Network General Equilibrium as a Threat to Identification: The Failure of Time Fixed Effects*

Bernard Herskovic† Gill Segal‡

August 2026

Abstract

Time fixed effects are not a general-equilibrium control. In network economies, aggregate shocks do not enter firm outcomes uniformly: they propagate through equilibrium prices, costs, and demand, leaving a residual network gradient after time demeaning. When an instrument, treatment, or exposure design even partially aligns with this gradient, standard panel estimators recover the direct effect plus an endogenous network-leakage component. We develop a tractable framework that signs and quantifies this Network-SUTVA bias in closed form, showing how network topology governs its magnitude. Weak instruments magnify the leakage. Difference-in-differences designs are similarly contaminated because controls can be untreated by assignment but treated in equilibrium. Calibrated to the U.S. production network, the distortion can exceed 130% of the true direct effect. We propose observed- and hidden-network diagnostics and remedies, and test them on canonical difference-in-differences designs, showing that standard estimators can confound direct policy responses with systemic network leakage.

JEL Codes: C23, D85, L14, G30

Keywords: Production Networks, Spillover Effects, Time Fixed Effects, Difference-in-Differences, Identification, SUTVA

*All errors are our own.

†UCLA Anderson School of Management and NBER. Email: bernard.herskovic@anderson.ucla.edu.

‡UNC Kenan-Flagler Business School. Email: gill.segal@unc.edu.

Empirical research in economics and finance increasingly studies macroeconomic and local shocks in settings where firms are linked through supply chains, common input markets, and financial intermediaries. In such settings, local shocks rarely remain local, and aggregate shocks rarely enter firm outcomes uniformly. Once equilibrium prices, costs, and constraints adjust, other firms' outcomes move even without a direct shock. That is the structural sense in which the stable unit treatment value assumption (SUTVA) fails in a connected economy: a firm's potential outcome depends not only on its own treatment or exposure, but also on shocks propagating through equilibrium linkages.

A common empirical design is to include time fixed effects and interpret the remaining variation as idiosyncratic. This paper shows why that response is generally incomplete: time fixed effects absorb only the cross-sectional average of the equilibrium response, not the heterogeneous propagation pattern induced endogenously by the network. Consider a simple vertical supply chain hit by an aggregate oil price shock. Even if all firms have the same direct sensitivity to oil prices, the supply shock cascades downstream cumulatively, so downstream firms absorb compounded cost increases from their upstream suppliers. The indirect equilibrium effect is therefore much larger downstream. A time fixed effect meant to absorb aggregate oil shocks removes only the average of this total effect across all firms. For outcomes that fall when costs rise, such as output, partialing out time fixed effects leaves upstream firms with artificially positive outcome residuals because their cost exposure is below average, while downstream firms are left with artificially negative outcome residuals because their cost exposure is above average. In a network economy, the omitted object is therefore not a simple macroeconomic scalar but a cross-sectional gradient of spillovers. Because standard panel IV, shift-share, difference-in-differences, and event-study designs are built from cross-sectional exposure patterns, this residual gradient can systematically overlap with the very identifying variation used to estimate the treatment effect.

Specifically, applied work often aims to recover a *direct* effect of an endogenous regressor X_{it} on an outcome Y_{it} while holding equilibrium conditions fixed. This is a controlled direct effect (Robins and Greenland, 1992; Pearl, 2001; Acharya et al., 2016). Network general equilibrium makes this parameter difficult to identify because the same initial shock that shifts X_{it} also shifts the equilibrium state, including prices, input costs, funding conditions, or other system-wide constraints. Nor is the problem avoided by reinterpreting the coefficient as a total equilibrium effect. Because two-way fixed effects absorb only the uniform average of the network response, the resulting estimate is only a relative cross-sectional equilibrium effect rather than the economically relevant total effect.

To bridge structural network macroeconomics and reduced-form panel identification, we develop a unified framework to diagnose, quantify, and address this network-induced equilibrium leakage under time fixed effects. We make five contributions.

First, we characterize the network state generated by shock propagation and show why time fixed effects need not absorb it. For example, in a production network model, aggregate and idiosyncratic shocks change equilibrium conditions across firms.[1] Time fixed effects absorb only the average of these changes, leaving a residual cross-sectional gradient that reflects differences in how firms are affected by network propagation. The gradient's shape is governed by where shocks enter and by the network multiplier. More broadly, we show that the residual gradient can arise beyond production networks, in equilibrium models where shocks to one unit affect the states of other units.

Second, we quantify and decompose Network-SUTVA bias in closed form. Every empirical design places identifying variation on a set of firms or industries. As explained above, propagation creates a residual network gradient showing where that variation reappears after the network adjusts and time fixed effects are removed. A scalar alignment statistic $\mathcal{K}$ measures how strongly the design lines up with that residual gradient. This score is governed jointly by network topology and the empirical design. Dense balanced networks leave little residual propagation, so $\mathcal{K}$ vanishes asymptotically, whereas hub-dominated networks create large residual directions and bias materializes when the treatment or instrument loads on those hubs. In a panel IV setting, the estimator's bias is exactly proportional to $\mathcal{K}$ divided by first-stage strength, so weaker first-stage relevance amplifies the same leakage. A diagonal/off-diagonal decomposition separates a relative own-unit component from a pure cross-unit Network-SUTVA component, and the same alignment logic yields contamination results for DiD and event studies.

Third, we propose diagnostics to assess ex ante whether Network-SUTVA threatens identification and control-based remedies designed to recover the direct effect. When the network is observed, the diagnostics and controls use its measured propagation structure. When the network is hidden, residual comovement is used to diagnose the problem and estimate the relevant equilibrium factors out of sample. In both cases, granularity measures whether enough identifying variation remains after the correction. The remedies have a sharp limit. If the instrument or treatment variation lies largely in the equilibrium component being controlled, the correction can eliminate the first stage or absorb the treatment variation used by DiD.

Fourth, we quantify the Network-SUTVA bias using empirically disciplined simulations. Turning the theory into a diagnostic laboratory, we build on the 2017 BEA detailed input–output matrix covering 402 industries and simulate standard panel IV and difference-in-differences designs directly on the physical U.S. production network. In our baseline BEA calibration of a panel IV design, the naive estimator is biased upward by 131.4% of the true direct effect, and a weaker

[1] Our benchmark maps empirical instruments directly to structural primitives, namely firm-level productivity shocks in a competitive production network. This strengthens the claim: even standard micro-founded shocks mechanically violate SUTVA in a connected economy. But the framework is not production-specific. The same first-order propagation object applies to financial contagion, demand linkages, credit-market spillovers, and other fixed-point environments in which local disturbances generate cross-sectional equilibrium gradients.

first stage can raise the bias severalfold. Difference-in-differences is contaminated by the same mechanism: in an event-study policy targeting core industries, pre-treatment coefficients remain essentially flat, yet post-treatment coefficients overshoot the true direct effect by 112.7% because the peripheral control group is untreated by assignment but treated in equilibrium. The remedies recover the direct effect when the controls span the omitted state and granular variation remains.

Finally, we illustrate the empirical relevance of Network-SUTVA bias using a canonical macro-finance template: target-window difference-in-differences exposure designs that interact aggregate fiscal shocks with predetermined exposure to construction and core manufacturing. Across two episodes, the 2010 bonus-depreciation expansion and the 2018 Tax Cuts and Jobs Act, conventional panel methods deliver statistically significant exposure coefficients. Once we add the observed-network remedy control, the coefficient attenuates by 12.9 percent in 2010 and 31.2 percent in 2018, and is no longer statistically significant at the 10 percent level in either episode. More than 80 percent of the residual identifying variation remains after the correction. Time fixed effects can leave a cross-sectional equilibrium gradient disguised as a direct exposure effect.

Importantly, the mechanism we examine differs from standard omitted-variable bias. The omitted network state is not a predetermined confounder that merely happens to correlate with treatment or instrument exposure: it is generated by the shock itself through equilibrium propagation. A firm's equilibrium environment is shifted by its own exposure and also by other firms' exposures through supplier prices, customer demand, input availability, funding conditions, and market clearing. This cross-unit, post-assignment propagation is precisely the mechanism by which SUTVA structurally fails, and it explains why naive omitted-variable fixes can be misleading: realized prices, quantities, markups, funding constraints, and downstream demand are often equilibrium mediators caused by the same shock, so conditioning on them can introduce textbook bad-control bias rather than restore the direct effect.

Moreover, the network-induced equilibrium state is fundamentally latent. It aggregates firm-specific shadow prices, input costs, markups, and financing conditions that are rarely observed in standard balance-sheet or market data. Reconstructing that state from an observed network is itself difficult in precisely the directions that matter for identification: small link-measurement errors are amplified by the network inverse, and coarse input–output controls leave within-sector propagation residuals. "Controlling for spillovers" is not an operational fix unless it is possible to span the relevant equilibrium factor space accurately while retaining independent identifying variation outside that space.

Related literature. Our paper lies at the intersection of the literature on networks, macroeconomics and applied microeconometrics. The common theme is propagation: a shock that begins in one part of the economy changes outcomes elsewhere. We treat this propagation as an identification object. After time fixed effects, general-equilibrium propagation does not disappear. It becomes a residual cross-sectional gradient that can align with the variation used for identification.

A large macroeconomic literature uses input–output linkages to study how microeconomic shocks aggregate into macroeconomic fluctuations (Domar, 1961; Acemoglu et al., 2012; Gabaix, 2011; Baqaee and Farhi, 2019). We repurpose these structural tools to ask the reverse econometric question: how does the equilibrium state generated by network propagation enter firm-level identification after fixed effects? The same structural objects that organize aggregate propagation also determine the residual directions that survive time demeaning and enter the exclusion restriction.

Our paper also relates to work on spillovers, interference, and exposure-based identification. Existing work emphasizes that one unit's outcome can depend on other units' treatments or shocks, violating SUTVA and complicating causal interpretation (e.g., Aronow and Samii, 2017; Berg et al., 2021; de Paula, 2017; Gao, 2025; Hudgens and Halloran, 2008; Manski, 1993; Sävje, Aronow, and Hudgens, 2021). Related applied work studies how shocks transmit through firm, bank, and local financial networks (e.g., Huber, 2018). The shift-share and spatial-inference literatures clarify when exposure designs are valid and how spatial or shock-level dependence affects inference (e.g., Conley, 1999; Adão et al., 2019; Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022; Borusyak and Hull, 2023). We depart from these approaches by embedding interference in a structural equilibrium network. In this setting, the omitted object is more than a local neighbor effect or a residual correlation structure: it is a network-generated equilibrium gradient that remains after time fixed effects. This distinction also creates an *estimand* problem, not merely an inference problem: network-robust standard errors can correct the size of tests, but they cannot remove the equilibrium component from the point estimate. Thus shock exogeneity is necessary but not sufficient here: even an exogenous shock can generate bias when its propagated network response remains aligned with the residualized design.[2] Relatedly, Donaldson (2026) studies a complementary failure of time fixed effects in cross-sectional exposure designs: the shock of interest moves a second aggregate variable, such as monetary policy, and bias arises when exposure to the shock also predicts sensitivity to that aggregate response. Our contribution is the cross-sectional topology channel. The identifying shock itself propagates through equilibrium linkages, and time demeaning leaves a topology-governed network gradient whose severity is summarized by the alignment statistic and the spectral structure of residual propagation. The two mechanisms are therefore complementary: her framework isolates a dynamic aggregate-contamination channel, while we isolate a cross-sectional network-propagation channel diagnosed and addressed using

[2]This distinction is not a claim that existing interference or exposure-design literatures ignore spillovers. Rather, the object differs. The reflection problem concerns identifying peer-effect coefficients when units and their peers are jointly determined. Shift-share and spatial-inference methods discipline exposure designs under shock-level orthogonality, nonrandom exposure, or spatial-dependence conditions. By contrast, the mechanism here is a population estimand problem: the empirical design can be correlated with the equilibrium state generated by the same shock after network propagation and time demeaning. Empirical network-spillover papers document particular transmission channels. The contribution here is the generic probability-limit map from such transmission to standard IV and DiD estimands.

network-based tools including alignment scores, eigenvectors, and latent factors.

The results also speak to the DiD and event-study literatures, which study inference, staggered timing, treatment-effect heterogeneity, and the interpretation of pre-trends (e.g., Bertrand et al., 2004; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Borusyak et al., 2024; Roth et al., 2023). Our mechanism is distinct from these concerns. The control group may be untreated by assignment but treated in equilibrium, because the treatment shock moves supplier prices, customer demand, input costs, funding conditions, or competitor constraints. Flat pre-trends need not rule out this channel, because the omitted equilibrium impulse response can be zero before treatment and turn on only afterward. In DiD, the same alignment problem appears in the treated–control comparison.

Finally, the paper relates to factor methods and spatial filtering. Interactive fixed effects and eigenvector spatial filtering provide natural tools for absorbing low-dimensional dependence (e.g., Bai, 2009; Griffith, 2003). We provide a structural foundation for when those tools help and when they fail. Factor controls are useful only when the omitted equilibrium component is sufficiently low-rank and the instrument retains variation outside that factor space. If the instrument itself is systemic, meaning its identifying variation lies in the dominant propagation direction, removing the dominant equilibrium factor destroys the first stage. If the hidden network factor is weak, PCA may fail to recover the relevant direction, consistent with the random-matrix limits in Baik et al. (2005). This leads to operational diagnostics and remedies based on observed-network alignment, granularity, residual eigenvalues, out-of-sample factor corrections, and bounds when point identification is not credible.

1 The Network Equilibrium Framework

In empirical applications, the same policy, technology, or credit shock typically moves many firms at once. The first object is where the shock enters the economy. The second is the cross-sectional state that appears after the network adjusts. Identification is fragile when the empirical design lines up with that state after fixed effects. This section formalizes that link. We fix notation in a production-network benchmark and then show the same first-order propagation object arises in any differentiable nonlinear equilibrium.

1.1 A production-network benchmark

Consider an economy with a representative final-good producer and N intermediate firms. The final-good producer combines intermediate goods into a final good using a constant-returns technology, $C_t = \prod_{i=1}^{N} c_{it}^{\theta_i}$, where c_{it} is the amount of good i used for final consumption and $\sum_{i=1}^{N} \theta_i = 1$. Each intermediate firm i produces output y_{it} using a Cobb–Douglas technology, combining labor l_{it} with a bundle of intermediate inputs x_{ijt} purchased from other firms j:

$$y_{it} \;=\; z_{it}\, l_{it}^{\alpha_i} \prod_{j=1}^{N} x_{ijt}^{A_{ij}}. \tag{1}$$

Here, z_{it} is a firm-specific productivity shifter (or, more generally, a primitive wedge in marginal cost), α_i is the firm-specific labor share of production, and A_{ij} is the elasticity of firm i's output with respect to the input from firm j. We assume constant returns to scale firm-by-firm, such that $\alpha_i + \sum_{j=1}^{N} A_{ij} = 1$ for all i. Allowing for heterogeneous labor shares α_i is important: it permits some firms to be pure primary-factor producers, while others combine primary factors with intermediate inputs. The matrix $A \in \mathbb{R}^{N\times N}$ with entries A_{ij} defines the input–output network of the economy.

General equilibrium requires that all markets clear. The labor market clears such that total labor demand equals exogenous supply L_t, $\sum_{i=1}^{N} l_{it} = L_t$. Goods market clearing requires that the total production of each intermediate good j equals final consumption demand plus the intermediate demand from all other firms:

$$y_{jt} \;=\; c_{jt} \;+\; \sum_{i=1}^{N} x_{ijt}. \tag{2}$$

Firms operate in competitive markets, taking the aggregate wage rate w_t and the vector of intermediate prices p_t as given. Cost minimization implies that firm i's marginal cost is proportional to $\frac{w_t^{\alpha_i}}{z_{it}} \prod_{j=1}^{N} p_{jt}^{A_{ij}}$. Because markets are competitive, price equals marginal cost ($p_{it} = mc_{it}$). Taking the natural logarithm of this pricing condition yields a linear system of equilibrium prices:

$$\log p_{it} \;=\; -\log z_{it} \;+\; \alpha_i \log w_t \;+\; \sum_{j=1}^{N} A_{ij} \log p_{jt} \;+\; \chi_i, \tag{3}$$

where χ_i collects firm-specific, time-invariant production constants. Stacking these conditions into vector form, we recover the standard duality representation of prices in a network economy:

$$\log p_t \;=\; u_t \;+\; A \log p_t \;+\; \chi, \tag{4}$$

where the primitive cost-shock vector u_t collects both the idiosyncratic productivity disturbances and the aggregate wage movements loaded through the heterogeneous labor shares. Letting $\boldsymbol{\alpha} \equiv (\alpha_1, \ldots, \alpha_N)'$, we define

$$u_t \;\equiv\; -\log z_t \;+\; \boldsymbol{\alpha} \log w_t, \tag{5}$$

and $\chi \in \mathbb{R}^N$ is the vector of firm-specific constants.[3] Under constant returns to scale, the row sum

[3] Our results to follow do not require that u_t depends on the aggregate wage component. One can set $\log w_t = 0$

of A for firm i is exactly $1 - \alpha_i$. We maintain the standard condition that the economy cannot reproduce itself entirely from intermediate goods without primary factors. Assuming that $(I - A)$ is invertible,[4] we obtain:

$$\log p_t \;=\; (I - A)^{-1}(u_t + \chi). \tag{6}$$

For each firm, define a *network state* as the log input-cost index:

$$s_t \;\equiv\; A \log p_t. \tag{7}$$

Substituting (6) into (7) yields an affine mapping from primitive shocks to equilibrium states:

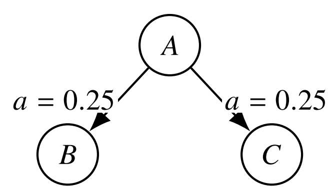
$$s_t \;=\; s_0 \;+\; M u_t, \qquad M \;\equiv\; A(I - A)^{-1}, \qquad s_0 \;\equiv\; M\chi. \tag{8}$$

Here s_0 is a time-invariant firm-specific constant, while $M u_t$ is the propagated cross-sectional equilibrium gradient. Let $L \equiv (I - A)^{-1}$ denote the standard Leontief inverse. Thus, $M = L - I$ maps primitive cost shocks into upstream input-cost states rather than full prices.

Two features matter for what follows. First, even a common aggregate source of u_t generally produces $M u_t \not\propto \mathbf{1}$. Equilibrium creates differential exposure because firms occupy different positions in the production network and have different labor shares. Second, M is typically dense: shocks load onto many firms through indirect paths.

Illustration 1 (Running example: the upstream fork)**.** We use a three-firm economy as the main running example. Firms are ordered as (A, B, C), and arrows point from a supplier to a customer. Firm A is an upstream supplier to both B and C. With input share $a = 0.25$, the network is

$$A^{\text{fork}} = \begin{pmatrix} 0 & 0 & 0 \\ a & 0 & 0 \\ a & 0 & 0 \end{pmatrix}, \qquad a = \frac{1}{4},$$



Upstream fork: A supplies both B and C

Firm A is a common supplier whose primitive cost disturbance enters multiple downstream production costs. The danger is that the same disturbance changes both A's outcome and the equilibrium environment of B and C.

and interpret $u_t = -\log z_t$ as a generic primitive marginal-cost wedge. We retain the wage term only to illustrate that even a scalar aggregate component can become a heterogeneous equilibrium state once loaded through the network.

[4] A sufficient economic condition is that every supply chain eventually requires at least some primary input, such as labor. This ensures that no subset of firms can operate as a closed, self-sustaining loop using only intermediate goods. Formally, this requires that every closed communicating class contains at least one firm with strictly positive primary-factor share, ensuring $\rho(A) < 1$ so that $I - A$ is invertible.

Under constant returns, firm A uses only primary factors ($\alpha_A = 1$), while B and C combine inputs from A with primary factors ($\alpha_B = \alpha_C = 0.75$). Thus the first row of A^{fork} sums to zero, while rows two and three sum to $a = 0.25$.

Since this network has no feedback loops, $(A^{\text{fork}})^2 = 0$, so $L^{\text{fork}} \equiv (I - A^{\text{fork}})^{-1} = I + A^{\text{fork}}$ and $M^{\text{fork}} = A^{\text{fork}}(I - A^{\text{fork}})^{-1} = A^{\text{fork}}L^{\text{fork}} = A^{\text{fork}}$. If a primitive cost shock of size u_A hits firm A, then $u_t = u_A e_A = (u_A, 0, 0)'$ and $M^{\text{fork}}u_t = (0, u_A/4, u_A/4)'$.

Firm A's primitive cost shock moves A's own price directly. It also changes the input-cost states of B and C because they purchase from A. The entire off-diagonal propagation channel is $s_B = s_C = u_A/4$. The rest of the paper asks what happens when an empirical design treats this propagated equilibrium component as if it had been absorbed by fixed effects.

1.2 Local linearization of a general nonlinear equilibrium

The identification argument does not depend on linear technologies or fixed input shares. With adjustment costs, substitution, or financial constraints, the relevant object is the local equilibrium derivative. Under differentiability, it is the Jacobian from primitive shocks to equilibrium states.[5]

Lemma 1 (Implicit function theorem linearization)**.** Let the equilibrium be characterized by $F(s_t, u_t) = 0$, where $F : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$ is continuously differentiable. Suppose that at (s_0, u_0) we have $F(s_0, u_0) = 0$ and $\nabla_s F(s_0, u_0)$ is invertible. Then there exists a neighborhood $\mathcal{U}$ of u_0 and a continuously differentiable function $s(\cdot)$ such that $F(s(u), u) = 0$ for all $u \in \mathcal{U}$. Moreover, the local propagation matrix is

$$\frac{\partial s}{\partial u}(u_0) \;=\; -\Big(\nabla_s F(s_0, u_0)\Big)^{-1} \nabla_u F(s_0, u_0) \;\equiv\; M.$$

The analysis requires neither Leontief production nor constant coefficients. All subsequent arguments use the Jacobian multiplier M that maps primitive shocks into equilibrium states. Thus network "edges" need not be physical links or distances. They are equilibrium derivatives measuring how a disturbance to one unit changes another's state. When M is approximately low-rank, equilibrium effects can sometimes be controlled for with factor methods. When M is high rank, contamination is fundamentally harder to remove.

[5] The matrix M should be interpreted as a total local equilibrium derivative, not as a set of fixed physical input weights. Supplier substitution, inventories, contract rigidities, capacity constraints, and financial rewiring change the entries of this derivative rather than eliminating it. The Leontief inverse is an exact benchmark under fixed cost shares. With substitution or other nonlinear adjustment, Lemma 1 shows that the same role is played by the equilibrium Jacobian. The identification problem is therefore governed by the residual component of this propagation operator and its alignment with the empirical design.

2 The Failure of Time Fixed Effects

Panel designs routinely absorb common shocks using time fixed effects. This is well-motivated when the omitted object is scalar: one time dummy removes one period-level number. Network equilibrium changes the omitted object: propagation produces a vector Mu_t whose entries differ across firms. A time fixed effect subtracts only its cross-sectional average, so the firm-specific deviations in relative exposure are left untouched.

Recall the equilibrium mapping $s_t = s_0 + Mu_t$, where s_0 is time-invariant and u_t collects the full vector of time-varying primitive shocks, including aggregate components. Firm fixed effects absorb s_0. Let

$$H \equiv I - \frac{1}{N}\mathbf{1}\mathbf{1}'$$

denote the within-time demeaning operator.

Proposition 1 (Time fixed effects remove only common levels)**.** After absorbing firm fixed effects and applying time demeaning, the residual equilibrium state is

$$\tilde{s}_t \equiv H(s_t - s_0) = HMu_t.$$

If Mu_t is not proportional to $\mathbf{1}$ (meaning the propagated equilibrium state is not perfectly flat across firms), then $\tilde{s}_t \neq 0$. Thus, purely aggregate shocks remain after time fixed effects whenever the network transforms a uniform initial shock into uneven equilibrium spillovers.

Proposition 1 is therefore a projection result: time fixed effects eliminate exactly one direction, $\text{span}\{\mathbf{1}\}$, while network propagation generally sends shocks elsewhere. Even scalar aggregate shocks can become heterogeneous through input shares, centrality, or financing links. The residual variation after time effects is not necessarily idiosyncratic: it can be the mean-zero part of general equilibrium itself. The next illustration shows this projection failure in the upstream-fork economy introduced in Section 1.1.

Illustration 2 (Time-demeaning and the residual cross-sectional gradient)**.** Return to the upstream fork from the running example and let firm A receive a primitive cost shock of -10, so $u_t = (-10, 0, 0)'$. Time fixed effects remove only equilibrium effects that are common across firms. Here, the shock first passes through M, producing a heterogeneous equilibrium state before time demeaning. In the production benchmark,

$$M^{\text{fork}}u_t = (0, -2.5, -2.5)'.$$

Thus A has no upstream input-cost change, while B and C inherit a -2.5 reduction from A.

Now apply the time-demeaning matrix $H = I - \frac{1}{3}\mathbf{1}\mathbf{1}'$. Since the cross-sectional average of $M^{\text{fork}}u_t$ is $-5/3$,

$$HM^{\text{fork}}u_t = \left(\frac{5}{3}, -\frac{5}{6}, -\frac{5}{6}\right)' \approx (1.667, -0.833, -0.833)'\,.$$

Because the propagated equilibrium state varies across firms, time demeaning removes only its cross-sectional average, leaving a zero-sum cross-sectional gradient.

Here B and C are not directly shocked in the primitive sense, yet inherit A's shock through the network. After time demeaning, those firms enter the residualized cross-section with negative equilibrium states, while A enters with a positive residual state. The regression therefore compares firms with different network exposure to the same shock rather than holding equilibrium fixed. What remains is not a common macro effect, but the network gradient across firms, HMu_t.

The same failure arises for scalar aggregate shocks.

Illustration 3 (Purely aggregate wage shocks and residual state gradients)**.** Consider a purely aggregate wage shock, $g_t \equiv \log w_t$, with no idiosyncratic productivity movements ($\log z_{it} = 0$). In the fork, A uses only primary factors, while B and C combine labor with inputs from A. Thus $\boldsymbol{\alpha} = (1, 3/4, 3/4)'$ and the direct cost exposure is $u_t = \boldsymbol{\alpha} g_t = (1, 3/4, 3/4)' g_t$.

Although the macroeconomic shock is scalar, applying the state multiplier gives $s_t = M^{\text{fork}}u_t = (0, 1/4, 1/4)' g_t$: firm A has no upstream input-cost state, while B and C inherit A's wage-cost increase through supplier prices. Time fixed effects remove only the scalar cross-sectional mean. Since the average of s_t is $(1/6)g_t$, time demeaning leaves

$$Hs_t = \left(-\frac{1}{6}, \frac{1}{12}, \frac{1}{12}\right)' g_t \neq 0.$$

Thus the propagated effect of the aggregate wage shock survives time effects as a residual cross-sectional gradient. A scalar control for $\log w_t$ has the same limitation: it imposes a uniform loading proportional to $\mathbf{1}$, leaving the network-propagated exposure $M^{\text{fork}}\boldsymbol{\alpha}$ in the error term. The issue is not whether the primitive shock is idiosyncratic or aggregate, but whether network propagation makes the resulting equilibrium state heterogeneous across firms.[6]

The next section maps this residual gradient into the asymptotic bias of panel estimators.

[6] Heterogeneous macro exposure is the reduced-form symptom and network propagation is its structural source. Observing the full equilibrium loading $M\boldsymbol{\alpha}$ and letting it enter with unrestricted time variation could span this particular aggregate channel, but standard time effects and scalar macro controls absorb only common time variation, and direct-exposure controls capture the primitive loading $\boldsymbol{\alpha}$ rather than the propagated loading $M\boldsymbol{\alpha}$. The network representation determines which firms lie above or below the period average, which designs align with that residual ordering, and how the resulting term enters the probability limit.

3 The Core Identification Failure

We connect equilibrium propagation to panel IV, shift-share instruments, and DiD. The key step writes the residualized outcome, after fixed effects, as a function of the endogenous regressor and the omitted equilibrium state. When that state depends on the instrument through M, exclusion fails in a quantifiable way.

3.1 Econometric setup and Network-SUTVA

Let Y_{it} be an outcome (investment, spreads, employment), X_{it} an endogenous regressor (borrowing, capital, input costs), and Z_{it} an instrument (policy exposure, Bartik shift-share, lender shocks). After absorbing firm fixed effects and time fixed effects, we work with within-time residuals $\tilde{Y}_{it}$, $\tilde{X}_{it}$, and $\tilde{s}_{it}$. The baseline residualized outcome equation is

$$\tilde{Y}_{it} \;=\; \beta\,\tilde{X}_{it} \;+\; \gamma\,\tilde{s}_{it} \;+\; \tilde{\varepsilon}_{it}. \tag{9}$$

The econometric problem is that $\tilde{s}_{it}$ is not predetermined. It is an equilibrium outcome determined jointly with X_{it}. Formally, define global potential outcomes $Y_i(\mathbf{Z}_t)$ and $X_i(\mathbf{Z}_t)$ depending on the entire instrument vector in period t. Network-SUTVA fails when equilibrium linkages make $Y_i(\mathbf{Z}_t)$ depend on more than Z_{it}. Equation (9) is a mediation equation: β is the *controlled direct effect* of X on Y holding the equilibrium state fixed, while γ loads the equilibrium state. Because the state is endogenous, recovering β with a single instrument generally requires that instrument to be orthogonal to the propagation gradient.[7]

Mapping to structural parameters. To map β and γ to primitives, take the special case of fixed aggregate wages so $u_{it} = -\log z_{it}$. Two outcomes illustrate:

- *Price pass-through:* Recall the competitive unit-cost condition (4). Defining the endogenous regressor as the cost shock, $X_{it} \equiv u_{it} = -\log z_{it}$, and substituting the network state s_{it} from (7), we obtain $\log p_{it} = X_{it} + s_{it} + \chi_i$. After residualizing with respect to firm and time fixed effects, this yields $\widetilde{\log p}_{it} = \tilde{X}_{it} + \tilde{s}_{it}$. Mapping to (9), $\beta = 1$ and $\gamma = 1$.

- *Revenue under CES demand:* Assume a standard CES demand system $y_{it} = \kappa_t (p_{it}/P_t)^{-\sigma}$ with demand elasticity $\sigma > 1$. Firm revenue $R_{it} \equiv p_{it} y_{it}$ satisfies $\log R_{it} = \text{const}_t + (1-\sigma)\log p_{it}$. Substituting the pricing relation yields $\widetilde{\log R}_{it} = (\sigma - 1)\widetilde{\log z}_{it} + (1-\sigma)\tilde{s}_{it}$. Now take the

[7] Appendix IA.A.14 formalizes this mediation-rank failure under the rank condition stated there: a nonzero mean-zero instrument generally cannot move X without also moving the propagated state HMz. Point identification of the CDE therefore requires controls that span the induced equilibrium state, restrictions that make the induced state orthogonal to the instrument, or an additional source of variation that moves X independently of that state.

regressor to be productivity itself, $X_{it} \equiv \log z_{it}$. Productivity then has a positive direct effect, $\beta = \sigma - 1 > 0$, while upstream inflation enters with a negative sensitivity, $\gamma = 1 - \sigma < 0$.

The econometric specification in equation (9) applies beyond this static production example. With financial frictions, for example, an exogenous credit-supply shock could move X_{it} while s_{it} captures the system-wide cost of capital. In all cases, β is confounded because the same underlying shock that moves X_{it} also generates the network equilibrium component Mu_t, violating SUTVA.

3.2 First stage and primitive shocks

Let the first stage after fixed effects be

$$\tilde{X}_{it} \;=\; \pi\,\tilde{Z}_{it} \;+\; \tilde{v}_{it}. \tag{10}$$

Here Z_t^{raw} denotes the exposure after removing firm fixed effects but before within-period demeaning. Time fixed effects remove each period's cross-sectional mean, so define $\tilde{Z}_t \equiv HZ_t^{raw}$ and use $\tilde{Z}_t$ throughout the identification analysis. To connect the instrument to equilibrium, we model primitive shocks as loading on this demeaned exposure:

$$u_t \;=\; \kappa\tilde{Z}_t \;+\; \xi_t, \tag{11}$$

where ξ_t collects primitive disturbances orthogonal to the instrument after fixed effects.

We impose standard panel IV exogeneity and first-stage relevance plus a network-specific orthogonality condition isolating propagation not mechanically driven by the instrument. These are Assumptions 1–3 listed below. For notation purposes, we stack firms within each period, with $\tilde{Z}_t \in \mathbb{R}^N$ collecting the entries $\tilde{Z}_{it}$, and $\tilde{X}_t, \tilde{Y}_t, \tilde{\varepsilon}_t$ defined analogously.

Assumption 1 (Source exogeneity)**.** $\mathbb{E}\left[\tilde{Z}_t'\tilde{\varepsilon}_t\right] = 0.$[8]

Assumption 2 (Relevance)**.** *The residualized first stage satisfies* (10) *with* $\pi \neq 0$ *and* $\mathbb{E}\left[\tilde{Z}_t'\tilde{v}_t\right] = 0$.

Assumption 3 (Primitive orthogonality)**.** *Letting* $u_t = \kappa\tilde{Z}_t + \xi_t$ *as in* (11), $\mathbb{E}\left[\tilde{Z}_t' HM\xi_t\right] = 0$.

Remark 1 (The role of primitive orthogonality). Assumption 3 requires that the unobserved primitive shock ξ_t not propagate in a way systematically correlated with the instrument. It is strong but purposeful: it isolates the pure equilibrium leakage from standard omitted-variable bias. It is

[8] Source exogeneity includes random assignment, but it is not equivalent to network exogeneity. A firm-level lottery can ensure $E[\widetilde{Z}_t'\widetilde{\varepsilon}_t] = 0$, yet it need not imply $E[\widetilde{Z}_t' HM\widetilde{Z}_t] = 0$, because the latter is a quadratic form in the realized assignment vector and the propagation operator. Under symmetry and independent assignment unrelated to network centrality, this term may vanish asymptotically. Clustered or hub-concentrated randomization can leave $\mathcal{K}_{\text{net}} \neq 0$ (see equation (14)) despite perfect source randomization.

not required for the core critique. Section 4 relaxes it and shows the IV estimator then inherits the classical invalid-IV bias in addition to the network leakage term.

As a special case, if the primitive shock has zero cross-sectional mean and equals the residualized exposure, $u_t = \tilde{Z}_t = Hu_t$, then $\kappa = 1$, $\xi_t = 0$, and Assumption 3 holds trivially.

3.3 Network-IV decomposition

In a panel IV regression of Y_{it} on X_{it} with firm and time fixed effects, the exclusion restriction fails when the instrument remains aligned with the equilibrium response it generates through the network. The design places identifying variation in $\tilde{Z}_t$. The network sends that variation into $HM\tilde{Z}_t$. If these two objects are orthogonal, the network state drops from the IV numerator. If they line up, the instrument correlates with the omitted equilibrium state even under primitive exogeneity. Thus projection failure is not yet identification failure: time effects may leave a residual state $HMu_t \neq 0$, but bias requires it to covary with the residualized treatment or instrument.[9]

Proposition 2 (Network-IV decomposition)**.** Define the *network loading factor*

$$\mathcal{K} \;\equiv\; \kappa \frac{\mathbb{E}\big[\tilde{Z}_t' HM\tilde{Z}_t\big]}{\mathbb{E}\big[\tilde{Z}_t'\tilde{Z}_t\big]}, \tag{12}$$

which is the normalized overlap between the instrument and the residual network state generated by the same instrument. Equivalently, $\mathcal{K}/\kappa$ is the population slope from regressing $HM\widetilde{Z}_t$ on $\widetilde{Z}_t$. Under Assumptions 1–3, the probability limit of the 2SLS coefficient using $(\tilde{Z}_t, \tilde{X}_t)$ is[10][11]

$$\beta^{IV} \;=\; \beta \;+\; \frac{\gamma\mathcal{K}}{\pi}. \tag{13}$$

Moreover, writing $M = M_{\text{diag}} + M_{\text{off}}$ where M_{diag} contains the diagonal entries of M, we can

[9] Thus, the claim is not that connectedness alone invalidates every design. A non-flat network state is harmless when the identifying variation is orthogonal to it. Conversely, even a source-exogenous instrument fails when the network-propagated equilibrium state remains aligned with the instrument after fixed effects. Formally, Proposition 1 establishes the residual-state condition $HMu_t \neq 0$, while Proposition 2 imposes the additional alignment condition $E[\widetilde{Z}_t' HM\widetilde{Z}_t] \neq 0$.

[10] To formally establish this probability limit, we assume standard regularity conditions: finite second moments and a cross-sectional law of large numbers such that sample moments converge to their population expectations, with a strictly non-zero first-stage denominator. For readability, we suppress the $1/N$ normalizations in all vector dot products.

[11] The probability limit in Proposition 2 differs from standard omitted-mediator IV bias. The difference is that the omitted state is not an arbitrary omitted variable. It is the residual equilibrium state generated by network propagation after time fixed effects, $HM\widetilde{Z}_t$ up to the scalar loading κ. The scalar $\mathcal{K}$ therefore turns a generic exclusion failure into a computable network-design diagnostic. Later, Proposition 3 links this diagnostic to network topology, the diagonal/off-diagonal decomposition isolates cross-unit Network-SUTVA leakage, and the remedy section asks whether the propagated state can be spanned while leaving granular identifying variation.

decompose[12]

$$\mathcal{K} = \underbrace{\kappa\,\frac{\mathbb{E}\big[\tilde{Z}_t' H M_{\text{diag}} \tilde{Z}_t\big]}{\mathbb{E}\big[\tilde{Z}_t'\tilde{Z}_t\big]}}_{\mathcal{K}_{\text{auto}}} + \underbrace{\kappa\,\frac{\mathbb{E}\big[\tilde{Z}_t' H M_{\text{off}} \tilde{Z}_t\big]}{\mathbb{E}\big[\tilde{Z}_t'\tilde{Z}_t\big]}}_{\mathcal{K}_{\text{net}}}. \tag{14}$$

Equation (13) shows that network bias vanishes only when $\gamma = 0$ or $\mathcal{K} = 0$. A sufficient condition is $HM\tilde{Z}_t = 0$ almost surely. Outside these cases, the estimator targets the direct effect plus an equilibrium-loading term.

The decomposition in (14) separates the bias generated by a firm's own equilibrium response from the bias transmitted through other firms. The term $\mathcal{K}_{\text{auto}}$ is the diagonal component as the instrument shifts firm i's own equilibrium mediator through M_{ii}. The term $\mathcal{K}_{\text{net}}$ is the off-diagonal component as other firms' shocks shift firm i's mediator through network propagation. This second term can survive even when $M_{\text{diag}} = 0$. Thus the exclusion restriction can fail even when the instrument is clean at the individual-firm level, because the network turns many individually exogenous shocks into a residual cross-sectional gradient that loads on $\tilde{Z}_t$.

For the worked examples and topology analysis below, it is useful to specialize $\mathcal{K}$ to a fixed cross-sectional exposure pattern. Proposition 2 defines $\mathcal{K}$ for stochastic, time-varying panel-IV instruments. For any residualized exposure vector z satisfying $Hz = z$, define

$$K_0(z) \equiv \frac{z'HMz}{z'z}. \tag{15}$$

When an instrument's cross-sectional exposure pattern is fixed over time up to a scalar shock, $\mathcal{K}/\kappa = K_0(z)$.

This specialization also makes the role of network topology transparent. The numerator of $K_0(z)$ is a scalar quadratic form, so only the symmetric part of residual propagation matters for alignment. Define

$$S \equiv \tfrac{1}{2}\left(HMH + HM'H\right). \tag{16}$$

For residualized z, $K_0(z) = z'Sz/(z'z)$. This representation lets us characterize which residualized design directions are most strongly aligned using the eigenvalues and eigenvectors of S.

Corollary 1 (Bounds on Equilibrium Propagation)**.** Let $\lambda_j(S)$ denote the eigenvalues of S on the mean-zero subspace, and define $\lambda^{\text{abs}}_{\max}(S) \equiv \max_j |\lambda_j(S)|$. Because $K_0(z)$ is a Rayleigh quotient, $|K_0(z)| \le \lambda^{\text{abs}}_{\max}(S)$. Consequently, $|\beta^{IV} - \beta| \le |\gamma\kappa/\pi|\,\lambda^{\text{abs}}_{\max}(S)$.

The bound says that the largest possible bias comes from designs that load most strongly on the

[12] The labels $\mathcal{K}_{\text{auto}}$ and $\mathcal{K}_{\text{net}}$ refer to the diagonal and off-diagonal components of the structural multiplier before the time-demeaning projection. Because the projection matrix H has dense off-diagonal entries, left-multiplication by H mechanically converts even purely diagonal propagation into relative cross-sectional comparisons.

leading residual propagation direction; designs that are diffuse or nearly orthogonal to that direction remain far below the bound.

Remark 2 (From IV covariance to a DiD alignment score). In a static DiD, the IV alignment logic has a direct fixed-design counterpart. Treatment has a fixed cross-sectional assignment that switches on after the policy. In a balanced two-period design, let $d_i \in \{0, 1\}$ denote assignment, $D_{it} = d_i \mathbf{1}\{t = 1\}$, and let $z \equiv Hd$ denote the demeaned treatment vector. Firm and time fixed effects give $-z/2$ in the pre period and $+z/2$ in the post period, so the same factor $1/4$ multiplies the numerator and denominator of the alignment ratio and cancels. Hence the normalized alignment ratio reduces to $K_0(Hd)$. With the policy indicator itself as the regressor, the residualized first-stage slope is one ($\widetilde{X}_{it} = \widetilde{D}_{it}$). Thus $K_0(Hd)$ is the DiD alignment score: it measures whether the demeaned treatment vector aligns with the residual network state generated by that vector.[13]

Illustration 4 (The $\mathcal{K}$ penalty in the upstream fork)**.** Consider price pass-through in the production benchmark. The outcome is log price, the endogenous regressor is the primitive own-cost shock, and the omitted equilibrium state is the input-cost index: $\tilde{Y}_{it} = \widetilde{\log p_{it}}$, $\tilde{X}_{it} = \tilde{u}_{it}$, and $\tilde{s}_{it} = \widetilde{A \log p_{it}}$. From the pricing equation (4), the structural coefficients are $\beta = \gamma = 1$. To isolate the network channel, take the idealized case in which the instrument equals the demeaned primitive shock, $\tilde{Z}_t = Hu_t$, so $\kappa = 1$, and set the first stage to $\pi = 1$. The remaining question is whether the instrument is aligned with the residual network state.

For a unit shock to firm A, $u_t = e_A$. The demeaned instrument, its propagated network state, and the residual state after time demeaning are, respectively,

$$\begin{aligned} \tilde{z}_A &= He_A = (2/3, -1/3, -1/3)', \\ M^{\text{fork}} \tilde{z}_A &= (0, 1/6, 1/6)', \\ HM^{\text{fork}} \tilde{z}_A &= (-1/9, 1/18, 1/18)'. \end{aligned}$$

The corresponding alignment score is

$$K_0 \equiv \frac{\tilde{z}_A' HM^{\text{fork}} \tilde{z}_A}{\tilde{z}_A' \tilde{z}_A} = \frac{-1/9}{2/3} = -\frac{1}{6} \approx -0.1667.$$

Because $diag(M^{\text{fork}}) = 0$, the entire loading comes from the off-diagonal network channel, $\mathcal{K}_{\text{auto}} = 0$ and $\mathcal{K}_{\text{net}} = \mathcal{K} = \kappa K_0 = -\frac{1}{6}$. Substituting into Proposition 2,

$$\beta^{IV} = \beta + \frac{\gamma \mathcal{K}}{\pi} = 1 + \frac{1 \cdot (-1/6)}{1} = \frac{5}{6} \approx 0.8333.$$

[13] Section 4.5 shows that the exact DiD post-period bias uses the original assignment d in the network state, whereas $K_0(Hd)$ is the symmetric diagnostic based on the demeaned treatment vector. The exact loading, denoted L_D, need not equal $K_0(Hd)$.

Thus a true one-for-one price pass-through coefficient is attenuated to 0.8333, which an applied reader would mistake for roughly 17% incomplete pass-through. The exclusion restriction fails even though there is no diagonal, own-firm propagation channel. The failure is entirely due to $\mathcal{K}_{\text{net}}$: the instrument isolates variation associated with firm A's shock, while network propagation carries that shock to the downstream firms, leaving residual equilibrium variation across firms after demeaning.

This example also clarifies what the proposition does *not* say. The problem is not generic connectedness. In a dense symmetric network, a shock that diffuses approximately evenly across all firms produces a propagated vector close to a common level, so time demeaning removes most of the propagated component and the $\mathcal{K}$ penalty approaches zero. By contrast, the upstream fork is asymmetric: the instrument is centered on a supplier hub, while the residual equilibrium state is borne by downstream customers. The bias is therefore driven by alignment between the identifying variation and the asymmetric propagation geometry of the network.

The point of Proposition 2 can be stated as a weighting problem. Standard IV estimators place statistical weight on observations with high identifying variance. In network economies, central hubs often generate precisely that variance: their shocks are large in the first stage and also create the largest propagated equilibrium disturbances. The regression therefore tends to overweight the observations most responsible for network contamination.

The upstream-fork example produces negative alignment. To see what governs the sign more generally, consider a raw deterministic design Z and let $\widetilde{Z} \equiv HZ$ denote its demeaned exposure vector. Define $c_i^{\text{out}} \equiv \sum_k M_{ki}$, the total equilibrium propagation generated by a shock to sector i, and $c_i^{\text{in}} \equiv \sum_k M_{ik}$, sector i's total exposure to shocks elsewhere. These are the Leontief out-degree and in-degree, respectively. Define $\bar{c} \equiv N^{-1} \sum_i c_i^{\text{out}} = N^{-1} \sum_i c_i^{\text{in}}$, and let Cov_N and σ_Z^2 denote finite-population covariance and variance, normalized by $1/N$. The appendix shows that the corresponding alignment score $K_0(\widetilde{Z})$ can be written in terms of the raw design Z as:

$$K_0(\widetilde{Z}) = \frac{1}{\sigma_Z^2} \left[\underbrace{\frac{1}{N} Z' M Z}_{\text{Internal Propagation}} \underbrace{- \bar{Z}\, \text{Cov}_N\Big(Z,\, c^{\text{out}} + c^{\text{in}}\Big)}_{\text{Systemic Leakage}} - \underbrace{\bar{Z}^2 \bar{c}}_{\substack{\text{Mean Propagation Removed} \\ \text{by Time Fixed Effects}}} \right]. \tag{17}$$

Equation (17) decomposes alignment into three pieces. Internal propagation is large when shocks assigned to high-Z sectors continue to affect high-Z sectors. Systemic leakage is large when high-Z sectors either transmit shocks broadly (high out-degree) or are strongly exposed to shocks from elsewhere (high in-degree). The final term is the mean propagation removed by time fixed effects. In particular, when identifying variation is concentrated in central sectors,

$\text{Cov}_N(Z, c^{\text{out}} + c^{\text{in}}) > 0$, increasing the magnitude of systemic leakage.

When the two negative terms dominate internal propagation, low-Z sectors carry relatively more of the residual network response and $K_0(\widetilde{Z}) < 0$; when propagation remains concentrated among high-Z sectors, $K_0(\widetilde{Z}) > 0$. Since IV bias is $(\gamma\kappa/\pi)K_0(\widetilde{Z})$, negative alignment attenuates a positive direct effect when $\gamma\kappa/\pi > 0$, while positive alignment amplifies it. If $\gamma\kappa/\pi < 0$, these directions reverse.

3.4 Network Topology and Asymptotic Bias

The preceding analysis shows that alignment depends jointly on the empirical design and the network. We now ask when network topology makes this alignment vanish or remain large as the number of firms grows. The network determines the residual propagation directions that survive time fixed effects, while the design determines how strongly its identifying variation aligns with those directions.

To separate network topology from propagation strength, consider a sequence of economies indexed by N. Let W_N describe network topology and let ρ denote a fixed scalar propagation strength. Let $H_N \equiv I_N - \frac{1}{N}\mathbf{1}\mathbf{1}'$ denote the time-demeaning projection, and consider residualized designs z_N satisfying $H_N z_N = z_N$. Define the corresponding spillover multiplier $\mathcal{M}_N(\rho)$ and centered symmetric propagation operator S_N,

$$\mathcal{M}_N(\rho) \equiv (I - \rho W_N)^{-1} - I = \sum_{\ell \geq 1} \rho^\ell W_N^\ell, \qquad S_N \equiv \tfrac{1}{2}\left(H_N \mathcal{M}_N(\rho) H_N + H_N \mathcal{M}_N(\rho)' H_N\right),$$

for $|\rho|\|W_N\|_{\text{op}} < 1$, where $\|\cdot\|_{\text{op}}$ denotes the operator norm. The fixed-design alignment score is $K_{0,N}(z_N) \equiv z_N' S_N z_N / (z_N' z_N)$. Proposition 3 asks how large this alignment can be across residualized designs as N grows.

Proposition 3 (Topology-design characterization of Network-SUTVA bias)**.** Consider a sequence of networks $\{W_N\}$ and fixed propagation strength ρ satisfying $|\rho|\|W_N\|_{\text{op}} < 1$. Let $\lambda_{\max}^{\text{abs}}(S_N) \equiv \max_j |\lambda_j(S_N)|$ denote the largest absolute eigenvalue of the centered residual propagation operator on the mean-zero subspace.

(i) *Uniform characterization.* The worst-case Network-SUTVA exposure over all residualized designs is exactly the largest absolute eigenvalue of the residual propagation operator:

$$\sup_{\substack{H_N z_N = z_N \\ z_N \neq 0}} \left|K_{0,N}(z_N)\right| = \lambda_{\max}^{\text{abs}}(S_N).$$

Thus time fixed effects eliminate equilibrium leakage uniformly over all mean-zero designs if and only if $\lambda_{\max}^{\text{abs}}(S_N) \to 0$. If the largest absolute eigenvalue is bounded away from zero, then there

exists a sequence of mean-zero designs with non-vanishing Network-SUTVA exposure.
(ii) *Dense balanced null.* Suppose the network is balanced, $W_N\mathbf{1} = \mathbf{1}$ and $\mathbf{1}'W_N = \mathbf{1}'$, and suppose that, after removing the common component, the remaining network variation vanishes: $\eta_N \equiv \|H_N W_N H_N\|_{\text{op}} \to 0$. Then, for all sufficiently large N,

$$\sup_{\substack{H_N z_N = z_N \\ z_N \neq 0}} |K_{0,N}(z_N)| \leq \frac{|\rho|\eta_N}{1 - |\rho|\eta_N} \to 0.$$

Thus, in dense balanced networks whose non-common propagation directions vanish, the Network-SUTVA penalty disappears asymptotically after time demeaning.
(iii) *Hub-dominated residual propagation.* Suppose instead that residual propagation has one dominant direction. Because S_N is symmetric on the mean-zero subspace, write its orthonormal eigen-decomposition as,

$$S_N = \sum_{j=1}^{N-1} \lambda_{j,N} v_{j,N} v'_{j,N}, \qquad |\lambda_{1,N}| \geq |\lambda_{2,N}| \geq \cdots,$$

and suppose the leading direction is asymptotically separated: $|\lambda_{2,N}| = o(|\lambda_{1,N}|)$. For any nonzero mean-zero design z_N, define its leading-direction loading share by

$$q_{1,N}(z_N) \equiv \frac{(z'_N v_{1,N})^2}{z'_N z_N}.$$

This scalar lies in $[0, 1]$ and equals the share of the design's residual cross-sectional variation loaded on the leading network direction. Then

$$K_{0,N}(z_N) = \lambda_{1,N} q_{1,N}(z_N) + O(|\lambda_{2,N}|).$$

The proposition separates the role of network topology from the role of the empirical design. Network topology determines the size and shape of S_N, while the design determines how strongly its identifying variation aligns with the directions where S_N is large. Part (i) gives the exact uniform condition: time fixed effects are safe for all residualized designs if and only if the largest absolute eigenvalue of S_N vanishes. Part (ii) gives a sufficient network condition: in dense balanced economies, shocks are spread almost evenly after propagation, so time demeaning removes the residual spillover component. Production networks with heterogeneous labor shares generally do not satisfy this balance condition. The BEA network studied later is therefore not in the dense-balanced null; its residual propagation is instead concentrated, closer to the case in part (iii).

Part (iii) gives the opposite case. In a hub-heavy economy, residual propagation can be dominated by a hub-versus-periphery direction. But topology alone does not jeopardize every design: a policy or instrument that loads heavily on this direction is more exposed, while one with little overlap is less exposed. The econometric failure is therefore the interaction between network topology and empirical design.[14]

4 Why Standard Empirical Defenses Fail

Once equilibrium creates structured exclusion failure, common empirical fixes are natural candidates. This section explains why three defenses, namely total-effect reinterpretation, spatial proxy controls, and weak-instrument robustness, do not eliminate Network-SUTVA bias.[15]

4.1 Total equilibrium effect versus two-way fixed effects

A natural rescue is to abandon the controlled direct effect and reinterpret the contaminated coefficient as a legitimate total general-equilibrium effect, which deliberately includes the network response. This rescue fails because TWFE identifies only demeaned cross-sectional variation. The total general-equilibrium response to the instrument combines the direct response $\beta\pi I$ and the network response $\gamma\kappa M$. Time effects apply H to their sum, leaving only $H(\beta\pi I + \gamma\kappa M)\tilde{Z}_t$. Any common component of the general-equilibrium response is therefore removed. Thus a TWFE coefficient does not recover the total general-equilibrium effect. At most, it captures a design-weighted relative equilibrium response, that is, the component that survives demeaning and covaries with the residualized treatment or instrument. Appendix IA.A.15 formalizes this observation.

[14] For stochastic panel-IV designs, the same logic applies after replacing the fixed loading share $q_{1,N}(z_N)$ with its average squared analogue. If $\widetilde{Z}_{N,t}$ is the residualized instrument and

$$q_{1,N}^{IV} \equiv \frac{E[(\widetilde{Z}_{N,t}' v_{1,N})^2]}{E[\widetilde{Z}_{N,t}' \widetilde{Z}_{N,t}]},$$

then under the leading-direction separation condition in part (iii),

$$\mathcal{K}_N = \kappa\lambda_{1,N} q_{1,N}^{IV} + O(|\kappa|\,|\lambda_{2,N}|).$$

Thus the same topology-design interaction governs both deterministic DiD designs and stochastic panel-IV designs.

[15] Two other frequent suggestions also fall short. Spatially robust standard errors can correct inference but not a biased point estimate. Nor can the network term generally be absorbed into a LATE-style heterogeneous-effect interpretation: because $s_i(Z) = [Mu(Z)]_i$ depends on the global assignment, reassigning other firms changes firm i's equilibrium environment. With heterogeneous direct effects, the IV estimand therefore contains an additive equilibrium-leakage term rather than only a reweighting of stable direct effects; Appendix IA.A.16 gives the full estimand and weights.

4.2 A generalized IV decomposition

The decomposition in Proposition 2 isolates equilibrium leakage by imposing source exogeneity and primitive orthogonality. The network term survives even when those clean-IV assumptions are relaxed: dropping them does not remove the equilibrium leakage but merely adds the familiar invalid-IV components on top of it. The estimator then deviates from β through classical error correlation, mechanical Network-SUTVA leakage, and endogenous primitive shocks:

$$\beta^{IV} = \beta + \underbrace{\frac{\mathbb{E}[\tilde{Z}_t'\tilde{\varepsilon}_t]}{\mathbb{E}[\tilde{Z}_t'\tilde{X}_t]}}_{\text{standard invalid-IV bias}} + \underbrace{\frac{\gamma\,\kappa\,\mathbb{E}[\tilde{Z}_t' H M \tilde{Z}_t]}{\mathbb{E}[\tilde{Z}_t'\tilde{X}_t]}}_{\text{Network-SUTVA leakage}} + \underbrace{\frac{\gamma\,\mathbb{E}[\tilde{Z}_t' H M \xi_t]}{\mathbb{E}[\tilde{Z}_t'\tilde{X}_t]}}_{\text{endogenous equilibrium shocks}} . \tag{18}$$

Standard invalid-IV fixes therefore do not remove the network term $\gamma\kappa\,\mathbb{E}[\tilde{Z}_t' H M \tilde{Z}_t]/\mathbb{E}[\tilde{Z}_t'\tilde{X}_t] = \gamma\mathcal{K}/\pi$.[16]

Remark 3 (Sign interaction). Equation (18) does not imply that standard invalid-IV bias and network leakage necessarily reinforce each other. Their signs depend on different objects: $\mathbb{E}[\tilde{Z}_t'\tilde{\varepsilon}_t]$ for the classical term and $\gamma\mathcal{K}$ for the equilibrium term. They may reinforce or offset one another.

4.3 Spatial controls and equilibrium proxies are often practically limited

A more ambitious response proxies directly for the equilibrium state in (9) to isolate β. In principle, this is the right target. In practice, the state is generated by the same shock that moves X_{it}, so realized prices, quantities, funding conditions, or neighbor outcomes are often post-shock mediators. Conditioning on them can introduce textbook bad-control bias rather than restore the direct effect.

The admissible strategy uses predetermined exposures, observed network loadings, or out-of-sample factor controls instead of realized post-treatment outcomes. Examples include predetermined spatial or neighbor exposures, sector input–output controls, and upstream or downstream exposure measures. The true state, s_t or Mu_t, is high-dimensional, while available proxies are often noisy, incomplete, or aggregated.[17] Proposition 4 formalizes the consequence: imperfect or coarse proxies can leave part of the equilibrium state in the error, where it can remain correlated with the

[16]The same distinction applies to overidentification and weak-instrument robust tests. These procedures address internal moment consistency or weak relevance; they do not remove the network term from the population estimand. When $\mathbb{E}[\widetilde{Z}_t' H M \widetilde{Z}_t] \neq 0$, the estimand still contains $\gamma\mathcal{K}/\pi$.

[17]A spatial autoregressive (SAR) control based on realized outcomes is not a generic remedy. If the same shock or instrument moves neighboring outcomes, a spatial lag of those outcomes is itself a post-shock equilibrium object rather than a predetermined exposure control. Moreover, the omitted state in Proposition 2 may be prices, input costs, funding conditions, markups, or demand states rather than neighboring outcomes. A SAR control removes leakage only if it spans the relevant residual state while preserving the target estimand.

instrument.

Proposition 4 (The inadequacy of spatial controls)**.** Consider the 2SLS estimator for β after augmenting (9) with a proxy $\widehat{s}_t$ for the equilibrium state.

(i) **Noisy equilibrium controls.** Let $\perp$ denote residualizing with respect to $\widehat{s}_t$. Then the population IV estimator in the controlled regression is

$$\beta^{IV|\widehat{s}} = \beta + \gamma \frac{\mathbb{E}[\tilde{Z}_t^{\perp\prime} \tilde{s}_t^{\perp}]}{\mathbb{E}[\tilde{Z}_t^{\perp\prime} \tilde{X}_t^{\perp}]}. \tag{19}$$

Hence any component of the true equilibrium state not spanned by $\widehat{s}_t$ remains in the error and can continue to correlate with the residualized instrument.

(ii) **Leontief amplification of network measurement error.** If the proxy is reconstructed from a measured network $\widehat{A} = A + \Omega$, then the associated propagation operator $\widehat{M} = \widehat{A}(I - \widehat{A})^{-1}$ satisfies the resolvent identity

$$\widehat{M} - M \;=\; (I - \widehat{A})^{-1}\Omega(I - A)^{-1}. \tag{20}$$

Thus small measurement errors in links are amplified by the dense inverse matrices that map primitive shocks into equilibrium states.

(iii) **Coarsening and aggregation bias.** Group firms into sectors. For any firm-level vector v, let Qv replace each firm's value with its sector average, and let $Rv \equiv v - Qv$ be the deviation from that sector average. Suppose the first stage using only this within-sector variation is $R\tilde{X}_t = \pi^{(Q)} R\tilde{Z}_t + \tilde{v}_t^{(Q)}$. Then the IV estimator after partialing out sector-level controls is

$$\beta^{IV|R} = \beta + \frac{\gamma\kappa}{\pi^{(Q)}} \frac{\mathbb{E}[(R\tilde{Z}_t)' RHM\tilde{Z}_t]}{\mathbb{E}[(R\tilde{Z}_t)'(R\tilde{Z}_t)]}. \tag{21}$$

So even perfect measurement of a coarse network proxy leaves within-sector equilibrium leakage whenever $RHM\tilde{Z}_t$ is correlated with $R\tilde{Z}_t$.

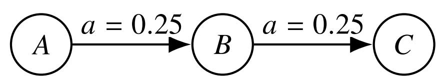
Corollary 2 (Sufficient condition for proxy-control success)**.** Assume finite second moments and suppose the controlled IV estimator in (19) is well defined. If $|\mathbb{E}[N^{-1}\tilde{Z}_t^{\perp\prime}\tilde{X}_t^{\perp}]| \geq c > 0$, then Cauchy–Schwarz bounds the controlled bias by $(|\gamma|/c)$ times the product of the root-mean-square residualized instrument and residual equilibrium-state components. Hence proxy controls are effective whenever the residual equilibrium component left outside the control space is small in mean square and the residualized first stage remains nontrivial.

The "industry-average" trap illustrates the difficulty. A common response is to control for sector averages using macro input–output (I–O) tables. If treatment and instrument vary at the *firm* level, this absorbs only the sector mean. The true Mu_t depends on firm-to-firm links. Unabsorbed within-sector variation remains in the error and mechanically correlates with the firm-level instrument. Macro I–O controls therefore leave a within-sector propagation residual that continues to structurally confound identification.[18]

A related failure occurs when firm-level proxies truncate the network at first-degree suppliers or customers. A simple chain formalizes this truncation bias.

Illustration 5 (The failure of local controls)**.** Truncation is easiest to see in a chain, where a shock can propagate beyond a direct supplier-customer link. Let firm A supply firm B, and let firm B supply firm C, with arrows again pointing from supplier to customer. With link weight $a = 0.25$,

$$A^{\text{chain}} = \begin{pmatrix} 0 & 0 & 0 \\ a & 0 & 0 \\ 0 & a & 0 \end{pmatrix}, \qquad a = \frac{1}{4}$$



Chain: A supplies B, and B supplies C

Because $(A^{\text{chain}})^3 = 0$, the price Leontief inverse is $L^{\text{chain}} \equiv (I - A^{\text{chain}})^{-1} = I + A^{\text{chain}} + (A^{\text{chain}})^2$, so the equilibrium state multiplier is

$$M^{\text{chain}} = A^{\text{chain}}(I - A^{\text{chain}})^{-1} = A^{\text{chain}} + (A^{\text{chain}})^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1/4 & 0 & 0 \\ 1/16 & 1/4 & 0 \end{pmatrix}.$$

For a unit primitive shock at A, $u_t = e_A = (1, 0, 0)'$, so the true equilibrium state is $M^{\text{chain}} u_t = (0, 1/4, 1/16)'$. Suppose instead that the researcher controls only for direct supplier exposure, using the first-degree proxy $\widehat{s}_t = A^{\text{chain}} u_t = (0, 1/4, 0)'$. This proxy captures the direct path $A \to B$ but misses the indirect path $A \to B \to C$, leaving the omitted component $(A^{\text{chain}})^2 u_t = (0, 0, 1/16)' = (0, 0, 0.0625)'$. After time demeaning, this omitted component remains nonzero:

$$H(A^{\text{chain}})^2 u_t = (-1/48, -1/48, 1/24)'.$$

Thus a control based only on direct links fails to absorb the higher-order path $A \to B \to C$.

If the instrument is related to shocks at A, it is therefore also related to the omitted higher-order propagation reaching C. Conditioning on direct supplier exposure does not restore exclusion,

[18] The same logic applies to sector-by-time fixed effects: in each period, they remove only the sector average of HMu_t, leaving any within-sector equilibrium variation. Identification therefore requires the remaining within-sector component to be orthogonal to the residualized treatment or instrument. Firm-level contracts, product specificity, inventories, geography, or substitute-supplier access can generate such within-sector variation; if it remains correlated with the residualized treatment or instrument, Proposition 4(iii) applies.

because the instrument still moves outcomes through this unabsorbed equilibrium component. Controlled IV bias can therefore survive first-degree supplier controls.

Proposition 4 and Corollary 2 show why "just control for spillovers" is rarely operational. Even with an observed network, the relevant state depends on long paths, and measurement noise or aggregation can distort the identifying directions. Ad hoc neighbor averages can therefore miss the omitted object: the full equilibrium gradient Mu_t.

4.4 Weak instruments amplify, rather than attenuate, equilibrium leakage

Weak first stages (small $|\pi|$) do not make equilibrium leakage negligible. Equation (13) shows why. The omitted equilibrium term enters the numerator, while the first stage enters the denominator. Weak relevance shrinks the denominator but not the propagated equilibrium state. An ideal IV design is strong for X but weakly aligned with the propagated equilibrium gradient: large π, small $\mathcal{K}$.

Proposition 5 (Weak IV amplification)**.** Consider a sequence of panels indexed by N in which the first-stage coefficient satisfies $\pi_N \to 0$ while the network loading factor $\mathcal{K}_N$ remains bounded away from zero and $\gamma \neq 0$. Then the IV estimator satisfies

$$\beta_N^{IV} = \beta + \frac{\gamma \mathcal{K}_N}{\pi_N},$$

so $|\beta_N^{IV}| \to \infty$, with the sign determined by the eventual sign of $\gamma\mathcal{K}_N/\pi_N$.

This amplification is a population bias, not merely a finite-sample weak-IV problem. The instrument can continue to move outcomes through the network even when its first-stage effect on X is small, and 2SLS scales that outcome response by the weak first stage. As in Bound, Jaeger, and Baker (1995), weak relevance is especially dangerous when exclusion is imperfect. Here the imperfection is the network term: when $\gamma\mathcal{K} \neq 0$, the instrument remains correlated with the omitted equilibrium state, so a smaller π magnifies the bias.

4.5 Application: DiD and dynamic event studies

DiD is often motivated as robust to omitted aggregate shocks after unit and time fixed effects. The network problem is that the control group may be untreated by assignment but treated in equilibrium. If treated firms' shocks move supplier prices, customer demand, competitor constraints, or lender balance sheets, untreated firms absorb part of the post-treatment response. DiD then compares treated firms to a counterfactual that has itself moved.

To formalize the same interference in DiD, let $D_{it} \in \{0,1\}$ denote treatment and let D_t collect treatment assignments across firms. Under Network-SUTVA, a firm's outcome can depend on the full assignment vector through its equilibrium state: $Y_{it}(D_t) = Y_{it}(D_{it}, s_{it}(D_t))$, with $s_t(D_t) = s_{0t} + Mu_t(D_t)$.

Define the period difference $\Delta Y_i \equiv Y_{i1} - Y_{i0}$, and analogously ΔX_i and $\Delta\varepsilon_i$. A two-period DiD compares these changes after removing their cross-sectional means, so let $\Delta\widetilde{Y} = H(Y_1 - Y_0)$, $\Delta\widetilde{X} = H(X_1 - X_0)$, and $\Delta\widetilde{\varepsilon} = H(\varepsilon_1 - \varepsilon_0)$. Proposition 6 asks whether this DiD comparison isolates the direct effect when treatment also shifts the network equilibrium state.

Proposition 6 (Two-period DiD / IV-DiD trap)**.** Suppose the demeaned outcome equation satisfies

$$\Delta\widetilde{Y} \;=\; \beta\,\Delta\widetilde{X} \;+\; \gamma\,\Delta\widetilde{s} \;+\; \Delta\widetilde{\varepsilon}, \tag{22}$$

where $\Delta\widetilde{s} = HM\Delta u$. Consider the canonical binary DiD design $D_{it} = d_i\mathbf{1}\{t = 1\}$, where $d_i \in \{0,1\}$ records firm i's treatment assignment. Let $z \equiv Hd$ denote the demeaned treatment vector, and define the post-period first-stage slope using this variation by $\pi_D \equiv z'\Delta\widetilde{X}/(z'z) \neq 0$. The network responds to the original assignment d before fixed effects demean it. Accordingly, suppose the primitive post-treatment shock satisfies $\Delta u = \kappa d + \Delta\xi$, and define the population DiD/IV-DiD estimand as the reduced-form slope of $\Delta\widetilde{Y}$ on z divided by the first-stage slope of $\Delta\widetilde{X}$ on z, that is, $\widehat{\beta}_{DiD} \equiv z'\Delta\widetilde{Y}/(z'\Delta\widetilde{X})$. Then, under source exogeneity $z'\Delta\widetilde{\varepsilon} = 0$ and the corresponding primitive-orthogonality condition $z'HM\Delta\xi = 0$,

$$\widehat{\beta}_{DiD} \;=\; \beta \;+\; \frac{\gamma\kappa}{\pi_D}\underbrace{\frac{(Hd)'HMd}{(Hd)'Hd}}_{L_D}. \tag{23}$$

When the policy indicator itself is the regressor, $\Delta\widetilde{X} = Hd = z$, so $\pi_D = 1$.

For the diagnostics in Section 5, we use $K_D \equiv K_0(Hd)$ as the symmetric DiD alignment score. Appendix IA.C.4 gives its relation to the exact bias loading L_D.

Proposition 6 is the DiD analogue of Proposition 2: the same treatment assignment can affect outcomes directly and generate a network equilibrium response.[19] Even if ε satisfies parallel trends, the DiD comparison need not isolate β: when HMd aligns with the demeaned treatment vector Hd, the post-treatment comparison includes the network response as well.[20]

[19] Although Proposition 6 uses two periods, the same logic extends to staggered designs: network contamination must be evaluated separately by cohort and event time because the relevant treatment assignment, and therefore the induced equilibrium response, can differ across those comparisons. This contamination arises before any TWFE weighting is applied.

[20] Dynamic event studies add a timing complication. A flat pre-trend tests whether the omitted equilibrium state

The off-diagonal network channel makes the control-group problem explicit: units with no direct assignment can still absorb the post-treatment network response $M_{\text{off}}\Delta u$, where M_{off} is the off-diagonal part of the propagation multiplier. Intuitively, in dense symmetric networks, this response is closer to a common shift and is more likely to be absorbed by time effects. In asymmetric networks, it can remain as a treated–control difference.

5 Diagnostics and Remedies

Subsection 5.1 provides ex-ante diagnostics for when network propagation threatens identification. Sections 5.2 and 5.3 provide remedies for the cases where the relevant propagation structure is observed or hidden. Section 5.4 states the common requirements for recovery. Appendix IA.B collects the formal derivations, implementation details, and sensitivity results.

5.1 Ex-ante diagnostics: alignment and residual comovement

Proposition 2 and Section 4.5 showed that Network-SUTVA bias is governed by alignment between the residualized design and the network equilibrium state. We now use that same alignment object as an ex-ante diagnostic and show how to compute it across common empirical designs. When the network topology, or a proxy for it, is observed, the score can be computed directly from the empirical design and the measured propagation operator. Its sign indicates positive or negative alignment; larger absolute values indicate greater potential exposure to Network-SUTVA, all else equal.

For a static DiD, for example, Section 4.5 introduced $K_D \equiv K_0(Hd)$. To compute this diagnostic, write $z = Hd$ and use

$$K_0(z) \equiv \frac{z'Sz}{z'z}, \tag{24}$$

where S is the centered symmetric propagation operator defined in equation (16). Because $Hz = z$, the same score can also be written $z'HMz/(z'z)$.[21]

The same diagnostic extends to staggered DiD and event studies (Appendix IA.B.2). For a time-varying panel-IV instrument $\widetilde{Z}_t$, compute the score period by period:

$$k_t \equiv \frac{\widetilde{Z}_t' S \widetilde{Z}_t}{\widetilde{Z}_t' \widetilde{Z}_t}. \tag{25}$$

moved before treatment, not whether treatment creates a post-treatment network response. If transmission is delayed, contamination may be small on impact but appear at later event times; Appendix IA.A gives the dynamic derivation.

[21] Recall that K_D is a symmetric alignment diagnostic, not generally the exact DiD bias loading L_D; the two can differ because the original assignment d enters the network before demeaning (Section 4.5).

Pooling across periods gives

$$K_Z^{\text{pool}} = \frac{\sum_t \widetilde{Z}_t' S \widetilde{Z}_t}{\sum_t \widetilde{Z}_t' \widetilde{Z}_t},$$

Equivalently, the pooled score averages the k_t's with more weight on periods containing more residualized instrument variation.

For interpretation, Appendix IA.B.1 gives the eigenvalue representation of K_0, while Appendix IA.B.2 uses this representation to provide an eigenvalue-based cutoff that signals potential exposure to sizable Network-SUTVA bias. The latter appendix also decomposes the score into own-unit and cross-unit components, $K_{\text{auto},0}$ and $K_{\text{net},0}$.

When the relevant propagation structure is not observed, the ex-ante diagnostic is residual comovement: after removing fixed effects and predetermined controls, do firms still move together through a stable factor with heterogeneous loadings? Residual comovement provides evidence that such a factor exists (a Network-SUTVA bias also requires the treatment or instrument to load on this factor). For brevity, Appendix IA.B.3 formalizes this screen using cross-sectional-dependence (CD) statistics, residual correlations, factor-count rules, and the overlap between the resulting factor controls and the empirical design.

5.2 Observed-network remedy

The observed-network remedy applies when the relevant network is observed, or when a useful proxy is available, such as an input–output table or supply-chain data. The goal is not to control for arbitrary spillovers. It is to use the measured network to construct the residual equilibrium state predicted by the empirical design and include that state as a control.[22]

To illustrate, consider a single-policy DiD design. Let $\widehat{A}$ be the measured network, let $\widehat{M} = \widehat{A}(I-\widehat{A})^{-1}$ be its propagation multiplier, let d denote policy assignment, and define $X_{it} = d_i \times \text{Post}_t$. The design-generated control is $q_C = H\widehat{M}d$. Raw assignment is propagated first, and the outer H removes the cross-sectional mean of the predicted state before it enters the fixed-effect regression. Because q_C is constructed from d, it may absorb some treatment variation; Section 5.4 reports how much remains. With the true network, $q_C = HMd$, so the control exactly captures the treatment-induced equilibrium state. With an imperfectly measured network, Appendix IA.B.5 characterizes the remaining component $H(M - \widehat{M})d$.

The corrected regression is

$$Y_{it} = \alpha_i + \delta_t + \beta^{\text{obs}} X_{it} + \Gamma' W_{it} + \psi\left(q_{C,i} \times \text{Post}_t\right) + u_{it},$$

[22] This is not a generic "control for spillovers" solution. Proposition 4 shows that noisy or coarse network measures can leave residual bias. The correction is most credible when the measured network captures the relevant equilibrium component and enough identifying variation remains after the control is added.

where W_{it} denotes predetermined controls.

The same construction extends to other empirical designs. In event studies, the design-induced state is interacted with event time. In panel-IV designs, the analogous state $H\widehat{M}\widetilde{Z}_t$ is constructed from the residualized instrument and included in both stages of 2SLS, or equivalently partialled out from Y_{it}, X_{it}, and the instrument before running IV. Section 5.4 states the recovery conditions.

5.3 Hidden-network remedy

The hidden-network remedy applies when the relevant propagation structure is not measured, is incomplete, or is not a physical supply-chain network.[23] The idea is to use residual comovement outside the target estimation to estimate a small set of common factors, then use the estimated factor loadings as controls. The method works only if the training sample contains enough equilibrium variation to reveal the firm-specific loadings of the hidden equilibrium component that matters in the target regression.

Consider a DiD or event-study design. Let $\mathcal{T}_0$ contain pre-treatment periods or other observations excluded from the target estimation, and let $\mathcal{T}_1$ contain the estimation periods. In the training sample, residualize outcomes using the same fixed effects and predetermined controls as in the baseline specification:

$$Y_{it} = \alpha_i + \delta_t + \Gamma' W_{it} + \widehat{e}_{it}, \qquad t \in \mathcal{T}_0.$$

The target treatment or policy variable is not included in this training regression.[24] If omitted equilibrium propagation remains in the residuals, $\widehat{e}_t \approx HMu_t + \eta_t$, where η_t collects residual idiosyncratic variation. Residual comovement can then reveal how firms load on the omitted equilibrium state.

Next, stack the training residuals and use the residual-comovement screen in Appendix IA.B.3 to choose the number of factors r. Extract the first r principal components and their cross-firm loading matrix $\widehat{\Lambda}_r$. Estimate these loadings using only $\mathcal{T}_0$, then hold them fixed in $\mathcal{T}_1$.[25] In $\mathcal{T}_1$,

[23] "Hidden network" need not mean an unobserved physical graph. It means any unobserved structure that makes firms respond heterogeneously to the same shock, such as balance-sheet similarity, depositor overlap, common lenders, funding constraints, or other equilibrium co-exposures. A purely common shock with identical loadings is absorbed by time fixed effects. Heterogeneous loadings remain as residual factors that can align with the design.

[24] The purpose of the training regression is to estimate the pre-treatment factor structure, not to partial out the causal channel before estimating it. Including the target causal regressor in the training residualization would risk removing part of the effect the final regression is meant to estimate.

[25] Full-sample estimation generally violates this requirement. If treatment creates a persistent common response among treated units, full-sample PCA may recover that response as a factor and mechanically partial out part of the causal effect. How firms load on the factors must also remain stable across training and estimation windows. If treatment changes network topology or substantially changes how firms load on the factors, the appropriate output is sensitivity or bounds, not a corrected point estimate. Pre-specified factor selection and granularity keep the hidden-network correction from becoming an ex post search for attenuation. Because the factor loadings are estimated rather than observed, inference must also account for this generated-regressor uncertainty. Appendix IA.B.7 does so by re-

form factor-by-time controls $\widehat{C}_{it}^{(k,\tau)} = \widehat{\Lambda}_{ik}\mathbf{1}\{t = \tau\}$. This gives each estimated factor a separate coefficient in every estimation period: its strength can vary freely over time, while which firms load more or less on it is fixed by the pattern estimated in $\mathcal{T}_0$.

In the DiD illustration, let $X_{it} = d_i \times \text{Post}_t$ and estimate

$$Y_{it} = \alpha_i + \delta_t + \beta^{\text{hid}} X_{it} + \Gamma' W_{it} + \sum_{k=1}^{r} \sum_{\tau \in \mathcal{T}_1} \psi_{k\tau} \widehat{C}_{it}^{(k,\tau)} + \nu_{it}.$$

Event-study designs replace X_{it} with the event-time treatment interactions. For panel IV, include the same frozen controls in both stages, or equivalently partial them out of Y_{it}, X_{it}, and the excluded instrument before running IV. Section 5.4 states the recovery conditions and the corresponding sensitivity analysis.

5.4 Scope and boundaries

Both remedies must satisfy two requirements. First, the controls must capture the equilibrium component that threatens identification. We call this *spanning*. For DiD, the relevant state is HMd. For panel IV, it is the corresponding propagated instrument state. Second, enough identifying variation must remain after those controls are added. We call this *granularity*. To measure granularity in IV, let $\widetilde{Z}$ denote the instrument after removing the same fixed effects and baseline controls used in the estimating equation. Let C denote the remedy controls. Define $\widehat{g}(C) = 1 - R_C^2(\widetilde{Z})$, where $R_C^2(\widetilde{Z})$ is the R^2 from regressing $\widetilde{Z}$ on C. Thus $\widehat{g}(C)$ is the share of residualized instrument variation left after adding the remedy controls. For DiD, use the treatment indicator after the same adjustment in place of $\widetilde{Z}$. Low granularity means the remedy absorbs most first-stage variation in IV or most treated–control variation in DiD.

Appendix IA.B.5 gives separate recovery results for observed-network and hidden-network corrections in panel IV and DiD, including network-measurement error and hidden-factor approximation. Appendix IA.B.4 gives the corresponding sensitivity intervals.

To state the large-sample limit of the hidden-network remedy, write $HM = \Lambda F' + R$, where Λ contains the population factor loadings, F' maps primitive shocks into the corresponding factors, and R is the equilibrium propagation not captured by this factor structure. Let P_Λ denote the projection onto the columns of Λ, and define $\tilde{Z}_t^{\perp} \equiv (I - P_\Lambda)\tilde{Z}_t$, the residualized instrument variation left after removing the factor-loading directions. Proposition 7 asks whether the factor controls can remove the factor component $\Lambda F'$ while leaving enough instrument variation to identify β.

estimating the training residuals, loadings, factor controls, and final regression inside the bootstrap or sample-splitting procedure. This generated-regressor issue concerns uncertainty in the estimated remedy, not the source of Network-SUTVA bias itself. In the observed-network implementation, the external measured network is treated as fixed for sampling inference; network-measurement error is addressed separately in Appendix IA.B.5.

Proposition 7 (Factor controls and first-stage collapse)**.** Under the post-control moment conditions stated in Appendix IA.A, suppose $HM = \Lambda F' + R$ holds with a fixed number of factors r, the residual propagation component R is asymptotically negligible, and the out-of-sample PCA estimator consistently recovers the span of Λ. The population granularity condition is

$$\liminf_{N\to\infty} \mathbb{E}\left[N^{-1}\|\tilde{Z}_t^{\perp}\|^2\right] > 0. \tag{26}$$

If (26) holds, enough instrument variation remains after the factor controls are added, and the factor correction removes the dominant equilibrium component, leaving only the residual propagation component R. If $\mathbb{E}[N^{-1}\|\tilde{Z}_t^{\perp}\|^2] \to 0$, controlling for those factors removes essentially all identifying variation in $\tilde{Z}_t$, so the first stage after the controls are added collapses and the controlled direct effect β is not recovered.

Although Proposition 7 is stated for IV, the same identification limit arises in DiD when factor controls absorb the treatment variation used for identification. A design can therefore look strong before correction yet leave no usable variation for estimating the direct effect afterward.

6 Quantitative Evidence from the BEA Network

This section uses the 2017 BEA input–output network as a calibrated laboratory. We first report features computed directly from the BEA matrix, then simulate panel-IV and DiD designs on that fixed network. Because the DGP is known, these are internal validation exercises: we can compare estimated coefficients with the true direct effect, quantify the bias formulas, and verify recovery when the spanning and granularity conditions are satisfied.

6.1 A Unified BEA-Calibrated Design

Data-generating processes. To isolate the alignment mechanism in a transparent finite-sample setting, the IV design uses a scalar shift-share analogue. In each period, a scalar aggregate shock g_t loads on a time-invariant cross-sectional exposure vector $z(a)$, giving

$$Z_t = g_t z(a),$$

where $Z_t \in \mathbb{R}^N$ is the residualized instrument vector. Here g_t supplies the aggregate shock, $z(a)$ determines how exposure varies across industries, and a indexes how strongly that exposure aligns with the network's main propagation direction, as defined below. A full Bartik design instead combines many sectoral shocks with an exposure matrix. Appendix IA.C.1 shows formally that the scalar construction preserves the same alignment object, $\mathcal{K}/\kappa = K_0(z(a))$, even when the conditional volatility of g_t varies over time.

After firm and time fixed effects, the IV design is as follows. To keep the simulation notation compact, we suppress tildes on the post-fixed-effect variables and write $s_t \equiv HMu_t$ for the residual equilibrium state:

$$X_{it} = \pi Z_{it} + \nu_{it}, \qquad u_t = \kappa Z_t + \xi_t, \qquad s_t = HMu_t, \qquad Y_{it} = \beta X_{it} + \gamma s_{it} + \varepsilon_{it}.$$

Here X_{it} is the endogenous regressor, Z_{it} is the residualized instrument, π is the first-stage coefficient, and ν_{it} is first-stage noise. The matrix H removes the cross-sectional mean, and M is the network multiplier. The vector $u_t \in \mathbb{R}^N$ is the primitive shock, and $s_t \in \mathbb{R}^N$ is the residual equilibrium state left after time demeaning. The coefficient β is the controlled direct effect, κ maps the instrument into primitive shocks, γ maps the equilibrium state into outcomes, and ε_{it} is idiosyncratic outcome noise. The unobserved component ξ_t is specified as:

$$\xi_t = \eta f_t \ell + \zeta_t, \qquad f_t = \rho_f f_{t-1} + e_t.$$

To allow the hidden-network remedy to be meaningfully evaluated in a known DGP, ξ_t includes a stable low-rank component, $\eta f_t \ell$, that appears in the training residuals. Without this component, there would be no latent low-rank structure for the PCA procedure to estimate. Note that $\ell \in \mathbb{R}^N$ is the factor's cross-firm loading vector, η scales its strength, ζ_t is idiosyncratic primitive-shock noise, ρ_f governs persistence, and e_t is the factor innovation. Both f_t and ζ_t are independent of the design shock g_t, so ξ_t does not create an additional source of instrument invalidity: $\mathbb{E}[Z_t' HM\xi_t] = 0$. Appendix IA.C.2 gives the full construction and verifies the primitive-orthogonality condition in the simulation.

The DiD design uses a fixed treatment vector rather than a continuous instrument. Let $d \in \{0,1\}^N$ indicate the treated nodes, and let Post_t indicate post-treatment periods. Treatment is

$$D_t = d\,\text{Post}_t, \qquad D_{it} = d_i\,\text{Post}_t, \qquad X_{it} = D_{it}.$$

The outcome and equilibrium state are generated by

$$u_t = \kappa_{\text{did}} D_t + \xi_t, \qquad s_t = HMu_t, \qquad Y_{it} = \beta D_{it} + \gamma s_{it} + \varepsilon_{it}.$$

Here κ_{did} maps treatment assignment into the primitive shock. As in Section 4.5, treatment affects outcomes directly through βD_{it} and indirectly through the network state $s_t = HMu_t$. The DiD exercise also includes the pre-treatment low-rank component in ξ_t introduced above. Because this factor is independent of treatment timing, it is neither an anticipation effect nor a second policy shock. Appendix IA.C.3 gives additional details on the DiD construction and examines what happens when the training sample does not contain the factor structure needed for recovery.

Alignment design. Let A denote the cost-share matrix and construct $M = A(I - A)^{-1}$ and $S = \frac{1}{2}(HMH + HM'H)$, as above. To isolate how alignment changes bias while holding the network and exposure scale fixed, we construct a one-parameter family of exposure vectors. Let $v_\star$ be the unit leading eigenvector of S, which serves as the high-alignment benchmark. Let $h \in \mathbb{R}^N$ be a unit mean-zero vector orthogonal to $v_\star$, chosen to have negligible network alignment, with $|K_0(h)| < 10^{-4}$. For $a \in [0, 1]$, define

$$z(a) = \sqrt{a}\, v_\star + \sqrt{1 - a}\, h.$$

The square-root weights keep $\|z(a)\| = 1$. At $a = 0$, $z(a) = h$ is the low-alignment design. As a rises, the instrument increasingly loads on the leading propagation direction. At $a = 1$, $z(a) = v_\star$. The network, propagation strength, and outcome equation remain fixed throughout.

The network. The network A is fixed throughout the simulation. We use the 2017 BEA detailed input–output matrix, so all variation across exercises comes from the empirical design placed on that fixed network.

In the data, the largest eigenvalue of A is $\rho(A) = 0.8975$. Since this value is below one, the equilibrium multiplier $M = A(I - A)^{-1}$ is well-defined. The BEA diagnostic operator is concentrated, but not one-dimensional. Let λ_j denote the eigenvalues of S on the mean-zero subspace, ordered by absolute value. The leading-to-second absolute-eigenvalue ratio is 1.21, so the alignment operator is not driven by a single eigenvalue. At the same time, the absence of a clean rank-one gap does not imply a diffuse economy. Because S is symmetric but has signed eigenvalues, we summarize concentration using squared-eigenvalue shares, $s_j = \lambda_j^2 / \sum_k \lambda_k^2$. These shares are nonnegative and sum to one, so they measure how concentrated the operator is across its eigenvalues. The largest share is 16.2%, the top three sum to 35.7%, and the top five to 45.4%. The inverse-Herfindahl effective rank, $1/\sum_j s_j^2$, is 17.9 out of 401 mean-zero directions, so the operator is far more concentrated than one spread evenly across all 401 dimensions left after demeaning.

The leading eigenvector is also concentrated across industries. Its inverse-Herfindahl effective support, computed from squared industry loadings, is 26 sectors. The ten largest industry loadings carry 42.3% of the loading mass, and the twenty-five largest carry 55.3%. High-out-degree industries with large loadings on one side of this eigenvector include other real estate, management of companies, monetary authorities and depository credit intermediation, employment services, and nondepository credit. Economically, the leading eigenvector summarizes upstream and service hubs where residual equilibrium propagation is concentrated. The U.S. production network is therefore concentrated, not evenly spread. This concentration is the empirical environment in which Proposition 3 becomes quantitatively relevant.

For the hidden-network recovery check, we set $\ell = v_\star$ in IV and $\ell = d$ in DiD. Here ℓ determines which industries load more heavily on the latent factor that PCA must estimate from the training residuals. The Network-SUTVA bias itself is generated by the instrument-driven shock κZ_t in IV and the treatment-driven shock $\kappa_{\text{did}} D_t$ in DiD.

Calibration and implementation. The calibration separates population-bias parameters from finite-sample noise and factor-recovery parameters. We normalize the controlled direct effect to $\beta = 1$, set $\kappa = 1$, and set $\kappa_{\text{did}} = 1$ in the DiD design. We set the outcome-state loading to its price pass-through benchmark from Section 3, $\gamma = 1$, so that a one-unit move in the residual equilibrium state passes one-for-one into the outcome.

Let $a_\star$ denote the baseline alignment design used as a reference point in the IV exercises. Applied designs often place identifying variation in economically concentrated parts of the economy. Trade shocks may load on manufacturing hubs, government spending shocks on defense hubs, and credit or procurement shocks on large upstream suppliers. We therefore set $a_\star = 0.70$ as a baseline with economically relevant, but not maximal, network alignment. The full $z(a)$ family is examined below.

The standard deviations σ_ν, σ_ξ, and σ_ε govern first-stage noise, idiosyncratic primitive-shock noise, and outcome noise. The latent-factor parameters ρ_f and η govern the persistence and strength of the low-rank component and therefore affect how well out-of-sample PCA recovers it. None of these parameters changes the population bias term $\gamma\kappa K_0/\pi$. They affect estimation precision, finite-sample first-stage diagnostics, and hidden-factor recovery. Appendix IA.C.5 reports the finite-sample parameter values and how they are disciplined.

The Monte Carlo constructs M, S, $v_\star$, h, and $z(a)$ from the fixed 2017 BEA matrix. It then generates 1000 Monte Carlo panels from the structural equations above, using $T = 200$ periods for the IV design and 40 training periods plus an event window with 8 pre-periods and 8 post-periods for the DiD design. Each replication estimates the naive IV and DiD specifications and then applies the observed-network and hidden-network corrections. Appendix IA.C gives the full simulation algorithm.

6.2 IV Network-SUTVA Bias

Figure 1 isolates the two forces in the IV bias term $\gamma\kappa K_0/\pi$. Panel A varies alignment through a for fixed values of π. Panel B fixes alignment at $a = a_\star$ and varies π.

Because h is orthogonal to the eigenvector $v_\star$, the cross term vanishes and $K_0(z(a)) = a\lambda_\star + (1-a)K_0(h)$, where $\lambda_\star$ is the eigenvalue of S associated with $v_\star$. The IV probability limit is

$$\beta^{IV}_{\text{plim}}(a, \pi) = \beta + \frac{\gamma\kappa K_0(z(a))}{\pi}.$$

Consequently, the probability limit of the bias is governed by $K_0(z(a))$, which is, in turn, a function of a.

Panel A shows the distribution and Monte Carlo mean of the signed IV bias, $100(\widehat{\beta} - \beta)/\beta$, as the alignment score $K_0(z(a))$ varies with the first-stage coefficient held fixed. Moving right in Panel A increases the instrument's alignment with the BEA network gradient that survives time fixed effects. When $K_0(z(a))$ is near zero, the instrument places little weight on that gradient, so IV bias is small and the simulation bands cross zero. The closed-form bias rises linearly with alignment, and the Monte Carlo means track that line. At the baseline $a = a_\star$, the implied alignment score is $K_0 = 1.31$, and the bias is about 131.4% of the true effect when $\pi = 1$. Because K_0 is a network loading rather than a correlation, values above one are possible when propagation amplifies exposure through multiple input–output layers. The diagonal/off-diagonal decomposition of Proposition 2 shows what this number represents:

$$K_0(z(a_\star)) = 1.3087 = K_{\text{auto},0} + K_{\text{net},0} = 0.1061 + 1.2025.$$

The relative own-unit component accounts for 8.1 percent of the alignment, while the cross-unit component accounts for 91.9 percent. The 131 percent benchmark therefore illustrates a sizable cross-unit Network-SUTVA distortion in the baseline BEA calibration. It is not generated by the finite-sample noise parameters.[26] Around the baseline, even the 5th percentile of the simulated bias remains above zero.

Panel B fixes the alignment at $a = a_\star$ and varies only the first-stage coefficient π. The horizontal axis is π, with stronger instruments on the left and weaker instruments on the right. At the same baseline alignment, a strong first stage with $\pi = 1.25$ yields a bias of roughly 105%. If the instrument is weak with $\pi = 0.15$, the same network alignment generates a bias of about 870%. Weak relevance therefore amplifies equilibrium leakage because the same reduced-form

[26] Varying a by ±30% around the baseline gives $a \in \{0.49, 0.70, 0.91\}$, $K_0(z(a)) = \{0.916, 1.309, 1.701\}$, and population IV biases of 91.6, 130.9, and 170.1 percent when $\gamma = \kappa = \pi = 1$. The broader formula grid over $\gamma \in \{0.5, 1, 2\}$ and $\pi \in \{0.5, 1, 1.5\}$ spans 30.5 to 680.5 percent. These are calibration sensitivities, not sampling confidence intervals.

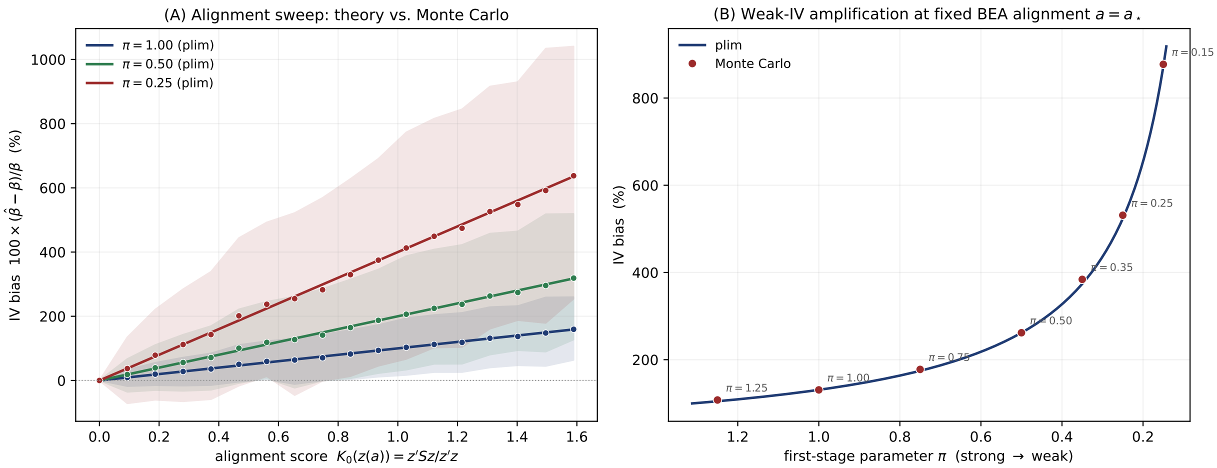


Figure 1: IV alignment and weak-IV amplification on the 2017 BEA network. Panel A varies the alignment score $K_0(z(a))$ with the network fixed. Panel B fixes $a = a_\star$ and varies the structural first-stage coefficient π, with the x-axis inverted so the design weakens to the right. Dots are Monte Carlo means, shaded bands are 5–95 percentiles, and solid lines are the closed-form bias curves for three values of π.

equilibrium component is divided by a weaker first stage.

6.3 DiD Network-SUTVA Bias

To study DiD, we construct a binary treatment vector d from the BEA network. The treated group contains 133 of the 402 BEA industries, or 33.1%, and the remaining industries form the comparison group.[27] The resulting design has the familiar broad-group DiD structure and is intended to mimic a setting in which treatment falls on industries that are also economically central in the production network.

Figure 2 reports the joint two-way fixed-effects DiD event study coefficients, relative to the reference point of $\tau = -1$, a period before Post_t turns to one. The pre-period coefficients are essentially flat because the treatment-induced component turns on only when $\text{Post}_t = 1$. The maximum absolute pre-treatment coefficient is 0.0004. The exercise is designed to show that this conventional pretrend pattern can coexist with large post-treatment contamination.

The post-period coefficients reveal the Network-SUTVA failure. Immediately after treatment begins, the event-study coefficients jump to 2.121, more than double the true direct effect of

[27] The assignment uses only information from the BEA network and is fixed before any outcomes are simulated. We form candidate treated groups from industries with the largest or smallest loadings on either the leading network direction $v_\star$ or the network response it generates, $q_\star = HMv_\star/\|HMv_\star\|$. We consider groups containing 10–40 percent of industries and require granularity above 0.25, so the network correction would still leave meaningful treatment variation. Among these candidates, we select the group with the largest absolute alignment score as a realistic stress test in which treatment is concentrated in industries with strong network propagation or exposure. The selected group has $K_D = 0.904$. Appendix IA.C.3 gives the exact grid and cutoffs. No simulated outcome or target coefficient enters the choice.

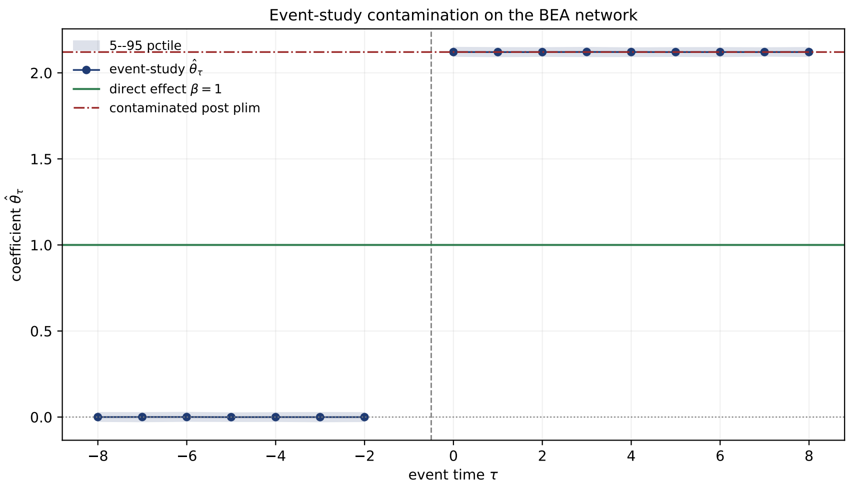


Figure 2: Joint two-way fixed-effects event study on the BEA network. The omitted event time is $\tau = -1$. The pre-period coefficients are flat, but the post-treatment coefficients overshoot the direct effect $\beta = 1$. The dotted reference line is the population post-treatment slope for the binary treatment assignment, $\beta + \gamma\kappa_{\text{did}}\text{cov}(d, HMd)/\text{var}(d)$.

one. The red reference line is 2.121, reflecting the population post-treatment slope implied by the treatment assignment.[28] The Monte Carlo post-treatment coefficients closely track this line. The event-study coefficient therefore combines the direct effect with the equilibrium response created by treatment. This bias is not a pre-treatment trend failure. Rather, the untreated sectors are untreated by assignment but treated in equilibrium. Once treatment changes the network state, untreated units inherit part of the equilibrium response.

6.4 Recovery of the Direct Effect

In this known-DGP exercise, the observed-network correction uses the BEA matrix to construct the relevant network controls. In IV, this is $q_\star = HM\ell/\|HM\ell\|$, and in DiD the corresponding control is $q_D = HMd/\|HMd\|$, using the chosen IV loading vector ℓ and DiD assignment vector d. The normalizations only fix scale. The hidden-network correction estimates factor loadings from training residuals, holds them fixed, and partials the resulting factor controls out of the target regression.

Table 1 applies the two correction procedures to the Monte Carlo simulated paths. The table checks two requirements. First, the corrected coefficient should move back toward the controlled direct effect, which is normalized to one. Second, enough identifying variation must remain after the control is added. This is reported by the granularity column and, for the IV rows, by $\log_{10} F_{\text{pooled}}$, the log first-stage F-statistic. A corrected coefficient near one is not meaningful if the first stage

[28] From Proposition 6, the exact DiD bias loading for this assignment is $L_D = 1.121$. With $\beta = \gamma = \kappa_{\text{did}} = 1$, the population post-treatment coefficient is $1 + 1.121 = 2.121$.

has been absorbed.

Table 1: Recovery of the direct effect under BEA network remedies. The table reports Monte Carlo means, bias, RMSE, granularity, and first-stage strength. In this known-DGP exercise, the naive IV and DiD estimators are biased, while the observed-network and hidden-network corrections recover β when the relevant equilibrium state is captured and identifying variation remains. Granularity, defined in Section 5.4, is the share of design variation surviving the correction controls. OOS denotes out-of-sample. The IV rows report $\log_{10} F_{\text{pooled}}$; DiD rows have no first stage.

Specification	Mean $\hat{\beta}$	Bias/β	RMSE	Granularity (share)	$\log_{10} F_{\text{pooled}}$
Naive IV	2.314	1.314	1.425	1.00	5.81
Observed BEA control	0.997	-0.003	0.003	0.60	5.59
Hidden OOS PCA control	0.963	-0.037	0.046	0.58	5.28
Naive DiD post-treatment average	2.127	1.127	3.704	1.00	–
Observed BEA-controlled DiD	1.000	-0.000	0.005	0.70	–
Hidden OOS PCA-controlled DiD	1.000	-0.000	0.005	0.70	–

The IV rows provide a direct recovery check against the known value $\beta = 1$. The naive estimate is 2.314, reflecting the roughly 130 percent bias documented above. With the observed-network remedy, the estimate is 0.997. With the hidden out-of-sample PCA remedy, it is 0.963. Thus the observed remedy essentially restores the estimate to the true direct effect of one, and the hidden remedy recovers it closely. Importantly, this recovery does not come from eliminating the identifying variation: granularity remains 0.60 and 0.58, and $\log_{10} F_{\text{pooled}}$ remains 5.59 and 5.28. The table reports the corresponding bias and RMSE.

The DiD rows provide the analogous recovery check using both the observed- and hidden-network remedies. The naive post-treatment average is 2.127 even though the true direct effect is one. Both corrected specifications return 1.000, with RMSE 0.005, while retaining 70 percent of treatment variation. Thus both remedies recover the direct effect without mechanically absorbing the treatment assignment. In this DiD DGP, both corrections remove the same HMd equilibrium component: the observed-network remedy constructs it from the BEA network, while the hidden-network remedy estimates the same cross-sectional loading pattern from pre-treatment residuals. This is why the two corrected rows coincide to displayed precision.

7 Empirical Evidence: Physical-Capital Exposure Designs

We illustrate the empirical relevance of Network-SUTVA bias with fiscal-shock DiD designs that interact corporate-tax or investment-incentive shocks with predetermined exposure to physical-capital sectors: firms in construction and core manufacturing. These sectors are natural exposure groups because core manufacturing supplies equipment, materials, and components, while construction produces structures and installed capital. Their supplier-customer linkages make the exposure design a stress test for whether standard DiD specifications absorb or leave behind network-generated equilibrium responses.

We study two fiscal episodes that directly changed incentives for physical capital adjustment: the 2010 bonus-depreciation expansion and the 2018 Tax Cuts and Jobs Act. The empirical template is broader than these tax shocks. The same exposure structure arises whenever physical-capital or capital-intensive firms are compared with other firms around aggregate policy or macroeconomic shocks, including investment subsidies, tariffs, credit disruptions, or commodity-price swings.[29]

7.1 The empirical template

The canonical specification is

$$Y_{it} = \alpha_i + \delta_t + \beta \left(\text{CapForm}_i \times Post_t\right) + W_{it}'\Gamma + u_{it}, \tag{27}$$

where Y_{it} is the firm outcome, α_i and δ_t are firm and year fixed effects, CapForm_i is a predetermined indicator equal to one for firms in construction or core manufacturing, defined by SIC codes 1500–1799 and 2000–3899, $Post_t$ is the post-period target indicator for the relevant fiscal shock episode, and u_{it} is the residual. We use the same CapForm_i definition in both episodes. The standard control vector W_{it} includes lagged firm size, lagged Tobin's Q, lagged leverage, cash flow scaled by lagged assets, lagged cash holdings, and sales growth. Firm and year fixed effects absorb the main effects of CapForm_i and $Post_t$. Thus β is the standard two-way fixed-effects DiD estimand: the post-episode change in construction and core manufacturing firms relative to other firms.

The DiD regressor is $X_{it} = \text{CapForm}_i \times Post_t$. The exposure group is not chosen because the statutes formally apply only to these firms. It is chosen because these sectors are directly tied to the physical-capital margin of the shocks: construction firms produce structures and installed capital, while core manufacturing firms supply equipment, materials, and components. A construction-and-core-manufacturing versus other-firms DiD therefore has an intuitive interpretation and a clear diagnostic risk: the estimated coefficient may combine the direct policy response with a production-network component.

In the examples below, the outcome variable is growth in property, plant, and equipment,

[29] Examples of related treatment structures include Zwick and Mahon (2017) and Ohrn (2018) for tax and investment incentives, Duchin, Ozbas, and Sensoy (2010) for crisis-period credit conditions, Dao, Minoiu, and Ostry (2021) for global and exchange-rate exposure, and manufacturing-exposure designs such as Flaaen and Pierce (2024) and Amiti, Redding, and Weinstein (2019). Our goal is not to formally replicate these studies. Rather, we use a physical-capital-sector exposure design as a laboratory for network diagnostics.
Because the true direct effect is not observed in these fiscal episodes, the application is an empirical diagnostic illustration rather than a known-DGP validation exercise. The known-DGP internal validation is the BEA Monte Carlo in Section 6.4, where the direct effect is fixed by construction.

$d\log(\text{PP\&E})$. This outcome is narrower than total assets, sales, or cash holdings. It directly captures the physical-capital margin targeted by accelerated depreciation, expensing, and corporate tax reforms. All data are from Compustat North America Fundamentals Annual firm-year observations, with firms mapped to BEA industries using SIC/NAICS information. We exclude financial firms and utilities (SIC 6000–6999 and 4900–4999).

7.2 Two fiscal episodes and conventional event-study evidence

The first is the 2010 bonus-depreciation expansion. In 2010, Congress extended and expanded accelerated depreciation for qualifying investment. The Small Business Jobs Act extended first-year bonus depreciation for property placed in service in 2010, and subsequent legislation raised it to 100 percent for qualifying investment placed in service after September 8, 2010, changing deductions for physical investment and linking the episode to PP&E growth. For this design, $Post_t$ is the impact-year target, equal to one in 2010. The second shock is the Tax Cuts and Jobs Act, enacted in December 2017. The TCJA lowered the corporate tax rate and temporarily allowed 100 percent expensing for qualified property placed in service after September 27, 2017, changing both the after-tax return to investment and the timing of deductions for physical capital. We treat 2018 as implementation and set $Post_t = 1$ in 2018 and 2019 for the cumulative target.

For the event-study diagnostics, we estimate the dynamic analogue of equation (27):

$$Y_{it} = \alpha_i + \delta_t + \sum_{j=-3,\, j\neq -1}^{3} \beta_j \left(\text{CapForm}_i \times \mathbf{1}\{\tau_t = j\}\right) + W_{it}'\Gamma + u_{it}, \tag{28}$$

where τ_t denotes event time relative to the fiscal episode. The event window is symmetric, $\tau = -3, \ldots, 3$, and event time -1 is the omitted reference period. The Wald pretrend test is the joint test that the pre-treatment coefficients at $\tau = -3$ and $\tau = -2$ are equal to zero.

Figure 3 displays the event-time paths. The conventional joint pretrend test does not reject in either episode, even though the target-window estimates may still contain a post-treatment network component.

Under conventional panel estimation, the two uncorrected exposure-design regressions appear to identify meaningful target-window effects: the 2010 coefficient has a two-way clustered t-statistic of -2.17, and the 2018 coefficient has $t = -2.10$. The negative signs have economically plausible ex-ante interpretations. Even taken at face value, a standard DiD reading would not mean that tax incentives mechanically reduced aggregate investment. Rather, construction and core manufacturing firms' PP&E growth fell relative to other firms. One channel runs through downstream demand: tax incentives raise the demand for capital goods among firms outside these sectors, and these other firms purchase new equipment, structures, and installed capital, while

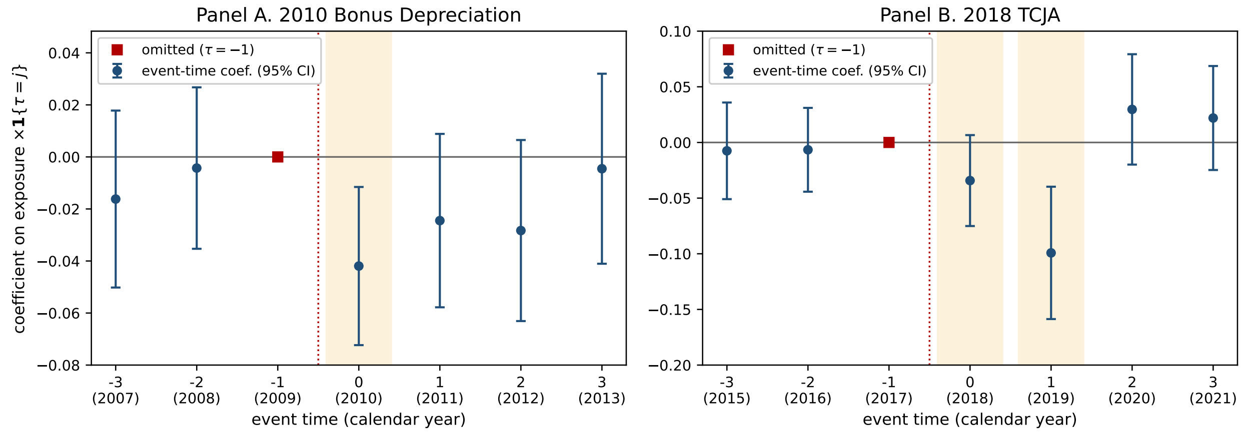


Figure 3: Event studies for physical-capital-sector exposure around the 2010 bonus-depreciation and 2018 TCJA shocks. CapForm_i identifies construction (SIC 1500–1799) and core manufacturing (SIC 2000–3899), and the outcome is $d\log(\text{PP\&E})$. The window is $\tau = -3, \ldots, 3$, with $\tau = -1$ omitted and shown at zero. Dots have 95-percent confidence intervals; shading marks the 2010 target in Panel A and the 2018–2019 target in Panel B. The joint pretrend null for $\tau = -3, -2$ is not rejected at the 5 percent level in either episode.

firms in the exposed sectors supply equipment, materials, or construction services. Higher sales in the exposed sectors therefore need not imply immediate capacity expansion, because the demand increase may be transitory or firms may adjust through utilization, markups, order backlogs, project pipelines, or cash flows rather than new plant investment. If policy-induced demand raises capital-goods or construction-input prices, firms in these sectors may face a higher user cost of expansion.

The diagnostic question is whether the interpretation above reflects a genuine direct effect or whether the estimates also contain an equilibrium propagation component.

7.3 Observed-network correction

Because the physical-capital-sector exposure is directly tied to the production network, we apply the observed-network correction from Section 5. Let d denote the cross-sectional exposure vector implied by CapForm_i. The control is $q_C = H\widehat{M}d$, constructed from the fixed 2017 BEA input–output multiplier $\widehat{M}$ and the original exposure vector d.[30] The corrected specification adds this control interacted with the same post-period target:

$$Y_{it} = \alpha_i + \delta_t + \beta^{BEA} X_{it} + \psi\left(q_{C,i} \times Post_t\right) + W_{it}'\Gamma + u_{it}. \tag{29}$$

The coefficient β^{BEA} is the physical-capital-sector DiD coefficient after controlling for the network response predicted by the BEA measure.

[30] We hold $\widehat{M}$ fixed across episodes. If $\widehat{M} = M$, then $q_C = HMd$ exactly equals the treatment-induced network state. With network measurement error, the remaining component is $H(M - \widehat{M})d$, as formalized in Corollary 5. The design-specific scores are $K_D = 0.326$, $L_D = 0.318$ in 2010 and $K_D = 0.329$, $L_D = 0.306$ in 2018.

Table 2 reports the regression results. The 2010 sample contains 31,309 firm-year observations from 5,709 firms, and the 2018 sample contains 26,077 observations from 4,809 firms. In 2010, the standard DiD coefficient is −0.0303, with $t = -2.17$; adding q_C gives −0.0264, with $t = -1.43$, an attenuation of 12.9 percent.

In 2018, the standard coefficient is −0.0696, with $t = -2.10$; the correction gives −0.0479, with $t = -1.13$, an attenuation of 31.2 percent. Thus, in both episodes, the conventional target-window DiD coefficient is statistically significant at the 5 percent level, while the corrected coefficient is not significant at the 10 percent level. The coefficient moves toward zero in each case, modestly in 2010 and by nearly one third in 2018. Granularity is 0.813 and 0.848, so more than 80 percent of residual treatment variation remains after the correction.

Table 2: Observed-network correction for physical-capital-sector DiD designs. CapForm_i identifies construction (SIC 1500–1799) and core manufacturing (SIC 2000–3899). The target is 2010 in the bonus-depreciation design and 2018–2019 in the TCJA design; the outcome is PP&E growth. "Observed BEA" adds $q_{C,i} \times Post_t$, with $q_C = H\widehat{M}d$. The t-statistics use two-way clustering by firm and BEA-industry-year. Attenuation is $1 - |\widehat{\beta}^{Corrected}|/|\widehat{\beta}^{Standard}|$; granularity is the residual treatment-variation share after adding q_C. All specifications include firm and year fixed effects and the controls in Section 7.1.

	2010 Bonus Depreciation		2018 TCJA	
	Standard DiD	Observed BEA	Standard DiD	Observed BEA
CapForm × $Post$ coefficient	−0.0303	−0.0264	−0.0696	−0.0479
t-statistic	−2.17	−1.43	−2.10	−1.13
Coefficient attenuation	–	12.9%	–	31.2%
Observed-BEA granularity	–	0.813	–	0.848

7.4 Interpretation

Because q_C is constructed from the fixed BEA network measure and predetermined exposure d, rather than realized post-treatment firm outcomes, the correction does not condition on post-treatment outcome variables. A separate concern is whether it absorbs the treatment variation. Granularity addresses that question: it is $1 - R^2$ from regressing the residualized treatment on the residualized BEA control. The values 0.813 and 0.848 imply that more than 80 percent of the residual identifying variation remains. The corrected coefficient is therefore the PP&E-growth difference between physical-capital-sector firms and other firms after removing the component predicted by the BEA network. It is not the policy's total general-equilibrium effect.

The positive K_D and L_D values show that the exposure aligns with a BEA propagation direction that time fixed effects do not absorb. Taken together, the attenuation documented above and the high granularity are consistent with the conventional DiD coefficients containing a downward PP&E component predicted by the measured BEA network. The correction removes only the component captured by that measured network.[31]

8 Conclusion

Network-SUTVA is an identification constraint for empirical work in connected economies. When shocks propagate through production, financial, or demand networks, two-way fixed effects absorb the common level but leave a heterogeneous cross-sectional gradient. Bias is governed by how the research design aligns with that residual equilibrium object, and the same interaction enters both instrumental-variables and difference-in-differences estimands.

The known-DGP BEA exercises show that Network-SUTVA bias can be economically large. In the aligned panel-IV design, the naive estimator is biased upward by about 130 percent of the true direct effect, and weak relevance amplifies rather than mitigates the problem. In the event-study design, pre-treatment coefficients remain essentially flat, yet the average post-treatment coefficient overstates the direct effect by about 113 percent because the control group is untreated by assignment but treated in equilibrium. Passing a pre-trends test therefore does not rule out Network-SUTVA contamination.

The physical-capital-sector application provides a firm-level empirical diagnostic of the same Network-SUTVA mechanism. Around the 2010 bonus-depreciation expansion and the 2018 Tax Cuts and Jobs Act, conventional target-window DiD estimates are statistically significant at the 5 percent level for construction and core manufacturing firms relative to other firms. After adding the observed-network remedy control, the corrected coefficient is insignificant at the 10 percent level in both episodes, while more than 80 percent of the residual treatment variation remains. Credible designs should therefore report alignment, granularity, and residual factor diagnostics. Together, these quantities help assess whether a design can sustain a point-identified direct effect or instead requires sensitivity analysis and bounds.

References

Acemoglu, D., Carvalho, V., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica* 80(5), 1977–2016.

Acharya, A., Blackwell, M., and Sen, M. (2016). Explaining causal findings without bias: detecting and assessing direct effects. *American Political Science Review* 110(3), 512–529.

Adão, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: theory and inference. *The Quarterly Journal of Economics* 134(4), 1949–2010.

[31] The interpretation concerns the component spanned by the observed BEA topology and therefore diagnoses production-network propagation rather than a dynamic aggregate-counterfactual channel. We focus on the observed-network correction to retain the full estimation sample. For these specific shocks and treatment groups, implementing the hidden-network remedy using complete firm histories in a separate eight-year training window would exclude nearly half of the firm-year observations.

Amiti, M., Redding, S. J., and Weinstein, D. E. (2019). The impact of the 2018 tariffs on prices and welfare. *Journal of Economic Perspectives* 33(4), 187–210.

Aronow, P. M., and Samii, C. (2017). Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics* 11(4), 1912–1947.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica* 77(4), 1229–1279.

Baik, J., Ben Arous, G., and Péché, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Annals of Probability* 33(5), 1643–1697.

Baqaee, D. R. and Farhi, E. (2019). The macroeconomic impact of microeconomic shocks: beyond Hulten's theorem. *Econometrica* 87(4), 1155–1203.

Berg, T., Reisinger, M., and Streitz, D. (2021). Spillover effects in empirical corporate finance. *Journal of Financial Economics* 142(3), 1109–1127.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics* 119(1), 249–275.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies* 89(1), 181–213.

Borusyak, K., and Hull, P. (2023). Nonrandom exposure to exogenous shocks. *Econometrica* 91(6), 2155–2185.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: robust and efficient estimation. *The Review of Economic Studies* 91(6), 3253–3285.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association* 90(430), 443–450.

Callaway, B., and Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics* 225(2), 200–230.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics* 92(1), 1–45.

Dao, M. C., Minoiu, C., and Ostry, J. D. (2021). Corporate investment and the real exchange rate. *Journal of International Economics* 131, 103437.

de Paula, Á. (2017). Econometrics of network models. In B. Honoré, A. Pakes, M. Piazzesi, and L. Samuelson (eds.), *Advances in Economics and Econometrics: Eleventh World Congress*, Vol. 1, 268–323. Cambridge University Press.

Donaldson, P. (2026). Cross-sectional identification with heterogeneous exposure to general equilibrium effects. Working paper, University of California, San Diego.

Domar, E. (1961). On the measurement of technological change. *Economic Journal* 71(284), 709–729.

Duchin, R., Ozbas, O., and Sensoy, B. A. (2010). Costly external finance, corporate investment, and the subprime mortgage credit crisis. *Journal of Financial Economics* 97(3), 418–435.

Flaaen, A. B., and Pierce, J. R. (2024). Disentangling the effects of the 2018–2019 tariffs on a globally connected U.S. manufacturing sector. *The Review of Economics and Statistics*, advance online publication. DOI: 10.1162/rest_a_01498.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica* 79(3), 733–772.

Gao, M. (2025). Endogenous interference in randomized experiments. Working paper.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review* 110(8), 2586–2624.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics* 225(2), 254–277.

Griffith, D. A. (2003). *Spatial Autocorrelation and Spatial Filtering: Gaining Understanding Through Theory and Scientific Visualization*. Springer.

Huber, K. (2018). Disentangling the effects of a banking crisis: evidence from German firms and counties. *American Economic Review* 108(3), 868–898.

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

Manski, C. F. (1993). Identification of endogenous social effects: the reflection problem. *The Review of Economic Studies* 60(3), 531–542.

Ohrn, E. (2018). The effect of corporate taxation on investment and financial policy: Evidence from the DPAD. *American Economic Journal: Economic Policy* 10(2), 272–301.

Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence (UAI)*.

Pesaran, M. H. (2021). General diagnostic tests for cross-sectional dependence in panels. *Empirical Economics* 60(1), 13–50.

Robins, J. M. and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 3(2), 143–155.

Roth, J., Sant'Anna, P. H. C., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics* 235(2), 2218–2244.

Sävje, F., Aronow, P. M., and Hudgens, M. G. (2021). Average treatment effects in the presence of unknown interference. *The Annals of Statistics* 49(2), 673–701.

Sun, L., and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225(2), 175–199.

Zwick, E., and Mahon, J. (2017). Tax policy and heterogeneous investment behavior. *The American Economic Review* 107(1), 217–248.

Internet Appendix for Network General Equilibrium as a Threat to Identification

IA.A Theoretical Extensions and Proofs

This appendix collects proofs for results stated without full proof in the main text. We keep notation from the paper. In particular, $H \equiv I - \frac{1}{N}\mathbf{1}\mathbf{1}'$ denotes the time-demeaning operator, and $\tilde{v}_t \equiv Hv_t$ denotes a time-FE residual.

IA.A.1 Proof of Lemma 1

Proof of Lemma 1. Fix a period t and suppress t subscripts to simplify notation. Let $(s^*, u^*) \in \mathbb{R}^N \times \mathbb{R}^N$ satisfy $F(s^*, u^*) = 0$. Assume F is continuously differentiable in a neighborhood of (s^*, u^*) and that $\nabla_s F(s^*, u^*)$ is invertible.

(1) Existence of a local equilibrium mapping. By the Implicit Function Theorem (IFT), there exist open neighborhoods $\mathcal{U}$ of u^* and $\mathcal{S}$ of s^* and a unique continuously differentiable function $g : \mathcal{U} \to \mathcal{S}$ such that

$$g(u^*) = s^* \qquad \text{and} \qquad F(g(u), u) = 0 \text{ for all } u \in \mathcal{U}.$$

This establishes that (locally) the equilibrium state is a well-defined mapping of the primitive shock vector: $s = g(u)$.

(2) The local propagation operator is the Jacobian. Differentiate the identity $F(g(u), u) \equiv 0$ with respect to u. By the chain rule,

$$\nabla_s F(g(u), u)\, Dg(u) \ + \ \nabla_u F(g(u), u) \ = \ 0,$$

where $Dg(u)$ denotes the Jacobian matrix of g at u. Evaluating at $u = u^*$ and using invertibility of $\nabla_s F(s^*, u^*)$ yields

$$Dg(u^*) = -\big(\nabla_s F(s^*, u^*)\big)^{-1} \nabla_u F(s^*, u^*).$$

Define the local propagation operator

$$M \ \equiv \ Dg(u^*) = -\big(\nabla_s F(s^*, u^*)\big)^{-1} \nabla_u F(s^*, u^*).$$

Since g is differentiable at u^*, it admits a first-order expansion:

$$g(u) = g(u^*) + M\,(u - u^*) + r(u), \qquad \frac{\|r(u)\|}{\|u - u^*\|} \to 0 \text{ as } u \to u^*,$$

so locally the equilibrium state obeys

$$s \;=\; s^* + M\,(u - u^*) + o(\|u - u^*\|).$$

This proves the claim that the linear fixed-point representation used in the paper is the first-order (Jacobian) characterization of an arbitrary nonlinear equilibrium.

(3) Leontief as a special case. In the Cobb–Douglas production-network example, the equilibrium in log prices $p \in \mathbb{R}^N$ can be written as

$$(I - A)\log p \;-\; u \;-\; \chi \;=\; 0, \qquad u \equiv -\log z + \alpha \log w,$$

where χ collects time-invariant firm-specific constants. Define

$$F(\log p, u) \equiv (I - A)\log p - u - \chi.$$

Then $\nabla_{\log p} F = I - A$ and $\nabla_u F = -I$. Applying the formula above gives

$$\frac{\partial \log p}{\partial u} = -\,(I - A)^{-1}(-I) = (I - A)^{-1}.$$

The equilibrium state used in the paper is the upstream price index $s = A\log p$. Therefore,

$$\frac{\partial s}{\partial u} = A\,\frac{\partial \log p}{\partial u} = A(I - A)^{-1} = M,$$

which is exactly the multiplier matrix from Section 1. □

IA.A.2 Proof of Proposition 1

Proof. From the equilibrium mapping $s_t = s_0 + Mu_t$, firm fixed effects remove the time-invariant component s_0. Applying the within-time demeaning projection $H = I - N^{-1}\mathbf{1}\mathbf{1}'$ gives

$$\tilde{s}_t = H(s_t - s_0) = HMu_t.$$

If $Mu_t \notin \operatorname{span}\{\mathbf{1}\}$, then $HMu_t \neq 0$, because H annihilates exactly the span of $\mathbf{1}$.

For a scalar aggregate shock $u_t = bg_t$, the residual equilibrium state is

$$\tilde{s}_t = HMb\, g_t.$$

Hence time effects eliminate the propagated aggregate shock if and only if $HMb = 0$, equivalently $Mb \in \mathrm{span}\{\mathbf{1}\}$. □

IA.A.3 Proof of Proposition 2

Proof of Proposition 2 (The Generalized Network Bias). We first expand the population IV estimator based on time-FE residuals,

$$\beta^{IV} = \frac{\mathbb{E}\left[\tilde{Z}_t'\tilde{Y}_t\right]}{\mathbb{E}\left[\tilde{Z}_t'\tilde{X}_t\right]}.$$

From (9),

$$\mathbb{E}\left[\tilde{Z}_t'\tilde{Y}_t\right] = \beta\,\mathbb{E}\left[\tilde{Z}_t'\tilde{X}_t\right] + \gamma\,\mathbb{E}\left[\tilde{Z}_t'\tilde{s}_t\right] + \mathbb{E}\left[\tilde{Z}_t'\tilde{\varepsilon}_t\right].$$

Assumption 1 implies $\mathbb{E}\left[\tilde{Z}_t'\tilde{\varepsilon}_t\right] = 0$. Assumption 2 implies

$$\mathbb{E}\left[\tilde{Z}_t'\tilde{X}_t\right] = \pi\,\mathbb{E}\left[\tilde{Z}_t'\tilde{Z}_t\right].$$

Next, write the equilibrium state as $\tilde{s}_t = HMu_t$ and use the primitive mapping $u_t = \kappa\tilde{Z}_t + \xi_t$ from (11). Then

$$\tilde{s}_t = HM(\kappa\tilde{Z}_t + \xi_t) = \kappa\,HM\tilde{Z}_t + HM\xi_t,$$

and therefore

$$\mathbb{E}\left[\tilde{Z}_t'\tilde{s}_t\right] = \kappa\,\mathbb{E}\left[\tilde{Z}_t'HM\tilde{Z}_t\right] + \mathbb{E}\left[\tilde{Z}_t'HM\xi_t\right].$$

Assumption 3 implies $\mathbb{E}\left[\tilde{Z}_t'HM\xi_t\right] = 0$, yielding

$$\mathbb{E}\left[\tilde{Z}_t'\tilde{s}_t\right] = \kappa\,\mathbb{E}\left[\tilde{Z}_t'HM\tilde{Z}_t\right].$$

Substituting these expressions into the IV ratio gives

$$\beta^{IV} = \beta + \frac{\gamma\,\kappa}{\pi} \cdot \frac{\mathbb{E}\left[\tilde{Z}_t'HM\tilde{Z}_t\right]}{\mathbb{E}\left[\tilde{Z}_t'\tilde{Z}_t\right]}.$$

Recognizing the loading factor defined in (12) yields (13).

Diagonal/off-diagonal decomposition: Recall the loading factor definition

$$\mathcal{K}(M,\{Z_t\}) = \kappa\,\frac{\mathbb{E}\big[\tilde{Z}_t' H M \tilde{Z}_t\big]}{\mathbb{E}\big[\tilde{Z}_t'\tilde{Z}_t\big]}.$$

Decompose the propagation operator into its diagonal and off-diagonal parts:

$$M = M_{\text{diag}} + M_{\text{off}}, \qquad M_{\text{diag}} \equiv \text{diag}(M), \qquad M_{\text{off}} \equiv M - \text{diag}(M).$$

By linearity of matrix multiplication,

$$\tilde{Z}_t' H M \tilde{Z}_t = \tilde{Z}_t' H (M_{\text{diag}} + M_{\text{off}}) \tilde{Z}_t = \tilde{Z}_t' H M_{\text{diag}} \tilde{Z}_t + \tilde{Z}_t' H M_{\text{off}} \tilde{Z}_t.$$

Taking expectations and dividing by the common denominator $\mathbb{E}[\frac{1}{N}\tilde{Z}_t'\tilde{Z}_t]$ yields the stated additive decomposition:

$$\mathcal{K}(M,\{Z_t\}) = \underbrace{\kappa\,\frac{\mathbb{E}[\tilde{Z}_t' H M_{\text{diag}} \tilde{Z}_t]}{\mathbb{E}[\tilde{Z}_t'\tilde{Z}_t]}}_{\equiv \mathcal{K}_{\text{auto}}} + \underbrace{\kappa\,\frac{\mathbb{E}[\tilde{Z}_t' H M_{\text{off}} \tilde{Z}_t]}{\mathbb{E}[\tilde{Z}_t'\tilde{Z}_t]}}_{\equiv \mathcal{K}_{\text{net}}}.$$

If $M_{\text{diag}} = 0$, then

$$\mathcal{K}(M,\{Z_t\}) = \mathcal{K}_{\text{net}}(M,\{Z_t\}).$$

A connected network with nonzero off-diagonal propagation therefore creates a potential pure Network-SUTVA channel. For a given design, this channel contaminates the IV estimator precisely when

$$\mathbb{E}[\tilde{Z}_t' H M_{\text{off}} \tilde{Z}_t] \neq 0.$$

Equivalently, there exist instrument directions aligned with the nonzero symmetric component of HM_{off} for which $\mathcal{K}_{\text{net}} \neq 0$. Connectedness alone does not imply contamination for every possible instrument.

To derive Equation (17), let $\tilde{Z} = HZ$ and unpack the deterministic alignment score $K_0(\tilde{Z}) = \frac{\tilde{Z}' H M \tilde{Z}}{\tilde{Z}'\tilde{Z}}$. Because $\tilde{Z} = HZ$ is mean-zero, $H\tilde{Z} = \tilde{Z}$ and $\tilde{Z}'H = \tilde{Z}'$. Thus the numerator simplifies to $\tilde{Z}' H M \tilde{Z} = \tilde{Z}' M \tilde{Z}$. The denominator is simply $\tilde{Z}'\tilde{Z} = N\sigma_Z^2$.

Substituting $\tilde{Z} = Z - \bar{Z}\mathbf{1}$ into the numerator yields:

$$\tilde{Z}' M \tilde{Z} = (Z - \bar{Z}\mathbf{1})' M (Z - \bar{Z}\mathbf{1}) = Z'MZ - \bar{Z}\,\mathbf{1}'MZ - \bar{Z}\,Z'M\mathbf{1} + \bar{Z}^2\,\mathbf{1}'M\mathbf{1}.$$

We evaluate the cross-terms using the centrality definitions $c_i^{\text{out}} = \sum_k M_{ki}$ and $c_i^{\text{in}} = \sum_k M_{ik}$. The left-multiplied term sums the columns of M, giving out-degree:

$$\mathbf{1}'MZ = \sum_{j=1}^{N} Z_j c_j^{\text{out}} = N\left[\text{Cov}_N(Z, c^{\text{out}}) + \bar{Z}\bar{c}\right].$$

The right-multiplied term sums the rows of M, giving in-degree:

$$Z'M\mathbf{1} = \sum_{i=1}^{N} Z_i c_i^{\text{in}} = N\left[\text{Cov}_N(Z, c^{\text{in}}) + \bar{Z}\bar{c}\right].$$

The scalar term sums all elements of the matrix: $\mathbf{1}'M\mathbf{1} = N\bar{c}$.

Substituting these components back into the expanded numerator gives:

$$\begin{aligned}\tilde{Z}'M\tilde{Z} &= Z'MZ - N\bar{Z}\left[\text{Cov}_N(Z, c^{\text{out}}) + \bar{Z}\bar{c}\right] - N\bar{Z}\left[\text{Cov}_N(Z, c^{\text{in}}) + \bar{Z}\bar{c}\right] + N\bar{Z}^2\bar{c} \\ &= Z'MZ - N\bar{Z}\,\text{Cov}_N\left(Z, c^{\text{out}} + c^{\text{in}}\right) - N\bar{Z}^2\bar{c}.\end{aligned}$$

Dividing the numerator by the denominator $N\sigma_Z^2$ perfectly isolates the terms:

$$K_0(\tilde{Z}) = \frac{1}{\sigma_Z^2}\left[\underbrace{\frac{1}{N}Z'MZ}_{\text{Internal Propagation}} \underbrace{-\bar{Z}\ \text{Cov}_N\left(Z,\ c^{\text{out}} + c^{\text{in}}\right) - \bar{Z}^2\bar{c}}_{\text{Systemic Leakage}}\right]. \tag{IA.A.1}$$

□

IA.A.4 Proof of Corollary 1

Proof of Corollary 1. For any residualized exposure vector z satisfying $Hz = z$, the definition (16) gives $K_0(z) = z'Sz/(z'z)$. This is a Rayleigh quotient of the symmetric operator S on the mean-zero subspace. By the Rayleigh–Ritz theorem, $|K_0(z)| \le \lambda_{\max}^{\text{abs}}(S)$, where $\lambda_{\max}^{\text{abs}}(S) \equiv \max_j |\lambda_j(S)|$ and the eigenvalues are computed on that subspace. Substituting this bound into Proposition 2 yields the inequality for $|\beta^{IV} - \beta|$. □

IA.A.5 Proof of Proposition 3

Proof. We prove each part in turn.

Part (i). Since H_N is an orthogonal projection and

$$S_N = \frac{1}{2}\left(H_N\mathcal{M}_N(\rho)H_N + H_N\mathcal{M}_N(\rho)'H_N\right),$$

we have

$$H_N S_N H_N = S_N.$$

Hence $S_N \mathbf{1} = 0$, and S_N maps the mean-zero subspace into itself. For any nonzero mean-zero z_N,

$$K_{0,N}(z_N) = \frac{z_N' S_N z_N}{z_N' z_N}$$

is a Rayleigh quotient of the symmetric operator S_N restricted to the mean-zero subspace. By the Rayleigh–Ritz theorem,

$$\sup_{\substack{H_N z_N = z_N \\ z_N \neq 0}} \left| \frac{z_N' S_N z_N}{z_N' z_N} \right| = \max_j |\lambda_j(S_N)| = \lambda_{\max}^{\text{abs}}(S_N),$$

where the eigenvalues are taken on the mean-zero subspace. The uniform equivalence follows immediately. If $\lambda_{\max}^{\text{abs}}(S_N) \to 0$, then the supremum of $|K_{0,N}(z_N)|$ over all nonzero mean-zero designs converges to zero. Conversely, if this supremum converges to zero, then the equality above implies $\lambda_{\max}^{\text{abs}}(S_N) \to 0$.

Finally, suppose $\lambda_{\max}^{\text{abs}}(S_N)$ is bounded away from zero along a subsequence. Along that subsequence, choose a unit eigenvector v_N associated with an eigenvalue whose absolute value equals $\lambda_{\max}^{\text{abs}}(S_N)$. Such an eigenvector exists because S_N is symmetric on the mean-zero subspace. For the design $z_N = v_N$,

$$|K_{0,N}(z_N)| = |v_N' S_N v_N| = \lambda_{\max}^{\text{abs}}(S_N).$$

Thus the Network-SUTVA exposure is bounded away from zero along that subsequence.

Part (ii). Let

$$P_N \equiv \frac{1}{N} \mathbf{1}\mathbf{1}', \qquad H_N = I_N - P_N.$$

The balance conditions $W_N \mathbf{1} = \mathbf{1}$ and $\mathbf{1}' W_N = \mathbf{1}'$ imply

$$W_N P_N = P_N, \qquad P_N W_N = P_N.$$

Therefore

$$H_N W_N H_N = (I_N - P_N) W_N (I_N - P_N) = W_N - P_N.$$

Define

$$B_N \equiv H_N W_N H_N = W_N - P_N.$$

Since $B_N P_N = 0$ and $P_N B_N = 0$, we have

$$W_N = B_N + P_N$$

with mutually annihilating components. Hence, for every integer $\ell \geq 1$,

$$W_N^\ell = B_N^\ell + P_N.$$

Multiplying by H_N on both sides gives

$$H_N W_N^\ell H_N = B_N^\ell = (H_N W_N H_N)^\ell.$$

Using the Neumann expansion,

$$\mathcal{M}_N(\rho) = \sum_{\ell \geq 1} \rho^\ell W_N^\ell,$$

we obtain

$$H_N \mathcal{M}_N(\rho) H_N = \sum_{\ell \geq 1} \rho^\ell H_N W_N^\ell H_N = \sum_{\ell \geq 1} \rho^\ell (H_N W_N H_N)^\ell.$$

Let

$$A_N \equiv H_N \mathcal{M}_N(\rho) H_N.$$

Since $\eta_N \to 0$, for all sufficiently large N, $|\rho|\eta_N < 1$. For such N,

$$\|A_N\|_{\text{op}} \leq \sum_{\ell \geq 1} |\rho|^\ell \|H_N W_N H_N\|_{\text{op}}^\ell = \sum_{\ell \geq 1} (|\rho|\eta_N)^\ell = \frac{|\rho|\eta_N}{1 - |\rho|\eta_N}.$$

Since

$$S_N = \frac{1}{2}(A_N + A_N'),$$

we have

$$\|S_N\|_{\text{op}} \leq \frac{1}{2}\left(\|A_N\|_{\text{op}} + \|A_N'\|_{\text{op}}\right) = \|A_N\|_{\text{op}}.$$

Because S_N is symmetric on the mean-zero subspace,

$$\|S_N\|_{\text{op}} = \lambda_{\max}^{\text{abs}}(S_N).$$

Therefore

$$\lambda_{\max}^{\text{abs}}(S_N) \leq \frac{|\rho|\eta_N}{1 - |\rho|\eta_N}.$$

Part (i) then gives

$$\sup_{\substack{H_N z_N = z_N \\ z_N \neq 0}} |K_{0,N}(z_N)| = \lambda_{\max}^{\text{abs}}(S_N) \leq \frac{|\rho|\eta_N}{1 - |\rho|\eta_N} \to 0.$$

Part (iii). Because S_N is symmetric on the mean-zero subspace, its eigenvectors form an orthonormal basis for that subspace. Write the normalized design as

$$\frac{z_N}{\|z_N\|} = \sum_{j=1}^{N-1} a_{j,N} v_{j,N}, \qquad \sum_{j=1}^{N-1} a_{j,N}^2 = 1.$$

Then

$$K_{0,N}(z_N) = \frac{z_N' S_N z_N}{z_N' z_N} = \sum_{j=1}^{N-1} \lambda_{j,N} a_{j,N}^2.$$

By definition,

$$a_{1,N}^2 = \frac{(z_N' v_{1,N})^2}{z_N' z_N} = q_{1,N}(z_N).$$

Separating the leading term gives

$$K_{0,N}(z_N) = \lambda_{1,N} q_{1,N}(z_N) + \sum_{j\geq 2} \lambda_{j,N} a_{j,N}^2.$$

Let

$$R_N \equiv \sum_{j\geq 2} \lambda_{j,N} a_{j,N}^2.$$

Since $|\lambda_{j,N}| \leq |\lambda_{2,N}|$ for all $j \geq 2$,

$$|R_N| \leq |\lambda_{2,N}| \sum_{j\geq 2} a_{j,N}^2 = |\lambda_{2,N}| \left(1 - q_{1,N}(z_N)\right).$$

This proves

$$K_{0,N}(z_N) = \lambda_{1,N} q_{1,N}(z_N) + R_N, \qquad |R_N| \leq |\lambda_{2,N}| \left(1 - q_{1,N}(z_N)\right),$$

and hence

$$K_{0,N}(z_N) = \lambda_{1,N} q_{1,N}(z_N) + O(|\lambda_{2,N}|).$$

The stochastic-panel-IV statement in the footnote follows from the same eigen-expansion. Let $\widetilde{Z}_{N,t}$ be a residualized stochastic instrument taking values in the mean-zero subspace, and assume

$$0 < E[\widetilde{Z}_{N,t}' \widetilde{Z}_{N,t}] < \infty.$$

Define

$$q_{j,N}^{IV} \equiv \frac{E[(\widetilde{Z}_{N,t}' v_{j,N})^2]}{E[\widetilde{Z}_{N,t}' \widetilde{Z}_{N,t}]}.$$

Since the $v_{j,N}$'s form an orthonormal basis of the mean-zero subspace, the weights satisfy $q^{IV}_{j,N} \geq 0$ and

$$\sum_{j=1}^{N-1} q^{IV}_{j,N} = 1.$$

The stochastic network loading factor is

$$\mathcal{K}_N = \kappa \frac{E[\widetilde{Z}'_{N,t} S_N \widetilde{Z}_{N,t}]}{E[\widetilde{Z}'_{N,t} \widetilde{Z}_{N,t}]} = \kappa \sum_{j=1}^{N-1} \lambda_{j,N} q^{IV}_{j,N}.$$

Under the leading-direction separation condition in part (iii), the leading term is $\kappa \lambda_{1,N} q^{IV}_{1,N}$. The remainder satisfies

$$\left| \kappa \sum_{j \geq 2} \lambda_{j,N} q^{IV}_{j,N} \right| \leq |\kappa| \, |\lambda_{2,N}| \sum_{j \geq 2} q^{IV}_{j,N} = |\kappa| \, |\lambda_{2,N}| (1 - q^{IV}_{1,N}).$$

Therefore

$$\mathcal{K}_N = \kappa \lambda_{1,N} q^{IV}_{1,N} + O(|\kappa| \, |\lambda_{2,N}|).$$

This completes the proof. □

IA.A.6 Proof of Proposition 4 (part (i))

Bias from noisy equilibrium controls. We work in population and suppress t to streamline notation. The structural equation is

$$\tilde{Y} = \beta \tilde{X} + \gamma \tilde{s} + \tilde{\varepsilon}.$$

The formula in Proposition 4(i) is conditional on the post-control exclusion moment $\mathbb{E}[\tilde{Z}^{\perp\prime} \tilde{\varepsilon}^{\perp}] = 0$, stated here because it need not follow automatically from the unprojected source-exogeneity condition. Suppose a noisy proxy $\tilde{\hat{s}} = \tilde{s} + \tilde{e}$ is observed and 2SLS controls for $\tilde{\hat{s}}$. By partialing out the proxy, this is equivalent (in population) to applying the residual-maker with respect to $\tilde{\hat{s}}$ to $(\tilde{Y}, \tilde{X}, \tilde{Z})$. Let

$$\mathcal{P}_{\hat{s}} \equiv \tilde{\hat{s}} \, (\tilde{\hat{s}}' \tilde{\hat{s}})^{-1} \tilde{\hat{s}}', \qquad \mathcal{M}_{\hat{s}} \equiv I - \mathcal{P}_{\hat{s}}.$$

Define residualized variables

$$\tilde{Y}^{\perp} \equiv \mathcal{M}_{\hat{s}} \tilde{Y}, \quad \tilde{X}^{\perp} \equiv \mathcal{M}_{\hat{s}} \tilde{X}, \quad \tilde{Z}^{\perp} \equiv \mathcal{M}_{\hat{s}} \tilde{Z}, \quad \tilde{s}^{\perp} \equiv \mathcal{M}_{\hat{s}} \tilde{s}, \quad \tilde{\varepsilon}^{\perp} \equiv \mathcal{M}_{\hat{s}} \tilde{\varepsilon}.$$

Applying $\mathcal{M}_{\hat{s}}$ to the structural equation yields

$$\tilde{Y}^{\perp} = \beta \tilde{X}^{\perp} + \gamma \tilde{s}^{\perp} + \tilde{\varepsilon}^{\perp}.$$

The population IV estimator with this control set is

$$\beta^{IV|\hat{s}} \equiv \frac{\mathbb{E}[\tilde{Z}^{\perp\prime}\tilde{Y}^{\perp}]}{\mathbb{E}[\tilde{Z}^{\perp\prime}\tilde{X}^{\perp}]}.$$

Substitute $\tilde{Y}^{\perp}$:

$$\mathbb{E}[\tilde{Z}^{\perp\prime}\tilde{Y}^{\perp}] = \beta\,\mathbb{E}[\tilde{Z}^{\perp\prime}\tilde{X}^{\perp}] + \gamma\,\mathbb{E}[\tilde{Z}^{\perp\prime}\tilde{s}^{\perp}] + \mathbb{E}[\tilde{Z}^{\perp\prime}\tilde{\varepsilon}^{\perp}].$$

Under this post-control exclusion condition,

$$\beta^{IV|\hat{s}} = \beta + \gamma\,\frac{\mathbb{E}[\tilde{Z}^{\perp\prime}\tilde{s}^{\perp}]}{\mathbb{E}[\tilde{Z}^{\perp\prime}\tilde{X}^{\perp}]}.$$

This is exactly (19) in the main text after noting that $\tilde{Z}^{\perp} = \mathcal{M}_{\hat{s}}\tilde{Z}$ and $\tilde{s}^{\perp} = \mathcal{M}_{\hat{s}}\tilde{s}$. The key implication is that unless $\mathcal{M}_{\hat{s}}\tilde{s} = 0$ (i.e., the proxy spans the true equilibrium component), the residual equilibrium state remains in the error and can remain correlated with the residualized instrument. □

IA.A.7 Proof of Proposition 4 (part (ii))

Leontief amplification of control error. Let the true adjacency matrix be A and the measured one be $\widehat{A} = A + \Omega$. Define the corresponding multipliers

$$M \equiv A(I - A)^{-1}, \qquad \widehat{M} \equiv \widehat{A}(I - \widehat{A})^{-1}.$$

A convenient identity is

$$A(I - A)^{-1} = (I - A)^{-1} - I,$$

because $(I - A)^{-1} = I + A(I - A)^{-1}$. Applying this to both A and $\widehat{A}$ yields

$$M = (I - A)^{-1} - I, \qquad \widehat{M} = (I - \widehat{A})^{-1} - I.$$

Therefore

$$\widehat{M} - M = (I - \widehat{A})^{-1} - (I - A)^{-1}.$$

Now apply the resolvent identity for invertible matrices B and C:

$$B^{-1} - C^{-1} = B^{-1}(C - B)C^{-1}.$$

Let $B = I - \widehat{A}$ and $C = I - A$. Then $C - B = (I - A) - (I - \widehat{A}) = \widehat{A} - A = \Omega$, hence

$$\widehat{M} - M = (I - \widehat{A})^{-1}\,\Omega\,(I - A)^{-1}.$$

This is the stated identity.

To obtain the amplification bound, take any submultiplicative matrix norm $\|\cdot\|$ and apply submultiplicativity:

$$\|\widehat{M} - M\| \le \|(I - \widehat{A})^{-1}\|\,\|\Omega\|\,\|(I - A)^{-1}\|.$$

If, in addition, $\|A\| < 1$ and $\|\widehat{A}\| < 1$ in the same induced norm, then applying the Neumann series bound yields:

$$\|\widehat{M} - M\| \le \frac{\|\Omega\|}{(1 - \|\widehat{A}\|)(1 - \|A\|)}.$$

For nonnormal networks, the relevant amplification is governed by the resolvent norms $\|(I-A)^{-1}\|$ and $\|(I - \widehat{A})^{-1}\|$, not by the spectral radius $\rho(A)$ alone. While the amplification may be moderate in sparse empirical networks, in dense networks with strong equilibrium feedback, these resolvent norms can be exceptionally large, meaning small measurement errors in the adjacency matrix severely distort the equilibrium control.

□

IA.A.8 Proof of Proposition 4 (part (iii))

Proof of Proposition 4 (part (iii)). Proposition 4 formalizes why *coarsened* network controls constructed from a $K \times K$ IO table cannot, in general, close the micro-level equilibrium backdoor.

Step 1: Sectoral projection and within-sector residual. Let $P \in \{0,1\}^{N\times K}$ be the partition matrix and $\Delta \equiv P'P = \mathrm{diag}(n_1, \ldots, n_K)$. The orthogonal projection onto the space of sector-constant vectors range(P) is

$$Q \equiv P\Delta^{-1}P', \qquad Q^2 = Q, \quad Q' = Q.$$

Let $R \equiv I - Q$ denote the residual-maker onto the orthogonal complement (within-sector deviations).

For any equilibrium state s_t, the best sector-level approximation (in least squares) is Qs_t, and the residual (within-sector component) is Rs_t. In particular, in the network model $s_t = s_0 + Mu_t$, the within-sector residual equilibrium component is

$$r_t \equiv RHMu_t,$$

because time FE remove s_0 (Proposition 1) and R removes sector means.

Step 2: IV after partialing out sector-level controls. Consider an IV regression that partials out a sector-level proxy $\hat{s}_t^{\text{agg}}$ that is constant within sectors, i.e., $\hat{s}_t^{\text{agg}} \in \text{range}(P)$. The displayed formula in Proposition 4(iii) is conditional on the projected source-exogeneity and primitive-orthogonality moments $\mathbb{E}[(R\tilde{Z}_t)' R\tilde{\varepsilon}_t] = 0$ and $\mathbb{E}[(R\tilde{Z}_t)' RHM\xi_t] = 0$. Augmenting the regression with the full set of sector-by-time controls is equivalent, by partialing out the controls, to applying the residual-maker R to the time-demeaned variables. (Equivalently: identification comes from within-sector cross-sectional variation.)

Define within-sector residualized variables:

$$\tilde{Y}_t^{(R)} \equiv R\tilde{Y}_t, \quad \tilde{X}_t^{(R)} \equiv R\tilde{X}_t, \quad \tilde{Z}_t^{(R)} \equiv R\tilde{Z}_t, \quad \tilde{s}_t^{(R)} = RHM(\kappa \tilde{Z}_t + \xi_t) = \kappa\, RHM\tilde{Z}_t + RHM\xi_t.$$

Applying R to the structural equation $\tilde{Y}_t = \beta\tilde{X}_t + \gamma\tilde{s}_t + \tilde{\varepsilon}_t$ gives

$$\tilde{Y}_t^{(R)} = \beta\tilde{X}_t^{(R)} + \gamma\tilde{s}_t^{(R)} + R\tilde{\varepsilon}_t.$$

The corresponding population IV estimator (based on within-sector variation) is

$$\beta^{IV|R} \equiv \frac{\mathbb{E}[\tilde{Z}_t^{(R)\prime}\tilde{Y}_t^{(R)}]}{\mathbb{E}[\tilde{Z}_t^{(R)\prime}\tilde{X}_t^{(R)}]}.$$

Substitute the residualized structural equation and use the projected source-exogeneity condition to obtain

$$\beta^{IV|R} = \beta + \gamma\, \frac{\mathbb{E}[\tilde{Z}_t^{(R)\prime}\tilde{s}_t^{(R)}]}{\mathbb{E}[\tilde{Z}_t^{(R)\prime}\tilde{X}_t^{(R)}]}.$$

Step 3: Express the bias term in terms of within-sector equilibrium leakage. Using $u_t = \kappa\tilde{Z}_t + \xi_t$ and $\tilde{s}_t = HMu_t$ after time FE, we have

$$\tilde{s}_t^{(R)} = RHM(\kappa\tilde{Z}_t + \xi_t) = \kappa\, RHM\tilde{Z}_t + RHM\xi_t.$$

Under the projected primitive-orthogonality condition, $\mathbb{E}[\tilde{Z}_t^{(R)\prime} RHM\xi_t] = 0$, hence

$$\mathbb{E}[\tilde{Z}_t^{(R)\prime}\tilde{s}_t^{(R)}] = \kappa\, \mathbb{E}[(R\tilde{Z}_t)' RHM\tilde{Z}_t].$$

Define $\pi^{(Q)}$ as the population slope from projecting $R\tilde{X}_t$ on $R\tilde{Z}_t$, and write $R\tilde{X}_t = \pi^{(Q)} R\tilde{Z}_t + \tilde{v}_t^{(Q)}$. By construction, $\mathbb{E}[(R\tilde{Z}_t)'\tilde{v}_t^{(Q)}] = 0$. Then

$$\mathbb{E}[\tilde{Z}_t^{(R)\prime}\tilde{X}_t^{(R)}] = \pi^{(Q)}\,\mathbb{E}[(R\tilde{Z}_t)'(R\tilde{Z}_t)].$$

Therefore

$$\beta^{IV|R} = \beta + \frac{\gamma\kappa}{\pi^{(Q)}} \underbrace{\frac{\mathbb{E}[(R\tilde{Z}_t)'RHM\tilde{Z}_t]}{\mathbb{E}[(R\tilde{Z}_t)'(R\tilde{Z}_t)]}}_{\equiv \mathcal{K}^{(Q)}(M,\{Z_t\})}.$$

This is the expression stated in Proposition 4, with $r_t \equiv RHMu_t$. Unless the within-sector residual loading $\mathcal{K}^{(Q)}$ is zero (which requires a knife-edge orthogonality between within-sector instrument variation and within-sector equilibrium propagation), coarse IO controls do not eliminate the equilibrium backdoor. □

IA.A.9 Proof of Corollary 2

Proof of Corollary 2. Apply Cauchy–Schwarz to the normalized numerator:

$$\left|\mathbb{E}\left[N^{-1}\tilde{Z}_t^{\perp\prime}\tilde{s}_t^{\perp}\right]\right| \le \left(\mathbb{E}\left[N^{-1}\tilde{Z}_t^{\perp\prime}\tilde{Z}_t^{\perp}\right]\right)^{1/2}\left(\mathbb{E}\left[N^{-1}\tilde{s}_t^{\perp\prime}\tilde{s}_t^{\perp}\right]\right)^{1/2}.$$

Since

$$\left|\mathbb{E}\left[N^{-1}\tilde{Z}_t^{\perp\prime}\tilde{X}_t^{\perp}\right]\right| \ge c > 0,$$

substituting this bound into (19) yields the result. □

IA.A.10 Proof of Proposition 5

Proof of Proposition 5. From Proposition 2, under the maintained assumptions the population IV estimator satisfies

$$\beta_N^{IV} = \beta + \frac{\gamma\,\mathcal{K}_N}{\pi_N},$$

where

$$\mathcal{K}_N = \kappa\,\frac{\mathbb{E}\big[\tilde{Z}_t' HM\tilde{Z}_t\big]}{\mathbb{E}\big[\tilde{Z}_t'\tilde{Z}_t\big]}.$$

If $|\mathcal{K}_N| \ge c > 0$ eventually and $\pi_N \longrightarrow 0$, then

$$\left|\beta_N^{IV} - \beta\right| = \left|\frac{\gamma\,\mathcal{K}_N}{\pi_N}\right| \ge |\gamma|\,\frac{c}{|\pi_N|} \longrightarrow \infty$$

whenever $\gamma \neq 0$. Hence $|\beta_N^{IV}| \to \infty$, with the sign determined by the eventual sign of $\gamma \mathcal{K}_N / \pi_N$.

The key point is that the equilibrium leakage term remains first order while the direct first stage vanishes, so the IV ratio explodes rather than attenuating toward the direct effect. $\square$

IA.A.11 Proof of Proposition 6

Proof of Proposition 6. Let $z \equiv Hd$. All differences are the cross-sectionally demeaned differences defined before Proposition 6. Starting from (22),

$$\Delta\widetilde{Y} = \beta\,\Delta\widetilde{X} + \gamma\,\Delta\widetilde{s} + \Delta\widetilde{\varepsilon},$$

take the inner product with the residualized treatment contrast z:

$$z'\Delta\widetilde{Y} = \beta\,z'\Delta\widetilde{X} + \gamma\,z'\Delta\widetilde{s} + z'\Delta\widetilde{\varepsilon}.$$

Source exogeneity gives $z'\Delta\widetilde{\varepsilon} = 0$. Using

$$\Delta\widetilde{s} = HM\Delta u \qquad \text{and} \qquad \Delta u = \kappa d + \Delta\xi,$$

we obtain

$$\Delta\widetilde{s} = \kappa\,HMd + HM\Delta\xi.$$

Therefore

$$z'\Delta\widetilde{s} = \kappa\,z'HMd + z'HM\Delta\xi.$$

Under primitive orthogonality in differences, $z'HM\Delta\xi = 0$, so

$$z'\Delta\widetilde{s} = \kappa\,z'HMd.$$

By definition of the residualized first-stage slope, $z'\Delta\widetilde{X} = \pi_D z'z$. Substituting into the DiD/IV-DiD ratio defined in Proposition 6 yields

$$\widehat{\beta}_{DiD} = \frac{z'\Delta\widetilde{Y}}{z'\Delta\widetilde{X}} = \beta + \frac{\gamma\kappa}{\pi_D}\frac{z'HMd}{z'z} = \beta + \frac{\gamma\kappa}{\pi_D}\frac{(Hd)'HMd}{(Hd)'Hd},$$

which is (23). Therefore, even if the untreated component $\Delta\widetilde{\varepsilon}$ satisfies parallel trends, outcomes need not do so when the treatment assignment creates a post-treatment equilibrium state aligned with the residualized treated-control contrast. $\square$

IA.A.12 Proof of Proposition 7

Proof of Proposition 7. The proof compares the feasible PCA correction with the infeasible correction based on the true factors, shows that the remaining network component becomes negligible, and then checks whether enough first-stage variation remains after the factor controls are added.

Step 0: Setup and moment convention. Let $\mathcal{F}_0$ be the sigma-field generated by the training sample used to estimate $\widehat{\Lambda}$. All target-sample expectations in this proof are conditional on $\mathcal{F}_0$, and we write $\mathbb{E}[\,\cdot\,]$ for $\mathbb{E}_1[\,\cdot \mid \mathcal{F}_0]$; a target-sample law of large numbers makes feasible 2SLS asymptotically equivalent to the corresponding conditional moment ratio. Randomness in $\widehat{\Lambda}$ therefore remains, the conditional moments below are random through the training sample, and all convergence statements are in probability over the training sample.

Starting from $HM = \Lambda F' + R$, define

$$f_t \equiv F' u_t, \qquad r_t \equiv R u_t,$$

so that $\tilde{s}_t = \Lambda f_t + r_t$. Let P_Λ denote the projection onto the columns of Λ, and define

$$\tilde{Z}_t^{\perp} \equiv (I - P_\Lambda)\tilde{Z}_t.$$

Let $\widehat{\Lambda}$ denote the PCA estimate of Λ, let $\widehat{P}_\Lambda$ be its loading-space projection, and define the feasible residual-maker

$$\Pi_{\widehat{\Lambda}}^{\perp} \equiv I - \widehat{P}_\Lambda.$$

Define feasible PCA-residualized variables $\tilde{Y}^{PCA} \equiv \Pi_{\widehat{\Lambda}}^{\perp}\tilde{Y}$, $\tilde{X}^{PCA} \equiv \Pi_{\widehat{\Lambda}}^{\perp}\tilde{X}$, and $\tilde{Z}^{PCA} \equiv \Pi_{\widehat{\Lambda}}^{\perp}\tilde{Z}$. The population IV estimator after partialing out the PCA factors is

$$\beta^{IV,PCA} \equiv \frac{\mathbb{E}[N^{-1}\tilde{Z}^{PCA\prime}\tilde{Y}^{PCA}]}{\mathbb{E}[N^{-1}\tilde{Z}^{PCA\prime}\tilde{X}^{PCA}]}.$$

The post-control moments used below are

$$\mathbb{E}[N^{-1}\tilde{Z}^{PCA\prime}\Pi_{\widehat{\Lambda}}^{\perp}\tilde{\varepsilon}] = 0, \qquad \mathbb{E}[N^{-1}\tilde{Z}'\Pi_{\widehat{\Lambda}}^{\perp}\tilde{\nu}] = 0,$$

conditional on $\mathcal{F}_0$. The statement that R is asymptotically negligible is used through

$$\mathbb{E}\left[\left|N^{-1}\tilde{Z}'(I - P_\Lambda)Ru_t\right|\right] = o_p(1),$$

with normalized instrument and state second moments bounded.

Step 1: PCA recovers the factor projection. Under standard pervasiveness conditions, there exists a rotation matrix G such that $N^{-1}\|\widehat{\Lambda} - \Lambda G\|_F^2 \xrightarrow{p} 0$, and consequently

$$\left\|\Pi_{\widehat{\Lambda}}^{\perp} - (I - P_\Lambda)\right\|_{op} \xrightarrow{p} 0.$$

Write $\Pi_\Lambda^{\perp} \equiv I - P_\Lambda$ for the infeasible residual-maker onto the true orthogonal complement.

Step 2: The controlled IV moment condition. Apply $\Pi_{\widehat{\Lambda}}^{\perp}$ to the structural equation and take the normalized inner product with $\tilde{Z}^{PCA}$. Under the stated post-control exclusion moment,

$$\beta^{IV,PCA} - \beta = \gamma\,\frac{\mathbb{E}[N^{-1}\tilde{Z}^{PCA\prime}\Pi_{\widehat{\Lambda}}^{\perp}\tilde{s}]}{\mathbb{E}[N^{-1}\tilde{Z}^{PCA\prime}\tilde{X}^{PCA}]}.$$

Step 3: The numerator is asymptotically negligible. Using $\tilde{s} = \Lambda f + r_t$ and adding and subtracting the infeasible projector,

$$\mathbb{E}[N^{-1}\tilde{Z}^{PCA\prime}\Pi_{\widehat{\Lambda}}^{\perp}\tilde{s}] = \mathbb{E}[N^{-1}\tilde{Z}'\Pi_{\Lambda}^{\perp}r_t] + \mathbb{E}[N^{-1}\tilde{Z}'(\Pi_{\widehat{\Lambda}}^{\perp} - \Pi_{\Lambda}^{\perp})\tilde{s}].$$

By the stated negligibility condition,

$$\mathbb{E}\left[\left|N^{-1}\tilde{Z}'\Pi_{\Lambda}^{\perp}r_t\right|\right] = o_p(1).$$

For the second term, Cauchy–Schwarz gives

$$\left|N^{-1}\tilde{Z}'(\Pi_{\widehat{\Lambda}}^{\perp} - \Pi_{\Lambda}^{\perp})\tilde{s}\right| \le (N^{-1}\|\tilde{Z}\|^2)^{1/2}\|\Pi_{\widehat{\Lambda}}^{\perp} - \Pi_{\Lambda}^{\perp}\|_{op}(N^{-1}\|\tilde{s}\|^2)^{1/2}.$$

Since the projection error is $o_p(1)$ and the normalized vector norms are $O_p(1)$, the numerator is $o_p(1)$.

Step 4: Granularity and the first stage after controls. Using $\tilde{X} = \pi\tilde{Z} + \tilde{v}$ and the stated post-control first-stage orthogonality, the denominator is

$$\pi\,\mathbb{E}[N^{-1}\tilde{Z}'\Pi_{\widehat{\Lambda}}^{\perp}\tilde{Z}].$$

If (26) holds, consistency of the estimated factor projection implies that this denominator remains bounded away from zero. If instead $\mathbb{E}[N^{-1}\|\tilde{Z}_t^{\perp}\|^2] \to 0$, the denominator converges to zero.

Step 5: Conclusion. Under (26), the numerator is $o_p(1)$ and the denominator remains bounded away from zero, so $\beta^{IV,PCA} - \beta = o_p(1)$. If $\mathbb{E}[N^{-1}\|\tilde{Z}_t^{\perp}\|^2] \to 0$, the factor controls absorb asymptotically all residual instrument variation and the first stage collapses. □

IA.A.13 Justification for Assumption 3 (Primitive Orthogonality)

Justification for Assumption 3. We show that Assumption 3 follows from the sufficient mean-independence condition $\mathbb{E}[\xi_t \mid Z_t] = 0$. Recall $u_t = \kappa \tilde{Z}_t + \xi_t$ and $\tilde{Z}_t = HZ_t$.

Compute the moment in Assumption 3:

$$\mathbb{E}\left[\tilde{Z}_t' HM\xi_t\right] = \mathbb{E}\left[Z_t' HM\xi_t\right],$$

since H is idempotent and symmetric and $\tilde{Z}_t = HZ_t$. Take iterated expectations conditioning on Z_t:

$$\begin{aligned}
\mathbb{E}\left[Z_t' HM\xi_t\right] &= \mathbb{E}\left[\mathbb{E}\left[Z_t' HM\xi_t \,\middle|\, Z_t\right]\right] \\
&= \mathbb{E}\left[Z_t' HM \, \mathbb{E}[\xi_t \mid Z_t]\right].
\end{aligned}$$

If $\mathbb{E}[\xi_t \mid Z_t] = 0$ (or more generally if $Z_t' HM \, \mathbb{E}[\xi_t \mid Z_t] = 0$ almost surely), then the inner term is identically zero, hence the unconditional expectation is zero. This is exactly Assumption 3. □

IA.A.14 Derivation: mediation rank failure for the controlled direct effect (Section 3)

Mediation rank failure for the controlled direct effect. We formalize the strongest route to exclusion: a relevant instrument direction that leaves the equilibrium mediator unchanged. From the structural equation after fixed effects,

$$\tilde{Y}_t = \beta \tilde{X}_t + \gamma \tilde{s}_t + \tilde{\varepsilon}_t, \qquad \tilde{s}_t = HMu_t,$$

and the maintained linear first stage $\tilde{X}_t = \pi \tilde{Z}_t + \tilde{v}_t$ with $\pi \neq 0$. In the linear mapping $u_t = \kappa \tilde{Z}_t + \xi_t$, an instrument direction $z \in \mathcal{V} \equiv \{z : \mathbf{1}' z = 0\}$ leaves the mediator unchanged only if

$$HMz = 0 \qquad \text{and} \qquad z \neq 0 \text{ (relevance)}. \tag{IA.A.2}$$

Thus the candidate instrument must lie in the kernel of the equilibrium transmission operator HM. Under the rank condition that HM is injective on $\mathcal{V}$, $HMz = 0$ and $z \in \mathcal{V}$ imply $z = 0$, so no nonzero mediator-invariant direction exists.

This result rules out a nonzero instrument that leaves the mediator unchanged under the stated rank condition. It does not rule out identification through the weaker moment condition $z' HMz = 0$, controls that span the induced equilibrium state, or additional instruments that separately identify the direct and equilibrium components. □

IA.A.15 Derivation: Time fixed effects and the total equilibrium effect (Section 4)

Derivation: Time fixed effects and the total equilibrium effect. We derive the total equilibrium effect of the instrument on outcomes and show why time fixed effects prevent identification of its level. For this derivation, impose the stronger conditional-mean restrictions

$$\mathbb{E}[\nu \mid Z] = 0, \qquad \mathbb{E}[\xi \mid Z] = 0, \qquad \mathbb{E}[\varepsilon \mid Z] = 0.$$

Start from the maintained linear structure (suppressing t):

$$Y = \beta X + \gamma s + \varepsilon, \qquad X = \pi Z + \nu, \qquad u = \kappa Z + \xi, \qquad s = s^{(0)} + Mu.$$

Hence

$$\mathbb{E}[X \mid Z] = \pi Z, \qquad \mathbb{E}[s \mid Z] = s^{(0)} + \kappa M Z,$$

and

$$\mathbb{E}[Y \mid Z] = \beta\pi Z + \gamma\kappa M Z + s_Y^{(0)},$$

where $s_Y^{(0)}$ is invariant to Z. The total general-equilibrium derivative of Y with respect to Z is therefore

$$\Theta \equiv \frac{\partial \mathbb{E}[Y \mid Z]}{\partial Z} = \beta\pi I + \gamma\kappa M.$$

Firm fixed effects first remove firm-specific components of $s_Y^{(0)}$. Time demeaning then maps the total-response operator into $H\Theta$. After both projections, the derivative of the Z-dependent conditional mean is

$$H(\beta\pi I + \gamma\kappa M) = H\Theta.$$

Because H has nontrivial nullspace span$\{\mathbf{1}\}$, the mapping $\Theta \mapsto H\Theta$ is not invertible. Many distinct total general-equilibrium responses generate the same demeaned image, and any common component is annihilated. A two-way-fixed-effects estimator can therefore recover at most a relative, design-weighted equilibrium response, not the absolute total general-equilibrium effect. □

IA.A.16 Derivation: Heterogeneous direct effects and centrality weighting (Section 4)

Heterogeneous direct effects and centrality weighting. Suppress t. The heterogeneous-coefficient equation is

$$\tilde{Y} = B\,\tilde{X} + \gamma\,\tilde{s} + \tilde{\varepsilon}, \qquad B \equiv \mathrm{diag}(\beta_1, \dots, \beta_N).$$

The population IV estimator using time-FE residuals is

$$\beta^{IV} = \frac{\mathbb{E}[\tilde{Z}'\tilde{Y}]}{\mathbb{E}[\tilde{Z}'\tilde{X}]} = \frac{\mathbb{E}[\tilde{Z}'\tilde{Y}]}{\mathbb{E}[\tilde{Z}'\tilde{X}]}.$$

Substitute $\tilde{Y} = B\tilde{X} + \gamma\tilde{s} + \tilde{\varepsilon}$:

$$\mathbb{E}\left[\tilde{Z}'\tilde{Y}\right] = \mathbb{E}\left[\tilde{Z}'B\tilde{X}\right] + \gamma\,\mathbb{E}\left[\tilde{Z}'\tilde{s}\right] + \mathbb{E}\left[\tilde{Z}'\tilde{\varepsilon}\right].$$

Under the maintained source exogeneity, $\mathbb{E}[\frac{1}{N}\tilde{Z}'\tilde{\varepsilon}] = 0$. Moreover,

$$\tilde{Z}'B\tilde{X} = \sum_{i=1}^{N} \tilde{Z}_i\,\beta_i\,\tilde{X}_i.$$

Therefore

$$\mathbb{E}\left[\frac{1}{N}\tilde{Z}'B\tilde{X}\right] = \frac{1}{N}\sum_{i=1}^{N} \beta_i\,\mathbb{E}[\tilde{Z}_i\tilde{X}_i].$$

Define weights

$$\omega_i \equiv \frac{\mathbb{E}[\tilde{Z}_i\tilde{X}_i]}{\sum_{j=1}^{N}\mathbb{E}[\tilde{Z}_j\tilde{X}_j]}, \qquad \sum_{i=1}^{N}\omega_i = 1,$$

so that

$$\frac{\mathbb{E}[\frac{1}{N}\tilde{Z}'B\tilde{X}]}{\mathbb{E}[\frac{1}{N}\tilde{Z}'\tilde{X}]} = \sum_{i=1}^{N}\omega_i\,\beta_i.$$

Hence

$$\beta^{IV} = \sum_{i=1}^{N}\omega_i\,\beta_i + \gamma\,\frac{\mathbb{E}[\tilde{Z}'\tilde{s}]}{\mathbb{E}[\tilde{Z}'\tilde{X}]}.$$

Finally, use the first-stage decomposition $\tilde{X} = \pi_0\tilde{Z} + \pi_1 HM\tilde{Z} + \tilde{v}$ and the orthogonality $\mathbb{E}[\tilde{Z}'\tilde{v}] = 0$ to compute the weights:

$$\mathbb{E}[\tilde{Z}_i\tilde{X}_i] = \pi_0\,\mathbb{E}[\tilde{Z}_i^2] + \pi_1\,\mathbb{E}[\tilde{Z}_i(HM\tilde{Z})_i].$$

In the shift-share special case $\tilde{Z}_i = b_i g_t$ with scalar shock g_t, this becomes

$$\mathbb{E}[\tilde{Z}_i\tilde{X}_i] = \mathbb{E}[g_t^2]\Big(\pi_0 b_i^2 + \pi_1 b_i(HMb)_i\Big),$$

so

$$\omega_i \propto \pi_0 b_i^2 + \pi_1 b_i(HMb)_i,$$

as claimed. ∎

IA.A.17 Derivation of the Generalized IV Decomposition (Section 4)

The generalized IV decomposition. Suppress t subscripts. Start from the maintained structural equation

$$\tilde{Y} = \beta\tilde{X} + \gamma\tilde{s} + \tilde{\varepsilon}, \qquad \tilde{s} = HMu, \qquad u = \kappa\tilde{Z} + \xi.$$

The population IV estimator based on $\tilde{Z}$ is

$$\beta^{IV} = \frac{\mathbb{E}[\tilde{Z}'\tilde{Y}]}{\mathbb{E}[\tilde{Z}'\tilde{X}]}.$$

Substituting the structural equation yields

$$\beta^{IV} = \beta + \frac{\mathbb{E}[\tilde{Z}'\tilde{\varepsilon}]}{\mathbb{E}[\tilde{Z}'\tilde{X}]} + \gamma\,\frac{\mathbb{E}[\tilde{Z}'\tilde{s}]}{\mathbb{E}[\tilde{Z}'\tilde{X}]}.$$

Next expand the equilibrium term using $u = \kappa\tilde{Z} + \xi$:

$$\tilde{s} = HM(\kappa\tilde{Z} + \xi) = \kappa\,HM\tilde{Z} + HM\xi.$$

Therefore

$$\mathbb{E}[\tilde{Z}'\tilde{s}] = \kappa\,\mathbb{E}[\tilde{Z}'HM\tilde{Z}] + \mathbb{E}[\tilde{Z}'HM\xi].$$

Substituting back gives

$$\beta^{IV} = \beta + \frac{\mathbb{E}[\tilde{Z}'\tilde{\varepsilon}]}{\mathbb{E}[\tilde{Z}'\tilde{X}]} + \frac{\gamma\kappa\,\mathbb{E}[\tilde{Z}'HM\tilde{Z}]}{\mathbb{E}[\tilde{Z}'\tilde{X}]} + \frac{\gamma\,\mathbb{E}[\tilde{Z}'HM\xi]}{\mathbb{E}[\tilde{Z}'\tilde{X}]},$$

which is (18). Finally, if the first stage satisfies

$$\mathbb{E}[\tilde{Z}'\tilde{X}] = \pi\,\mathbb{E}[\tilde{Z}'\tilde{Z}],$$

then the middle term becomes

$$\frac{\gamma\kappa\,\mathbb{E}[\tilde{Z}'HM\tilde{Z}]}{\mathbb{E}[\tilde{Z}'\tilde{X}]} = \frac{\gamma\mathcal{K}}{\pi},$$

giving the network term stated in Section 4. ∎

IA.A.18 Derivation: Dynamic Event-Study Contamination (Section 4.5)

Dynamic event-study contamination. Consider the dynamic network equilibrium state

$$s_t = \Phi s_{t-1} + \Psi u_t, \qquad \rho(\Phi) < 1.$$

Let the policy path have $u_t^{(1)} = \kappa D \mathbf{1}\{t \geq T_0\}$ and the no-policy path have $u_t^{(0)} = 0$, with the same initial state $s_{T_0-1}^{(1)} = s_{T_0-1}^{(0)}$. For $\tau \geq 0$, subtracting the two recursively generated paths gives

$$s_{T_0+\tau}^{(1)} - s_{T_0+\tau}^{(0)} = \kappa \left(\sum_{k=0}^{\tau} \Phi^k \Psi \right) D.$$

This is the cumulative step response to the persistent policy shock; the matrices $\{\Phi^k \Psi\}_{k \geq 0}$ are the incremental impulse-response weights.

With structural outcome equation

$$Y_t = \beta X_t + \gamma s_t + \varepsilon_t,$$

define Δ^p as the policy path minus the no-policy path. Then

$$\Delta^p Y_{T_0+\tau} = \beta\, \Delta^p X_{T_0+\tau} + \gamma \kappa \left(\sum_{k=0}^{\tau} \Phi^k \Psi \right) D.$$

The treatment-induced network component is zero before T_0 and turns on only afterward, so flat pre-treatment coefficients do not rule out dynamic post-treatment contamination. If transmission is delayed, that contamination can be small on impact and appear at later event times. □

IA.B Formal Details for Diagnostics and Remedies

This appendix collects the formal details behind Section 5: the eigenvalue representation of alignment, observed-network diagnostic formulas, hidden-network factor-count screens, residual sensitivity bounds, remedy-validity conditions, and the weak-factor limit.

IA.B.1 Symmetric representation and eigen-alignment

Section 3 defines

$$S \equiv \frac{1}{2}(HMH + HM'H). \tag{IA.B.1}$$

For any exposure vector z, $z'Sz = (Hz)'HM(Hz)$. In particular, for any residualized exposure vector satisfying $Hz = z$,

$$z'HMz = z'Sz. \tag{IA.B.2}$$

The eigenvalue expansion below shows how each network direction contributes to the alignment score.

On the mean-zero subspace, let $v_1, \ldots, v_{N-1}$ be the eigenvectors of S (defined in (16)) with eigenvalues $\lambda_1, \ldots, \lambda_{N-1}$. For any nonzero residualized static policy assignment z satisfying $Hz = z$, write

$$z = \sum_{j=1}^{N-1} a_j v_j. \tag{IA.B.3}$$

Then

$$K_0(z) = \frac{z'Sz}{z'z} = \sum_{j=1}^{N-1} \lambda_j \omega_j(z), \qquad \omega_j(z) = \frac{a_j^2}{\sum_{\ell=1}^{N-1} a_\ell^2}.$$

The weights are nonnegative and sum to one. Consequently,

$$|K_0(z)| \leq \lambda_{\max}^{\text{abs}}(S), \qquad \lambda_{\max}^{\text{abs}}(S) \equiv \max_j |\lambda_j(S)|,$$

and Proposition 2 implies

$$|\beta^{IV} - \beta| \leq \left|\frac{\gamma\kappa}{\pi}\right| \lambda_{\max}^{\text{abs}}(S).$$

IA.B.2 Observed-network diagnostic formulas

Let $\widehat{A}$ be the observed proxy network. In the linear benchmark,

$$\widehat{M} = \widehat{A}(I - \widehat{A})^{-1}, \qquad \widehat{S} = \frac{1}{2}(H\widehat{M}H + H\widehat{M}'H).$$

Let $|\hat{\lambda}_{(1)}| \geq |\hat{\lambda}_{(2)}| \geq \cdots$ be the ordered absolute eigenvalues of $\widehat{S}$ after applying the same mean-zero restriction used in the empirical design. Choose an upper search bound J_A and set

$$\hat{r}_{AH} \in \arg\max_{1 \leq j \leq J_A} \frac{|\hat{\lambda}_{(j)}|}{|\hat{\lambda}_{(j+1)}|}.$$

The idiosyncratic eigenvalue benchmark is

$$\hat{\lambda}_{\text{idio}} \equiv |\hat{\lambda}_{(\hat{r}_{AH}+1)}|. \tag{IA.B.4}$$

For a static DiD assignment or policy indicator, let z denote the residualized assignment vector and compute $\widehat{K}_0(z) = \frac{z'\widehat{S}z}{z'z}$. For a panel-IV instrument, compute

$$\widehat{k}_t = \frac{\tilde{Z}_t'\widehat{S}\tilde{Z}_t}{\tilde{Z}_t'\tilde{Z}_t}, \qquad \widehat{K}_Z^{\text{pool}} = \frac{\sum_t \tilde{Z}_t'\widehat{S}\tilde{Z}_t}{\sum_t \tilde{Z}_t'\tilde{Z}_t}.$$

For a deterministic residualized design z, define

$$K_{\text{auto},0}(z) = \frac{z'HM_{\text{diag}}z}{z'z}, \qquad K_{\text{net},0}(z) = \frac{z'HM_{\text{off}}z}{z'z}.$$

Then $K_0(z) = K_{\text{auto},0}(z) + K_{\text{net},0}(z)$. The first is the relative own-unit component after demeaning; the second is the cross-unit Network-SUTVA component. For a measured network, replace M with $\widehat{M}$.

For staggered DiD or event-study designs, let $z_{g,\ell}$ be the residualized cohort-by-event-time policy indicator and compute

$$\widehat{K}_{0,g,\ell} = \frac{z_{g,\ell}'\widehat{S}z_{g,\ell}}{z_{g,\ell}'z_{g,\ell}}.$$

Optional observed-network screen. Define

$$D_A \equiv \begin{cases} |\widehat{K}_0(z)|, & \text{static DiD or policy indicator,} \\ \max_{g,\ell} |\widehat{K}_{0,g,\ell}|, & \text{staggered DiD or event-study design,} \\ |\widehat{K}_Z^{\text{pool}}| \text{ or } \max_t |\widehat{k}_t|, & \text{panel-IV design.} \end{cases}$$

If the researcher adopts this prespecified screen, the design is flagged as observed-network exposed when $D_A > \hat{\lambda}_{\text{idio}}$.

Raw versus de-meaned assignment in DiD. For a static DiD design the symmetric diagnostic and the exact bias loading are different objects. The following identity connects them and shows which measured control spans the state that DiD loads on.

Lemma 2 (Raw versus de-meaned DiD propagation)**.** Let $d \in \mathbb{R}^N$, $z \equiv Hd$, $\bar{d} \equiv N^{-1}\mathbf{1}'d$, and $S \equiv \frac{1}{2}(HMH + HM'H)$, and set $K_D \equiv z'Sz/(z'z)$ and $L_D \equiv z'HMd/(z'z)$. Then, with $q_{\text{raw}} \equiv H\widehat{M}d$, $q_{\text{res}} \equiv H\widehat{M}(Hd)$, and $q_{\mathbf{1}} \equiv H\widehat{M}\mathbf{1}$ for a measured multiplier $\widehat{M}$,

$$L_D = K_D + \bar{d}\,\frac{z'HM\mathbf{1}}{z'z}, \qquad q_{\text{raw}} = q_{\text{res}} + \bar{d}\,q_{\mathbf{1}}. \tag{IA.B.5}$$

Hence the included control space containing q_{res} spans q_{raw} if and only if $\bar{d}\,q_{\mathbf{1}}$ lies in that joint control space. In particular, this condition holds if the omitted term vanishes.

Proof. Because $z = Hd$ is mean-zero, $Hz = z$ and $z'H = z'$, so $z'Sz = \frac{1}{2}(z'HMHz + z'HM'Hz) = \frac{1}{2}(z'Mz + z'M'z) = z'Mz = z'HMz$, the third equality because $z'M'z = (z'Mz)'$ is a scalar. Since $Hd = d - \bar{d}\mathbf{1}$ we have $d = z + \bar{d}\,\mathbf{1}$, so $z'HMd = z'HMz + \bar{d}\,z'HM\mathbf{1}$; dividing by $z'z$ gives the first identity. Applying $H\widehat{M}$ to $d = Hd + \bar{d}\,\mathbf{1}$ gives the second. □

Thus K_D and L_D coincide only when $\bar{d} = 0$ or $z'HM\mathbf{1} = 0$; for a binary treatment with a substantial treated share, neither holds in general. For panel-IV designs the distinction is vacuous, because the primitive map is written in terms of the residualized instrument, which already satisfies $H\widetilde{Z}_t = \widetilde{Z}_t$ and hence has zero cross-sectional mean.

IA.B.3 Hidden-network residual diagnostics

These diagnostics are computed from residual series formed before estimating the target causal coefficient, after removing the baseline fixed effects and predetermined controls.

Let $\hat{\rho}_{ij}$ be the sample correlation between residual series i and j. The signed dependence statistic is

$$CD = \sqrt{\frac{2T}{N(N-1)}} \sum_{i<j} \hat{\rho}_{ij}.$$

Let p_{CD} denote the corresponding p-value for this statistic. Stack residuals as $\widehat{E} = [\hat{e}_1, \ldots, \hat{e}_T]$ and form the residual correlation matrix. Let $\hat{\mu}_{(1)} \geq \hat{\mu}_{(2)} \geq \cdots$ be its ordered eigenvalues. Choose J_e and a separation threshold $\tau_{\text{sep}} > 1$. Define $\hat{q}_e \equiv \max_{1 \leq j \leq J_e} \frac{\hat{\mu}_{(j)}}{\hat{\mu}_{(j+1)}}$. If $\hat{q}_e < \tau_{\text{sep}}$, set $\hat{r}_e = 0$. If $\hat{q}_e \geq \tau_{\text{sep}}$, set

$$\hat{r}_e \in \operatorname*{arg\,max}_{1 \leq j \leq J_e} \frac{\hat{\mu}_{(j)}}{\hat{\mu}_{(j+1)}}.$$

Optional low-rank screen. If the researcher adopts this prespecified low-rank screen at size α, the residuals are flagged as containing a hidden low-rank component when $p_{CD} < \alpha$ and $\hat{r}_e \geq 1$. Design relevance is $1 - \widehat{g}(C_H) = R^2_{C_H}(\widetilde{Z})$, computed with the frozen factor-by-time controls used in the final regression and with the residualized DiD treatment in DiD.

IA.B.4 Residual loading and sensitivity bounds

After factor controls, let P_r project onto them and define $\tilde{Z}_t^{\perp} = (I - P_r)\tilde{Z}_t$ as the remaining instrument variation. The residual equilibrium loading is

$$\mathcal{K}_{\text{res}}(\tilde{Z}_t) \equiv \kappa \frac{\mathbb{E}[(\tilde{Z}_t^{\perp})'(I - P_r)HM\tilde{Z}_t]}{\mathbb{E}[(\tilde{Z}_t^{\perp})'\tilde{Z}_t^{\perp}]}. \tag{IA.B.6}$$

Define $g_r \equiv \frac{\mathbb{E}\|\tilde{Z}_t^{\perp}\|^2}{\mathbb{E}\|\tilde{Z}_t\|^2}$, the population analogue of the granularity diagnostic in Section 5.4. Let m_r capture the maximum amplification of any unit-variance shock under the remaining propagation operator, defined formally as $m_r \equiv \sup_{\|x\|=1} \|(I - P_r)HMx\|$. If $g_r > 0$, then Cauchy-Schwarz implies

$$|\mathcal{K}_{\text{res}}| \leq |\kappa| \frac{m_r}{\sqrt{g_r}}. \tag{IA.B.7}$$

Proof. Starting from (IA.B.6), Cauchy-Schwarz gives

$$\left|(\tilde{Z}_t^{\perp})'(I - P_r)HM\tilde{Z}_t\right| \leq \|\tilde{Z}_t^{\perp}\| \, \|(I - P_r)HM\tilde{Z}_t\|.$$

By the definition of m_r,

$$\|(I - P_r)HM\tilde{Z}_t\| \leq m_r \|\tilde{Z}_t\|.$$

Taking expectations and applying Cauchy-Schwarz again,

$$\left|\mathbb{E}[(\tilde{Z}_t^{\perp})'(I - P_r)HM\tilde{Z}_t]\right| \leq m_r \sqrt{\mathbb{E}\|\tilde{Z}_t^{\perp}\|^2} \sqrt{\mathbb{E}\|\tilde{Z}_t\|^2}.$$

Dividing by $\mathbb{E}\|\tilde{Z}_t^{\perp}\|^2$ and multiplying by $|\kappa|$ gives (IA.B.7). This establishes formally that the factor remedy bounds the remaining equilibrium leakage. The residual bias becomes small provided that the chosen factor space successfully captures the bulk of the network spillovers (driving m_r toward zero) while leaving the instrument sufficiently granular (keeping the denominator $\sqrt{g_r}$ bounded away from zero). □

Combining (IA.B.7) with Proposition 2 yields the residual sensitivity interval:

$$\beta \in \left[\beta_{\text{corr}}^{IV} - \left|\frac{\gamma\kappa}{\pi_r}\right| \frac{m_r}{\sqrt{g_r}}, \quad \beta_{\text{corr}}^{IV} + \left|\frac{\gamma\kappa}{\pi_r}\right| \frac{m_r}{\sqrt{g_r}}\right],$$

where π_r is the post-control first-stage coefficient. This interval applies to panel-IV designs.

DiD sensitivity interval. The corresponding DiD interval follows from Proposition 10. Let C_D denote the predetermined DiD control set, let $g_D(C_D)$ be DiD granularity, and let $m_D(C_D) = \|R_D HMd\|/\|Hd\|$ be the residual-state scale defined there. Any maintained bound $m_D(C_D) \leq \bar{m}_D$ gives

$$\beta \in \left[\beta_{D,C} - \left|\frac{\gamma\kappa}{\pi_{D,C}}\right| \frac{\bar{m}_D}{\sqrt{g_D(C_D)}}, \quad \beta_{D,C} + \left|\frac{\gamma\kappa}{\pi_{D,C}}\right| \frac{\bar{m}_D}{\sqrt{g_D(C_D)}}\right]. \tag{IA.B.8}$$

When the network is measured, Corollary 5 supplies $\bar{m}_D$ in terms of network measurement error. When the relevant network is hidden, $\bar{m}_D$ is a sensitivity parameter unless external information or a validated factor model bounds the part of the design-induced state left outside the estimated loading space.

IA.B.5 When the proposed remedies restore identification

This subsection provides the formal sense in which the remedies in Section 5 can restore the controlled direct effect. The results do not claim that arbitrary network or factor controls *fully* solve Network-SUTVA. Rather, once the selected controls capture the relevant equilibrium state, the remaining bias is bounded by the network propagation left outside the controls divided by the post-control first stage.

Panel IV and DiD are treated separately, because the two designs induce different equilibrium states and therefore have different recovery formulas. In the panel-IV model the primitive shock is driven by the residualized instrument, $u_t = \kappa\widetilde{Z}_t + \xi_t$ with $H\widetilde{Z}_t = \widetilde{Z}_t$, so the induced state is built from an already mean-zero object. In the two-period DiD model of Proposition 6 the primitive shock is driven by raw assignment, $\Delta u = \kappa d + \Delta\xi$, so the induced state is HMd and the relevant control must span that vector. Propositions 8 and 9 and Corollary 4 are panel-IV results. Proposition 10 and Corollaries 5 and 6 are the corresponding DiD results.

Panel-IV remedies.

Proposition 8 (Panel IV: observed-network correction and residual bound)**.** Consider the residualized model

$$\tilde{Y}_t = \beta\tilde{X}_t + \gamma\tilde{s}_t + \tilde{\varepsilon}_t, \qquad \tilde{s}_t = HMu_t, \qquad u_t = \kappa\tilde{Z}_t + \xi_t.$$

Let $C_A \in \mathbb{R}^{N\times r_A}$ be a predetermined observed-network loading space, constructed from the measured topology before the realization of the outcome and entered with unrestricted period coefficients. Let P_A be the corresponding within-period orthogonal projection onto $\operatorname{span}(C_A)$, and define

$$\tilde{Z}_{A,t} \equiv (I - P_A)\tilde{Z}_t, \qquad \tilde{X}_{A,t} \equiv (I - P_A)\tilde{X}_t, \qquad \tilde{Y}_{A,t} \equiv (I - P_A)\tilde{Y}_t.$$

Suppose the post-control first stage satisfies

$$\mathbb{E}\left[N^{-1}\tilde{Z}'_{A,t}\tilde{X}_{A,t}\right] = \pi_A\,\mathbb{E}\left[N^{-1}\tilde{Z}'_{A,t}\tilde{Z}_{A,t}\right], \qquad |\pi_A| \geq \underline{\pi} > 0,$$

and the post-control source-exogeneity and primitive-orthogonality conditions hold:

$$\mathbb{E}\left[N^{-1}\tilde{Z}'_{A,t}(I - P_A)\tilde{\varepsilon}_t\right] = 0, \qquad \mathbb{E}\left[N^{-1}\tilde{Z}'_{A,t}(I - P_A)HM\xi_t\right] = 0.$$

Then the population 2SLS estimand after including the observed-network controls is

$$\beta_A^{IV} = \beta + \frac{\gamma\kappa}{\pi_A}\frac{\mathbb{E}\left[N^{-1}\tilde{Z}'_{A,t}(I - P_A)HM\tilde{Z}_t\right]}{\mathbb{E}\left[N^{-1}\tilde{Z}'_{A,t}\tilde{Z}_{A,t}\right]}.$$

Define the residual propagation norm and the granularity ratio

$$m_A \equiv \|(I - P_A)HM\|_{\mathrm{op}}, \qquad g_A \equiv \frac{\mathbb{E}\|\tilde{Z}_{A,t}\|^2}{\mathbb{E}\|\tilde{Z}_t\|^2}.$$

If $g_A > 0$, then

$$|\beta_A^{IV} - \beta| \le \left|\frac{\gamma\kappa}{\pi_A}\right| \frac{m_A}{\sqrt{g_A}}.$$

Consequently, if $m_A = O(\delta_A)$, $g_A \ge \underline{g} > 0$, and $|\pi_A| \ge \underline{\pi} > 0$, then

$$\beta_A^{IV} = \beta + O(\delta_A).$$

If $(I - P_A)HM = 0$, the observed-network correction restores the controlled direct effect exactly, provided the post-control first stage remains nonzero.

Proof. Apply $I - P_A$ to the structural equation:

$$\tilde{Y}_{A,t} = \beta\tilde{X}_{A,t} + \gamma(I - P_A)HMu_t + (I - P_A)\tilde{\varepsilon}_t.$$

Taking the inner product with $\tilde{Z}_{A,t}$, using source exogeneity, and substituting $u_t = \kappa\tilde{Z}_t + \xi_t$, we obtain

$$\mathbb{E}[N^{-1}\tilde{Z}'_{A,t}\tilde{Y}_{A,t}] = \beta\mathbb{E}[N^{-1}\tilde{Z}'_{A,t}\tilde{X}_{A,t}] + \gamma\kappa\mathbb{E}[N^{-1}\tilde{Z}'_{A,t}(I - P_A)HM\tilde{Z}_t],$$

where the primitive-orthogonality condition removes the term involving ξ_t. Since $\mathbb{E}[N^{-1}\tilde{Z}'_{A,t}\tilde{X}_{A,t}] = \pi_A\mathbb{E}[N^{-1}\|\tilde{Z}_{A,t}\|^2]$, division gives the result.

For the bound, Cauchy–Schwarz and the definition of m_A imply

$$\left|\mathbb{E}[\tilde{Z}'_{A,t}(I - P_A)HM\tilde{Z}_t]\right| \le m_A\sqrt{\mathbb{E}\|\tilde{Z}_{A,t}\|^2}\sqrt{\mathbb{E}\|\tilde{Z}_t\|^2}.$$

Dividing by $\mathbb{E}\|\tilde{Z}_{A,t}\|^2$ gives

$$\left|\frac{\mathbb{E}[\tilde{Z}'_{A,t}(I - P_A)HM\tilde{Z}_t]}{\mathbb{E}[\tilde{Z}'_{A,t}\tilde{Z}_{A,t}]}\right| \le \frac{m_A}{\sqrt{g_A}}.$$

Multiplying by $|\gamma\kappa/\pi_A|$ proves the result. □

Corollary 3 (Panel IV: direct design-generated state control)**.** Let $q_t \equiv H\widehat{M}\tilde{Z}_t$ be included in both stages of the stacked panel-IV regression, together with the baseline controls. Let superscript $\perp q$ denote population residualization with respect to this final control set. Under the corresponding post-control source-exogeneity and primitive-orthogonality conditions and a nonzero post-control first stage π_q,

$$\beta_q^{IV} - \beta = \frac{\gamma\kappa}{\pi_q} \frac{\mathbb{E}[N^{-1}(\tilde{Z}_t^{\perp q})'\{H(M - \widehat{M})\tilde{Z}_t\}^{\perp q}]}{\mathbb{E}[N^{-1}(\tilde{Z}_t^{\perp q})'\tilde{Z}_t^{\perp q}]}.$$

Thus if $\widehat{M} = M$, the direct state control recovers β exactly whenever the post-control first stage remains nonzero. More generally, with $g_q \equiv \mathbb{E}\|\tilde{Z}_t^{\perp q}\|^2 / \mathbb{E}\|\tilde{Z}_t\|^2$,

$$|\beta_q^{IV} - \beta| \leq \left|\frac{\gamma\kappa}{\pi_q}\right| \frac{\|H(M - \widehat{M})\|_{\text{op}}}{\sqrt{g_q}}.$$

Corollary 4 (Panel IV: measured topology and Leontief-amplified error)**.** Suppose C_A is chosen as the leading r_A left singular space of $H\widehat{M}$, where $\widehat{M} = \widehat{A}(I - \widehat{A})^{-1}$. Then

$$m_A = \|(I - P_A)HM\|_{\text{op}} \leq \sigma_{r_A+1}(H\widehat{M}) + \|H(M - \widehat{M})\|_{\text{op}}.$$

If $\widehat{A} = A + \Omega$, then

$$\|H(M - \widehat{M})\|_{\text{op}} \leq \|(I - \widehat{A})^{-1}\|_{\text{op}} \|\Omega\|_{\text{op}} \|(I - A)^{-1}\|_{\text{op}}.$$

Therefore the observed-network correction restores identification up to

$$O\left(\frac{\sigma_{r_A+1}(H\widehat{M}) + \|(I - \widehat{A})^{-1}\|_{\text{op}} \|\Omega\|_{\text{op}} \|(I - A)^{-1}\|_{\text{op}}}{\sqrt{g_A}}\right),$$

provided the post-control first stage remains bounded away from zero.

Proposition 9 (Panel IV: out-of-sample factor correction)**.** Suppose the residual propagation operator admits the approximate factor representation

$$HM = \Lambda F' + R,$$

where $\Lambda \in \mathbb{R}^{N \times r}$ has fixed r, P_Λ denotes the projection onto $\text{span}(\Lambda)$, and R is the residual high-rank or weak-factor component. Let $\widehat{P}_\Lambda$ be an out-of-sample or cross-fitted PCA estimate of P_Λ, constructed from periods excluded from the identifying comparison, and suppose

$$\|\widehat{P}_\Lambda - P_\Lambda\|_{\text{op}} = a_{NT} = o_p(1).$$

Assume also that $\|\Lambda F'\|_{\text{op}}$ is bounded, or more generally that $a_{NT}\|\Lambda F'\|_{\text{op}} = O_p(\delta_{P,NT})$. Let

$$\tilde{Z}_{L,t} \equiv (I - \widehat{P}_\Lambda)\tilde{Z}_t$$

be the factor-residualized instrument. Define analogously $\tilde{X}_{L,t} \equiv (I - \widehat{P}_\Lambda)\tilde{X}_t$ and $\tilde{Y}_{L,t} \equiv (I - \widehat{P}_\Lambda)\tilde{Y}_t$. Suppose the post-control first stage satisfies

$$\mathbb{E}[N^{-1}\tilde{Z}'_{L,t}\tilde{X}_{L,t}] = \pi_L \mathbb{E}[N^{-1}\tilde{Z}'_{L,t}\tilde{Z}_{L,t}], \qquad |\pi_L| \geq \underline{\pi} > 0,$$

and define

$$g_L \equiv \frac{\mathbb{E}\|\tilde{Z}_{L,t}\|^2}{\mathbb{E}\|\tilde{Z}_t\|^2}.$$

Assume $g_L \geq \underline{g} > 0$ with probability approaching one, and impose the post-control exogeneity and primitive-orthogonality conditions analogous to Proposition 8. Then the out-of-sample factor-controlled IV estimand satisfies

$$|\beta_L^{IV} - \beta| \leq \left|\frac{\gamma\kappa}{\pi_L}\right| \frac{\|R\|_{\text{op}} + a_{NT}\|\Lambda F'\|_{\text{op}}}{\sqrt{g_L}} + o_p(1).$$

Consequently, if $\|R\|_{\text{op}} = O(\delta_{R,N})$, $a_{NT}\|\Lambda F'\|_{\text{op}} = O_p(\delta_{P,NT})$, $g_L \geq \underline{g} > 0$, and $|\pi_L| \geq \underline{\pi} > 0$, then

$$\beta_L^{IV} = \beta + O_p(\delta_{R,N} + \delta_{P,NT}).$$

In the exact low-rank case $R = 0$, the remaining bias is governed only by the out-of-sample factor-space estimation error. If the factor is weak so that $a_{NT} \not\to 0$, the PCA correction need not restore identification.

Proof. Let $\widehat{P} = \widehat{P}_\Lambda$. As in the observed-network proof, after partialing out the estimated factor space,

$$\beta_L^{IV} - \beta = \frac{\gamma\kappa}{\pi_L} \frac{\mathbb{E}[N^{-1}\tilde{Z}'_{L,t}(I - \widehat{P})HM\tilde{Z}_t]}{\mathbb{E}[N^{-1}\tilde{Z}'_{L,t}\tilde{Z}_{L,t}]}.$$

Using $HM = \Lambda F' + R$,

$$(I - \widehat{P})HM = (I - \widehat{P})\Lambda F' + (I - \widehat{P})R.$$

Since $(I - P_\Lambda)\Lambda = 0$,

$$(I - \widehat{P})\Lambda = (P_\Lambda - \widehat{P})\Lambda,$$

so

$$\|(I - \widehat{P})\Lambda F'\|_{\text{op}} \leq \|\widehat{P} - P_\Lambda\|_{\text{op}}\|\Lambda F'\|_{\text{op}} = a_{NT}\|\Lambda F'\|_{\text{op}}.$$

Also,

$$\|(I - \widehat{P})R\|_{\text{op}} \leq \|R\|_{\text{op}}.$$

Therefore

$$\|(I - \widehat{P})HM\|_{\text{op}} \leq a_{NT}\|\Lambda F'\|_{\text{op}} + \|R\|_{\text{op}}.$$

Applying the same Cauchy–Schwarz argument as in Proposition 8 yields

$$|\beta_L^{IV} - \beta| \le \left|\frac{\gamma\kappa}{\pi_L}\right| \frac{\|R\|_{\mathrm{op}} + a_{NT}\|\Lambda F'\|_{\mathrm{op}}}{\sqrt{g_L}} + o_p(1).$$

The $o_p(1)$ term reflects the out-of-sample generated-regressor approximation and vanishes under the maintained cross-fitting or pre-treatment estimation design. □

Difference-in-differences remedies. The results above are panel-IV statements. The DiD design of Proposition 6 induces the state HMd from raw assignment, so its remedy requires a separate result. The next proposition gives the generic controlled-DiD estimand and bound; the two corollaries specialize it to a measured network and to estimated factors.

Proposition 10 (DiD: predetermined controls in two-period difference-in-differences)**.** Consider the two-period DiD model of Proposition 6, with $z \equiv Hd$, $\Delta u = \kappa d + \Delta\xi$, and $\Delta s = HM\Delta u$. Let $C_D \in \mathbb{R}^{N\times r_D}$ be a predetermined control matrix, already residualized with respect to the same fixed effects and baseline controls as the treatment, so $HC_D = C_D$; let P_D project onto $\mathrm{span}(C_D)$, $R_D \equiv I - P_D$, and set $z_D \equiv R_D z$, $\Delta X_D \equiv R_D H\Delta X$, $\Delta Y_D \equiv R_D H \Delta Y$. Define DiD granularity and the design-specific residual-state scale by

$$g_D(C_D) \equiv \frac{\|z_D\|^2}{\|z\|^2}, \qquad m_D(C_D) \equiv \frac{\|R_D HMd\|}{\|z\|}, \tag{IA.B.9}$$

and assume $g_D(C_D) > 0$. Suppose the post-control first stage satisfies $z_D'\Delta X_D = \pi_{D,C} z_D' z_D$ with $\pi_{D,C} \neq 0$, and impose the post-control source-exogeneity and primitive-orthogonality conditions $z_D' R_D H\Delta\varepsilon = 0$ and $z_D' R_D HM\Delta\xi = 0$ (because $z_D' R_D = z_D'$, the outer R_D is redundant and is kept only for symmetry with Proposition 8). Then the population DiD estimand after including C_D, and its bound, are

$$\beta_{D,C} \equiv \frac{z_D'\Delta Y_D}{z_D'\Delta X_D} = \beta + \frac{\gamma\kappa}{\pi_{D,C}} \frac{z_D' R_D HMd}{z_D' z_D}, \qquad |\beta_{D,C} - \beta| \le \left|\frac{\gamma\kappa}{\pi_{D,C}}\right| \frac{m_D(C_D)}{\sqrt{g_D(C_D)}}. \tag{IA.B.10}$$

Consequently, if $HMd \in \mathrm{span}(C_D)$, the controlled DiD estimand recovers β exactly. When the policy indicator itself is the regressor and the same fixed effects and controls are used in the treatment and outcome equations, $\pi_{D,C} = 1$. With no controls, $R_D = I$ and (IA.B.10) reduces to the bias $\gamma\kappa L_D/\pi_D$ of Proposition 6.

Proof. Applying $R_D H$ to the differenced outcome equation and using $H^2 = H$ gives $\Delta Y_D = \beta\Delta X_D + \gamma R_D HM(\kappa d + \Delta\xi) + R_D H\Delta\varepsilon$. Take the inner product with z_D; source exogeneity removes the last term and primitive orthogonality the term in $\Delta\xi$, leaving $z_D'\Delta Y_D = \beta z_D'\Delta X_D +$

$\gamma\kappa z_D' R_D HMd$. Dividing by $z_D'\Delta X_D = \pi_{D,C} z_D' z_D$, nonzero because $g_D(C_D) > 0$ and $\pi_{D,C} \neq 0$, gives the estimand. Cauchy–Schwarz gives $|z_D' R_D HMd| \leq \|z_D\| \|R_D HMd\|$, so the bias factor is at most $\|R_D HMd\|/\|z_D\| = m_D(C_D)/\sqrt{g_D(C_D)}$, using $\|z_D\| = \|z\|\sqrt{g_D(C_D)}$. If $HMd \in \operatorname{span}(C_D)$ then $R_D HMd = 0$. If $\Delta X = d$ then $\Delta X_D = R_D Hd = z_D$, so $\pi_{D,C} = 1$. Setting C_D empty gives $R_D = I$, $z_D = z$, and $z'HMd/(z'z) = L_D$. □

Corollary 5 (DiD: measured raw-assignment control)**.** Let the control be the measured design-induced state $q_C \equiv H\widehat{M}d$, so $C_D = q_C$. Then $R_D HMd = R_D H(M - \widehat{M})d$ and $m_D(C_D) \leq \|H(M - \widehat{M})d\|/\|Hd\|$; if $\widehat{M} = M$, the control spans the treatment-induced equilibrium state exactly, so the controlled DiD estimand equals β whenever $g_D(C_D) > 0$. If $\widehat{A} = A + \Omega$, then

$$\|H(M - \widehat{M})d\| \leq \|(I - \widehat{A})^{-1}\|_{\mathrm{op}} \|\Omega\|_{\mathrm{op}} \|(I - A)^{-1}\|_{\mathrm{op}} \|d\|.$$

The residual bias is therefore governed by network measurement error in the specific design-induced direction d, divided by surviving DiD granularity. For comparison, let $q_{\mathrm{res}} = H\widehat{M}(Hd)$, let P_D^{res} project onto $\operatorname{span}(q_{\mathrm{res}})$, and set $R_D^{\mathrm{res}} = I - P_D^{\mathrm{res}}$. By Lemma 2, when $\widehat{M} = M$ the alternative control leaves $R_D^{\mathrm{res}} HMd = \bar{d}\, R_D^{\mathrm{res}} HM\mathbf{1}$, so it does not in general span HMd.

Proof. Because $R_D q_C = 0$ we have $R_D H\widehat{M}d = 0$, so $R_D HMd = R_D H(M - \widehat{M})d$; an orthogonal residual maker has operator norm one, giving the bound on m_D after dividing by $\|z\| = \|Hd\|$. The displayed inequality follows from the resolvent identity $\widehat{M} - M = (I - \widehat{A})^{-1}\Omega(I - A)^{-1}$, submultiplicativity, and $\|H\|_{\mathrm{op}} = 1$. Exact recovery when $\widehat{M} = M$ follows from Proposition 10, and the last claim is (IA.B.5) with $R_D^{\mathrm{res}} q_{\mathrm{res}} = 0$. □

Corollary 6 (DiD: factor correction)**.** In the setting of Proposition 10, suppose the design-induced state decomposes as $HMd = \Lambda a_D + r_D$, where $\Lambda \in \mathbb{R}^{N\times r}$ contains stable population factor loadings. Let P_Λ project onto $\operatorname{span}(\Lambda)$ and let $\widehat{P}_\Lambda$ be an out-of-sample or cross-fitted estimate constructed from observations excluded from the target estimation, with $\|\widehat{P}_\Lambda - P_\Lambda\|_{\mathrm{op}} = a_{NT} = o_p(1)$. Using the estimated loadings as C_D, so $R_D = I - \widehat{P}_\Lambda$, and setting $g_{D,L} \equiv \|(I - \widehat{P}_\Lambda)Hd\|^2/\|Hd\|^2$, the post-control conditions of Proposition 10 give

$$|\beta_{D,L} - \beta| \leq \left|\frac{\gamma\kappa}{\pi_{D,L}}\right| \frac{\|r_D\| + a_{NT}\|\Lambda a_D\|}{\|Hd\|\sqrt{g_{D,L}}}.$$

Consequently, if $\|r_D\| + a_{NT}\|\Lambda a_D\| = o_p(\|Hd\|)$, and $g_{D,L}$ and the post-control first stage are bounded away from zero, then $\beta_{D,L} \to_p \beta$. If $g_{D,L} \to 0$, the factor controls absorb the DiD treatment variation and the direct effect is not separately identified. The requirement is that the estimated loading space span the design-induced state HMd; the operator HM need not be globally low rank.

Proof. Because $(I - P_\Lambda)\Lambda a_D = 0$ we have $\Lambda a_D = P_\Lambda \Lambda a_D$, so $(I - \widehat{P}_\Lambda)HMd = (P_\Lambda - \widehat{P}_\Lambda)\Lambda a_D + (I - \widehat{P}_\Lambda)r_D$ and hence $\|(I - \widehat{P}_\Lambda)HMd\| \leq a_{NT}\|\Lambda a_D\| + \|r_D\|$, using $\|I - \widehat{P}_\Lambda\|_{\mathrm{op}} = 1$. Substituting into (IA.B.10) with $\|z_D\| = \|Hd\|\sqrt{g_{D,L}}$ gives the bound; the consistency and collapse statements follow from the stated conditions on the numerator, granularity, and post-control first stage. □

IA.B.6 Weak-factor fragility

Proposition 11 (Weak-factor fragility)**.** Suppose the residual panel has the one-spike covariance representation $\Sigma_e = \sigma_\eta^2(I_N + \theta v v')$, with $\|v\| = 1$, $Hv = v$, and $N/T \to c \in (0, \infty)$. Assume the residual observations satisfy the standard spiked-Wishart eigenvector-overlap conditions, for example independent Gaussian or suitable sub-Gaussian isotropic noise. Let $\hat{v}$ be the leading sample eigenvector. If $\theta \leq \sqrt{c}$, then $|\hat{v}'v|^2 \xrightarrow{p} 0$. If $\theta > \sqrt{c}$, then

$$|\hat{v}'v|^2 \xrightarrow{p} \frac{1 - c/\theta^2}{1 + c/\theta}.$$

Consequently, below the weak-factor threshold, PCA does not recover the latent equilibrium direction. Above the threshold, PCA is informative, but the recovered factor is still imperfect unless the signal is sufficiently strong.

If the residualized instrument is asymptotically aligned with the true factor,

$$\frac{\|(I - P_v)\tilde{Z}_t\|}{\|\tilde{Z}_t\|} \xrightarrow{p} 0,$$

and $v'Sv \to \lambda_S \neq 0$, then a PCA control based on $\hat{v}$ leaves an $O(1)$ equilibrium-loading term when $\theta \leq \sqrt{c}$.

Proof. The BBP phase transition for the fixed-spike covariance model implies that the sample eigenvector separates from the random-matrix bulk if and only if $\theta > \sqrt{c}$. When $\theta \leq \sqrt{c}$,

$$|\hat{v}'v|^2 \xrightarrow{p} 0.$$

When $\theta > \sqrt{c}$, the squared overlap converges to

$$\frac{1 - c/\theta^2}{1 + c/\theta}.$$

Now suppose $\tilde{Z}_t = a_t v + o_p(\|\tilde{Z}_t\|)$, with $a_t \neq 0$ in probability. Below the threshold,

$$(I - P_{\hat{v}})\tilde{Z}_t = \tilde{Z}_t - \hat{v}(\hat{v}'\tilde{Z}_t) = a_t v + o_p(\|\tilde{Z}_t\|),$$

because $\hat{v}'v \xrightarrow{p} 0$. Therefore

$$\frac{((I - P_{\hat{v}})\tilde{Z}_t)'S((I - P_{\hat{v}})\tilde{Z}_t)}{((I - P_{\hat{v}})\tilde{Z}_t)'((I - P_{\hat{v}})\tilde{Z}_t)} = v'Sv + o_p(1) = \lambda_S + o_p(1).$$

If $\lambda_S \neq 0$, the residual equilibrium-loading term remains $O(1)$. Thus PCA has not removed the equilibrium direction. It has projected out an essentially orthogonal noise direction. □

IA.B.7 Implementation of the remedies

This subsection gives implementation details for the remedies summarized in Section 5. The alignment and residual-comovement diagnostics are formalized in Appendices IA.B.2 and IA.B.3; granularity and the remedy-validity conditions are stated in Section 5.4 and Appendix IA.B.5.

IA.B.7.1 Observed-network remedy

When physical links are measured, form $\widehat{M} = \widehat{A}(I - \widehat{A})^{-1}$ and $\widehat{S} = \frac{1}{2}(H\widehat{M}H + H\widehat{M}'H)$. Appendix IA.B.2 gives the alignment scores for static DiD, event-study, and panel-IV designs and, if desired, the optional observed-network exposure screen.

For a single-policy DiD with assignment d, the design-induced equilibrium state is HMd, because the DiD data-generating process sends raw assignment through the network and the fixed effects demean the resulting state only afterwards. The remedy therefore adds $q_C = H\widehat{M}d$ to the corrected regression, interacted with Post_t or with event time. This is the measured-network analogue of the design-induced state: Corollary 5 shows that it spans that state exactly when $\widehat{M} = M$, and characterizes the residual component left by network measurement error otherwise. For panel IV, construct $q_t = H\widehat{M}\widetilde{Z}_t$ from the residualized instrument and include it in both stages, or partial it out before IV estimation; Corollary 3 gives the corresponding population result.

When more than one propagated direction is needed, let $\widehat{v}_1, \ldots, \widehat{v}_{r_A}$ denote residualized directions selected by the rule in Appendix IA.B.2 and define $\widehat{q}_j = H\widehat{M}\widehat{v}_j / \|H\widehat{M}\widehat{v}_j\|$. A DiD specification then includes $\widehat{q}_{j,i} \times \text{Post}_t$ for $j = 1, \ldots, r_A$. Granularity is computed using the same propagated controls included in the corrected regression. For sampling inference, the external measured network $\widehat{M}$ is treated as fixed; network-measurement error is addressed separately by the measurement-error results in Appendix IA.B.5.

IA.B.7.2 Hidden-network remedy

When propagation is unobserved, the remedy learns a stable low-rank equilibrium component from residual comovement outside the identifying comparison. Let $\mathcal{T}_0$ denote pre-treatment or held-out training periods and $\mathcal{T}_1$ the estimation periods. In $\mathcal{T}_0$, remove the baseline fixed effects and predetermined controls, excluding the target causal regressor, and stack the resulting residuals as $\widehat{E}_0$. Select the factor count r using Appendix IA.B.3, extract the first r principal components, and freeze the loadings $\widehat{\Lambda}_r$ for the estimation sample. The training window must also excite the design-induced state: by Corollary 6, the correction is point-identifying only if the estimated loading space spans HMd up to a remainder small relative to $\|Hd\|$. Report the selected factor count, training window, factor-selection output, design-factor overlap, and granularity against the full time-interacted factor-control set.

For $k = 1, \ldots, r$ and $\tau \in \mathcal{T}_1$, define the factor-by-time controls $\widehat{C}_{it}^{(k,\tau)} = \widehat{\Lambda}_{ik}\mathbf{1}\{t = \tau\}$. DiD and event-study specifications include these controls directly. For panel IV, partial out the same frozen controls from the outcome, endogenous regressor, and excluded instrument, or equivalently include them in both stages. Let $\tilde{Z}_{r,it}$ denote the excluded instrument after this projection. The resulting 2SLS system is

$$\tilde{X}_{it} = \pi_r \tilde{Z}_{r,it} + \sum_{k=1}^{r} \sum_{\tau \in \mathcal{T}_1} \theta_{k\tau} \widehat{C}_{it}^{(k,\tau)} + \tilde{\nu}_{it}^{(r)},$$

$$\tilde{Y}_{it} = \beta_r \tilde{X}_{it} + \sum_{k=1}^{r} \sum_{\tau \in \mathcal{T}_1} \psi_{k\tau} \widehat{C}_{it}^{(k,\tau)} + \tilde{\varepsilon}_{it}^{(r)}.$$

with $\tilde{Z}_{r,it}$ as the excluded instrument. Granularity is the surviving instrument or treatment variation after projection on this frozen factor-control space. Inference should re-estimate the training residuals, loadings, controls, and final regression inside the bootstrap or sample-splitting procedure. This re-estimation accounts for generated-regressor uncertainty in the estimated factor loadings and controls; it is an inference issue for the estimated remedy, not the source of Network-SUTVA bias.

IA.C BEA Simulation Design and Calibration Details

This appendix collects the implementation and calibration details for the BEA-calibrated Monte Carlo in Section 6. It gives the shift-share derivation behind the scalar design, the full data-generating process including the hidden low-rank factor, the per-parameter calibration, and the simulation procedure. Throughout, the BEA topology is held fixed and all notation matches Section 6.

In the BEA cost-share matrix, row i is the customer or input-using industry and column j is the supplier or input-source industry, matching the orientation of A in the main text. The multiplier is $M = A(I - A)^{-1}$. The construction gives $\rho(A) = 0.8975$, a leading eigenvalue of S of approximately 1.870, a leading-to-second absolute-eigenvalue ratio of 1.21, a leading squared-spectrum share of 16.2%, an effective rank of 17.9, leading-eigenvector effective support of 26, and top-ten squared-loading mass of 42.3%. The low-alignment direction h is unit norm, mean zero, orthogonal to $v_\star$, and selected numerically to satisfy $|K_0(h)| < 10^{-4}$.

IA.C.1 Scalar Shift-Share Derivation

The IV design is a finite-sample scalar shift-share analogue. A scalar aggregate structural shock g_t loads on the unit-norm exposure direction $z(a)$. For finite-sample generation we use

$$z^{MC}(a) \equiv \sqrt{N}\, z(a),$$

which has unit cross-sectional RMS, and set $Z_t = g_t z^{MC}(a)$ with $\mathbb{E}[g_t^2] = 1$. This rescaling leaves $K_0(z(a))$ unchanged because the alignment score is scale invariant.

A standard shift-share instrument instead has raw exposure

$$Z_{it}^{raw} = \sum_{k=1}^{K} w_{ik} g_{kt}.$$

Stacking exposures in W and shocks in g_t, the demeaned instrument is $\widetilde{Z}_t = HWg_t$. Under the primitive-shock loading in (11), the equilibrium loading factor is

$$\mathcal{K} = \kappa \frac{\mathbb{E}[g_t' W' HMHW g_t]}{\mathbb{E}[g_t' W' HW g_t]}.$$

Equivalently, if g_t has covariance Σ_g, the numerator and denominator can be written as trace expressions involving $W'HMHW$ and $W'HW$. In the scalar simulation, W collapses to one exposure direction. Because $z(a)$ is mean-zero and the common $\sqrt{N}$ scale cancels,

$$\mathcal{K} = \kappa \frac{\mathbb{E}[g_t^2] z^{MC}(a)' HM z^{MC}(a)}{\mathbb{E}[g_t^2] z^{MC}(a)' z^{MC}(a)} = \kappa \frac{z(a)' S z(a)}{z(a)' z(a)} \equiv \kappa K_0(z(a)).$$

Thus the scalar Monte Carlo is the one-shock version of a Bartik design in which the stochastic covariance collapses to the deterministic alignment score $K_0(z(a))$.

IA.C.2 The Hidden Low-Rank Factor and Orthogonality

The DGP nests the IV and DiD exercises, with notation chosen to make the connection to the primitive-shock equation in Section 3 explicit. Recall $u_t = \kappa \widetilde{Z}_t + \xi_t$, where ξ_t is orthogonal to the instrument after fixed effects. In the simulation, write $\xi_t^{MC} \equiv \xi_t$ and decompose this orthogonal component into low-rank and idiosyncratic parts:

$$\xi_t^{MC} \equiv \eta f_t \ell + \zeta_t.$$

The residualized IV equations, after firm and time fixed effects, are

$$X_{it} = \pi Z_{it} + \nu_{it}, \qquad u_t = \kappa Z_t + \xi_t^{MC}, \qquad s_t = HMu_t, \qquad Y_{it} = \beta X_{it} + \gamma s_{it} + \varepsilon_{it}.$$

Here X_{it} is endogenous, π is the first-stage coefficient, and ν_{it} is first-stage noise. The vectors $u_t \in \mathbb{R}^N$ and $s_t \in \mathbb{R}^N$ are the primitive shock and residual equilibrium state, and ε_{it} is idiosyncratic outcome noise. The scalar β is the controlled direct effect, κ maps exposure into primitive shocks, and γ maps the state into outcomes. The vector $\zeta_t \in \mathbb{R}^N$ is the idiosyncratic component of the primitive disturbance. The scalar f_t is an unobserved aggregate macroeconomic shock, and $\ell \in \mathbb{R}^N$ is its stable spatial loading. In the IV rotation, $\ell = v_\star$.

The latent factor follows

$$f_t = \rho_f f_{t-1} + e_t, \qquad |\rho_f| < 1,$$

where e_t is a mean-zero innovation. The term $\eta f_t \ell$ is the low-rank component of ξ_t^{MC} used to evaluate the PCA remedy, not a second instrument or policy shock. The known-DGP recovery exercise includes this stable component so that the hidden-network procedure has a loading pattern to estimate from training residuals. If $\eta = 0$, there is no corresponding latent low-rank component to recover.

This decomposition preserves the primitive-orthogonality condition. In the IV design, $Z_t = g_t z^{MC}(a)$, and the latent factor f_t and idiosyncratic component ζ_t are both independent of the aggregate instrument shock g_t. Hence

$$\mathbb{E}[Z_t' HM \xi_t^{MC}] = \eta\, \mathbb{E}[g_t f_t] z^{MC}(a)' HM\ell + \mathbb{E}[g_t z^{MC}(a)' HM\zeta_t] = 0.$$

Thus the latent component can align with an important BEA direction without violating orthogonality. The orthogonality is temporal: the latent factor is not systematically related to the instrument realization.

IA.C.3 The DiD Data-Generating Process

For DiD, the equations below describe the DGP, and the Monte Carlo estimator applies the same unit and time fixed effects as the event-study specification. The continuous instrument and first stage are replaced by a binary treatment design. Let $d \in \{0,1\}^N$ denote the spatial treatment-assignment vector, where $d_i = 1$ for treated industries. Let t^* denote the treatment date, define event time $\tau = t - t^*$, and let $\text{Post}_t = \mathbf{1}\{\tau \geq 0\}$. The treatment vector and unit-level treatment indicator are

$$D_t = d\, \text{Post}_t, \qquad D_{it} = d_i\, \text{Post}_t, \qquad X_{it} = D_{it}.$$

For the DiD exercise, the latent-factor loading is $\ell = d$, so

$$\xi_t^{MC} \equiv \eta f_t d + \zeta_t, \qquad u_t = \kappa_{\text{did}} D_t + \xi_t^{MC}, \qquad s_t = HMu_t, \qquad Y_{it} = \beta D_{it} + \gamma s_{it} + \varepsilon_{it}.$$

As in the IV design, treatment affects outcomes directly and also changes the equilibrium state through the BEA network. The latent factor is present before treatment and, with $\ell = d$, its propagated loading is HMd, the same equilibrium direction generated by treatment. It is independent of treatment timing, so it is neither an anticipation effect nor a second policy shock. This construction evaluates the hidden-network procedure when the target loading is present in the training sample. If that loading is absent, the procedure cannot recover it from those residuals, and the appropriate output is sensitivity or bounds rather than a corrected point estimate.

The event-study contamination exercise reported in Figure 2 sets $\eta = 0$ so that the display isolates the treatment-induced Network-SUTVA channel. The Table 1 recovery exercise uses $\eta = 4.5$ so that the hidden-network correction can be evaluated. In both exercises the population treatment-induced post coefficient remains 2.121, because the latent factor is independent of treatment timing.

Selection of the treated set. The binary treatment vector d is selected from the BEA matrix alone, before any outcome is simulated, by the following pre-specified rule. Let $v_\star$ be the unit leading eigenvector of S and $q_\star = HMv_\star/\|HMv_\star\|$ its BEA-predicted equilibrium state. For each candidate direction $b \in \{v_\star, q_\star, -v_\star, -q_\star\}$ and each cutoff τ on an equally spaced grid of 31 points over $[0.60, 0.90]$, form

$$d_i^{(b,\tau)} = \mathbf{1}\{b_i \geq Q_\tau(b)\},$$

where $Q_\tau(b)$ is the τ quantile of b. Retain a candidate only if its treated share lies in $[0.10, 0.40]$ and its granularity against $q_\star$ exceeds 0.25. Among the retained candidates, select the one with the largest $|K_D|$. Applied to the 2017 BEA matrix, the rule selects $b = -v_\star$ at $\tau = 0.67$, giving 133 treated industries out of 402 (33.1%) and $|K_D| = 0.904$. Equivalently, the treated set is the third of industries with the most negative loadings on $v_\star$. The sign of an eigenvector is arbitrary, so the rule places treatment at whichever pole of the leading eigenvector yields the larger $|K_D|$. This is a disclosed BEA-only network stress-design search. No simulated outcome or target coefficient enters the choice.

IA.C.4 Why the DiD calibration reports both K_D and L_D

Section 6.3 reports two related DiD quantities. Let $d \in \mathbb{R}^N$ denote the cross-sectional treatment vector, and let H remove the cross-sectional mean. The first quantity is

$$K_D = K_0(Hd) = \frac{(Hd)'S(Hd)}{(Hd)'Hd}, \qquad S = \frac{1}{2}(HMH + HM'H).$$

It asks whether the demeaned DiD regressor Hd is aligned with the network object built from that same demeaned regressor. The second quantity is

$$L_D = \frac{(Hd)'HMd}{(Hd)'Hd}.$$

It applies because the simulated network state is generated from raw assignment d and demeaned before entering the fixed-effect regression, giving HMd. Thus K_D scores the demeaned DiD regressor itself, while L_D scores the network response generated by raw assignment. This is why Section 6.3 reports both numbers. In the BEA calibration, $K_D = 0.904$ and $L_D = 1.121$.

IA.C.5 Calibration of the Finite-Sample and Factor Parameters

The population bias is closed form, so holding K_0 and π fixed, its elasticity with respect to γ and κ is exactly one. Varying a by ±30% around the baseline gives $a \in \{0.49, 0.70, 0.91\}$, $K_0(z(a)) \in \{0.916, 1.309, 1.701\}$, and population IV biases of 91.6%, 130.9%, and 170.1% when $\gamma = \kappa = \pi = 1$. The broader pre-specified formula grid over $\gamma \in \{0.5, 1, 2\}$ and $\pi \in \{0.5, 1, 1.5\}$ spans 30.5% to 680.5%. These are calibration sensitivities, not sampling confidence intervals. This formula grid is distinct from the disclosed BEA-only treated-set search above, which uses no simulated outcome or target coefficient.

The finite-sample parameters $\sigma_\nu, \sigma_\xi, \sigma_\varepsilon, \eta$, and ρ_f do not change $\gamma\kappa K_0/\pi$. They affect standard errors, finite-sample dispersion, factor recovery, and first-stage diagnostics. The innovations and idiosyncratic shocks are mutually independent and mean zero, while f_t follows the AR(1) process above. We set $\sigma_\nu = 0.350$, $\sigma_\xi = 0.050$ for the idiosyncratic component ζ_t, and $\sigma_\varepsilon = 0.1000$. With unit-RMS Monte Carlo exposure and unit-variance g_t, $\sigma_\nu = 0.350$ gives per-cell $F_{\text{eff}} \approx 1/0.350^2 = 8.16$, consistent with the reported value 8.12 and the pooled first-stage scale. The recovery experiment uses $\rho_f = 0.50$ and $\eta = 4.500$. These parameters govern the persistence and strength of the low-rank component and therefore affect its recovery by out-of-sample PCA without making the factor observed.

Simulation procedure. The Monte Carlo proceeds in five steps. First, using the BEA matrix A, construct the equilibrium objects from Section 3 and the leading eigenvector $v_\star$ of S. Second, use the unit, mean-zero direction $h \perp v_\star$ with $|K_0(h)| < 10^{-4}$ and form $z(a)$ by rotating between h and $v_\star$. Third, generate $R = 1000$ Monte Carlo panels using the structural equations above. The IV panel has $T = 200$ periods. The DiD recovery design uses 40 training periods and event times $\tau = -8, \ldots, 8$, with $\tau = -1$ omitted. Fourth, estimate the naive IV and DiD designs. Fifth, apply the remedies.

The observed IV correction uses $q^\star = HM\ell/\|HM\ell\|$ as a cross-sectional loading with unrestricted time variation, while observed DiD includes $q_D \times \mathrm{Post}_t$ with $q_D = HMd/\|HMd\|$. The hidden correction estimates loadings from training or held-out residuals, freezes them, and enters factor-by-time controls. In IV, the same final control space is partialled from Y, X, and Z. Granularity is computed relative to the actual final control space. The granularity threshold used to select candidate DiD treatment groups is distinct from the final remedy granularity reported in Table 1.