum_k p_k c_{ck}}$ is consumer c's expenditure share on good i. In this scenario, computing the responsiveness weights $\kappa_i^{P_c}$ requires measures of the exposure of goods prices to productivity shocks $\mathrm{d} \log p_k / \mathrm{d} \log A_i$.

A second potential outcome of interest is the (nominal) income of representative agent c, which is the sum of factor incomes from all sources: $I_c \equiv \sum_{f \in F} \Phi_{cf} w_f L_f$. If factor supplies are inelastic, we have

$$\kappa_i^{I_c} \equiv \frac{\mathrm{d} \log I_c}{\mathrm{d} \log A_i} = \sum_f \omega_{cf} \frac{\mathrm{d} \log w_f}{\mathrm{d} \log A_i},$$

where $\omega_{cf} \equiv \Phi_{cf} w_f L_f / I_c$ is c's income share from factor f. In this scenario, computing the responsiveness weights $\kappa_i^{I_c}$ requires measures of the exposure of factor prices to productivity shocks $\frac{\mathrm{d} \log w_f}{\mathrm{d} \log A_i}$.

Finally, one may be interested in measuring the effect of treatment on an inequality measure that summarizes the distribution of real income levels across agents. To do so, denote real consumption per capita by $C_c \equiv I_c / (P_c N_c)$, where N_c is the number of households in group c. Without loss of generality, sort consumers such that $C_1 \leq \ldots \leq C_C$, denote by $\nu_c \equiv N_c / (\sum_d N_d)$ the share of group c in the population, and define cumulative weights $R_c = \sum_{i \leq c} \nu_i$, with $R_0 = 0$. The Gini coefficient can then be defined as

$$G = \sum_{c=1}^{C} s_c \left(R_{c-1} + R_c - 1 \right),$$

where $s_c = \frac{\nu_c C_c}{\sum_d \nu_d C_d}$.[8] Differentiating then yields

$$\mathrm{d} \log G = \sum_{c=1}^{C} \underbrace{\frac{s_c}{G} \Big(R_{c-1} + R_c - (1 + G) \Big)}_{\varpi_c} \, \mathrm{d} \log C_c,$$

[8]The Gini index is defined as one minus twice the polygonal area under the discrete Lorenz curve formed by $\{(R_c, S_c)\}_{c=0}^{C}$, with $S_c \equiv \left(\sum_{i=1}^{c} \nu_i C_i \right) / \left(\sum_{i=1}^{C} \nu_i C_i \right)$ and $S_0 = 0$. That area equals $\sum_c \frac{S_{c-1} + S_c}{2} (R_c - R_{c-1})$ and regrouping terms yields the given expression for G.

and with $\mathrm{d}\log C_c = \mathrm{d}\log I_c - \mathrm{d}\log P_c$, we finally obtain the following responsiveness weights:

$$\kappa_i^G \equiv \frac{\mathrm{d}\log G}{\mathrm{d}\log A_i} = \sum_{c=1}^{C} \varpi_c \left(\frac{\mathrm{d}\log I_c}{\mathrm{d}\log A_i} - \frac{\mathrm{d}\log P_c}{\mathrm{d}\log A_i} \right)$$
$$= \sum_{c=1}^{C} \varpi_c \left[\left(\sum_f \omega_{cf} \frac{\mathrm{d}\log w_f}{\mathrm{d}\log A_i} \right) - \left(\sum_k b_{ck} \frac{\mathrm{d}\log p_k}{\mathrm{d}\log A_i} \right) \right]. \quad (8)$$

Naturally, this corresponds to a particular weighted sum of (log) real income changes across consumer groups c, with weights ϖ_c given by the initial income distribution and the particular functional form implied by the Gini index.

Taking stock. In a flexible model of the economy, responsiveness weights can be computed for a variety of aggregate outcomes of interest as a function of observed objects (such as Domar weights and the input-output matrix Ω), and the exposure of quantities and prices to productivity shocks (e.g., measures of $\frac{\mathrm{d}\log \Lambda_f}{\mathrm{d}\log A_j}$, $\frac{\mathrm{d}\log p_i}{\mathrm{d}\log A_j}$, or $\frac{\mathrm{d}\log w_f}{\mathrm{d}\log A_j}$). In the next section, we describe how to compute these exposure measures in a range of contexts.

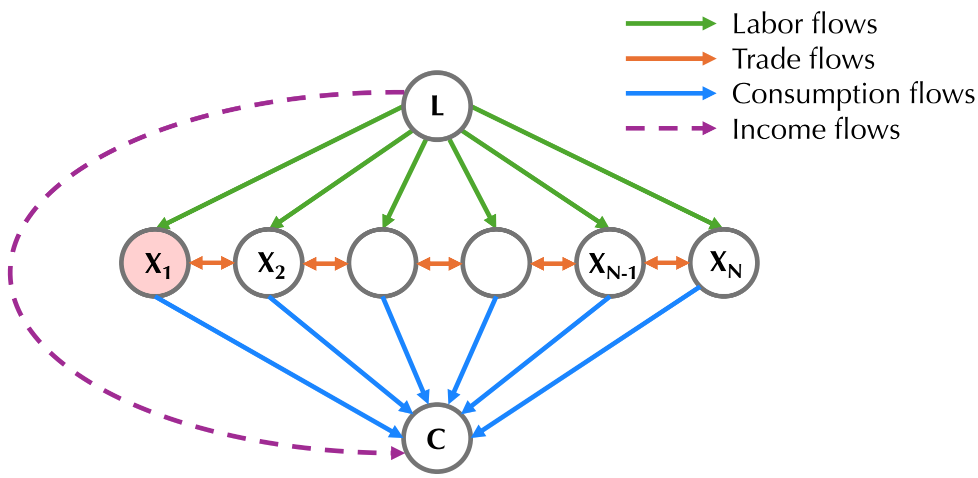
4.3 Computing exposure measures

4.3.1 Introductory example

To illustrate the measurement difficulties caused by spillovers and the construction of the exposure matrix, we first consider the simple "island" economy depicted in Figure 1. The island is populated by a household that inelastically supplies a unit of labor to any of the $N = 10$ firms on the island. Firms are symmetric and each produces competitively a differentiated good that can be used for final consumption or as an intermediate input. Each firm uses labor and all varieties of the differentiated goods in production. These inputs are substitutable with a constant elasticity of substitution denoted by σ. At time $t = 0$, each firm spends a share α of its costs on labor and a share $(1-\alpha)/N$ on each differentiated input. The representative household, which receives all labor income, consumes a Cobb-Douglas bundle of the differentiated products, placing equal weight on each variety.

In this economy, the input-output matrix has 12 rows and 12 columns: the first row/column represents the representative household's consumption, the next ten rows/columns represent firms, and the last row/column represents labor supply. The

Figure 1: A simple example: The island economy



Notes: This figure represents the simple economy described in Section 4.3.1. We consider the effects of a 10% drop in the productivity of firm 1, shaded in red.

pre-shock input-output matrix is given by:

$$\Omega = \begin{array}{c} \\ \\ \end{array}\begin{array}{cccccc} C & X_1 & \cdots & X_N & L & \\ \end{array}$$

$$\Omega = \begin{pmatrix} 0 & 1/N & \cdots & 1/N & 0 \\ 0 & (1-\alpha)/N & \cdots & (1-\alpha)/N & \alpha \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & (1-\alpha)/N & \cdots & (1-\alpha)/N & \alpha \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \begin{array}{c} C \\ X_1 \\ \vdots \\ X_N \\ L \end{array}$$

Consider a scenario in which firm 1 is hit at $t = 1$ by an environmental disaster (e.g., a flood) that leads to a 10% Hicks-neutral productivity decrease. The shock does not directly affect the productivity of other firms in the economy. However, it will indirectly affect their activities because they share inputs and customers with firm 1 and are part of the same trade network.

Suppose a researcher seeks to evaluate the aggregate impact of this event on the island's real GDP. They know that the disaster damaged the productivity of firm 1 (and only firm 1), but not to what extent. Equipped with data on sales, they may estimate the following regression equation:

$$\Delta \log(\text{sales}_i) = \delta + \rho T_i + \varepsilon_i, \tag{9}$$

where $\Delta \log(\text{sales}_i)$ is the change in the sales of firm i between time 0 and time 1, and T_i is a treatment dummy that is equal to 1 for firm 1 and to zero for all other firms. In the absence of any further shocks to the economy, we report in the last row of Table 1 the results this researcher would obtain (on average) for different values of σ (the firms' elasticity of substitution across inputs) and α (the relative importance of labor in firms' costs).

Table 1: Effects of a 10% productivity drop of firm 1

σ (firms' EoS between inputs)	0.8	1.5	2	2
α (firms' share of labor costs in production)	0.5	0.5	0.5	0.8
True effect of the shock on real GDP (W)	-2.0%	-2.0%	-2.0%	-1.3%
Estimated effect based on ρ using equation (9)	+1.1%	-2.6%	-5.1%	-2.0%

Notes: This table shows, for different values of σ (the firms' elasticity of substitution across inputs) and α (the relative importance of labor in firms' costs) the effect of a 10% productivity shock to a firm in the island economy (depicted in Figure 1) on that firm's sales, measured using regression equation (9). In all scenarios, the first-order effect of the productivity shock on GDP is given by Hulten's theorem as 10% (the size of the productivity shock) times firm 1's Domar weight.

This table illustrates the difficulty of using regressions of sales on direct exposure to a shock to understand that shock's economic effect. The first three columns of the table describe the shock's effects when the share α of labor costs in total production costs is 0.5, for varying values of the firms' elasticity of substitution σ. Despite the shock having the same effect on real GDP in all three scenarios (-2%, per Hulten's theorem), the effect of the shock on sales that the researcher would recover (on average) when estimating equation (9) varies widely due to general equilibrium forces. When σ is below 1, a drop in the productivity of firm 1 even leads to an increase in its sales as a fraction of GDP. In columns 3 and 4, we compare the effects of the shock for two different values of α, keeping σ fixed. When α is lower, materials play a more important role in production, productivity shocks hitting a firm have larger spillover effects on other firms, and productivity shocks to one firm have larger effects on real GDP.

The effect of the shock on GDP can be recovered by applying Proposition 1 using Domar weights as responsiveness weights, and elements of the exposure matrix are given by

$$E_{ij} = \frac{\mathrm{d}\log\lambda_i}{\mathrm{d}\log A_j} = (\sigma-1)(1-\alpha)\left(\mathbf{1}_{i=j} + \frac{1-\alpha}{\alpha N}\right).$$

4.3.2 Representative agent and single factor

To expand on this simple example, now consider an economy with a representative agent who supplies a single factor (e.g., labor) inelastically to a fixed number of firms connected through arbitrary input–output linkages. We further assume that there are no distortions in the economy, and that all production and consumption nodes are CES, i.e., that

$$\frac{y_i}{\bar{y}_i} = \frac{A_i}{\bar{A}_i}\left(\sum_{j\in N,F}\omega_{ij}\left(\frac{x_{ij}}{\bar{x}_{ij}}\right)^{\frac{\sigma_i-1}{\sigma_i}}\right)^{\frac{\sigma_i}{\sigma_i-1}},$$

with variables with a bar representing pre-treatment values. In this scenario, the elements of the exposure matrix E are given by

$$E_{ij} = \frac{\mathrm{d}\log\lambda_i}{\mathrm{d}\log A_j} = \sum_{k\in C,N}\frac{\lambda_k}{\lambda_i}(\sigma_k-1)\,\mathrm{Cov}_{\Omega^{(k)}}\left(\Psi_{(i)},\Psi_{(j)}\right), \tag{10}$$

where σ_k is the elasticity of substitution between the inputs of node k, and $\mathrm{Cov}_{\Omega^{(k)}}\left(\Psi_{(i)},\Psi_{(j)}\right)$ is the input-output covariance operator, defined by:

$$\mathrm{Cov}_{\Omega^{(k)}}\left(\Psi_{(i)},\Psi_{(j)}\right) = \sum_{l\in N,F}\Omega_{kl}\Psi_{li}\Psi_{lj} - \left(\sum_{l\in N,F}\Omega_{kl}\Psi_{li}\right)\left(\sum_{l\in N,F}\Omega_{kl}\Psi_{lj}\right).$$

Hence, it corresponds to the covariance between the ith and jth columns of Ψ, using the kth row of Ω as weights.

Beyond the exposure matrix, the Leontief inverse also allows us to compute the effects of productivity shocks on prices, with $\frac{\mathrm{d}\log p_i}{\mathrm{d}\log A_j} = -\Psi_{ij} + \lambda_j$. Indeed, when a node j receives a positive productivity shock, the prices of goods downstream that use j as an input decrease to the extent of their exposure Ψ_{ij}. However, in general equilibrium, the shock also raises (real) factor prices, which increases the prices of all goods proportionally to λ_j.

4.3.3 Multiple consumers and factors

Expanding on the previous example, it is straightforward to add an arbitrary number of final consumers and factors of production to this model economy. To do so, define the network-adjusted consumption share of good i for agent c as $\lambda_i^c = \sum_{j \in N} \Omega_{cj} \Psi_{ji}$, reflecting the direct and indirect consumption of good i by consumer c. We can similarly compute Λ_f^c as the reliance of consumer c on factor f. Furthermore, let $\Phi_{cf} = \frac{w_f L_{cf}}{w_f L_f}$ denote the share of factor f's income accruing to consumer c. With these additional objects in hand, Baqaee and Farhi (2019) show that in efficient economies, we can express the elements of the exposure matrix as follows:

$$\frac{\mathrm{d}\log\lambda_i}{\mathrm{d}\log A_j} = \sum_{k \in C,N} \frac{\lambda_k}{\lambda_i} (\sigma_k - 1) \operatorname{Cov}_{\Omega^{(k)}} \left(\Psi_{(i)}, \Psi_{(j)} \right) - \sum_{f \in F} \sum_{k \in C,N} \frac{\lambda_k}{\lambda_i} (\sigma_k - 1) \operatorname{Cov}_{\Omega^{(k)}} \left(\Psi_{(i)}, \Psi_{(f)} \right) \frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log A_j} + \frac{1}{\lambda_i} \sum_{f \in F} \sum_{c \in C} (\lambda_i^c - \lambda_i) \Phi_{cf} \Lambda_f \frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log A_j}. \quad (11)$$

The first term of this equation is the same as in equation (10) and reflects how productivity shocks propagate through the input-output network, keeping relative factor prices fixed. The second term adjusts exposure for changes in relative factor prices caused by the shock. Finally, the third term provides a further adjustment to account for the changes in consumers' relative incomes.

To compute the exposure matrix using equation (11), we need to know how factor prices react to productivity shocks, as summarized by the $\frac{\mathrm{d}\log\Lambda}{\mathrm{d}\log A}$ matrix. This matrix can be obtained by solving the following system of equations for each value of j:

$$\frac{\mathrm{d}\log\Lambda_n}{\mathrm{d}\log A_j} = \sum_{k \in C,N} \frac{\lambda_k}{\Lambda_n} (\sigma_k - 1) \operatorname{Cov}_{\Omega^{(k)}} \left(\Psi_{(n)}, \Psi_{(j)} \right) - \sum_{f \in F} \sum_{k \in C,N} \frac{\lambda_k}{\Lambda_n} (\sigma_k - 1) \operatorname{Cov}_{\Omega^{(k)}} \left(\Psi_{(n)}, \Psi_{(f)} \right) \frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log A_j} + \frac{1}{\Lambda_n} \sum_{f \in F} \sum_{c \in C} (\Lambda_n^c - \Lambda_n) \Phi_{cf} \Lambda_f \frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log A_j}. \quad (12)$$

These expressions describing how productivity shocks affect relative factor prices allow us to describe their effects on the prices of factors and goods, with

$$\frac{\mathrm{d}\log w_f}{\mathrm{d}\log A_j} = \frac{\mathrm{d}\log \Lambda_f}{\mathrm{d}\log A_j} + \frac{\mathrm{d}\log Y}{\mathrm{d}\log A_j}, \qquad \frac{\mathrm{d}\log p_i}{\mathrm{d}\log A_j} = -\Psi_{ij} + \sum_{g\in F} \Psi_{ig} \frac{\mathrm{d}\log w_g}{\mathrm{d}\log A_j}. \tag{13}$$

4.3.4 Elastic factor supplies

While the expressions of the exposure measures derived above assume inelastic factor supplies, they can be extended to the case where factor supplies are elastic. For instance, Baqaee and Farhi (2020) show that in the case of a representative agent and in the absence of distortions, when the supply of each factor is a function of its price and the agent's income, i.e., when $L_f = G_f(w_f, Y)$, the elements of the exposure matrix are given by

$$\begin{aligned}\frac{\mathrm{d}\log \lambda_i}{\mathrm{d}\log A_k} &= \sum_j (\sigma_j - 1)\frac{\lambda_j}{\lambda_i}\mathrm{Cov}_{\Omega^{(j)}}\left(\Psi_{(k)}, \Psi_{(i)}\right) \\ &- \sum_j (\sigma_j - 1)\frac{\lambda_j}{\lambda_i}\mathrm{Cov}_{\Omega^{(j)}}\left(\sum_g \Psi_{(g)}\frac{1}{1+\zeta_g}\frac{\mathrm{d}\log \Lambda_g}{\mathrm{d}\log A_k} + \sum_g \Psi_{(g)}\frac{\gamma_g - \zeta_g}{1+\zeta_g}\frac{\mathrm{d}\log Y}{\mathrm{d}\log A_k}, \Psi_{(i)}\right),\end{aligned} \tag{14}$$

where $\zeta_g = \frac{\partial \log G_g}{\partial \log w_g}$ (resp., $\gamma_g = -\frac{\partial \log G_g}{\partial \log Y}$) is the elasticity of the supply of factor g to its wage (resp., to income). As before, the exposure of factor Domar weights to productivity shocks can be obtained by solving a system of equations:

$$\begin{aligned}\frac{\mathrm{d}\log \Lambda_f}{\mathrm{d}\log A_k} &= \sum_j (\sigma_j - 1)\frac{\lambda_j}{\Lambda_f}\mathrm{Cov}_{\Omega^{(j)}}\left(\Psi_{(k)}, \Psi_{(f)}\right) \\ &- \sum_j (\sigma_j - 1)\frac{\lambda_j}{\Lambda_f}\mathrm{Cov}_{\Omega^{(j)}}\left(\sum_g \Psi_{(g)}\frac{1}{1+\zeta_g}\frac{\mathrm{d}\log \Lambda_g}{\mathrm{d}\log A_k} + \sum_g \Psi_{(g)}\frac{\gamma_g - \zeta_g}{1+\zeta_g}\frac{\mathrm{d}\log Y}{\mathrm{d}\log A_k}, \Psi_{(f)}\right).\end{aligned} \tag{15}$$

When factor supplies are elastic, $\frac{\mathrm{d}\log Y}{\mathrm{d}\log A_k}$ is no longer given by Hulten's theorem. Instead, we have $\frac{\mathrm{d}\log Y}{\mathrm{d}\log A_k} = \varrho\left[\lambda_k - \sum_g \frac{1}{1+\zeta_g}\Lambda_g \frac{\mathrm{d}\log \Lambda_g}{\mathrm{d}\log A_k}\right]$, where $\varrho \equiv 1/\left(\sum_f \Lambda_f \frac{1+\gamma_f}{1+\zeta_f}\right)$.

These results can be extended to settings with multiple consumers and where the supply of a factor is given by the wages of all factors. This allows for the inclusion of phenomena such as migration and commuting—settings in which the labor supply in each location is elastic because of mobility.

4.3.5 Extensions

In Appendix B, we describe how to build the matrices $\frac{\mathrm{d}\log\lambda}{\mathrm{d}\log A}$ and $\frac{\mathrm{d}\log\Lambda}{\mathrm{d}\log A}$, describing first-order changes in Domar weights following productivity shocks in a wide range of additional settings, including inefficient economies in which there are arbitrary wedges μ between the costs of production of each good and the price at which it is sold. We also describe how to build exposure matrices when production and consumption nodes are not CES.

Finally, while we have focused so far on the effects of productivity shocks on Domar weights, the framework of Baqaee and Farhi (2019) allows us to compute how Domar weights shift in response to changes in factor supplies, to the factor ownership matrix, or to wedges. In Appendix B, we provide expressions for $\frac{\mathrm{d}\log\lambda}{\mathrm{d}\log L}$, $\frac{\mathrm{d}\log\lambda}{\mathrm{d}\log\Phi}$, and $\frac{\mathrm{d}\log\lambda}{\mathrm{d}\log\mu}$.

5 Application: the 2010 Chilean earthquake

The results described above suggest a simple procedure to estimate the effect on GDP of a treatment shifting the productivity of some segments of the economy.

Summary of the estimation procedure

- **Step 1**: **Compute exposure matrix E of nodes to shocks** using the results of Section 4.
- **Step 2**: **Compute the responsiveness weights κ** using equation (7).
- **Step 3**: **Compute the weighted exposure to treatment z**, as described in equation (5).
- **Step 4**: **Regress the changes in Domar weights on z and controls**. The coefficient on z is an estimate of the target of interest.

We now apply this methodology to evaluate the economic consequences of a major natural disaster: the earthquake that struck Chile on February 27, 2010.

5.1 Context and data

With a magnitude of 8.8, the 2010 Chilean earthquake was the seventh strongest earthquake ever recorded. For comparison, it was 500 times as powerful as the earthquake that struck Haiti in January 2010, and was associated with a tsunami and landslides. The Chilean earthquake caused 525 deaths, the destruction or severe damage of about 370,000 homes, and, based on estimates from insurance companies, inflicted total economic losses of between 15 and 30 billion USD. For reference, the GDP of Chile in 2010 was 217 billion USD. Following the earthquake, the Chilean government announced plans to spend 8.4 billion USD from 2010 to 2014 to assist the areas where the disaster had struck.

In this section, we apply the methodology described above to estimate the effect of this earthquake on the Chilean economy. We do so by using administrative data covering the universe of formal firms in Chile. Through tax forms, we can track their yearly sales, employment levels, and labor costs.[9] VAT data allows us to map firm-to-firm transactions and build an input-output matrix for the Chilean economy, which we represent in Appendix Figure S.1. In the spirit of the "disaggregated national accounts" described in Andersen, Huber, Johannesen, Straub and Vestergaard (2022), we build our input-output matrices so to be consistent with the figures of the national accounts tables published by the Chilean government, yet more detailed. We further gathered data on the allocation of government spending from the Ministry of Finance. Finally, we gathered from government decrees data on projects funded through the National Reconstruction Fund, established in the aftermath of the disaster through law 20.444 to facilitate reconstruction efforts. This data gives us information on the spatial distribution of the government transfers allocated following the shock. Appendix Figure S.2 shows the share of government spending allocated to the regions hit by the earthquake during our period of study.

Consistent with the model of the Chilean economy we develop below, we aggregate data on firms' economic activity at the region $\times$ sector level. There were 15 administrative regions in Chile at the time of the earthquake, and we group firms in 14 sectors (e.g.,

[9]To secure the privacy of workers and firms, the Central Bank of Chile mandates that the development, extraction and publication of the results should not allow the identification, directly or indirectly, of natural or legal persons. Officials of the Central Bank of Chile processed the disaggregated data. All the analysis was implemented by the authors and did not involve nor compromise the *Servicios de Impuestos Internos de Chile*. The information contained in the databases of the Chilean IRS is of a tax nature originating in self-declarations of taxpayers presented to the Service; therefore, the veracity of the data is not the responsibility of the Service.

mining; transportation and storage; financial and insurance activities). Following the earthquake, three regions were declared in a state of catastrophic emergency: O'Higgins, Maule, and Biobío (see Appendix Figure S.3). We consider these regions to be those "treated" by the earthquake.

5.2 Reduced-form evidence

To study the effect of the earthquake on economic activity, we start by estimating the following event study equation:

$$Y_{rst} = \sum_{\substack{\tau \in [-5,6] \\ \tau \neq -1}} \beta_\tau T_r \mathbb{1}_{t=\tau} + \gamma_{rs} + \delta_{st} + \varepsilon_{rst}, \tag{16}$$

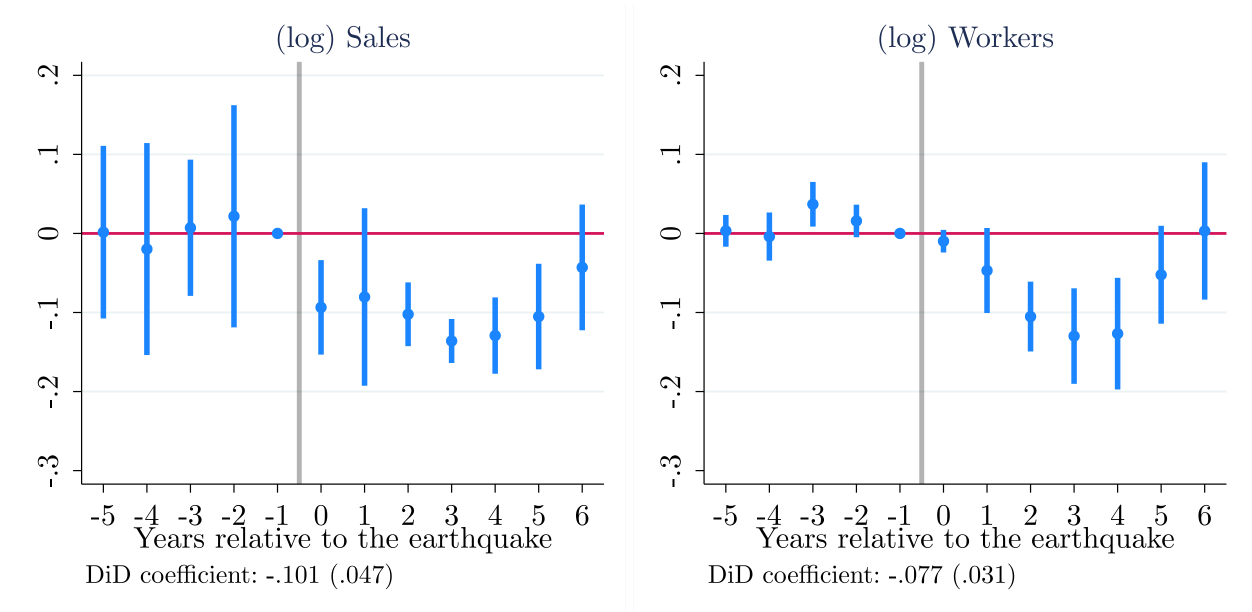
where Y_{rst} denotes an outcome of interest (such as sales or employment) for sector s in region r at time t (with t measured relative to the year of the earthquake, 2010), γ_{rs} is a region $\times$ sector fixed effect, and δ_{st} is a sector $\times$ year fixed effect. Standard errors are two-way clustered at the region and sector level, and observations are weighted proportionally to region $\times$ sector sales in the pre-period.

The results, shown in Figure 2, indicate that following the earthquake, sales in the treated segments of the economy (relative to the non-treated ones) immediately dropped by about 10 pp relative to trend, and continued to decline until 2013. Then, sales progressively returned to their pre-shock trend. The earthquake also appears to have caused a decline in affected firms' employment, with effects reaching -14 pp three years after the shock. This effect on employment, however, took longer to materialize, and there is no discernible effect in the year of the disaster. As with sales, employment levels eventually returned to trend.

5.3 A model of the Chilean economy

To measure the effect of the 2010 earthquake on the Chilean economy, we explicitly model the input-output links between each of its parts, which we operationalize as 210 nodes corresponding to each region $\times$ sector pair. In our model of the Chilean economy, we consider 15 representative households, each located in one of the country's regions and supplying their labor there. This labor supply corresponds to 15 separate factors in the economy. The government imposes a uniform tax of $\tau = 17\%$ on labor income. Because labor supply is inelastic, this tax is nondistortionary. The revenue raised through the tax is

Figure 2: Economic effects of the earthquake, event study



Notes: This figure shows event study estimates (following equation 16) of the effect of the earthquake on (log) sales and employment. We include one observation per region, sector, and year relative to the earthquake, control for region $\times$ sector and sector $\times$ year fixed effects, and cluster standard errors at the region and sector level. Observations are weighted proportionally to region $\times$ sector sales in the pre-period.

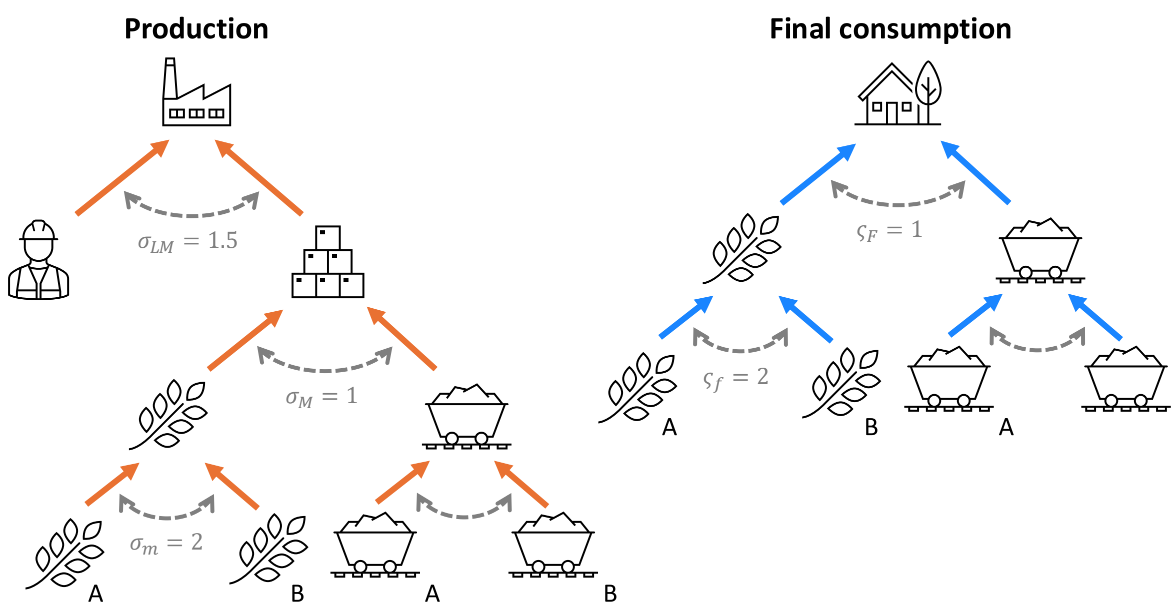
redistributed to households, with a share t_c of government revenue accruing to consumers in region c. We allow t_c to change over time; in particular, it will increase for the areas affected by the earthquake in the immediate aftermath of the disaster (and decrease for the unaffected regions). We do not consider any further taxes or markups in the economy, so this is an undistorted economy.

Government redistribution can be modeled in the framework of Section 4 by adding a wedge symbolizing the labor income tax as well as a fictitious factor to which tax income accrues, and then distributing the income from this fictitious factor across households. However, a convenient way to represent these government transfers is to directly incorporate them into the factor ownership matrix without adding wedges or a fictitious factor. In the absence of redistribution, the factor ownership matrix is simply the identity matrix: households in one region receive in full the income generated by their labor supply, and it is the only income they receive. Redistribution is equivalent to reassigning a share τ of the ownership of factors to households throughout the country, proportionally to t_c. In the context of our model, we can write $\Phi_{cf} = (1-\tau)\mathbb{1}_{c=f} + \tau t_c$.

Each of the production nodes in our model of the Chilean economy uses up to 211 different inputs: all of the goods produced by production nodes, and labor supplied in the region in which it is located. Firms' production is characterized by a nested aggregation of inputs, represented in Figure 3. At the highest level, firms combine labor and a composite bundle of intermediate inputs with an elasticity of substitution $\sigma_{LM} = 1.5$. The intermediate inputs bundle is, in turn, a Cobb-Douglas aggregate of composite inputs from different sectors. Finally, sector-specific inputs are CES aggregates of products from different regions, with an elasticity of substitution $\sigma_m = 2$.

Similarly, we assume that the representative household in each region consumes a bundle of all of the 210 goods produced in the economy, aggregated in a nested structure: the final consumption bundle is a Cobb-Douglas aggregate of composite goods from different sectors, and each sector-specific composite is a CES aggregate of goods from different regions, with elasticity of substitution $\varsigma_f = 2$.

Figure 3: Substitution patterns in our model of the Chilean economy



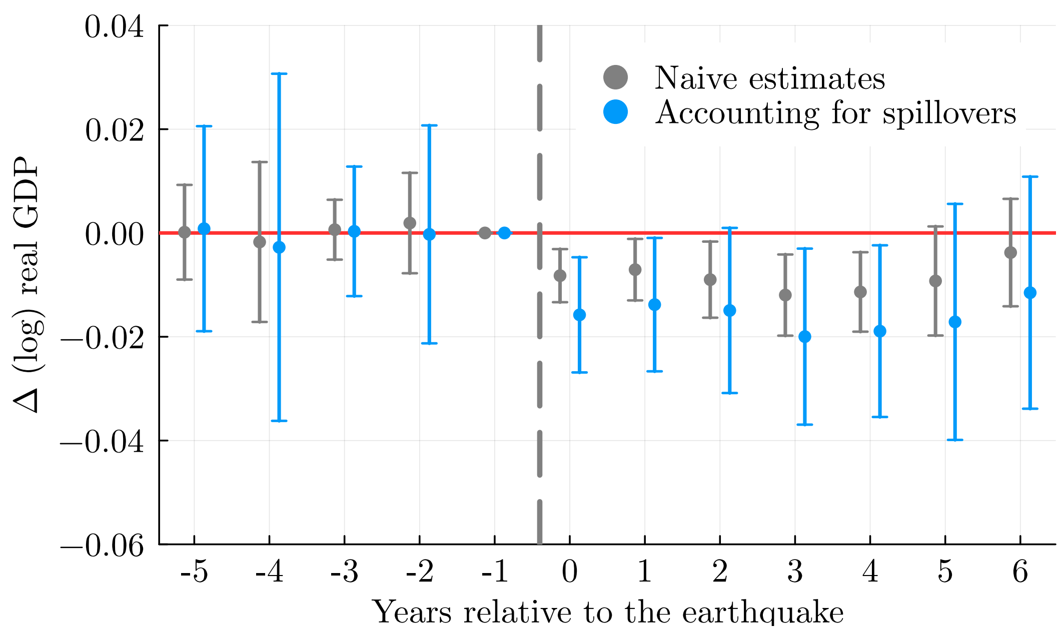
Notes: This figure illustrates substitution patterns in our model of the Chilean economy. Firms' production follows a nested input structure. At the top level, firms combine labor and a composite bundle of intermediate inputs with elasticity of substitution σ_{LM}. The intermediate bundle is a Cobb–Douglas aggregate of sectoral composites, and each sectoral input is a CES aggregate of regional varieties with elasticity σ_m. Similarly, households consume a nested bundle of goods: the final consumption bundle is a Cobb-Douglas aggregate across sectors, and each sectoral composite is a CES aggregate of regional goods with elasticity ς_f.

5.4 Results

In Sections 2–4, we focused on changes between two periods. To estimate the effect of the earthquake in its immediate aftermath, we can compare the year in which it took place (2010) with the last year before the shock (2009). To show its effects in the longer run, we can compare 2011, 2012, etc., with 2009. Furthermore, comparing 2009 with other years before the earthquake provides a set of placebo tests.

We benchmark our methodology against a more naive estimation that ignores spillover effects. In this specification, we regress changes in the (log) value added of a node on treatment and sector fixed effects, and multiply the estimated coefficient by the share of the affected regions in GDP. The results of these benchmark regressions and the estimation procedure we develop are shown in Figure 4. We find that accounting for spillover effects substantially increases the estimated magnitude of GDP losses caused by the disaster.

Figure 4: Estimated effect of the 2010 earthquake on Chilean real GDP



Notes: This figure reports estimates of θ—effects of the 2010 earthquake on Chile's real GDP—at different horizons. The naive estimation corresponds to coefficients of a regression of (log) value added on treatment and fixed effects, multiplied by the share of the treated regions in GDP. The alternative estimation strategy (accounting for spillovers) corresponds to the methodology described in this paper.

Reassuringly, all of the placebo coefficients that we estimate are close to zero and statistically insignificant, suggesting that the nodes hit by the earthquake and those that were not were on similar trends before the shock.

We can estimate the losses in real GDP caused by productivity drops in the six years following the earthquake by multiplying the estimated percentage-point loss associated with each year by that year's GDP. This back-of-the-envelope calculation yields an estimate of 23 billion USD in losses, a figure that is comparable to the estimates of 15 to 30 billion USD produced by insurance companies.

Taking spillovers into account in the setting studied here magnifies estimated economic effects relative to a more naive estimation strategy. There is no theoretical guarantee that this will be the case in general. Indeed, depending on the context, adjusting estimates for spillovers may either magnify or shrink them. In some cases, regressions that do not account for spillover effects may even fail to recover the correct sign of the treatment effect, as illustrated in Table 1.

5.5 Extensions

To provide a more credible estimation of the economic effects of the 2010 earthquake, we are in the process of enriching our model of the Chilean economy. In particular:

- While we have assumed elasticities of substitution in the Chilean economy to be known, a full estimation of the effects of the shock would involve estimating these elasticities. This can be achieved with data on prices. Indeed, within any CES nest with an elasticity of substitution σ, we have $\mathrm{d}\log(s_i/s_j) = -(\sigma - 1)\,\mathrm{d}\log(p_i/p_j)$, where s_i is the share of spending in the nest going to good i. Estimating θ requires exogenous variation in relative prices p_i/p_j. In our study, exposure to the earthquake can be used to provide a source of such variation.

- Until now, we have considered the economy of Chile to be closed, while its imports and exports represented about a third of its GDP at the beginning of the 2010s. Expanding our model to incorporate international trade would increase its accuracy. This can be done by adding the rest of the world as an outside location that trades with producers and consumers in Chile.

- We have modeled the Chilean economy as efficient. Data on measures of distortions (e.g. taxes and markups) in the economy can be easily included in the model to evaluate how these wedges interacted with the shock (i.e. inform the reallocation term in 7).

- Finally, in our evaluation of the effects of the earthquake, we have focused on the effects of the shock on real GDP stemming from productivity losses. However, the earthquake may have further affected real GDP through migration or the reallocation of government resources. Slight adjustments to our methodology can account for these additional consequences of the disaster.

6 Conclusion

In recent decades, applied researchers have developed a range of tools that allow them to exploit quasi-experimental variation to credibly estimate the economic effects of a "treatment". These methods compare treated units (such as regions in which the treatment occurred) with appropriate controls (such as other regions) and usually rely on the assumption that the treatment of one unit has no spillover effects on other units—an assumption often expressed as the Stable Unit Treatment Value Assumption, or SUTVA. While this assumption may be innocuous in some studies, spillovers are a natural feature of most economic settings because of the numerous trade, financial, and labor flows that connect different regions and segments of the economy.

This paper develops a methodology that provides a correction to the bias caused by such spillovers. It is designed to offer a middle ground between typical adjustments to reduced-form approaches (such as "buffer" methods) and full-fledged model simulations. It retains the transparency and ease of use of standard policy evaluation tools while leveraging increasingly accessible data on economic linkages to measure the extent of spillovers and appropriately account for them in estimation.

We have applied this new methodology in the context of a major natural disaster and recover an estimate of its aggregate cost. Such measures of the economic effect of environmental damage are of important economic interest, as they inform mitigation policies ex-ante. Naturally, an important area of further environmental applications is to natural disasters whose likelihood is evolving due to climate change. However, our approach can also be used to study a wide range of wider policy and evaluation settings where spillover effects are suspected to be substantial.

References

Acemoglu, Daron, Vasco M Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "The Network Origins of Aggregate Fluctuations," *Econometrica*, 2012, *80* (5), 1977–2016.

Adao, Rodrigo, Costas Arkolakis, and Federico Esposito, "General Equilibrium Effects in Space: Theory and Measurement," 2019.

Andersen, Asger L, Kilian Huber, Niels Johannesen, Ludwig Straub, and Emil Toft Vestergaard, "Disaggregated Economic Accounts," *National Bureau of Economic Research Working Paper*, 2022.

Arkhangelsky, Dmitry and Wisse Rutgers, "Evaluating Local Policies in Centralized Markets," 2025. Working paper.

Aronow, Peter M and Cyrus Samii, "Estimating Average Causal Effects Under General Interference, with Application to a Social Network Experiment," *The Annals of Applied Statistics*, 2017, *11* (4), 1912.

Baqaee, David Rezza and Emmanuel Farhi, "Macroeconomics with Heterogeneous Agents and Input-Output Networks," 2019.

_ and _ , "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 2020, *135* (1), 105–163.

Barrot, Jean-Noël and Julien Sauvagnat, "Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks," *The Quarterly Journal of Economics*, 2016, *131* (3), 1543–1592.

Boehm, Christoph E, Aaron Flaaen, and Nitya Pandalai-Nayar, "Input Linkages and the Transmission of Shocks: Firm-Level Evidence from the 2011 Tōhoku earthquake," *Review of Economics and Statistics*, 2019, *101* (1), 60–75.

Borusyak, Kirill and Peter Hull, "Non-Random Exposure to Exogenous Shocks: Theory and Applications," 2020.

_ , Rafael Dix-Carneiro, and Brian Kovak, "Understanding migration responses to local shocks," 2022. Working paper.

Botzen, WJ Wouter, Olivier Deschenes, and Mark Sanders, "The Economic Impacts of Natural Disasters: A Review of Models and Empirical Studies," *Review of Environmental Economics and Policy*, 2019.

Boustan, Leah Platt, Matthew E Kahn, Paul W Rhode, and Maria Lucia Yanguas, "The Effect of Natural Disasters on Economic Activity in US Counties: A Century of Data," *Journal of Urban Economics*, 2020, *118*, 103257.

Bradford, David F, "Factor Prices May Be Constant but Factor Returns Are Not," *Economics Letters*, 1978, *1* (3), 199–203.

Cai, Jing and Adam Szeidl, "Indirect effects of access to finance," *American Economic Review*, 2024, *114* (8), 2308–2351.

Carvalho, Vasco M, Makoto Nirei, Yukiko U Saito, and Alireza Tahbaz-Salehi, "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake," *The Quarterly Journal of Economics*, 2021, *136* (2), 1255–1321.

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy, and Juan Pantano, "Catastrophic Natural Disasters and Economic Growth," *Review of Economics and Statistics*, 2013, *95* (5), 1549–1561.

Cox, David Roxbee, "Planning of Experiments," 1958.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment," *The Quarterly Journal of Economics*, 2013, *128* (2), 531–580.

de Andrade Lima, Ricardo Carvalho and Antonio Vinícius Barros Barbosa, "Natural Disasters, Economic Growth and Spatial Spillovers: Evidence From a Flash Flood in Brazil," *Papers in Regional Science*, 2019, *98* (2), 905–924.

Deryugina, Tatyana, "The Fiscal Cost of Hurricanes: Disaster Aid Versus Social Insurance," *American Economic Journal: Economic Policy*, 2017, *9* (3), 168–98.

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

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker, "General equilibrium effects of cash transfers: experimental evidence from Kenya," *Econometrica*, 2022, *90* (6), 2603–2643.

Faridani, Stefan and Paul Niehaus, "Linear Estimation of Global Average Treatment Effects," 2024.

Felbermayr, Gabriel J, Jasmin Gröschl, Mark Sanders, Vincent Schippers, and Thomas Steinwachs, "Shedding Light on the Spatial Diffusion of Disasters," *Available at SSRN 3237340*, 2018.

Franklin, Simon, Clement Imbert, Girum Abebe, and Carolina Mejia-Mantilla, "Urban Public Works in Spatial Equilibrium: Experimental Evidence from Ethiopia," *American Economic Review*, 2024, *114* (5), 1382–1414.

Giovanni, Julian Di, Andrei A Levchenko, and Isabelle Mejean, "Firms, Destinations, and Aggregate Fluctuations," *Econometrica*, 2014, *82* (4), 1303–1340.

Hornbeck, Richard and Enrico Moretti, "Who Benefits From Productivity Growth? Direct and Indirect Effects of Local TFP Growth on Wages, Rents, and Inequality," 2018.

Hsiang, Solomon M and Amir S Jina, "The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth: Evidence From 6,700 Cyclones," 2014.

Hudgens, Michael G and M Elizabeth Halloran, "Toward Causal Inference with Interference," *Journal of the American Statistical Association*, 2008, *103* (482), 832–842.

Hulten, Charles R, "Growth Accounting with Intermediate Inputs," *The Review of Economic Studies*, 1978, *45* (3), 511–518.

Korovkin, Vasily and Alexey Makarin, "Production Networks and War: Evidence from Ukraine," 2022. Working paper.

Leung, Michael P, "Treatment and Spillover Effects Under Network Interference," *Review of Economics and Statistics*, 2020, *102* (2), 368–380.

Long, John B and Charles I Plosser, "Real Business Cycles," *Journal of Political Economy*, 1983, *91* (1), 39–69.

Manski, Charles F, "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

Miguel, Edward and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 2004, *72* (1), 159–217.

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg, "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 2018, *108* (12), 3855–90.

Munro, Evan, Xu Kuang, and Stefan Wager, "Treatment Effects in Market Equilibrium," *American Economic Review*, 2025, *115* (10), 3273–3321.

Redding, Stephen and Anthony J Venables, "Economic geography and international inequality," *Journal of international Economics*, 2004, *62* (1), 53–82.

Rotemberg, Martin, "Equilibrium effects of firm subsidies," *American Economic Review*, 2019, *109* (10), 3475–3513.

Rubin, Donald B, "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.," *Journal of Educational Psychology*, 1974, *66* (5), 688.

_ , "Randomization Analysis of Experimental Data: The Fisher Randomization Test, Comment," *Journal of the American Statistical Association*, 1980, *75* (371), 591–593.

Sävje, Fredrik, Peter Aronow, and Michael Hudgens, "Average treatment effects in the presence of unknown interference," *Annals of statistics*, 2021, *49* (2), 673.

Strobl, Eric, "The Economic Growth Impact of Hurricanes: Evidence From US Coastal Counties," *Review of Economics and Statistics*, 2011, *93* (2), 575–589.

Zigler, Corwin M and Georgia Papadogeorgou, "Bipartite Causal Inference with Interference," *Statistical Science*, 2020, *36* (1), 109.

Appendix

A Proofs and additional theoretical results

A.1 Proof of Proposition 1

Unbiasedness. By the Frisch–Waugh–Lovell theorem, the OLS coefficient on z in regression (6) equals

$$\widehat{\theta} = \frac{z^\top M_X y}{z^\top M_X z}.$$

With $w = \left(\kappa_T^\top G^{-1} \kappa_T\right)^{-1} G^{-1} \kappa_T$,

$$z^\top M_X z = w^\top G w = \frac{\kappa_T^\top G^{-1} \kappa_T}{\left(\kappa_T^\top G^{-1} \kappa_T\right)^2} = \frac{1}{\kappa_T^\top G^{-1} \kappa_T},$$

and

$$\widehat{\theta} = \left(\kappa_T^\top G^{-1} E_T^\top M_X\right) y.$$

Using $y = E_T \beta_T + X\gamma + \varepsilon$ and $M_X X = 0$,

$$\begin{aligned}
\widehat{\theta} &= \kappa_T^\top G^{-1} E_T^\top M_X E_T \beta_T + \kappa_T^\top G^{-1} E_T^\top M_X \varepsilon \\
&= \kappa_T^\top \beta_T + \kappa_T^\top G^{-1} E_T^\top M_X \varepsilon \\
&= \theta + \kappa_T^\top G^{-1} E_T^\top M_X \varepsilon.
\end{aligned}$$

Under Assumption 1, $M_X \mathbb{E}[\varepsilon \mid X, T] = 0$, hence

$$\mathbb{E}\left[\widehat{\theta} \mid X, T\right] = \theta.$$

So $\widehat{\theta}$ is unbiased.

Consistency. As ε^a and ε^y are orthogonal, $\operatorname{Var}(\varepsilon \mid X, E) = E\,\Sigma_a\,E^\top + \Sigma_y$. Then, as $\hat{\theta} - \theta = \kappa_T^\top G^{-1} E_T^\top M_X \varepsilon$,

$$\begin{aligned}
\operatorname{Var}(\hat{\theta} - \theta \mid X, E_T) &= (\kappa_T^\top G^{-1} E_T^\top M_X) \operatorname{Var}(\varepsilon \mid X, E)(M_X E_T G^{-1} \kappa_T) \\
&= \kappa_T^\top G^{-1} E_T^\top M_X E \Sigma_a E^\top M_X E_T G^{-1} \kappa_T + \kappa_T^\top G^{-1} E_T^\top M_X \Sigma_y M_X E_T G^{-1} \kappa_T.
\end{aligned}$$

The first term converges to zero. Indeed, set $A = E\Sigma_a E^\top$ and $B = M_X E_T G^{-1/2}$ ($G^{-1/2}$ exists because G is positive definite). A is positive semidefinite. Indeed, Σ_a is positive semidefinite because it is a covariance matrix, and for any vector v, $v^\top E \Sigma_a E^\top v = \left(E^\top v\right)^\top \Sigma_a \left(E^\top v\right) \geq 0$. Furthermore, as $G = E_T^\top M_X E_T$, we have $B^\top B = I$. Let $x = G^{-1/2} \kappa_T$.

Then

$$
\begin{aligned}
\kappa_T^\top G^{-1} E_T^\top M_X E \Sigma_a E^\top M_X E_T G^{-1} \kappa_T &= \kappa_T^\top G^{-1} E_T^\top M_X A M_X E_T G^{-1} \kappa_T \\
&= x^\top B^\top A B x \\
&\le \|A\|_{\mathrm{op}}\, x^\top B^\top B x \\
&= \|A\|_{\mathrm{op}}\, \kappa_T^\top G^{-1} \kappa_T.
\end{aligned}
$$

By Assumption 3(ii), $\|A\|_{\mathrm{op}} \le \|E\|_{\mathrm{op}}^2 \|\Sigma_a\|_{\mathrm{op}} \le C_E^2 C_a < \infty$. By Assumption 3(i), $\kappa_T^\top G^{-1}\kappa_T \to_p 0$. Hence the product converges to zero in probability by Slutsky's theorem.

The second term also converges to zero, as

$$
\begin{aligned}
\kappa_T^\top G^{-1} E_T^\top M_X \Sigma_y M_X E_T G^{-1} \kappa_T &\le \|\Sigma_y\|_{\mathrm{op}}\, \|M_X E_T G^{-1} \kappa_T\|_2^2 \\
&= \|\Sigma_y\|_{\mathrm{op}} \kappa_T^\top G^{-1} \kappa_T \\
&\to_p 0.
\end{aligned}
$$

Conditional Chebyshev yields $\hat{\theta} - \theta \to_p 0$ given (X, E_T).

A.2 Efficiency

Proposition 2 (BLUE under homoskedasticity)**.** *Under homoskedasticity and Assumptions 1-2, the estimator*

$$
\widehat{\theta} = \kappa_T^\top \widehat{\beta}_T = \kappa_T^\top (E_T^\top M_X E_T)^{-1} E_T^\top M_X y \tag{S.1}
$$

is the Best Linear Unbiased Estimator (BLUE) of $\theta = \kappa_T^\top \beta_T$ among all linear unbiased estimators that are functions of (X, T). Moreover,

$$
\widehat{\theta} = a^{\star\top} y \quad \textit{with} \quad a^\star = M_X E_T (E_T^\top M_X E_T)^{-1} \kappa_T, \tag{S.2}
$$

and

$$
\mathrm{Var}(\widehat{\theta} \mid X, T) = \sigma^2 \kappa_T^\top (E_T^\top M_X E_T)^{-1} \kappa_T, \tag{S.3}
$$

where $\mathrm{Var}(\varepsilon \mid X, T) = \sigma^2 I_N$*.*

Proof. Consider linear estimators of the form $\tilde{\theta} = a^\top y$ with $a = a(X, T) \in \mathbb{R}^N$. Unbiasedness for all γ and β_T requires

$$
a^\top X = 0 \qquad \text{and} \qquad a^\top E_T = \kappa_T^\top, \tag{S.4}
$$

since $\mathbb{E}[y \mid X, T] = E_T \beta_T + X\gamma$. Under homoskedasticity, $\mathrm{Var}(\tilde{\theta} \mid X, T) = \sigma^2 a^\top a$. Thus, the BLUE estimator can be found by solving the following problem:

$$
\min_{a \in \mathbb{R}^n} a^\top a \quad \text{s.t.} \quad a^\top X = 0, a^\top E_T = \kappa_T^\top.
$$

Form the Lagrangian $\mathcal{L}(a, \mu, \nu) = a^\top a + \mu^\top (a^\top E_T - \kappa_T^\top) + \nu^\top a^\top X$. The first-order condition is $2a + E_T\mu + X\nu = 0$, so

$$a = -\tfrac{1}{2}(E_T\mu + X\nu). \tag{S.5}$$

Premultiplying by $X^\top$ and imposing $a^\top X = 0$ yields $X^\top X\nu = -X^\top E_T\mu$, hence $\nu = -(X^\top X)^{-1}X^\top E_T\mu$. Substituting into (S.5) gives $a = -\frac{1}{2}\left[I - X(X^\top X)^{-1}X^\top\right]E_T\mu = -\frac{1}{2}M_X E_T\mu$. Let $b \equiv -\frac{1}{2}\mu$. Then any optimizer has the form

$$a = M_X E_T b. \tag{S.6}$$

Imposing the second unbiasedness restriction in (S.4) gives $\kappa_T^\top = a^\top E_T = b^\top E_T^\top M_X E_T = b^\top G$. By Assumption 2, G is invertible, so the unique solution is $b = G^{-1}\kappa_T$. Plugging this into (S.6) yields

$$a^\star = M_X E_T G^{-1}\kappa_T,$$

which proves (S.2). The variance at the optimizer is $\mathrm{Var}(\widehat{\theta} \mid X, T) = \sigma^2 a^{\star\top}a^\star = \sigma^2\kappa_T^\top G^{-1}\kappa_T$, which is (S.3). □

Proposition 3 (GLS estimator)**.** *Assume a known error covariance,* $\mathrm{Var}(\varepsilon \mid X, T) = \Omega$, *where* Ω *is symmetric positive definite and known. Define the* Ω*-weighted residual-maker and Gram matrix*

$$M_{X,\Omega} \equiv I - X(X^\top\Omega^{-1}X)^{-1}X^\top\Omega^{-1}, \qquad G_\Omega \equiv E_T^\top\Omega^{-1}M_{X,\Omega}E_T. \tag{S.7}$$

Assume $X^\top\Omega^{-1}X$ *and* G_Ω *are nonsingular. Define*

$$w_\Omega \equiv \frac{G_\Omega^{-1}\kappa_T}{\kappa_T^\top G_\Omega^{-1}\kappa_T}, \qquad z_\Omega \equiv E_T w_\Omega. \tag{S.8}$$

Let $\widehat{\theta}_{GLS}$ *be the GLS coefficient on* z_Ω *from*

$$y = \theta z_\Omega + X\gamma + \varepsilon, \quad \textit{estimated with weight matrix } \Omega^{-1}. \tag{S.9}$$

It is equal to

$$\widehat{\theta}_{\mathrm{GLS}} = \left(z_\Omega^\top\Omega^{-1}M_{X,\Omega}z_\Omega\right)^{-1}z_\Omega^\top\Omega^{-1}M_{X,\Omega}y. \tag{S.10}$$

$\widehat{\theta}_{\mathrm{GLS}}$ *is finite-sample unbiased for* $\theta = \kappa_T^\top\beta_T$ *and is BLUE among linear unbiased estimators of* θ*. Moreover,*

$$\mathrm{Var}\left(\widehat{\theta}_{\mathrm{GLS}} \mid X, T\right) = \kappa_T^\top G_\Omega^{-1}\kappa_T. \tag{S.11}$$

In the special case of WLS with weights ω, $\Omega = \sigma^2 W^{-1}$, *where* $W = \mathrm{diag}(\omega_i)$.

Proof. Set $y^\star = \Omega^{-1/2}y$, $X^\star = \Omega^{-1/2}X$, $E_T^\star = \Omega^{-1/2}E_T$, and $\varepsilon^\star = \Omega^{-1/2}\varepsilon$. Then $\mathrm{Var}(\varepsilon^\star \mid X, T) = I$. With $M_{X^\star} = I - X^\star(X^{\star\top}X^\star)^{-1}X^{\star\top}$,

$$G^\star \equiv E_T^{\star\top}M_{X^\star}E_T^\star = E_T^\top\Omega^{-1}\Big(I - X(X^\top\Omega^{-1}X)^{-1}X^\top\Omega^{-1}\Big)E_T = G_\Omega. \tag{S.12}$$

Apply the homoskedastic result to the transformed problem $(y^\star, X^\star, E_T^\star)$, with weight w_Ω from (S.8) (since $G^\star = G_\Omega$). OLS on $y^\star = \theta z_\Omega^\star + X^\star\gamma + \varepsilon^\star$, where $z_\Omega^\star \equiv E_T^\star w_\Omega$, is unbiased and BLUE. GLS on (S.9) equals that OLS, so $\widehat{\theta}_{\text{GLS}}$ is unbiased and BLUE. Finally, in the transformed regression with unit error variance,

$$\begin{aligned}\operatorname{Var}\left(\widehat{\theta}_{\text{GLS}} \mid X, T\right) &= \left(z_{\Omega}^{\star\top} M_{X^\star} z_{\Omega}^{\star}\right)^{-1} = \left(w_{\Omega}^{\top} G^{\star} w_{\Omega}\right)^{-1} \\ &= \left(\frac{\kappa_T^{\top} G_{\Omega}^{-1} \kappa_T}{\left(\kappa_T^{\top} G_{\Omega}^{-1} \kappa_T\right)^2}\right)^{-1} = \kappa_T^{\top} G_{\Omega}^{-1} \kappa_T,\end{aligned}$$

using $w_{\Omega} = G_{\Omega}^{-1}\kappa_T / (\kappa_T^{\top} G_{\Omega}^{-1} \kappa_T)$ and $G^{\star} = G_{\Omega}$. □

B Measures of exposure

In this section, we describe how to extend the results of Section 4 to a larger class of models, including inefficient economies and economies where production and consumption nodes are not CES. We also show how to compute first-order changes to Domar weights following shocks to wedges, factor supplies, and the factor ownership matrix. These results are either derived in Baqaee and Farhi (2019) or are natural extensions of their framework.

B.1 Inefficient economies

The profits generated by wedges in the economy can be viewed as income from an additional factor that Baqaee and Farhi (2019) describe as "fictitious". In what follows, we denote the set of "real" factors by F and the set of all factors (both real and fictitious) by F^*.

In the presence of distortions, elements of the exposure matrix E can be expressed as:

$$\begin{aligned} E_{ij} \equiv \frac{\mathrm{d}\log\lambda_i}{\mathrm{d}\log A_j} &= \sum_{k\in C,N} \frac{\lambda_k}{\lambda_i}\frac{(\sigma_k - 1)}{\mu_k} \operatorname{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(i)}, \tilde{\Psi}_{(j)}\right) \\ &\quad - \sum_{f\in F}\sum_{k\in C,N} \frac{\lambda_k}{\lambda_i}\frac{(\sigma_k - 1)}{\mu_k} \operatorname{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(i)}, \tilde{\Psi}_{(f)}\right) \frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log A_j} \\ &\quad + \frac{1}{\lambda_i}\sum_{f\in F^*}\sum_{c\in C}\left(\lambda_i^c - \lambda_i\right)\Phi_{cf}\Lambda_f \frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log A_j} \end{aligned} \tag{S.13}$$

Again, computing these entries necessitates the inversion of the following F^*-dimensional system of equations for different values of j to obtain the $\frac{\mathrm{d}\log\Lambda}{\mathrm{d}\log A}$ matrix.

$$\begin{aligned}\frac{\mathrm{d}\log\Lambda_n}{\mathrm{d}\log A_j} = \sum_{k\in C,N}\frac{\lambda_k}{\Lambda_n}\frac{(\sigma_k-1)}{\mu_k}\mathrm{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(n)},\tilde{\Psi}_{(j)}\right)\\ -\sum_{f\in F}\sum_{k\in C,N}\frac{\lambda_k}{\Lambda_n}\frac{(\sigma_k-1)}{\mu_k}\mathrm{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(n)},\tilde{\Psi}_{(f)}\right)\frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log A_j}\\ +\frac{1}{\Lambda_n}\sum_{f\in F^*}\sum_{c\in C}\left(\Lambda_n^c-\Lambda_n\right)\Phi_{cf}\Lambda_f\frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log A_j}\end{aligned} \tag{S.14}$$

B.2 Shocks to factor supplies

While Baqaee and Farhi (2019) focus on the effects of productivity shocks, their results can be extended to other types of shocks. Consider, for instance, changes to the supply of a given factor (e.g., inflows of capital through FDI, increases in the amount of agricultural land through deforestation, or increases in the labor supply following migration shocks). The following equations give the first-order effects of such changes in factor supply:

$$\begin{aligned}\frac{\mathrm{d}\log\lambda_i}{\mathrm{d}\log L_j} = \sum_{k\in C,N}\frac{\lambda_k}{\lambda_i}\frac{(\sigma_k-1)}{\mu_k}\mathrm{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(i)},\tilde{\Psi}_{(j)}\right)\\ -\sum_{f\in F}\sum_{k\in C,N}\frac{\lambda_k}{\lambda_i}\frac{(\sigma_k-1)}{\mu_k}\mathrm{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(i)},\tilde{\Psi}_{(f)}\right)\frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log L_j}\\ +\frac{1}{\lambda_i}\sum_{f\in F^*}\sum_{c\in C}\left(\lambda_i^c-\lambda_i\right)\Phi_{cf}\Lambda_f\frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log L_j}\end{aligned} \tag{S.15}$$

$$\begin{aligned}\frac{\mathrm{d}\log\Lambda_n}{\mathrm{d}\log L_j} = \sum_{k\in C,N}\frac{\lambda_k}{\Lambda_n}\frac{(\sigma_k-1)}{\mu_k}\mathrm{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(n)},\tilde{\Psi}_{(j)}\right)\\ -\sum_{f\in F}\sum_{k\in C,N}\frac{\lambda_k}{\Lambda_n}\frac{(\sigma_k-1)}{\mu_k}\mathrm{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(n)},\tilde{\Psi}_{(f)}\right)\frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log L_j}\\ +\frac{1}{\Lambda_n}\sum_{f\in F^*}\sum_{c\in C}\left(\Lambda_n^c-\Lambda_n\right)\Phi_{cf}\Lambda_f\frac{\mathrm{d}\log\Lambda_f}{\mathrm{d}\log L_j}\end{aligned} \tag{S.16}$$

B.3 Shocks to the allocation of factor income

Similarly, the framework of Baqaee and Farhi (2019) allows us to study shocks to the factor ownership matrix Φ (e.g., because of a shock to the allocation of government spending). The first-order effect of a shock to an entry Φ_{dg} of the factor ownership matrix is given by the following equations:

$$\frac{\mathrm{d}\log\lambda_i}{\mathrm{d}\Phi_{cf}} = -\sum_{g\in F}\sum_{k\in C,N}\frac{\lambda_k}{\lambda_i}\frac{(\sigma_k-1)}{\mu_k}\operatorname{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(i)},\tilde{\Psi}_{(g)}\right)\frac{\mathrm{d}\log\Lambda_g}{\mathrm{d}\Phi_{cf}} + \lambda_i^c\frac{\Lambda_f}{\lambda_i} \tag{S.17}$$

$$\frac{\mathrm{d}\log\Lambda_n}{\mathrm{d}\Phi_{cf}} = -\sum_{g\in F}\sum_{k\in C,N}\frac{\lambda_k}{\Lambda_n}\frac{(\sigma_k-1)}{\mu_k}\operatorname{Cov}_{\tilde{\Omega}^{(k)}}\left(\Psi_{(i)},\tilde{\Psi}_{(g)}\right)\frac{\mathrm{d}\log\Lambda_g}{\mathrm{d}\Phi_{cf}} + \Lambda_n^c\frac{\Lambda_g}{\Lambda_n} \tag{S.18}$$

B.4 Shocks to wedges

Changes to the vector of wedges μ also affect Domar weights. To first order, we have

$$\begin{aligned}\frac{\mathrm{d}\log\lambda_i}{\mathrm{d}\log\mu_j} = &\sum_{k\in C,N}(1-\sigma_k)\frac{\lambda_k}{\lambda_i}\mu_k^{-1}\operatorname{Cov}\left(\tilde{\Psi}_{(j)},\Psi_{(i)}\right)\\ &+\sum_{g\in F}\sum_{k\in C,N}(1-\sigma_k)\frac{\lambda_k}{\lambda_i}\mu_k^{-1}\operatorname{Cov}\left(\tilde{\Psi}_g,\Psi_{(i)}\right)\frac{\mathrm{d}\log\Lambda_g}{\mathrm{d}\log\mu_j}\\ &-\frac{\lambda_j}{\lambda_i}\left(\Psi_{ji}-\mathbb{1}_{i=j}\right)+\sum_{c\in C}\sum_{g\in F^*}\chi_c\Phi_{cg}\Lambda_g\frac{\lambda_i^c}{\lambda_i}\frac{\mathrm{d}\log\Lambda_g}{\mathrm{d}\log\mu_j},\end{aligned} \tag{S.19}$$

where $\chi_c = \frac{\sum_{i=1}^N p_i c_{ci}}{\sum_{j=1}^N\sum_{d=1}^C p_j c_{di}} = \sum_{f\in F^*}\Phi_{cf}\Lambda_f$ is consumer c's share in aggregate revenue. Changes in factors' Domar weights $\frac{\mathrm{d}\log\Lambda_g}{\mathrm{d}\log\mu_j}$ are given for real factors by solving the following system of equations:

$$\begin{aligned}\frac{\mathrm{d}\log\Lambda_g}{\mathrm{d}\log\mu_j} = &\sum_{k\in C,N}(1-\sigma_k)\frac{\lambda_k}{\Lambda_g}\mu_k^{-1}\operatorname{Cov}\left(\tilde{\Psi}_{(j)},\Psi_{(g)}\right)\\ &+\sum_{g\in F}\sum_{k\in C,N}(1-\sigma_k)\frac{\lambda_k}{\Lambda_g}\mu_k^{-1}\operatorname{Cov}\left(\tilde{\Psi}_g,\Psi_{(g)}\right)\frac{\mathrm{d}\log\Lambda_g}{\mathrm{d}\log\mu_j}\\ &-\frac{\lambda_j}{\Lambda_g}\left(\Psi_{jg}-\mathbb{1}_{g=j}\right)+\frac{1}{\Lambda_g}\sum_{c\in C}\sum_{g\in F^*}\chi_c\Phi_{cg}\Lambda_g\left(\Lambda_g^c-\Lambda_g\right)\frac{\mathrm{d}\log\Lambda_g}{\mathrm{d}\log\mu_j},\end{aligned} \tag{S.20}$$

and for fictitious factors by

$$\frac{\mathrm{d}\log\Lambda_{i^*}}{\mathrm{d}\log\mu_j} = \frac{\mathrm{d}\log\lambda_i}{\mathrm{d}\log\mu_j} + \frac{1}{\mu_i-1}\mathbb{1}_{i=j}, \tag{S.21}$$

where Λ_{i^*} is the fictitious factor associated with the markup placed on the ith good.

B.5 Non-CES production and consumption nodes

While we have focused so far on consumption and production nodes that aggregate inputs with a constant elasticity of substitution, the formulas to compute the exposure matrix can be adjusted to accommodate arbitrary production functions. To do so, parameters of the form $(\sigma_j - 1)\,\mathrm{Cov}_{\Omega^{(j)}}\left(\Psi_{(k)}, \Psi_{(l)}\right)$ need to be replaced with the input-output substitution operator that Baqaee and Farhi (2019) define as

$$\Xi_j\left(\Psi_{(k)}, \Psi_{(l)}\right) = -\sum_{x,y} \Omega_{jx}\left[\mathbb{1}_{xy} + \Omega_{jy}\left(\vartheta_j(x,y) - 1\right)\right]\Psi_{xk}\Psi_{yl}. \tag{S.22}$$

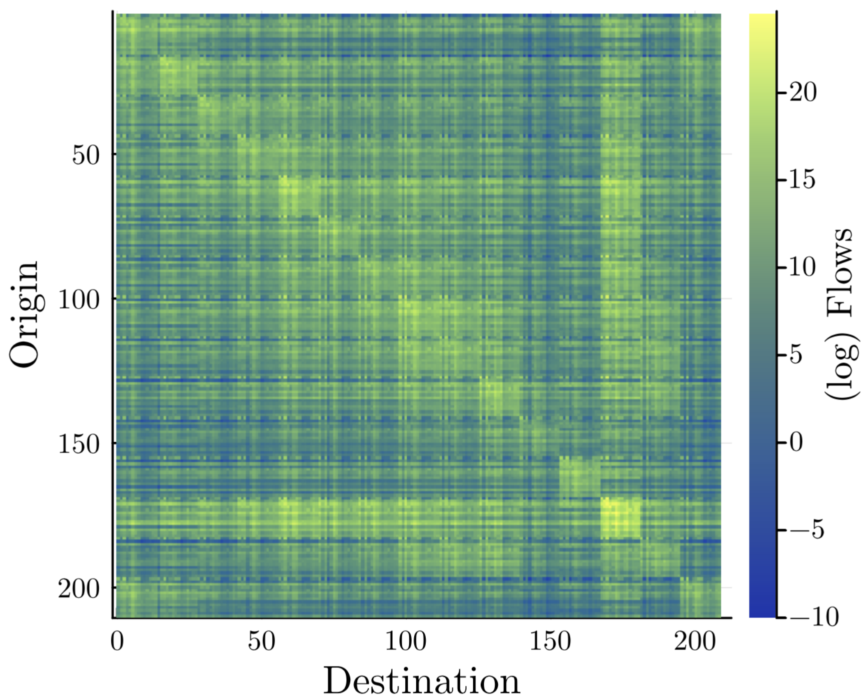
In this expression, $\vartheta_j(x,y)$ is the Allen-Uzawa elasticity of substitution between inputs x and y. Given the cost function C_j of producer j, this elasticity is given by

$$\vartheta_j(x,y) = \frac{C_j \frac{\mathrm{d}^2 C_j}{\mathrm{d}p_x \mathrm{d}p_y}}{\frac{\mathrm{d}C_j}{\mathrm{d}p_x} \cdot \frac{\mathrm{d}C_j}{\mathrm{d}p_y}} = \frac{\epsilon_j(x,y)}{\Omega_{jy}}, \tag{S.23}$$

where $\epsilon_j(x,y)$ denotes the elasticity of producer j's demand for input x with respect to p_y.

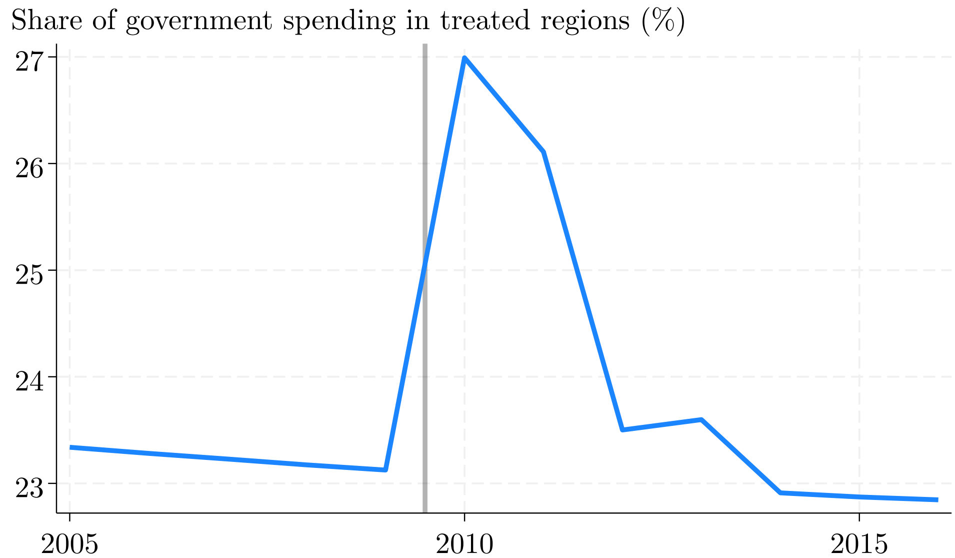
C Additional Figures

Figure S.1: IO matrix of the Chilean economy



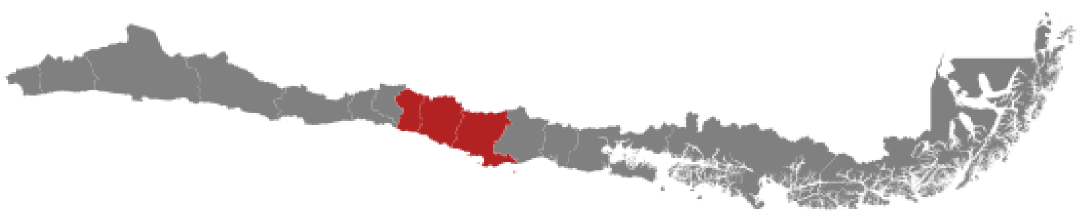
Notes: In this figure, we represent the IO matrix of our model of the Chilean economy used to implement the estimation procedure of Section 5.3. Each row (resp., column) of the matrix represents a separate node in our representation of the Chilean economy, i.e., a region × sector. When numbering nodes, we group nodes by region, and order them by sector. Node 1 hence corresponds to sector 1 (Agriculture, livestock, forestry and fishing) in region 1 (Tarapacá), and node 16 corresponds to sector 2 (Exploitation of mines and quarries) in region 2 (Antofagasta).

Figure S.2: Changes in government transfers after the shock



Notes: This figure shows the share of government spending that is allocated to treated regions.

Figure S.3: Treated regions



Notes: This figure shows a map of the Chilean regions – those colored in red are the three that were declared in a state of catastrophic emergency in the aftermath of the shock and that we consider to be the treated regions.