The Bullwhip: Time-to-Build and Sectoral Fluctuations*

Yan Leng — *UT Austin*

Ernest Liu — *Princeton and NBER*

Yifei Ren — *Princeton*

Aleh Tsyvinski — *Yale and NBER*

First draft: August 2024
This draft: August 2026

Abstract

We develop a theory of sectoral fluctuations driven by the propagation of demand shocks along supply chains with heterogeneous time-to-build production, and we solve the model in closed form. Downstream producers respond to current demand; upstream producers, who commit production in advance due to time-to-build, respond to expected demand at the horizon of their accumulated delays. The cross-section of upstream responses therefore traces out the expected time path of downstream demand: shocks that decay monotonically are dampened as they propagate upstream, while hump-shaped shocks—under which higher demand today signals higher demand ahead—are amplified, creating the bullwhip effect. Empirically, hump-shaped, amplified upstream impulse responses appear in supply chains attached to downstream sectors whose demand carries momentum, and not in those with monotone-decay demand. Quantitatively, hump-shaped shocks account for 75 percent of supply chain volatility in U.S. manufacturing.

*We thank Pol Antras, Ariel Burstein, Vasco Carvalho, Chad Jones, Aubhik Khan, Pete Klenow, Jean-Paul L'Huillier, Ezra Oberfield, Stephen Redding, Alireza Tahbaz-Salehi, Mathieu Taschereau-Dumouchel, Julia K. Thomas, Christopher Tonetti, our discussants Ben Golub, Glenn Magerman, and seminar participants at Cambridge, EAGLS, Princeton, Chicago Fed, Imperial College London, Penn State, Ohio State, UIUC, UCLA, UVA, and NBER ITI for very helpful comments and suggestions. We are particularly grateful to George-Marios Angeletos for a number of insightful suggestions. Ernest Liu acknowledges the support from the Sloan Foundation and the NSF.

1 Introduction

When a wrangler flicks a bullwhip, the motion propagates along its length, growing stronger and faster until the tip, or cracker, moves so fast that it breaks the sound barrier. Fluctuations in supply chains exhibit a similar phenomenon, with demand shocks originating downstream amplifying significantly as they propagate upstream.

We develop a theory of sectoral fluctuations driven by the dynamic propagation of demand shocks in a production network with heterogeneous time-to-build production delays, and we solve the model in closed form. Downstream producers respond immediately to current demand changes, whereas upstream producers, due to time-to-build, react based on expectations of future demand. The central result of the paper is a mapping from the temporal profile of downstream demand shocks into the cross-sectional pattern of responses along the supply chain: a supplier whose output takes three periods to reach the downstream sector must commit production three periods in advance, and so responds today to demand expected three periods ahead; the cross-section of upstream responses therefore traces out the expected time path of downstream demand. Hence demand shocks that decay monotonically are dampened as they propagate upstream, whereas hump-shaped demand shocks, under which higher demand today signals even higher demand in the near future, elicit amplified responses from upstream producers; this amplification is the bullwhip effect. Empirically and quantitatively, we show that this mapping is reflected in U.S. sectoral data: supply chains attached to downstream sectors with hump-shaped demand display amplified upstream fluctuations, whereas those attached to sectors with monotone-decay demand do not.

Our model builds on two canonical models: the time-to-build environment of Kydland and Prescott (1982) as a theory of aggregate fluctuations, and the neoclassical sectoral model of Long and Plosser (1983), in which all inputs require exactly one period to build, in the presence of either aggregate or sectoral productivity shocks. Methodologically, we develop and solve in closed form a production network model with sectoral demand shocks that explicitly incorporates heterogeneous time-to-build frictions. Substantively, our focus is on when sectoral demand shocks are amplified, rather than dampened, as they propagate through supply chains. The result is a tractable framework for analyzing how the temporal structure of demand interacts with heterogeneous time-to-build dynamics along supply chains.

Our theory of sectoral fluctuations features two core elements: sectoral demand shocks, arising from consumers' time-varying and stochastic preferences for different final goods, and heterogeneous time-to-build delays throughout the supply chain. Demand shocks take two forms—monotone-decay and hump-shaped—and each sector's demand innovation can be a mixture of the two. We fully characterize the equilibrium, analyzing the interplay between demand shocks, heterogeneous delays, and the cross-sectoral linkages of the economy.

The primary difficulty in characterizing the effects of sector-specific demand shocks in a time-to-build environment is as follows. Producers choose inputs in anticipation of future demand at every horizon, and expected future demand reflects both the exogenous impulse response of consumer demand and the endogenous responses of other producers—each situated at a different position in the supply chain, and each potentially facing distinct time-to-build delays. Each producer must therefore account simultaneously for the entire time path of demand and for the actions of every other producer along the chain. Formally, the planner's problem is a path-dependent stochastic optimal control problem: the value function depends on input decisions made at all previous dates across heterogeneous delays and cannot be summarized by a low-dimensional state variable.

Our first contribution is an analytical characterization of the solution to this model for general sectoral demand shock processes, expressed explicitly in terms of the model's primitives: the demand shocks, the production network structure, and the heterogeneity in time-to-build delays. The central object of the analytical characterization is the impulse response of sectoral revenue to sectoral demand shocks affecting various parts of the supply chain. Specifically, Theorem 1 provides a closed form solution for each sector's revenue at a given time as an explicit function of expected consumer demands across all future time horizons and across all different goods that directly or indirectly rely on that sector as an input supplier.

Formally, sectoral revenue is a summation over future horizons, each term the product of two objects derived from primitives: the conditional expectation of future consumer demand across goods, and a horizon-specific weight matrix that collects all direct and indirect network paths whose cumulated time-to-build delays equal that horizon. Unlike the Leontief inverse, which aggregates direct and indirect linkages equally across horizons, these weights distinguish linkages by the time at which they operate—precisely because delays differ across supplier–user pairs. Both the weights and the resulting impulse responses of sectoral revenue are derived in closed form from the demand process, the sectoral linkages, and the time-to-build structure.

We then specialize the demand shock process into two distinct stochastic components. The first component captures transient changes in consumer preferences, modeled as monotone-decay AR(1) shocks, where a demand innovation today is expected to diminish monotonically over time. The second component captures slower-developing changes, such as those arising from habit formation, consumer fads, or delayed adoption processes. These are modeled as hump-shaped AR(2) shocks, where a current increase in demand initially leads to increased demand at future horizons before eventually declining.

Theorem 2 characterizes the revenue response to each component. A monotone-decay shock propagates upstream with each round of network effects discounted both by the discount rate and by the rate at which the shock dissipates; its impact therefore declines with supply chain distance. A hump-shaped shock instead generates a non-monotone impulse response—the difficult part of the analysis—which we derive explicitly as the product of an arithmetically increasing and an exponentially decaying term.

In order to illustrate the intuition, consider an increase in consumer demand for automobiles. How producers respond to this shock depends crucially on their positions along the automobile supply chain. Downstream producers—those that are close in distance to the shock, such as automakers and manufacturers of tires or glass windows—respond by ramping up production to meet current or near-future expected demand. In contrast, producers located further upstream—such as manufacturers of semiconductors or steel alloy, which are indirectly used in auto manufacturing as intermediate components—face significant time-to-build delays. As a result, these upstream producers adjust their production levels primarily based on anticipated future demand rather than current demand.

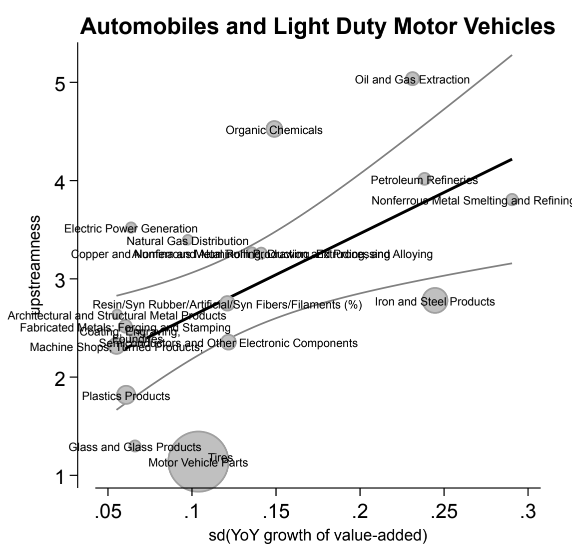
When a demand shock is hump-shaped—meaning higher automobile demand today signals even stronger demand in the future—upstream producers have significantly amplified responses, creating the bullwhip effect. This happens because distant suppliers respond more strongly to anticipated future peaks in demand, compared to downstream producers that mainly respond to current or short-term demand. Conversely, when demand shocks monotonically decline over time, downstream producers are always more responsive than upstream producers, preventing any amplified response along the supply chain.

The positions of producers in the supply chain, and the distances between them, play an important role in our analysis. A key feature of our model is the notion of delay—how long it takes for a given sector's value-added to reach a specific user of its input along the supply chain. With time-to-build, the distance from an input supplier to a producer translates to the relative delays of producing goods between producers. We show that the ratio between a producer's revenue response to a hump-shaped demand shock downstream and the response to a monotone-decay shock is a formal notion of supply chain distance in our model. This ratio captures the weighted average of delays over all possible direct and indirect routes between the input supplier and input user, where the weights are given by the discounted share of the supplier's output sent along each route. The remaining theoretical challenge is to characterize the weights, which are themselves endogenous. Our closed-form solution of the impulse responses of producers' actions to the hump-shaped shocks allows us to determine these endogenous weights and thus determine the supply chain distance measure in closed form over the primitives of the economy. Our definition of supply chain distance from one sector to another extends the influential upstreamness measures by Antràs et al. (2012) and Alfaro et al. (2019).

In general, a sector's demand innovation may be a mixture of the two shocks. We therefore characterize the model under a one-innovation demand process: a single innovation per sector drives both components, with the mixture weights as parameters of the demand process, and the impulse response to a demand innovation is the corresponding mixture of the two responses of Theorem 2. We show that the same process also describes observed demand when agents cannot distinguish the two shocks and forecast from realized demand alone: their forecasts follow a Kalman filter, and observed demand follows exactly this process, with the Kalman gain as the weights.

Empirically, we use our bilateral supply chain distance measure to document the bullwhip effect across U.S. sectors. As motivating evidence, we show in the cross section that volatility increases upstream within supply chains for goods important to final consumption. Figure 1 illustrates this for automobiles: each circle represents a sector supplying the auto industry (sized by input intensity), with supply chain distance on the Y-axis and value-added growth volatility on the X-axis. The pattern is consistent with the bullwhip effect: more upstream sectors exhibit greater volatility. In Section 4.2, we show that this pattern generalizes across all final goods supply chains.

Figure 1. Upstreamness and volatility for the supply chain of automobiles



Our theory predicts more than higher upstream volatility in the cross section: it pins down the dynamics of impulse responses. Following a demand innovation in a downstream sector, value-added responses along the supply chain are hump-shaped and initially amplified in the supply chains of sectors whose demand carries momentum—higher demand today signaling higher demand ahead—and dampened in those with monotone-decay demand.

We empirically validate the predictions by examining impulse-response functions for upstream value-added following downstream shocks. In order to do this, we first isolate downstream innovations in sectoral value-added—which respond to demand shocks—and classify downstream sectors according to the relative importance of AR(2) versus AR(1) shocks using the partial autocorrelation function (PACF). We then show how sectoral responses propagate through supply chains using the estimated innovations.

The observed heterogeneity in impulse-response functions provides a sharp test of our theory, aligning with four key theoretical predictions when demand innovations are mixtures of the two shocks. First, downstream sectors dominated by AR(2) shocks exhibit a distinctly hump-shaped response in their own value-added following demand shocks. Second, the value-added response along the supply chain is also hump-shaped and initially amplified as shocks propagate upstream, generating pronounced volatility. Third, this amplification eventually diminishes in sectors located sufficiently far upstream. Fourth and most importantly, the downstream hump-shaped responses and upstream amplification effects are markedly stronger in sectors where AR(2) shocks dominate, while such patterns are weak or absent in sectors characterized primarily by AR(1) shocks.

In the final part of the paper we conduct a model-based exercise. We estimate each downstream sector's mixture weights—identified from the autocovariance structure of realized demand alone—and perform variance decompositions for the supply chain associated with each downstream good. We find that on average, hump-shaped AR(2) shocks account for 75.44 percent of supply chain volatility in US manufacturing. The fraction ranges from 25.85 percent on the lower end (for the supply chain associated with "Sugar and Confectionery Products") to 99.98 percent on the higher end (for the supply chain associated with "Navigational, Measuring, Electromedical, and Control Instruments"). In a counterfactual world where all demand shocks are hump-shaped—while holding overall demand volatility constant—supply chain volatility increases substantially: by 155.5 percent overall, and by up to 241.6 percent for sectors supplying transportation goods. Conversely, when all demand shocks decay monotonically, supply chain volatility declines by 61.9 percent overall. Furthermore, the standard static production network model, which neglects time-to-build dynamics, understates the supply chain volatility attributable to downstream demand shocks by 22.2 percent overall, and by close to 44 percent for supply chains of non-food household goods (such as apparel).

Our paper relates to the growing literature on production networks (Carvalho, 2010; Gabaix, 2011; Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi, 2012; Jones, 2011, 2013; Oberfield, 2018; Liu, 2019; Baqaee and Farhi, 2019, 2020; Bigio and La'O, 2020; Elliott et al., 2022; Taschereau-Dumouchel, 2020; Pellet and Tahbaz-Salehi, 2023; Kopytov et al., 2024; Liu and Ma, 2024; Liu and Tsyvinski, 2024; Nikolakoudis, 2024). In contrast to the existing literature, which includes static models (e.g., Acemoglu et al., 2012), steady-states of dynamic models with one-period time-to-build delays (e.g., Long and Plosser, 1983), and transition dynamics with time-invariant consumer demand (e.g., vom Lehn and Winberry, 2022; Liu and Tsyvinski, 2024), we provide a full characterization of shock propagation in a dynamic production network with time-varying consumer demand and arbitrary and heterogeneous time-to-build.

Two recent working papers, Bizzarri et al. (2025) and Schaal and Taschereau-Dumouchel (2025), offer complementary studies of how heterogeneous time-to-build affects the propagation and persistence of shocks in input-output networks. While their analysis focuses on productivity shocks, similar in spirit to Liu and Tsyvinski (2024), our study instead centers on demand shocks. Both papers highlight asynchronous and oscillatory dynamics, offering insights into the cyclical mechanisms and time-series properties of sectoral adjustment.[1]

The time-to-build delays in our environment represent periods during which inputs have been produced but not yet sold to consumers, or completed but not immediately contributing value to production. Empirically, these goods are captured as inventories (Antràs and Tubdenov, 2025). Unlike models with an active inventory management margin to hedge against shocks (e.g., Blinder and Maccini, 1991; Kahn, 1992; Khan and Thomas, 2007; Wen, 2011), inventories in our model have specific shelf-lives. Our central theoretical results depend only on the first moments (expected values) of future demand and are thus invariant to higher-order moments, such as variance. Consequently, our theoretical characterizations remain valid even when analyzing unanticipated shocks where the variance of shocks approaches zero, thus eliminating any hedging role of inventories.

The study of supply chains as dynamic systems goes back to Forrester (1958), whose simulations of production–distribution systems showed orders and production fluctuating ever more strongly as one moves away from final demand—the phenomenon later termed the bullwhip effect. A large literature in operations research and management science has traced the effect to demand-signal processing, order batching, price fluctuations, and rationing at the level of individual firms (Lee et al., 1997, 2004), documented it experimentally in the beer distribution game (Sterman, 1989), quantified the amplification generated by demand forecasting under delivery lead times (Chen et al., 2000), and developed its measurement (Chen and Lee, 2012), alongside case-study evidence from firms at different positions along supply chains. More recently, Ferrari (2025) provides the first macroeconomic account of the bullwhip effect, embedding the procyclical-inventory mechanism from the operations tradition in a production network—orders track expected sales plus an inventory adjustment, so each stage passes a magnified fluctuation upstream—and documenting upstream amplification of final demand shocks using a shift-share

[1]Time-to-build here also plays a role similar to the adjustment costs that firms face in their investment decisions (Khan and Thomas, 2008, 2013): both frictions generate rich industry dynamics and complex responses of the economy to shocks.

design. In our model, the bullwhip arises without an inventory management margin, from time-to-build interacting with hump-shaped demand. It predicts amplification only when expected demand has momentum—whereas inventory and forecasting mechanisms (Lee et al., 1997; Chen et al., 2000) amplify after any persistent demand shock—and this shape dependence is the basis of our empirical tests.

Our paper also relates closely to the literature studying how expectations about future economic fundamentals can generate economic fluctuations independently of current realizations of shocks. Beaudry and Portier (2006) show that news about future productivity leads to significant macroeconomic responses well before productivity itself changes. Lorenzoni (2009) develops a model where noisy signals about future productivity induce agents to adjust current consumption and employment, creating fluctuations without contemporaneous fundamental shocks. Our multi-sector general equilibrium framework relates to this literature by introducing heterogeneous time-to-build delays, and shows that upstream producers respond predominantly to anticipated future demand rather than current conditions. These delays generate a bullwhip effect, wherein hump-shaped demand shocks amplify volatility upstream. This mechanism aligns closely with a key insight highlighted in Lorenzoni (2011): real frictions, such as adjustment costs or production delays, are essential for anticipated shocks to drive present-day economic fluctuations. Empirically, our analysis provides detailed sector-level evidence consistent with the aggregate findings of Blanchard et al. (2013), who document the quantitative importance of expectation-driven shocks at the aggregate level. Specifically, we show how anticipation-driven sectoral demand shocks propagate through production networks and amplify due to heterogeneous delays. Closer to our paper, Tian (2021) studies a production network in which inputs take one period to build, so that news about future demand moves current production through chains of input orders. In her setting, demand itself is unpredictable and anticipation comes entirely from outside signals; news makes upstream producers respond earlier than downstream producers, but never more strongly. In our framework, demand shocks themselves carry news—a hump-shaped shock today signals higher demand ahead—and delays accumulate along the supply chain, so upstream responses can be amplified rather than dampened, generating the bullwhip effect that we test directly in the data.

2 Model

In this section we describe the environment and characterize the equilibrium. The model builds on Long and Plosser (1983) and is a discrete-time dynamic production network with heterogeneous time-to-build, i.e., intermediate inputs are required to be supplied in advance. The main departures from Long and Plosser (1983) are the introduction and analysis of sectoral demand shocks and of the heterogeneous time-to-build across inputs.

2.1 Economic Environment

Preferences There is a representative consumer with preferences

$$V(\mathcal{S}_t) = \mathbb{E}\left[\sum_{s=t}^{\infty} \beta^{s-t}\left(\sum_{i=1}^{N} \theta_{is} \ln c_{is} - v\left(\bar{\ell}_s\right)\right) |\mathcal{S}_t\right] \tag{1}$$

where $\mathcal{S}_t$ is the set of state variables at the beginning of time t, c_{it} is consumption of good i, $\{\theta_{it}\}_{\geq 0}$ are the demand shocks, and $\bar{\ell}_t$ is total labor supplied.

Technology There are N sectors, each with constant-returns-to-scale production that uses labor and intermediate goods from other sectors. There is heterogeneous time-to-build across inputs: each input j must be supplied d_{ij} periods before production takes place in sector i, where $0 < d_{ij} < \infty$. The time-to-build is supplier-user specific; the mathematical formulation can be equivalently interpreted either as the supplier's order-to-delivery lag—the time input j takes to reach producer i—or as the duration required for input j to complete the production process for good i.

The production function of sector i is

$$y_{it} = z_{it}\ell_{it}^{\alpha_i}\prod_{j=1}^{N} m_{ij,t-d_{ij}}^{\omega_{ij}}, \qquad \alpha_i + \sum_j \omega_{ij} = 1, \tag{2}$$

where z_{it} is sectoral productivity, ℓ_{it} is the labor input, and $m_{ij,t-d_{ij}}$ is the intermediate good j used in production of good i. The elasticities of intermediate inputs $\{\omega_{ij}\}$ define the production network $\boldsymbol{\Omega} \equiv [\omega_{ij}]'$, where each column of $\boldsymbol{\Omega}$ is an input user and each row is an input supplier.[2]

Market Clearing The goods and labor market clearing conditions are

$$y_{jt} = c_{jt} + \sum_{i=1}^{N} m_{ijt}, \qquad \bar{\ell}_t = \sum_i \ell_{it}. \tag{3}$$

Sectoral Demand and Productivity Shocks We assume the economy is affected by sectoral demand shocks, as θ_{it} may fluctuate from period to period. For expositional simplicity, we assume that consumer demand is mean-reverting in the long run, i.e., $\lim_{s\to\infty}\mathbb{E}_t\left[\boldsymbol{\theta}_{t+s}\right] = \bar{\boldsymbol{\theta}}$ where $\bar{\boldsymbol{\theta}}$ represents steady-state levels of consumer demand. The extension to time-varying long-run demand is conceptually straightforward but complicates exposition.

[2] We assume that $\boldsymbol{I} - \boldsymbol{\Omega}$ is invertible, which is ensured when the production of any good eventually involves labor along the supply chain (i.e., for any j, $\sum_i \alpha_i \left[\boldsymbol{\Omega}^k\right]_{ij} > 0$ for some k).

The economy is also affected by sectoral productivity shocks z_{it}.

Demand shocks are the key objects of this study. As we show below, most of our results are invariant to the process of productivity shocks.

2.2 Equilibrium

In order to analyze the competitive equilibrium, it is more transparent to study the social planner's problem. Because there are no distortions or externalities in the model—time-to-build is a technological feature of the economic environment and is also part of the social planner's problem—the competitive equilibrium is efficient: allocations are identical to the planner's solution, and equilibrium prices correspond to the Lagrange multipliers of the planner's problem up to a choice of numeraire (and this choice is inconsequential for our analysis). The concordance between the planner's Lagrange multipliers and equilibrium prices enables us to interpret the planner's solution as equilibrium objects involving prices, such as labor income, value-added, sales, and the Domar weight (i.e., sectoral size normalized by aggregate labor income). For expositional clarity, we analyze the planner's solution in the main text; we describe below how the allocation decentralizes, and we formally demonstrate the mapping to a decentralized equilibrium in Appendix B.1.

The planner solves

$$
\begin{aligned}
V_t\left(\{m_{ij,t-q}\}_{q=1,\ldots,d_{ij}}\right) \quad = \quad & \max_{\{\ell_{it},c_{it},m_{ijt},\bar{\ell}_t\}} \sum_i \theta_{it}\ln c_{it} - v\left(\bar{\ell}_t\right) + \beta\mathbb{E}_t\left[V_{t+1}\left(\{m_{ij,t+1-q}\}_{q=1,\ldots,d_{ij}}\right)\right] \\
& + w_t\left[\bar{\ell}_t - \sum_j \ell_{jt}\right] + \sum_j p_{jt}\left[z_{jt}\ell_{jt}^{\alpha_j}\prod_k m_{jk,t-d_{jk}}^{\omega_{jk}} - c_{jt} - \sum_i m_{ijt}\right],
\end{aligned}
$$

where w_t is the Lagrange multiplier for the labor market clearing constraint, and p_{jt} is the Lagrange multiplier for the market clearing constraint of output from sector j.

We now discuss the difficulty of analyzing this problem. The key complication lies in the fact that this is an optimal control problem with endogenous state variables $\{m_{ij,t-q}\}_{q=1,\ldots,d_{ij}}$ that are matrices of all of the past choices of the suppliers' inputs at all horizons of the past time delays. That is, the optimal control problem is path-dependent through the dependence on the choices of the previous inputs at all previous delay horizons that cannot be summarized in a low-dimensional state variable.[3] The key difficulty is that the delays at all horizons also affect all of the future value functions and represent the complicated endogenous responses of input suppliers to all future stochastic demand shocks at all time horizons. In Theorem 1, we show that the solution to this problem can be represented in a strikingly simple form.

[3]The analysis of these problems, especially in stochastic environments, is in its nascency. See Bayraktar and Keller (2018) for a recent mathematical treatment of path-dependent Hamilton-Jacobi equations and Boerma et al. (2024) for the use of stochastic calculus of variations to provide a non-recursive approach.

The first-order conditions with respect to c_{jt}, ℓ_{jt}, and $\bar{\ell}_t$ are:

$$\{c_{jt}\} \qquad \theta_{jt} = p_{jt}c_{jt} \tag{4}$$

$$\{\ell_{jt}\} \qquad w_t\ell_{jt} = \alpha_j p_{jt} y_{jt} \tag{5}$$

$$\{\bar{\ell}_t\} \qquad v'\left(\bar{\ell}_t\right) = w_t. \tag{6}$$

The first-order condition with respect to m_{jkt} and the envelope condition jointly imply

$$\{m_{jkt}\} \qquad p_{kt}m_{jkt} = \omega_{jk}\beta^{d_{jk}}\mathbb{E}_t\left[p_{j,t+d_{jk}}y_{j,t+d_{jk}}\right] \tag{7}$$

As we show in Appendix B.1, the Lagrange multipliers correspond to prices at the time when input decisions are made in a decentralized competitive equilibrium when the numeraire is chosen such that the marginal utility of income in each period is normalized to one. We henceforth refer to p_{jt} as the price of good j and w_t as the wage rate. Because prices are dated at the time of sale rather than the time of use, all units of good j sold at t command the same price p_{jt}, whether they go to the household for consumption or to a producer as an input. Equation (4) shows that total expenditure on consumption of good j equals the concurrent demand shock θ_{jt}. Equation (5) shows that each producer's labor expenditure is α_j fraction of the revenue, where α_j is the value-added intensity in the sector. Equation (7) states that producer j's expenditure on input k at time t, $p_{kt}m_{jk,t}$, equals a fraction ω_{jk} of the producer's expected revenue at future time $t+d_{jk}$, discounted at rate β per period. The expectation in (7) encapsulates the measurability constraint in the planner's solution.

The delay admits two readings that deliver the same allocation. Under the first, d_{ij} is a delivery delay, and the purchase is a contract that fixes the quantity and a delivery date $t + d_{ij}$, with the input produced and sold at date t and delivered d_{ij} periods later. Under the second, d_{ij} is an installation lag: the input is bought on the spot at t and becomes productive only at $t + d_{ij}$, so ordinary spot markets suffice.[4] In either case the quantity is chosen d_{ij} periods before the input is used and cannot be revised, so the past orders that enter the planner's problem as state variables are, in the decentralized economy, commitments producers have already made. When a demand shock arrives, these commitments cannot adjust; what adjusts is current labor, the current prices and wage, and new orders, which govern production at future dates. Producers are competitive, and equation (7) is their optimality condition; because demand is realized after orders are placed, they earn ex post profits or losses, which accrue to the household and have zero expected present value under its stochastic discount factor.

Let $\gamma_{jt} \equiv p_{jt}y_{jt}$ denote the sector j's revenue at time t. This is an important object for our

[4]The two readings correspond to the two components of the delay that we measure in Section 4: the backlog ratio captures the supplier's order-to-delivery lag, and the inventory-to-COGS ratio of Antràs and Tubdenov (2025) captures the buyer's delivery-to-use lag.

analysis, as we can express the equilibrium allocation in shares (i.e., the fraction of workers in each sector $\ell_{it}/\bar{\ell}_t$; the share of each sector j's output used for consumption, c_{jt}/y_{jt}; and the share used as production inputs, m_{ijt}/y_{jt}) all as functions of the demand shocks and the current and expected future revenues:

$$\frac{\ell_{jt}}{\bar{\ell}_t} = \frac{\alpha_j \gamma_{jt}}{\sum_i \alpha_i \gamma_{it}}, \qquad \frac{c_{jt}}{y_{jt}} = \frac{\theta_{jt}}{\gamma_{jt}}, \qquad \frac{m_{ijt}}{y_{jt}} = \omega_{ij} \beta^{d_{ij}} \frac{\mathbb{E}_t \left[\gamma_{i,t+d_{ij}}\right]}{\gamma_{jt}}. \tag{8}$$

Aggregate labor income is $\sum_i \alpha_i \gamma_{it}$.[5] The Domar weight, defined as the sectoral revenue relative to the aggregate labor income, is

$$\text{(Domar weight)} \qquad \zeta_{it} \equiv \frac{p_{it} y_{it}}{\sum_i w_t \ell_{it}} = \frac{\gamma_{it}}{\sum_i \alpha_i \gamma_{it}}.$$

We now solve for sectoral sales $\{\gamma_{it}\}$ as a function of expected future demand. First multiply the goods market clearing condition (3) by p_{jt} on both sides and then substitute out c_{jt} and m_{ijt} using (4) and (7), we obtain

$$\begin{aligned} \gamma_{jt} &= \theta_{jt} + \sum_{i=1}^{N} \beta^{d_{ij}} \omega_{ij} \mathbb{E}_t \left[\gamma_{i,t+d_{ij}}\right] \\ &= \theta_{jt} + \sum_{d=1}^{\infty} \sum_{i=1}^{N} \beta^{d} \omega_{ij} \mathbf{1}_{d_{ij}=d} \mathbb{E}_t \left[\gamma_{i,t+d}\right]. \end{aligned}$$

That is, the revenue of sector j at time t consists of the concurrent time-t consumer demand for good j and the demand from other producers. Because of the heterogeneous time-to-build input delays by d_{ij} periods, the expenditure on input j at time t by producer i is equal to the input cost share ω_{ij} times the expected revenue at time $t + d_{ij}$, discounted by $\beta^{d_{ij}}$.

We then re-write the sectoral revenue equation in vector form. Throughout the paper, bold-face letters represent matrices (uppercase) and column vectors (lowercase). Let the $N \times N$ matrix $\mathbf{\Omega}_d \equiv \left[\omega_{ij} \mathbf{1}_{d_{ij}=d}\right]'$ denote the input cost shares for inputs with delay d, and let $\boldsymbol{\theta}_t \equiv [\theta_{it}]$ be the column vector of demand shocks. The collection of matrices $\{\mathbf{\Omega}_d\}_{d=1}^{\infty}$ defines the time-to-build structure over the production network. Note that each column of $\mathbf{\Omega}_d$ represents an input user and each row represents an input supplier.

Using the matrix notation, the vector of revenue can be written as

$$\boldsymbol{\gamma}_t = \boldsymbol{\theta}_t + \sum_{d=1}^{\infty} \beta^d \mathbf{\Omega}_d \mathbb{E}_t \left[\boldsymbol{\gamma}_{t+d}\right]. \tag{9}$$

The theorem below solves for the sectoral revenue at time t as a function of the expected future

[5] With advance purchases at discounted prices, value added exceeds labor compensation by the imputed interest on outstanding input orders—a wedge of order $1 - \beta$ per period of delay that vanishes as $\beta \to 1$; up to this wedge, ζ_{it} coincides with the conventional sales-to-GDP Domar weight.

demand across sectors.

Theorem 1. *Let Φ_s denote the set of sequences of positive integers $\phi \equiv (\phi_1, \ldots, \phi_n)$ that sum to s (i.e., $\sum_{k=1}^{n} \phi_k = s$ with $\phi_k > 0$ for all k). Sectoral revenue follows*

$$\boldsymbol{\gamma}_t = \boldsymbol{\theta}_t + \sum_{s=1}^{\infty} \boldsymbol{G}_s \mathbb{E}_t \left[\boldsymbol{\theta}_{t+s}\right], \qquad \text{where } \boldsymbol{G}_s \equiv \beta^s \sum_{\phi \in \Phi_s} \prod_{\phi_j \in \phi} \boldsymbol{\Omega}_{\phi_j}, \tag{10}$$

where each product is taken in the order of the sequence, $\boldsymbol{\Omega}_{\phi_1} \boldsymbol{\Omega}_{\phi_2} \cdots \boldsymbol{\Omega}_{\phi_n}$.

Proof. See appendix A.1. □

The infinite series (10) always converges given that demand is mean-reverting in the long run.

The main difficulty that the theorem overcomes is to express the complicated interactions of the suppliers' endogenous responses at all future time horizons to demand shocks at all future time horizons. The theorem shows that these interactions can be represented in a strikingly simpler form as a sum of products of only two components over the primitives of the economy, $\mathbb{E}_t \left[\boldsymbol{\theta}_{t+s}\right]$ and $\boldsymbol{G}_s$. We now discuss the object $\boldsymbol{G}_s$ in detail.

The ij-th entry of the matrix G_s captures how expected consumer demand for good j at time $t+s$ affects the time-t revenue of sector i through *all possible paths of direct and indirect linkages* that have a cumulative delay of s periods. To interpret, consider writing out $\boldsymbol{G}_s$ explicitly for $s = 1, 2, 3$. There is one positive sequence that sums to $s = 1$: $\Phi_1 = \{(1)\}$. There are two positive sequences that sum to $s = 2$: $\Phi_2 = \{(1,1), (2)\}$. There are four positive sequences that sum to $s = 3$: $\Phi_3 = \{(1,1,1), (1,2), (2,1), (3)\}$. Hence, the formula in (10) implies that $\boldsymbol{G}_1 = \beta \boldsymbol{\Omega}_1$, $\boldsymbol{G}_2 = \beta^2 \left(\boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_1^2\right)$, and $\boldsymbol{G}_3 = \beta^3 \left(\boldsymbol{\Omega}_3 + \boldsymbol{\Omega}_2 \boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_1 \boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_1^3\right)$. The power of the matrices captures the number of walks (i.e., the number of edges) associated with a network path: $\boldsymbol{\Omega}_1$, $\boldsymbol{\Omega}_2$, and $\boldsymbol{\Omega}_3$ capture one-walks (i.e., direct connection between producer pairs) with one, two, and three-period delays, respectively. $\boldsymbol{\Omega}_1^2$ captures two-walks (i.e., paths with two edges that connect producer pairs) where each walk has one-period delay, $\boldsymbol{\Omega}_2 \boldsymbol{\Omega}_1$ is a two-walk where the first walk has two-period delay and the second walk has one-period delay, and so on.

The matrix $\boldsymbol{G}_s$ summarizes how the vector of expected demand at $t+s$ affects the vector of revenue at time t, through *all possible paths of direct and indirect linkages* that have cumulative delays of s periods. While the concurrent consumer demand $\boldsymbol{\theta}_t$ contributes to a sector's revenue directly, the one-period ahead demand $\boldsymbol{\theta}_{t+1}$ contributes to a sector's revenue only indirectly through other producers to which the focal sector supplies with one period delay (captured by the matrix $\boldsymbol{\Omega}_1$ in $\boldsymbol{G}_1$). The contribution of demand two-period ahead ($\boldsymbol{\theta}_{t+2}$) to a focal sector i's revenue (captured by $\boldsymbol{G}_2$) is through both (a) producers to which the focal sector supplies with a delay of two periods (captured by the matrix $\boldsymbol{\Omega}_2$), and (b) any producer k that uses an input j

with one period delay, and where sector j uses input i also with one period delay (captured by the matrix $\boldsymbol{\Omega}_1^2$). The contribution of $\boldsymbol{\theta}_{t+3}$ is captured by enumerating all possible paths that have three periods of cumulative delays.

Theorem 1 states that the time-t revenue of a sector is the summation across contributions from expected demand of different consumer goods at different time in the future, linked by *all possible paths of direct and indirect linkages* in the production network with different delays.

In contrast to the existing literature on production networks, which includes static models (e.g., Acemoglu et al., 2012), steady-states of dynamic models with one-period time-to-build delays (e.g., Long and Plosser, 1983), and transition dynamics with time-invariant consumer demand (e.g., Liu and Tsyvinski, 2024), Theorem 1 provides a full characterization that accounts for time-varying consumer demand in a dynamic production network model with arbitrary and heterogeneous time-to-build. We now specialize Theorem 1 to time-invariant demand and discuss how this special case generalizes known results in the literature to heterogeneous time-to-build.

Implication: Time-Invariant Demand with Heterogeneous Time-to-Build Specializing Theorem 1 to time-invariant demand $\boldsymbol{\theta}_t = \bar{\boldsymbol{\theta}}$, we can simplify the summation over $\boldsymbol{G}_s$ and obtain the following result.

Corollary 1. *When demand is time-invariant, sectoral revenues are*

$$\bar{\boldsymbol{\gamma}} = \left(\boldsymbol{I} - \sum_{d=1}^{\infty} \beta^d \boldsymbol{\Omega}_d\right)^{-1} \bar{\boldsymbol{\theta}}. \tag{11}$$

Proof. See appendix A.2. □

We can contrast this result with existing literature. When all input delays are exactly one period ($\boldsymbol{\Omega} = \boldsymbol{\Omega}_1$), Corollary 1 recovers as a special case the steady-state result of Long and Plosser (1983): $\bar{\boldsymbol{\gamma}} = (\boldsymbol{I} - \beta\boldsymbol{\Omega})^{-1} \bar{\boldsymbol{\theta}}$. Similarly, the static model of Acemoglu et al. (2012), which yields sectoral revenue $\bar{\boldsymbol{\gamma}} = (\boldsymbol{I} - \boldsymbol{\Omega})^{-1} \bar{\boldsymbol{\theta}}$, emerges as a special case of our model when we allow inputs to have zero delays.[6]

Corollary 1 extends these results on time-invariant sectoral size to a production network with an arbitrary time-to-build structure. The corollary is a special case of Theorem 1, which

[6]For expositional clarity, our baseline model features a time-to-build of at least one period. We could extend the baseline model to incorporate static inputs without time-to-build (i.e., $d_{ij} = 0$), thereby nesting the static production network model as a special case. In that case, the formula in Theorem 1 needs to be modified as

$$\boldsymbol{\gamma}_t = (\boldsymbol{I} - \boldsymbol{\Omega}_0)^{-1} \left\{\boldsymbol{\theta}_t + \sum_{s=1}^{\infty} \boldsymbol{G}_s \mathbb{E}_t [\boldsymbol{\theta}_{t+s}]\right\}, \qquad \text{where } \boldsymbol{G}_s \equiv \beta^s \sum_{\phi\in\Phi_s} \prod_{\phi_j\in\phi} \left[\boldsymbol{\Omega}_{\phi_j} (\boldsymbol{I} - \boldsymbol{\Omega}_0)^{-1}\right],$$

where $\boldsymbol{\Omega}_0 \equiv \left[\omega_{ij} \mathbf{1}_{d_{ij}=0}\right]'$ captures the component of the input-output network with zero delay (proof is available upon request). Corollary 1 extends naturally to $\bar{\boldsymbol{\gamma}} = \left(\boldsymbol{I} - \sum_{d=0}^{\infty} \beta^d \boldsymbol{\Omega}_d\right)^{-1} \bar{\boldsymbol{\theta}}$.

fully characterizes the sectoral revenue away from steady-state, given any sequences of expected future consumer demand across different goods.

Implication: When Sectoral Demand Follows $ARMA\,(p,q)$ Theorem 1 shows that the equilibrium revenue in sector i at time t reflects the expected future demand $\mathbb{E}_t\left[\theta_{k,t+s}\right]$ at each time horizon in every other sector k, through each possible walk of direct and indirect input-output linkages that delivers output of i to k, accounting for the time-to-build associated with each walk.

We now provide a characterization of equilibrium revenue when sectoral demand follows a stationary and invertible $ARMA\,(p,q)$. The process determines the expected future demand $\mathbb{E}_t\left[\theta_{k,t+s}\right]$ under rational expectations. The analysis reveals an interesting parallel between (1) how a current demand shock affects the expected future demand through the autoregression of the demand shock process, and (2) how the time-to-build and the network structure determines the importance of demand at different horizons.

Specifically, we consider an $AR\,(\infty)$ representation of the ARMA shock process:[7]

$$\theta_{it} - \bar{\theta}_i = \sum_{s=1}^{\infty} \pi_s \left(\theta_{it-s} - \bar{\theta}_i\right) + e_{it}. \tag{12}$$

where $\boldsymbol{e}_t$ has mean zero and is independent of past shocks, and $\{\pi_s\}$ are autoregressive coefficients.

In this setting, the response of expected future demand to current demand innovation is

$$\frac{\partial \mathbb{E}_t\left[\theta_{i,t+s}\right]}{\partial e_{it}} = \sum_{\phi\in\Phi_s} \prod_{\phi_j\in\phi} \pi_{\phi_j}, \tag{13}$$

where, as in Theorem 1, Φ_s is the set of sequences of positive integers $\phi \equiv (\phi_1, \ldots, \phi_n)$ that sum to s. That is, e_{it}'s impact on $\mathbb{E}_t\left[\theta_{i,t+1}\right]$ is π_1, its impact on $\mathbb{E}_t\left[\theta_{i,t+2}\right]$ is $\pi_1^2 + \pi_2$, and on $\mathbb{E}_t\left[\theta_{i,t+3}\right]$ is $(\pi_1^3 + \pi_2\pi_1 + \pi_1\pi_2 + \pi_3)$, and so on. In other words, the impact of a demand shock e_{it} on the expected future demand at time $t+s$ has a recursive structure: it is the sum of its impact on demand $\mathbb{E}_t\left[\theta_{i,t+k}\right]$ for all $0 \le k \le s-1$, and then the impact of $\mathbb{E}_t\left[\theta_{i,t+k}\right]$ on $\mathbb{E}_t\left[\theta_{i,t+s}\right]$ through the autoregressive demand process.

Note in particular the parallel between equation (13) and the object $\boldsymbol{G}_s$ in Theorem 1. Both constructs enumerate all time gaps that sum to s. In $\boldsymbol{G}_s$, the gap is due to the time-to-build delays; in equation (13), the gap is due to the autoregressive structure of demand.

When demand exhibits cross-sector dependence and heterogeneous dynamics, the natural

[7] The representation exists for all stationary and invertible $ARMA\,(p,q)$ processes (see, e.g., Box et al., 2015).

generalization of (12) is the vector process

$$\boldsymbol{\theta}_t - \bar{\boldsymbol{\theta}} = \sum_{s=1}^{\infty} \boldsymbol{\Delta}_s \left(\boldsymbol{\theta}_{t-s} - \bar{\boldsymbol{\theta}}\right) + \boldsymbol{e}_t, \tag{14}$$

where $\boldsymbol{\Delta}_s$ are $N \times N$ coefficient matrices and the innovations $\boldsymbol{e}_t$ are mean zero, mean independent over time, and possibly correlated across sectors. The response of expected future demand to a current innovation is the matrix analogue of (13): $\partial \mathbb{E}_t \left[\boldsymbol{\theta}_{t+s}\right] / \partial \boldsymbol{e}_t = \boldsymbol{H}_s$, where $\boldsymbol{H}_0 = \boldsymbol{I}$ and $\boldsymbol{H}_s = \sum_{k=1}^{s} \boldsymbol{\Delta}_k \boldsymbol{H}_{s-k}$; equivalently, $\boldsymbol{H}_s = \sum_{\phi \in \Phi_s} \prod_{\phi_j \in \phi} \boldsymbol{\Delta}_{\phi_j}$ enumerates the lag sequences that sum to s, with products of matrices replacing products of scalar coefficients. The scalar process (12) is the special case $\boldsymbol{\Delta}_s = \pi_s \boldsymbol{I}$, and diagonal $\boldsymbol{\Delta}_s$ allows sector-specific dynamics.

Proposition 1. *When demand follows the stationary process (14), the revenue response to demand innovations is*

$$\frac{\partial \gamma_t}{\partial \boldsymbol{e}_t} = \sum_{s=0}^{\infty} \boldsymbol{G}_s \boldsymbol{H}_s = \boldsymbol{I} + \sum_{s=1}^{\infty} \beta^s \left(\sum_{\phi \in \Phi_s} \prod_{\phi_j \in \phi} \boldsymbol{\Omega}_{\phi_j} \right) \left(\sum_{\phi \in \Phi_s} \prod_{\phi_j \in \phi} \boldsymbol{\Delta}_{\phi_j} \right),$$

where Φ_s is the set of sequences of positive integers $\phi \equiv (\phi_1, \ldots, \phi_n)$ that sum to s, and each product is taken in the order of the sequence, as in Theorem 1.

Proof. See Appendix A.3. □

Quantities and Aggregate Implications Theorem 1 analytically solves for sectoral revenue as a function of model primitives: the network structure, time-to-build, and the expected future demand. The focus of our paper is on sectoral dynamics over supply chains, and subsequent sections of this paper build on Theorem 1. For completeness, Proposition 2 characterizes equilibrium quantities and the aggregate implications. This characterization is the discrete-time analogue of Liu and Tsyvinski (2024).

Proposition 2. *Given the endogenous state variables $\{m_{ij,t-q}\}_{q=1,\ldots,d_{ij}}$, exogenous productivities $\{z_{it}\}$, and the path of expected demand $\{\mathbb{E}_t \left[\boldsymbol{\theta}_{t+s}\right]\}_{s=0}^{\infty}$, sectoral output $\{y_{it}\}$, and consumption $\{c_{it}\}$ at each time is*

$$y_{it} = z_{it} \left(\frac{\alpha_i \gamma_{it}}{\sum_k \alpha_k \gamma_{kt}} \bar{\ell}_t \right)^{\alpha_i} \prod_j m_{ij,t-d_{ij}}^{\omega_{ij}}, \qquad c_{it} = \frac{\theta_{it}}{\gamma_{it}} y_{it}, \tag{15}$$

where γ_{it} is characterized by Theorem 1, and the total labor supplied $\bar{\ell}_t$ at time t satisfies:

$$v'\left(\bar{\ell}_t\right) \bar{\ell}_t = \sum_i \alpha_i \gamma_{it}. \tag{16}$$

The endogenous state variables evolve according to

$$m_{ijt} = \frac{\omega_{ij} \beta^{d_{ij}} \mathbb{E}_t \left[\gamma_{i,t+d_{ij}}\right]}{\gamma_{jt}} y_{jt}. \tag{17}$$

The dynamics over the aggregate consumption index $c_t \equiv \prod_i c_{it}^{\theta_{it}}$ follow directly from (15).

Proof. See Appendix A.4. $\square$

Inventories Our model features time-to-build delays in intermediate inputs, which implies periods during which inputs have been produced but not yet sold to consumers, or completed but not immediately contributing value to production. Empirically, such goods are recorded as inventories (Antràs and Tubdenov, 2025). Unlike models with an active inventory management margin to hedge against shocks (e.g., Blinder and Maccini, 1991; Kahn, 1992; Khan and Thomas, 2007; Wen, 2011), inventories in our model have specific shelf-lives determined by the delay parameter d_{ij}. Nevertheless, our central theoretical results—Theorem 1 above and Theorem 2 in the next section—depend only on the first moments of future demand. Hence, our theoretical characterizations remain valid even in scenarios where the variance of shocks approaches zero, effectively eliminating any hedging role of inventories.

Finally, the fixed-delivery-date structure can be relaxed. Suppose each input arrives spread across delivery lags in constant shares rather than at a single lag.[8] Such an economy maps exactly into the present one: relabel each input–lag pair as a distinct, fictitious producer that supplies the same good at a single fixed delay, so that production is Cobb-Douglas over these pairs. All results apply once the delay-specific matrices $\mathbf{\Omega}_d$ are expanded so that each pair appears at every lag it uses, weighted by the constant shares.

3 The Supply Chain Response to Demand Shocks

In this section, we analyze the equilibrium response to demand shocks along the supply chain. We begin by applying our characterization to a stochastic process for sectoral demand that has two components: a monotone-decay AR(1) component and a hump-shaped AR(2) component. This stochastic process, empirically motivated as shown later, defines the conditional expectation over future demand based on current information $\mathbb{E}_t[\boldsymbol{\theta}_{t+s}]$. Specifically, we use Theorem 1 to analyze how different sectoral demand shocks and their implied paths of expected future demand interact with the heterogeneous time-to-build and shape the equilibrium response along the supply chain. Section 3.1 describes the demand shock process and derives the impulse responses. Section 3.2

[8] Constant shares obtain when the buyer allocates its purchase across delivery dates through a CES aggregator over vintages, microfounded by storage technologies with heterogeneous efficiencies, at unit elasticity of substitution across vintages: the vintage aggregator is then Cobb-Douglas, and the share committed to each delivery lag is pinned down by the storage technology and independent of the state of the economy. That elasticity indexes how far an environment departs from ours. When vintages are gross substitutes, firms tilt deliveries toward dates with high expected marginal revenue and inventory allocation begins to reshape the upstream response; in the limit of perfect substitutability, units bought at different dates are interchangeable in use and the storage efficiencies act as depreciation, recovering the standard inventory technology. We leave these cases to future work.

discusses equilibrium implications. Section 3.3 develops the one-innovation representation of demand, in which a single innovation per sector, rather than two shocks, drives demand as a mixture of the two shocks; the mixture weights can be microfounded as Kalman gains when agents filter the components from realized demand (Appendix A.8).

3.1 Two Components of Demand Shocks

We assume sectoral demand is mean-reverting and has two distinct components. The first component captures transient consumer preference changes, modeled as monotone-decay AR(1) shocks, where a demand innovation today is expected to monotonically diminish over time. The second component captures slower-developing changes, such as those arising from habit formation, consumer fads, or delayed adoption processes. These are modeled as hump-shaped AR(2) shocks, where a current increase in demand initially leads to further demand increases in future horizons before eventually declining. Specifically, sectoral demand θ_{it} follows:

$$\theta_{it} - \bar{\theta}_i = \rho \left(\theta_{it-1} - \bar{\theta}_i \right) + x_{it} + \epsilon_{it}, \tag{18}$$

$$x_{it} = \rho x_{i,t-1} + u_{it}. \tag{19}$$

The parameter $\rho \in (0, 1)$ captures the rate at which sectoral demand mean-reverts to the steady-state or long-run level $\bar{\theta}_i$. There are two sources of demand shocks: ϵ_{it} and u_{it}, both are mean zero, mean independent over time, and the two shocks are uncorrelated with each other, $\mathbb{E}\left[\boldsymbol{\epsilon}_t \boldsymbol{u}_t'\right] = \mathbf{0}$. These shocks may be correlated across sectors, and their variances may be time-varying (so that, together with appropriately bounded innovations, sectoral demand stays non-negative). In our baseline specification, the two types of demand shocks share the same persistence parameter ρ. Appendix B.2 extends our results in this section to an environment with heterogeneous ρ's for the two types of shocks.

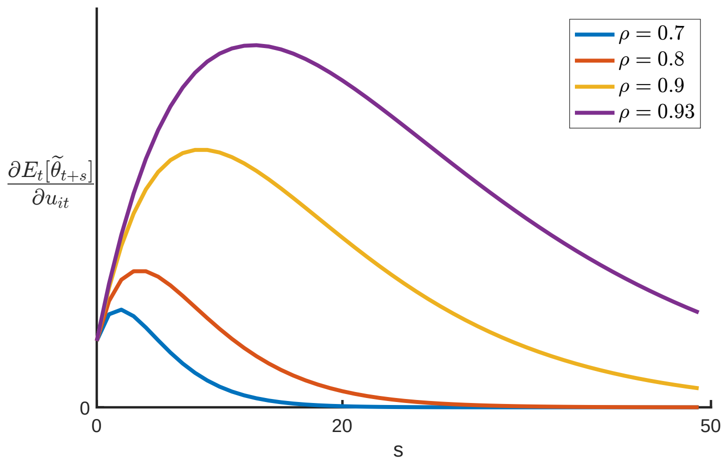
The monotone-decay component of demand shocks is captured by ϵ_{it}, and the hump-shaped component is captured by u_{it}. In the absence of u_{it}, the monotone-decay shock ϵ_{it} is AR(1). In the absence of ϵ_{it}, the hump-shaped shock u_{it} is AR(2).

Consider the response of expected demand $\mathbb{E}_t \left[\theta_{i,t+s}\right]$ to an impulse (ϵ_{it} or u_{it}). We have:

$$\text{For all } s \geq 0: \qquad \underbrace{\frac{\partial \mathbb{E}_t \left[\theta_{i,t+s}\right]}{\partial \epsilon_{it}}}_{\substack{\text{response to a} \\ \text{monotone-decay shock}}} = \rho^s, \qquad \underbrace{\frac{\partial \mathbb{E}_t \left[\theta_{i,t+s}\right]}{\partial u_{it}}}_{\substack{\text{response to a} \\ \text{hump-shaped shock}}} = (s+1)\,\rho^s. \tag{20}$$

The expected demand response to an ϵ_{it} shock declines monotonically and exponentially over time at rate ρ. In contrast, for any $\rho > 0.5$, the expected demand response to a u_{it} shock is hump-shaped: $(s+1)\,\rho^s u_{it}$ increases initially over time s, peaking at $\bar{s} \equiv \lceil \frac{2\rho - 1}{1-\rho} \rceil$, and subsequently declines for $s \geq \bar{s}$. Figure 2 illustrates the impulse-response of demand following a hump-shaped shock at different persistence levels ρ. As shown, the impulse-response is hump-shaped, and greater persistence (larger ρ) results in longer-lasting effects with peaks occurring later in time.

Figure 2. The impulse response $\partial \mathbb{E}_t\left[\theta_{i,t+s}\right] \Big/ \partial u_{it}$ of future demand following a hump-shaped shock, for different levels of ρ



Impulse Responses when the Two Components Are Observed We begin by assuming agents can separately observe the hump-shaped and monotone-decay components of demand. The conditional expectation is $\mathbb{E}_t\left[\cdot\right] \equiv \mathbb{E}\left[\cdot \middle| \left\{\boldsymbol{\theta}_{t-s}, \boldsymbol{x}_{t-s}\right\}_{s\geq 0}\right]$. Given the stochastic process of demand shocks, the information set consisting of the history of state variables $\left\{\boldsymbol{\theta}_{t-s}, \boldsymbol{x}_{t-s}\right\}_{s\geq 0}$ is identical to the information set consisting of the history of shocks $\left\{\boldsymbol{\epsilon}_{t-s}, \boldsymbol{u}_{t-s}\right\}_{s\geq 0}$, and we use the two information sets interchangeably.

We let $\widetilde{\cdot}$ denote deviations from steady-state levels, e.g., $\widetilde{\boldsymbol{\theta}}_t \equiv \boldsymbol{\theta}_t - \bar{\boldsymbol{\theta}}$.[9] We can then derive the characterization of conditional expectations as

$$\mathbb{E}_t\left[\widetilde{\boldsymbol{\theta}}_{t+s}\right] = \rho^s\left(\widetilde{\boldsymbol{\theta}}_t - \boldsymbol{x}_t\right) + (s+1)\,\rho^s \boldsymbol{x}_t. \tag{21}$$

On the right-hand side, the first term captures the monotone-decay component of demand deviation from the steady-state levels; this component decays exponentially at rate ρ. The second term captures the hump-shaped component: for $\rho > \frac{1}{2}$, $(s+1)\,\rho^s$ first increases and then decreases in s, peaking at time $\bar{s} \equiv \lceil \frac{2\rho-1}{1-\rho} \rceil$ periods into the future.

Substituting the conditional expectation into Theorem 1, we obtain how the two types of demand shocks explicitly affect sectoral revenue. We first state and prove the result under general network and time-to-build structure and then provide additional interpretation through a series of examples.

[9] Since $\boldsymbol{x} = 0$ in steady state, we have $\boldsymbol{x}_t = \widetilde{\boldsymbol{x}}_t$.

Theorem 2. *When agents separately observe the hump-shaped and monotone-decay components of demand shocks up to the current period, with demand characterized by equations (18) and (19), sectoral revenue at time t follows*

$$\widetilde{\gamma}_t = \boldsymbol{G}^{\epsilon}_{\infty}\left(\widetilde{\boldsymbol{\theta}}_t - \boldsymbol{x}_t\right) + \boldsymbol{G}^{x}_{\infty}\boldsymbol{x}_t, \tag{22}$$

where $\boldsymbol{G}^{\epsilon}_{\infty}$ is the impact of demand's monotone-decay component on revenue, and $\boldsymbol{G}^{x}_{\infty}$ is the impact of demand's hump-shaped component, and

$$\boldsymbol{G}^{\epsilon}_{\infty} \equiv \left(\boldsymbol{I} - \sum_{d=1}^{\infty}(\rho\beta)^d\,\boldsymbol{\Omega}_d\right)^{-1}, \qquad \boldsymbol{G}^{x}_{\infty} \equiv \boldsymbol{G}^{\epsilon}_{\infty}\left(\sum_{d=1}^{\infty} d\,(\rho\beta)^d\,\boldsymbol{\Omega}_d\right)\boldsymbol{G}^{\epsilon}_{\infty} + \boldsymbol{G}^{\epsilon}_{\infty}.$$

The impulse-response functions are:

$$\frac{\partial\widetilde{\gamma}_{t+s}}{\partial\boldsymbol{\epsilon}_t} = \rho^s\boldsymbol{G}^{\epsilon}_{\infty}, \qquad \frac{\partial\widetilde{\gamma}_{t+s}}{\partial\boldsymbol{u}_t} = s\rho^s\boldsymbol{G}^{\epsilon}_{\infty} + \rho^s\boldsymbol{G}^{x}_{\infty}.$$

Proof. See Appendix A.5. □

3.2 Implications of Theorem 2

We now discuss several implications of Theorem 2.

3.2.1 Upstream's Amplified Response to Expectations of Heightened Future Demand

We define the bullwhip effect as the phenomenon where a small downstream demand shock leads to an amplified response in upstream sectors. Throughout, we measure amplification horizon by horizon, relative to the response of the downstream sector experiencing the shock. Theorem 2 implies that the bullwhip effect does not occur when demand shocks decay monotonically, but it may arise when demand shocks are hump-shaped. Intuitively, the difference stems from producers' positions along the supply chain: those closer to the demand shock immediately increase production in response to current or near-future expected demand. On the other hand, because of time-to-build delays, upstream producers respond primarily to anticipated future demand. Therefore, when demand shocks monotonically diminish over time, upstream producers' responses are muted relative to downstream producers. In contrast, hump-shaped shocks generate anticipated future demand increases, eliciting an amplified response in upstream sectors and thereby creating a pronounced bullwhip effect.

Formally, the revenue response to a concurrent monotone-decay shock is

$$\partial\widetilde{\gamma}_t \Big/ \partial\boldsymbol{\epsilon}_t = \boldsymbol{I} + \rho\beta\boldsymbol{\Omega}_1 + (\rho\beta)^2\left(\boldsymbol{\Omega}_1^2 + \boldsymbol{\Omega}_2\right) + (\rho\beta)^3\left(\boldsymbol{\Omega}_1^3 + \boldsymbol{\Omega}_1\boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_2\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_3\right) + \cdots \tag{23}$$

and the response to a concurrent hump-shaped shock is

$$\partial\widetilde{\gamma}_t \Big/ \partial \boldsymbol{u}_t = 1\times\boldsymbol{I}+2\times\rho\beta\boldsymbol{\Omega}_1+3\times(\rho\beta)^2\left(\boldsymbol{\Omega}_1^2+\boldsymbol{\Omega}_2\right)+4\times(\rho\beta)^3\left(\boldsymbol{\Omega}_1^3+\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2+\boldsymbol{\Omega}_2\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_3\right)+\cdots \tag{24}$$

Each successive term in these summations captures the revenue response due to demand in a subsequent period. For a monotone-decay shock, the weights associated with each term form an exponentially decaying sequence: $1, \rho\beta, (\rho\beta)^2, \ldots$ The further distant a sector is from the shock, the lower is the impact of the shock on sectoral revenue. By contrast, the weights for a hump-shaped shock have a time profile $1, 2\rho\beta, 3(\rho\beta)^2, \ldots$, resulting in the potentially amplified response in upstream sectors.

In the special case where each input has one period delay only, $\boldsymbol{G}^{\epsilon}_{\infty} = (\boldsymbol{I}-\rho\beta\boldsymbol{\Omega}_1)^{-1}$ and $\boldsymbol{G}^{x}_{\infty} = (\boldsymbol{I}-\rho\beta\boldsymbol{\Omega}_1)^{-2}$, and equation (24) becomes

$$\partial\widetilde{\gamma}_t \Big/ \partial \boldsymbol{u}_t = (\boldsymbol{I}-\rho\beta\boldsymbol{\Omega}_1)^{-2} = \boldsymbol{I}+2\rho\beta\boldsymbol{\Omega}_1+3(\rho\beta\boldsymbol{\Omega}_1)^2+4(\rho\beta\boldsymbol{\Omega}_1)^3+\cdots$$

where the Leontief-inverse-squared is the matrix analogous to the scalar summation formula $\sum_{t=0}^{\infty}(t+1)\rho^t = (1-\rho)^{-2}$.

The key takeaway from the characterizations is that how producers respond to a downstream demand shock depends on their positions along the supply chain. Downstream producers respond directly to current demand, while upstream producers, due to time-to-build delays, respond to anticipated future demand. This insight explains why hump-shaped shocks can propagate upstream and become significantly amplified, generating pronounced volatility in upstream sectors.

3.2.2 Examples: Vertical and Horizontal Supply Chains

To gain further intuition, we consider two example networks that isolate the mapping from delays to responses: a downstream sector facing a vertical supply chain, in which delays to the downstream sector accumulate through successive production stages, and a horizontal supply chain, in which the downstream sector sources all inputs directly and delays differ across suppliers' order horizons.

In the vertical chain, there are $N = 4$ sectors: sector 1 produces with labor only, each successive sector $i \geq 2$ produces using input $i-1$ with intermediate elasticity $\omega_{i,i-1} = 1$, each good takes one period to build before being used by sector $(i+1)$ as an input, and all goods may be consumed. A supplier's delay to sector four equals the number of production stages between them: sector three commits production one period before its output serves sector four's demand; sector two, two periods; sector one, three periods. The revenue responses to monotone-decay and hump-shaped demand shocks are

$$\frac{\partial \widetilde{\boldsymbol{\gamma}}_t}{\partial \boldsymbol{\epsilon}_t} \times \boldsymbol{\epsilon}_t = \begin{bmatrix} 1 & \rho\beta & (\rho\beta)^2 & (\rho\beta)^3 \\ 0 & 1 & \rho\beta & (\rho\beta)^2 \\ 0 & 0 & 1 & \rho\beta \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{bmatrix}, \qquad \frac{\partial \widetilde{\boldsymbol{\gamma}}_t}{\partial \boldsymbol{u}_t} \times \boldsymbol{u}_t = \begin{bmatrix} 1 & 2\rho\beta & 3(\rho\beta)^2 & 4(\rho\beta)^3 \\ 0 & 1 & 2\rho\beta & 3(\rho\beta)^2 \\ 0 & 0 & 1 & 2\rho\beta \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}.$$

Consider shocks to the most downstream sector four and examine the last columns. Following a monotone-decay shock ϵ_4, sector four's revenue responds one-for-one, sector three's response is dampened by $\rho\beta < 1$, and sector two's is further dampened to $(\rho\beta)^2$: responses decline with upstream distance. Following a hump-shaped shock u_4, the responses of sectors three, two, and one are $2\rho\beta$, $3(\rho\beta)^2$, and $4(\rho\beta)^3$, each potentially amplified relative to the sector downstream of it.

In the horizontal supply chain, sector four instead sources all three inputs directly at different horizons, with production function

$$y_{4,t} = m_{1,t-1}^{1/3} m_{2,t-2}^{1/3} m_{3,t-3}^{1/3},$$

where input m_k must be committed k periods in advance, and each input supplier produces with labor only. Each supplier's current sales load on expected downstream demand at the horizon of its own delivery delay, and the responses to sector four's demand shocks are proportional to those in the vertical chain: the supplier at horizon k responds in proportion to $(\rho\beta)^k$ following a monotone-decay shock and to $(1+k)(\rho\beta)^k$ following a hump-shaped shock.

For $\rho\beta = 0.9$, the hump-shaped profile terms for the input suppliers equal 1.8, 2.43, and 2.92, and the supplier that must commit production furthest in advance responds most strongly—the bullwhip effect in its simplest form. Both examples convey the same mechanism: the bullwhip effect arises from suppliers' differential responses to intertemporal demand shocks under heterogeneous delays, as delays determine each supplier's exposure to expected future demand at different horizons. Network structure is a natural reason why delays are heterogeneous: more upstream suppliers naturally face longer delays, as their output must traverse more production stages before serving final demand. The cross-section of supplier responses then traces out the expected time path of downstream demand.

Note also that, for a sector sufficiently upstream, the response is dampened and eventually converges to zero (as $\lim_{s\to\infty}(s+1)(\rho\beta)^s = 0$): with a long enough cumulated delay, the sector does not respond even to hump-shaped demand shocks.

3.2.3 A Bilateral Measure of Supply Chain Delays

We now introduce a measure of the pairwise input delay between an input supplier and an input user. Consider sector i's revenue response to sector j's hump-shaped demand shock relative to that of a monotone-decay shock and define

$$\xi_{ij} \equiv \begin{cases} \left[\frac{\partial \widetilde{\gamma}_t}{\partial \boldsymbol{u}_t}\right]_{ij} \Big/ \left[\frac{\partial \widetilde{\gamma}_t}{\partial \boldsymbol{\epsilon}_t}\right]_{ij} - 1 & \text{if } \left[\frac{\partial \widetilde{\gamma}_t}{\partial \boldsymbol{\epsilon}_t}\right]_{ij} \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

ξ_{ij} is the element-wise ratio between the matrices capturing the response to hump-shaped and monotone-decay shocks; $\xi_{ij} = 0$ if $\left[\frac{\partial \widetilde{\gamma}_t}{\partial \boldsymbol{\epsilon}_t}\right]_{ij} = 0$. In the vertical supply chain example, the matrix $\boldsymbol{\Xi} \equiv \left[\xi_{ij}\right]$ is

$$\boldsymbol{\Xi} = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

ξ_{ij} equals the delay between the production of input i and its eventual use in sector j, and is set to zero if input i is never used (even indirectly) by sector j.

We now show that the notion of ξ_{ij} as a bilateral measure of input delays extends naturally to a general network structure. When multiple routes with different delays connect an input supplier i to an input user j, ξ_{ij} is the average delay across all of them:

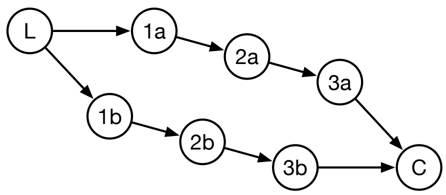
$$\xi_{ij} = \frac{\sum_{s \geq 1} s \left(\rho\beta\right)^s \left[\boldsymbol{\Omega}^{(s)}\right]_{ij}}{\sum_{s \geq 0} \left(\rho\beta\right)^s \left[\boldsymbol{\Omega}^{(s)}\right]_{ij}}, \qquad \boldsymbol{\Omega}^{(s)} \equiv \sum_{\phi \in \Phi_s} \prod_k \boldsymbol{\Omega}_{\phi_k}, \quad \boldsymbol{\Omega}^{(0)} \equiv \boldsymbol{I},$$

where $\left[\boldsymbol{\Omega}^{(s)}\right]_{ij}$ collects all routes from i to j with cumulative delay s (recall Φ_s is the set of sequences of positive integers that sum to s, as defined in Theorem 1). Each route is weighted by the value of sector i's output sent along it in response to a monotone-decay demand shock in j; the weights are the terms of the response in equation (23), normalized by their sum, and therefore sum to one. Routes with longer delays are discounted by $\left(\rho\beta\right)^s$: demand decay and time preference make distant expected demand matter less, so the distance that equilibrium responses load on is the discounted average delay.

Consider an example network below. The network has two supply chains labeled a and b, each with three sectors. Consider $3a$ as the focal sector. The delay from $2a$ to $3a$ is d_{2a}. The delay from $1a$ to $3a$ is $d_{1a} + d_{2a}$, as input $1a$ has to first go through producer $2a$. The delay from any sector along the b chain is unconnected to $3a$ and thus has a delay measure of zero. The matrix $\boldsymbol{\Xi}$ therefore reveals the bilateral supply chain distance measured by production delays.

It is important to note that our measure ξ_{ij} differs from the standard Leontief inverse. The

Figure 3. Two production chains



$$\Xi = \begin{bmatrix} 0 & d_{1a} & d_{1a}+d_{2a} & 0 & 0 & 0 \\ 0 & 0 & d_{2a} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & d_{1b} & d_{1b}+d_{2b} \\ 0 & 0 & 0 & 0 & 0 & d_{2b} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

latter captures a notion of network dependence—the network-adjusted value of each input used to produce a unit value of each good. Instead, our bilateral measure ξ_{ij} captures a notion of network distance due to delays. Our measure ξ_{ij} can therefore be used to identify the ordering of inputs along each good-specific supply chains, as evident from the example in Figure 3 of two production chains. We implement this measure empirically in Section 4.

Taking the final consumer as the destination converts the bilateral measure into a sector-level notion of upstreamness: aggregating ξ_{ij} over destinations with final-use weights and removing the discounting,

$$U_i^{\text{delay}} \equiv \lim_{\beta,\rho\to 1} \sum_j \xi_{ij} \frac{\left[(\boldsymbol{I}-\boldsymbol{\Omega})^{-1}\right]_{ij}\bar{\theta}_j}{\sum_k \left[(\boldsymbol{I}-\boldsymbol{\Omega})^{-1}\right]_{ik}\bar{\theta}_k} = 0\cdot\frac{\bar{\theta}_i}{\bar{\gamma}_i} + \sum_j d_{ji}\,\omega_{ji}\frac{\bar{\theta}_j}{\bar{\gamma}_i} + \sum_{j,k}(d_{kj}+d_{ji})\,\omega_{kj}\omega_{ji}\frac{\bar{\theta}_k}{\bar{\gamma}_i} + \cdots,$$

the average cumulative delay between sector i's production and its eventual consumption.

U_i^{delay} is the delay-weighted counterpart of the upstreamness measure ψ_i of Antràs et al. (2012), the average number of production stages from a sector's output to final consumption: it obtains when each walk contributes its number of stages instead of its cumulative delay, and with one-period delays throughout, $U_i^{\text{delay}} = \psi_i - 1$, direct final sales carrying delay zero rather than distance one. Alfaro et al. (2019) extend ψ_i to a bilateral, pair-level measure; relative to these, our measure counts time rather than stages, discounts distant routes, and is derived from equilibrium responses rather than constructed from the input–output table.

3.2.4 Hump-shaped Demand Shocks and Amplified Volatility in Upstream Sectors

In sections 3.2.1 and 3.2.2 we characterized the bullwhip effect in terms of impulse responses. Volatility follows as a corollary: conditional on current information, the only source of variation in one-period-ahead revenue is next period's demand shocks, on which revenue loads through the same matrices $\boldsymbol{G}_\infty^\epsilon$ and $\boldsymbol{G}_\infty^x$ of Theorem 2. Amplified responses to hump-shaped demand shocks therefore translate directly into amplified conditional volatility in upstream sectors.

Corollary 2. *In the environment of Theorem 2, the sectoral revenue volatility is:*

$$Var_t\left(\widetilde{\gamma}_{t+1}\right) = \boldsymbol{G}_\infty^\epsilon \boldsymbol{\Sigma}_t^\epsilon \boldsymbol{G}_\infty^{\epsilon\prime} + \boldsymbol{G}_\infty^x \boldsymbol{\Sigma}_t^u \boldsymbol{G}_\infty^{x\prime},$$

where $\boldsymbol{\Sigma}_t^\epsilon$ and $\boldsymbol{\Sigma}_t^u$ denote the conditional variance-covariance matrices $Var_t\left(\boldsymbol{\epsilon}_{t+1}\right)$ and $Var_t\left(\boldsymbol{u}_{t+1}\right)$ of next period's shocks, respectively, and $\boldsymbol{G}_\infty^\epsilon, \boldsymbol{G}_\infty^x$ are as defined in Theorem 2.

Proof. See Appendix A.6. □

Intuitively, the two terms decompose the conditional variance of sectoral revenue into the contributions of the monotone-decay and hump-shaped components of demand: in parallel to sections 3.2.1 and 3.2.2, monotone-decay demand shocks from downstream sectors create dampened volatility in upstream revenue, whereas hump-shaped demand shocks create amplified volatility—even when only a single input-output path connects the two sectors.[10]

3.3 One Demand Innovation that Loads on both Time Profiles

Sections 3.1 and 3.2 characterize the responses to two types of demand innovations: those whose expected path decays monotonically and those whose expected path is hump-shaped. In general, a sector's demand innovations need not be of either type—an innovation may load partly on each time profile. We therefore consider a demand process in which each sector's demand innovation is a mixture of the two shocks. We show that the same process also describes observed demand in the two-shock environment of Section 3.1 when agents cannot distinguish the two shocks and forecast from realized demand alone: their forecasts follow a Kalman filter, and observed demand follows exactly this process, with the Kalman gain as the weights.

Formally, suppose sectoral demand follows

$$\widetilde{\boldsymbol{\theta}}_t = \rho\widetilde{\boldsymbol{\theta}}_{t-1} + \boldsymbol{x}_t + \left(\boldsymbol{I} - \boldsymbol{D}\right)\boldsymbol{v}_t, \qquad \boldsymbol{x}_t = \rho\boldsymbol{x}_{t-1} + \boldsymbol{D}\boldsymbol{v}_t, \tag{26}$$

where $\boldsymbol{v}_t$ is the vector of mean-zero, serially uncorrelated demand innovations, which may be correlated across sectors, and the $N \times N$ matrix $\boldsymbol{D}$ collects these weights: each innovation enters the persistent state $\boldsymbol{x}_t$ through $\boldsymbol{D}$, generating the hump-shaped part of the expected path, and enters demand directly through $\boldsymbol{I} - \boldsymbol{D}$, generating the monotone-decay part. Agents observe the state $\boldsymbol{x}_t$ and the innovations $\boldsymbol{v}_t$ directly, as in the environment of Theorem 2. The process nests the two pure cases of Section 3.1: setting $\boldsymbol{D}$ to zero yields purely monotone-decay demand and setting $\boldsymbol{D} = \boldsymbol{I}$ yields purely hump-shaped demand. In our estimation we restrict $\boldsymbol{D}$ to be diagonal with entries $\delta_i \in [0, 1]$, so each sector's demand is a convex mixture of the two pure

[10] In the vertical supply chain example in Section 3.2.2, when only the downstream sector four faces uncertain demand—with variance $\left(\sigma^\epsilon\right)^2$ and $\left(\sigma^u\right)^2$ for the monotone-decay and hump-shaped shocks, respectively—the conditional variance of one-period-ahead revenue $Var_t\left(\widetilde{\gamma}_{i,t+1}\right)$ of sector i is

$$Var_t\left(\widetilde{\gamma}_{i,t+1}\right) = \left(\rho\beta\right)^{2(4-i)}\left(\sigma^\epsilon\right)^2 + \left(5-i\right)^2\left(\rho\beta\right)^{2(4-i)}\left(\sigma^u\right)^2.$$

The effect of monotone-decay demand shocks on upstream revenue is dampened by a factor of $\left(\rho\beta\right)^{2(4-i)}$, whereas the effect of hump-shaped demand shocks may be magnified.

profiles (Section 4.4). The expected path of demand following an innovation is the weighted average of the two pure profiles in equation (20): $\partial \mathbb{E}_t \left[\widetilde{\boldsymbol{\theta}}_{t+s}\right] \Big/ \partial \boldsymbol{v}_t = \rho^s \left(\boldsymbol{I} - \boldsymbol{D}\right) + \left(s+1\right) \rho^s \boldsymbol{D}$. Substituting into Theorem 1, the response of expected revenue interpolates linearly between the two responses of Theorem 2:

$$\frac{\partial \mathbb{E}_t \left[\widetilde{\gamma}_{t+s}\right]}{\partial \boldsymbol{v}_t} = \rho^s \left(\boldsymbol{G}^{\epsilon}_{\infty} \left(\boldsymbol{I} + \left(s-1\right) \boldsymbol{D}\right) + \boldsymbol{G}^{x}_{\infty} \boldsymbol{D}\right). \tag{27}$$

Sectors whose innovations load predominantly onto the persistent state behave as AR(2)-dominant and generate the upstream amplification of Section 3.2.1; sectors whose innovations load predominantly onto the monotone-decay component behave as AR(1)-dominant.

We now formally state the equivalence of the demand process (26) to a setting in which agents must forecast demand without separately observing the monotone-decay and hump-shaped components.

Proposition 3. *Suppose demand follows the two-shock process in equations (18) and (19) with Gaussian innovations and time-invariant covariance matrices, and agents observe only the history of realized demand* $\left\{\widetilde{\boldsymbol{\theta}}_{t-s}\right\}_{s\geq 0}$*, forecasting by Kalman filtering with steady-state gain* $\boldsymbol{K}$*. Then observed demand follows the one-innovation process (26) with* $\boldsymbol{D} = \boldsymbol{K}$*, in which the state is the agents' nowcast* $\hat{\boldsymbol{x}}_t = \mathbb{E}\left[\boldsymbol{x}_t \mid \left\{\widetilde{\boldsymbol{\theta}}_{t-s}\right\}_{s\geq 0}\right]$ *of the hidden component—not the hidden component itself—and* $\boldsymbol{v}_t$ *is the one-step-ahead forecast error of demand. For every history of demand, conditional expectations of future demand—and hence sectoral revenues—coincide in the two environments, with the response to a demand innovation given by (27); and, with the innovation covariance implied by the filter, the two environments generate identical autocovariances and an identical Gaussian likelihood for observed demand.*

Proof. See Appendix A.7. □

Intuitively, the feature that both components load on the same innovation is the filter's attribution of a single forecast error rather than an assumption. Observing only realized demand thus introduces no new propagation force: it determines which combination of the two responses of Theorem 2 the economy exhibits. We accordingly refer to the process as the one-innovation representation of demand. Sectoral revenue volatility in this process is

$$Var_t\left(\widetilde{\gamma}_{t+1}\right) = \left[\boldsymbol{G}^{\epsilon}_{\infty}\left(\boldsymbol{I} - \boldsymbol{D}\right) + \boldsymbol{G}^{x}_{\infty}\boldsymbol{D}\right] Var_t\left(\widetilde{\boldsymbol{\theta}}_{t+1}\right) \left[\boldsymbol{G}^{\epsilon}_{\infty}\left(\boldsymbol{I} - \boldsymbol{D}\right) + \boldsymbol{G}^{x}_{\infty}\boldsymbol{D}\right]', \tag{28}$$

where $Var_t\left(\widetilde{\boldsymbol{\theta}}_{t+1}\right) = \boldsymbol{\Sigma}_v$, since $\boldsymbol{v}_{t+1}$ is the one-step-ahead forecast error of demand.

The representation facilitates our empirical analysis: in the data, only realized demand is observed, and the two components are not separately observable; because the representation is written entirely in terms of observed demand and its innovations, it can be taken to the data directly. Section 4.4 estimates it on the panel of realized downstream demand.

4 Empirical Evidence and Quantification

We now turn to the empirical analysis with three goals. First, in Section 4.2 we document a motivating cross-sectional pattern. That is, within a sectoral supply chain, upstream sectors tend to exhibit higher volatility than downstream sectors. Second, we extract innovations to downstream value-added and estimate the impulse-response functions of upstream sectors along the supply chain in Section 4.3. We demonstrate a number of features of the estimated impulse-response functions that are consistent with our model's predictions, thereby providing a sharp test of our theory. Third, in Section 4.4 we conduct a model-based exercise. We infer the empirical process of demand shocks using the model and demonstrate that hump-shaped demand shocks for downstream sectors account for a quantitatively significant share of sectoral fluctuations along the supply chains.

4.1 Data

We use data from three main sources. First is the Industrial Production (IP) data published by the Federal Reserve Board. The data set has panel information on the sectoral value-added for manufacturing industries over the period 1972-2019, with industries classified by the North American Industry Classification System (NAICS). To smooth out measurement errors and high-frequency shocks, we average the monthly data to construct quarterly sectoral value-added in our baseline analysis, while monthly specifications are presented as robustness checks in Appendix D.10.

Second, to construct our bilateral measure of supply chains, we need information on input-output linkages and input time-to-build delays. For input-output data, we utilize the 2007 BEA input-output (IO) use table,[11] from which we construct input cost shares ω_{ij} and sectoral value-added intensities (value-added to revenue ratio).

We follow Foerster et al. (2011)'s algorithm to concord between the IP and IO datasets. This procedure results in 114 manufacturing industries appearing in both datasets, as well as 127 non-IP sectors present only in the IO table. Among the IP industries, the majority (65 out of 114) corresponds to four-digit NAICS codes, with the remaining mapping into either three- (4 industries) or five-digit (45 industries) NAICS codes. We format the IP data ("Value-Added Proportions") to represent sectoral value-added shares—each sector's nominal value-added as a proportion of the total nominal value-added across all 114 IP sectors. This format aligns with our analysis, which centers on sectoral and supply chain dynamics rather than aggregate effects.[12]

[11]The BEA IO tables are available every five years. We use the 2007 version as it is in the second half of our sample but before the Great Financial Crisis. Our results are not sensitive to this choice.

[12]We impute missing data according to the Federal Reserve Board's recommended methodology: if sector C is made up of sector A less sector B, then the value added of sector C is the value added of A less that of B.

Third, we need information on time-to-build delays. Our baseline strategy is to use the backlog ratio for each input, i.e., the ratio between the stock value of unfilled orders and the flow value of goods delivered, as a measure of supply chain delays (Liu and Tsyvinski, 2024). The strategy yields a supplier-specific measure of delay. We rely on the U.S. Census M3 survey of manufacturers' shipments, inventories, and orders. The data set provides broad-based, monthly statistical data on economic conditions in the manufacturing sector. For each industry, we compute the average backlog ratio between the years 2015 and 2019.[13] We match industries in the M3 survey to the finest partition possible in the IO table.[14] We impute the backlog ratio using the sample averages respectively for durable goods sectors and non-durable goods sectors that are not in the M3 survey. Because the backlog ratio captures only the average time between orders placed and orders received but not delays due to the production process of the input-buyer, we round up the backlog ratios when converting from months to quarters. Based on this procedure, twenty-three sectors have time-to-build delays of two quarters, two sectors have delays of three quarters, four sectors have delays of four quarters, and one sector ("Ship and Boat Building") has five-quarter delays.

As robustness checks, we implement two alternative strategies for measuring time-to-build delays. The first alternative strategy follows the recent work of Antràs and Tubdenov (2025), which infers an input-buyer-specific measure of time-to-build based on the ratio between the value of inventories and cost of goods sold (COGS) using the COMPUSTAT data and then aggregates to the sector level using a cost-weighted average across firms. This measure captures production delays by input-buyers but does not reflect shipping or production delays originating from input-suppliers. As such, it is complementary to our baseline measure. The second alternative strategy specifies that the total time-to-build delay from an input producer j to input user i is the sum of the backlog ratio of good j and the inventory-to-COGS ratio of good i. Our main results remain robust across these alternative measures of time-to-build—network structure consistently plays a more significant role, as detailed in the Appendix D.6.

4.2 Higher Volatility in Upstream Sectors Across Supply Chains

We first document a motivating cross-sectional pattern across supply chains of the U.S. production network: upstream sectors within a supply chain tend to have higher volatility.

We measure sectoral volatility as the standard deviation of the year-on-year (YoY) growth in sectoral value-added shares (calculated relative to total value-added across IP industries) over

[13]Most of the variation in the backlog ratio is cross-sectional. Regressing the backlog ratio between 1992 and 2019 on sector fixed effects yields an R^2 of 0.87, indicating limited variation over time.

[14]IP sectors without a direct M3 match are assigned the residual backlog ratio of their broad M3 group; for transportation, where the residual is dominated by multi-year aerospace procurement backlogs that overstate production lags, we assign the group-wide ratio instead—a conservative choice.

Table 1. The list of twenty-five downstream sectors

Apparel	Coffee and Tea
Ship and Boat Building	Other Transportation Equipment
Newspaper Publishers	Heavy Duty Trucks
Tobacco	Office And Other Furniture
Soft Drinks and Ice	Aerospace Products and Parts
Bakeries and Tortilla	Breweries
Motor Vehicle Bodies and Trailers	Other Food Except Coffee and Tea
Carpet and Rug Mills	Pharmaceuticals and Medicines
Soap, Cleaning Compounds, and Toilet Preparation	Navigational/Measuring/Electromedical/Control Instruments
Periodical, Book, and Other Publishers	Sugar and Confectionery Products
Automobiles and Light Duty Motor Vehicles	Animal Slaughtering and Processing
Major Electrical Household Appliances	Medical Equipment and Supplies
Fruit and Vegetable Preserving and Specialty Foods	

our sample. Using YoY growth in value-added offers several advantages: it removes seasonal fluctuations and normalizes for intrinsic size differences across sectors, allowing for meaningful volatility comparisons. Moreover, YoY growth serves as a convenient reduced-form proxy for the conditional variance, as last year's value-added levels effectively summarize the relevant time-varying information set. Finally, although we have formulated our theory in terms of sectoral revenue, the YoY growth measure removes the distinction between revenue and value-added to the extent that sectoral value-added intensity stays constant.

The main challenge in documenting this pattern is to identify the supply chains in the US production network. To do so, we rely on our bilateral measure of supply chain delays ξ_{ij}, following the procedure described in Section 3.2.3.

We start by identifying a set $\mathcal{D}$ of 25 downstream sectors that are important in the consumption bundle, following Antràs et al. (2012).[15] These sectors are listed in Table 1; they supply most of their output directly to the final consumer instead of other producers. Broadly speaking, these sectors fall into the categories of transportation, food, and other household consumption goods.

We then identify the supply chain for each of these downstream sectors. First, for each downstream sector $j \in \mathcal{D}$, we keep a set of sectors $\mathcal{U}_j$ that are considered as direct or indirect suppliers to j: a sector $i \neq j$ is in the set $\mathcal{U}_j$ if, in the steady-state of the model, producing 100 dollars of j's output involves at least one dollar of input i, directly or indirectly. Formally, this means that the ij-th entry of the Leontief inverse L_{ij}—computed from equation (11) with a realistic quarterly discount factor $\beta = 0.98$, used throughout the paper—is greater than 1 percent. Under this

[15] Specifically, we first select the 30 most downstream sectors according to the upstreamness measure by Antràs et al. (2012). We then drop "Other Miscellaneous Manufacturing", "Construction Machinery", "Agricultural Implements", "Industrial Machinery" and "Support Activities for Mining", as these five sectors do not account for significant consumer expenditure and are thus not important for the consumer demand.

threshold, each downstream good has on average 16 (median 16) sectors along its supply chain.

Next, for each downstream sector $j \in \mathcal{D}$, we order the set of suppliers to j using our bilateral supply chain delay measure ξ_{ij}, following the procedure described in Section 3.2.3. In the baseline, we set $\rho = 1$ when calculating ξ_{ij}. Our results are robust to using alternative parameter values: as Appendix Figure A.10 shows, ξ_{ij} remains highly correlated across a wide range of ρ, including the rate of decay parameter $\rho = 0.829$ estimated from the data (see Section 4.4). The resulting measure captures the average number of quarters in the time-to-build delays for input i to reach producer j, accounting for all the direct and indirect network linkages from i to j. Within the supply chain of each good j, we interpret sector $i \in \mathcal{U}_j$ as more upstream to $k \in \mathcal{U}_j$ if there is a longer delay from input i to producer j than from input k.

Take "Automobiles and Light Duty Motor Vehicles" ("automobiles" for short) and "Motor Vehicle Bodies and Trailers" ("trailers" for short) as two examples of downstream sectors. Despite the fact that these two sectors share many common suppliers—such as "Machine Shops; Turned Products; and Screws, Nuts, and Bolts", "Paperboard Containers", "Semiconductors", "Iron and Steel Products", and "Electric Power Generation, Transmission and Distribution"—our measure can isolate the differences in their supply chains. Specifically, our measure identifies that the sectors most immediately upstream to automobiles are "Motor Vehicle Bodies and Trailers", "Motor Vehicle Parts" and "Tires". By contrast, those that are most immediately upstream to trailers include "Major Electrical Household Appliances" and "Millwork".

Figure 4 shows the relationship between upstreamness, as measured by supply chain delays, and volatility, along the sector-specific supply chain for 9 of the 25 downstream sectors (we produce the same figure for each of the remaining 16 downstream sectors in Appendix Figure A.1).[16] Specifically, each panel represents the supply chain associated with a downstream sector, as indicated by the subtitle (e.g., the top-left panel is for the supply chain of automobiles). Within each panel, a circle represents an input supplier along the supply chain (e.g., the bottom circle in the top-left panel is "Motor Vehicle Parts"). The Y-axis captures the time-to-build delays from the supplier to the downstream producer, i.e., a notion of upstreamness. The X-axis is the volatility in the YoY value-added growth of the supplier. The size of each circle represents the share of the total downstream cost that is spent on each input, directly or indirectly.

The key takeaway from Figure 4 is the significant positive relationship between sectoral volatility and upstreamness, an empirical pattern consistently observed across supply chains for all 25 downstream sectors. In terms of magnitude, pooling across 25 supply chains, we find that going from the 25th to the 75th percentile in terms of upstreamness is associated with 6.1 per-

[16] We include only non-durable inputs—goods for which no more than 50% of its final use goes into fixed capital formation in the BEA IO table—when investigating this cross-sectional relationship, as the volatility of durable goods may be driven by important factors not considered in the model.

centage points higher volatility in value-added growth. This effect is pronounced across transportation, food, household consumption goods, and all other categories of downstream goods.

One potential concern is that the observed positive relationship between sectoral volatility and upstreamness might be driven by the volatility of commodity prices, as four commodity sectors occupy upstream positions across many supply chains.[17] However, we find that this is not the case. Appendix Figure A.2 shows that the positive relationship between sectoral volatility and upstreamness remains qualitatively unchanged even after excluding these commodity sectors.

4.3 Supply Chain Response to Downstream Value-Added Shocks

Crucially, our theory predicts more than higher upstream volatility in the cross section: it pins down the dynamics of impulse responses. Specifically, equation (27) implies that following an innovation to value-added in downstream sectors dominated by AR(2) shocks:

1. The downstream sectors' own value-added exhibits a hump-shaped response over time.
2. Value-added responses along the supply chain are also hump-shaped and initially amplified as shocks propagate upstream.
3. Moving along the supply chain, the initial amplification eventually diminishes in sectors located sufficiently far upstream.
4. The downstream hump-shaped responses and upstream amplification effects should be stronger in sectors where AR(2) shocks dominate, while such patterns could be absent in sectors characterized primarily by AR(1) shocks.

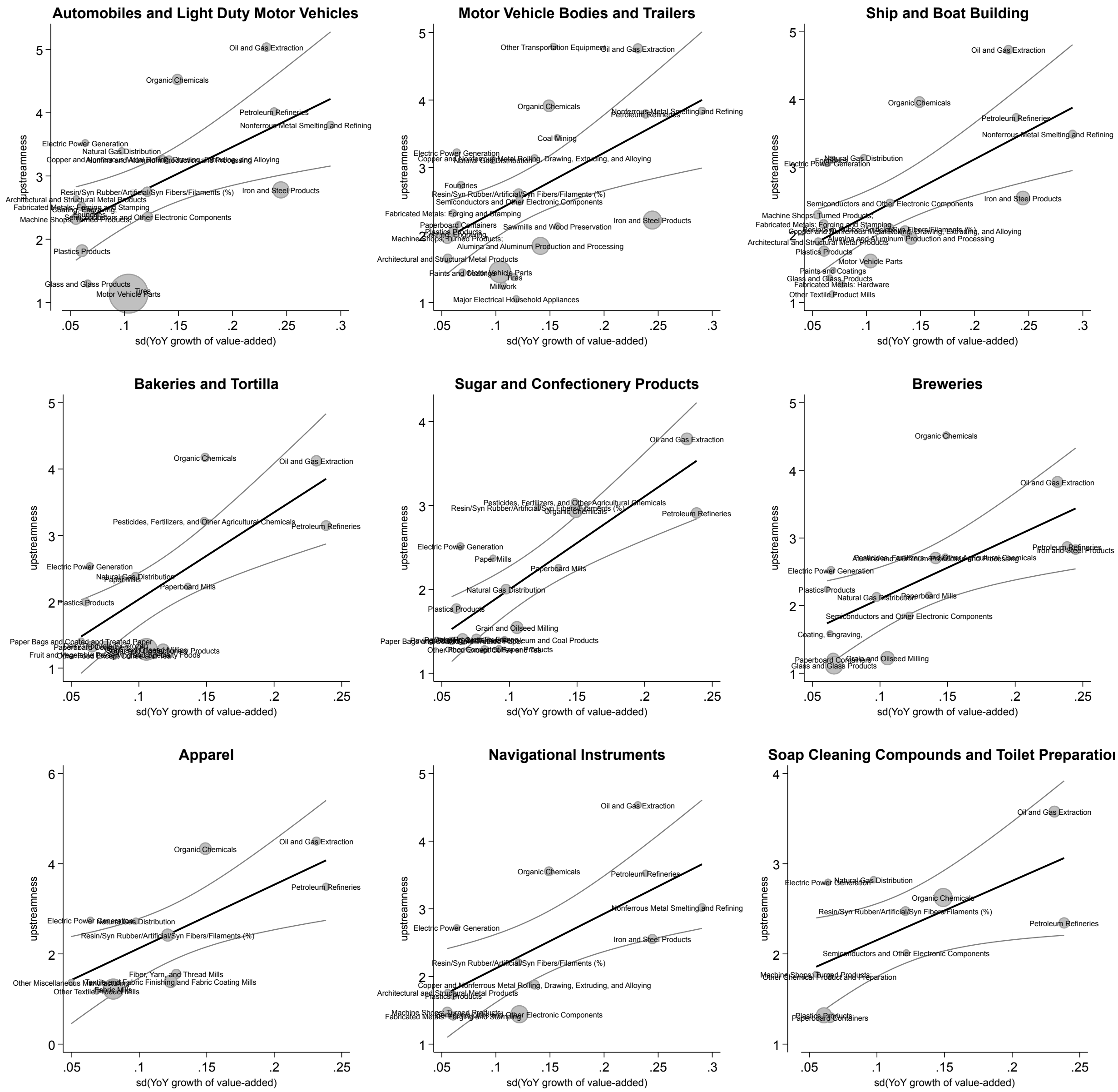
We empirically test these predictions by first removing linear time trends from each sector's value-added series, standardizing the detrended data, and using these standardized z-scores in our analysis to account for heterogeneous sectoral sizes.

Although in our theoretical framework, fluctuations in value-added (i.e., revenue multiplied by sectoral value-added intensity) originate exclusively from demand shocks—since productivity shocks have offsetting effects on prices and quantities—our empirically testable predictions directly concern fluctuations in sectoral value-added.

In the main text, we directly test the predictions based on extracting shocks to downstream value-added and estimating their impulse response functions along supply chains. Additionally, in Appendix D.12, we discuss a robustness check using a generalized model with constant-elasticity-of-substitution (CES) production functions to explicitly control for sector-specific cost

[17] These sectors are "Oil and Gas Extraction", "Coal Mining", "Copper, Nickel, Lead, and Zinc Mining", and "Iron, Gold, Silver and Other Metal Ore Mining".

Figure 4. Along supply chains, upstream sectors exhibit higher volatility



shocks. Specifically, we solve the CES model with one-period delays, linearize it around a steady-state, and derive separate implications of cost shocks and demand shocks. This approach enables us to nonparametrically control for cost shocks when estimating impulse-response functions. Our empirical findings remain robust after incorporating these controls.

Downstream Value-Added Has Both AR(1) and AR(2) Shocks We begin by using a reduced-form test to show that the time-series of detrended downstream value-added is well captured by a stochastic process with AR(1) and AR(2) shocks. Specifically, we follow the econometrics literature and use partial autocorrelation function (PACF) to select appropriate lags in an $AR\,(p)$ model (Box et al., 2015). The PACF coefficient ϕ_k of a time series is defined as the autocorrelation at lag k after conditioning on all previous lags 1 through $k-1$. According to the Box–Jenkins method (Box et al., 2015), insignificance of the PACF coefficient at lag $p+1$ indicates that the time-series can be modeled as an $AR(p)$ process.

We estimate the PACF coefficient sector-by-sector for the 25 downstream sectors $i \in \mathcal{D}$. We find that in our sample, all twenty-five downstream sectors have statistically significant (5%) PACF coefficients at lag 1 for the time-series of their value-added. Twenty-four out of twenty-five downstream sectors have statistically significant PACF coefficients at lag 2. Only five sectors ("Tobacco", "Other Food Except Coffee and Tea", "Navigational/Measuring/Electromedical/Control Instruments", "Apparel", "Major Electrical Household Appliances") have statistically significant PACF coefficients at lag 3. These findings suggest that downstream value-added is well approximated by a stochastic process with AR(1) and AR(2) shocks.

Moreover, the PACF coefficient at lag 2 provides a natural reduced-form index of the relative importance of hump-shaped shocks: treating each sector's demand as a univariate process with a single mixture weight δ_i—the diagonal specification we estimate in Section 4.4—$|\phi_{i,2}|$ is increasing in δ_i.[18] Conversely, sectors with smaller $|\phi_{i,2}|$ predominantly experience monotone-decay AR(1) shocks.

To test our model's prediction on the differential supply chain response following shocks to these sectors, we define P_i as the dummy variable for whether $|\phi_{i,2}|$ is above the median among the 25 downstream sectors. For expositional simplicity we refer to these sectors as having relatively large AR(2) components.

Extracting Innovations in Downstream Value-Added We use local projections to estimate the impulse-response along the supply chain following downstream value-added shocks. To estimate the value-added shock of the 25 downstream sectors, we project each sector's value-added on its own lags, as well as sector fixed effects μ_i and time fixed effects μ_t, and we extract the residuals. Specifically, the value-added shock r_{it} in a downstream sector i is the innovation in the following:

$$VA_{it}^{down} = \sum_{s=1}^{p} a_s VA_{i,t-s}^{down} + \sum_{s=1}^{p} \tilde{a}_s\, P_i \cdot VA_{i,t-s}^{down} + \mu_i + \mu_t + r_{it}, \quad i \in \mathcal{D}. \tag{29}$$

The autoregressive coefficients are a_s for the below-median group ($P_i = 0$) and $a_s + \tilde{a}_s$ for the above-median group ($P_i = 1$), so the residual r_{it} is the innovation relative to a group-specific AR(p) process. In the baseline specification, we include $p = 2$ lags, consistent with our model

[18] Formally, under the diagonal specification of Section 4.4, sector i's observable demand is the ARMA(2,1) process $(1-\rho L)^2 \widetilde{\theta}_{it} = (1-\rho\,(1-\delta_i)\,L)\, v_{it}$, with autocorrelation function $r_{i,k} = \rho^k\,(1+kb_i)$, where $b_i \equiv \frac{\delta_i\left(1-\rho^2\right)\left[\left(1-\rho^2\right)+\delta_i\rho^2\right]}{(1-\rho^2)^2+2\delta_i\rho^2(1-\rho^2)+\delta_i^2\rho^2(1+\rho^2)}$. Hence $\phi_{i,1} = \rho\,(1+b_i)$ and $\phi_{i,2} = \frac{r_{i,2}-r_{i,1}^2}{1-r_{i,1}^2} = -\frac{\rho^2 b_i^2}{1-\rho^2(1+b_i)^2}$.

specification of AR(1) and AR(2) shocks. Our results are not sensitive to the choice of lags; Appendix D.8 replicates our results for $p = 1$ and $p = 3$ as robustness checks.

Impulse-Response of Downstream Value-Added to Its Own Innovations Equipped with downstream's value-added shocks, we first estimate the impact of a shock on the downstream sector's own value-added in subsequent periods:

$$VA^{down}_{i,t+h} - VA^{down}_{i,t-1} = \alpha^h_0 r_{it} + \alpha^h_1 r_{it} \times P_i + X_{it} + \zeta_{it}, \text{ with controls } X_{it} \in \left\{ \left\{ VA^{down}_{i,t-s} \right\}^p_{s=1}, \left\{ P_i \cdot VA^{down}_{i,t-s} \right\}^p_{s=1}, \mu_i, \mu_t \right\}. \tag{30}$$

In the set of regressions, value-added in each quarter $t + h$ is regressed on the estimated value-added shock in quarter t, as well as controls X_{it}. Recall P_i is the dummy for whether sector i has a dominant AR(2) component (i.e., the PACF coefficient $|\phi_{i,2}|$ is above the median among all downstream sectors). The estimate α^h_0 represents the impulse response function at horizon h for downstream sector i's value-added when the dummy P_i is turned off, and the estimate $\alpha^h_0 + \alpha^h_1$ represents the impulse response function when the dummy is turned on.

All impulse-response figures display 90% percentile confidence intervals (as in Ramey, 2016) from a cluster bootstrap with 2000 replications. Because the innovations entering the local projections are estimated in the first stage (29), each bootstrap replication re-estimates the first stage, so that this estimation uncertainty propagates into the intervals.[19]

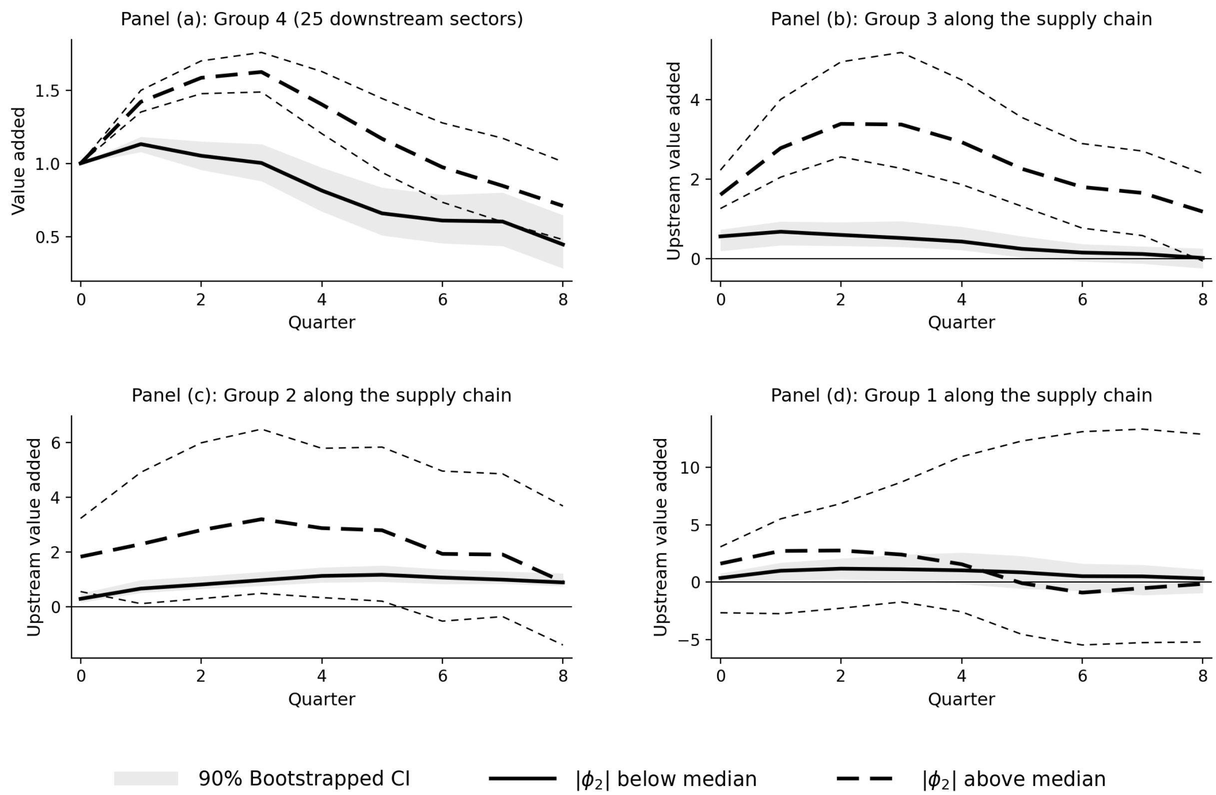
The top-left panel (a) in Figure 5 reports the coefficients α^h_0 (solid line) and $\alpha^h_0 + \alpha^h_1$ (dashed line), displaying the estimates from equation (30) of the impact of a demand shock from downstream sectors in quarter t on their own value-added share in quarter $t+h$ for a 2-year (8 quarters) horizon. Panel (a) in Figure 5 shows hump-shaped, mean-reverting impulse response functions, especially those sectors with relatively large AR(2) components. This is consistent with our model specification of AR(2) shocks, which implies hump-shaped response in downstream value-added following its own innovation. In terms of magnitude, for those sectors with large AR(2) components, a one standard deviation innovation in value-added to a downstream sector raises its own concurrent value-added by one unit by our construction. The expected value added increases over time, peaking in the third quarter at 1.6 standard deviations above the mean. The impact starts to decline after three quarters, but the impact is persistent and only decays to about half of the initial impact after two years. For those sectors with small AR(2) components, the hump-shaped response is present but significantly less pronounced. The impact of the shock peaks in

[19] Each replication resamples the 25 downstream sectors with replacement (each sector's full series as one cluster), re-estimates the first-stage slopes, and applies them to the full original panel, re-absorbing sector and quarter fixed effects, to reconstruct downstream innovations. For the projections of upstream suppliers' responses, introduced later in this section, the input-output network is held fixed, supplier-level innovation regressors are rebuilt from these reconstructed innovations, and suppliers are separately resampled with replacement (each supplier's full series as one cluster) to re-estimate the response equations. Across 2000 simulated panels from the estimated one-innovation process, empirical coverage rates for nominal 90% intervals average 87% across horizons and groups.

one quarter but is nevertheless persistent with a half-life of about two years.

Appendix Figure A.3 further confirms the presence of a pronounced hump-shaped impulse response function when all downstream sectors are grouped together.

Figure 5. Cumulative impulse response along the supply chain



Impulse-Response Along the Supply Chain We now use the innovations in downstream value-added to estimate the impulse-responses of upstream sectors along each supply chain. Since each supply chain involves many sectors, we group suppliers based on their supply chain positions. Specifically, for each downstream consumer good j, we classify input suppliers (indexed by i) into three groups according to terciles of our upstreamness/production delay measure ξ_{ij}. To maintain consistency with our four-sector vertical supply chain example in Section 3.2.2, we number groups in decreasing order of upstreamness: the most upstream suppliers are assigned to group $g = 1$, while suppliers closest to the downstream sector belong to group $g = 3$. We refer to the 25 downstream sectors themselves as group $g = 4$. Note that since the same sector can supply multiple downstream goods, it may occupy different positions and thus belong to different groups across various supply chains.

We estimate the impulse-response of sectors along the supply chain to innovations in downstream value-added using the following specification:

$$VA_{i,t+h}^{up,g} - VA_{i,t-1}^{up,g} = \beta^{h,g} \sum_{j\,:\,i \in \mathcal{U}_j} w_{ij} r_{jt} + \delta^{h,g} \sum_{j\,:\,i \in \mathcal{U}_j} w_{ij} r_{jt} \times P_j + Z_{it}^{g} + \epsilon_{it}, \tag{31}$$

$$\text{with controls} \quad Z_{it}^{g} \in \left\{ \left\{ \sum_{j\,:\,i \in \mathcal{U}_j} w_{ij} VA_{j,t-s}^{down}, \sum_{j\,:\,i \in \mathcal{U}_j} w_{ij} VA_{j,t-s}^{down} \times P_j,\, VA_{i,t-s}^{up,g} \right\}_{s=1}^{p}, \mu_i, \mu_t \right\}.$$

That is, we project the revenue of each sector along the supply chain of downstream consumption good j onto value-added innovations r_{jt}, estimating group-specific coefficients $\beta^{h,g}$ and $\delta^{h,g}$ for each time h after the shock.[20] Because an upstream sector i may supply multiple downstream goods j, we aggregate the downstream innovations using weights proportional to each sector's input usage intensity: $w_{ij} = L_{ij} \times \frac{\text{std}(VA_{j,t})/\alpha_j}{\text{std}(VA_{i,t})/\alpha_i}$. The term L_{ij} is computed from equation (11), representing the importance of supplier i for producing good j. The ratio of standard deviations and value-added intensities adjusts for the fact that both the downstream innovations r_{jt} and the upstream outcome variables are standardized in the regression.

We plot the coefficients $\beta^{h,g}$ and $\beta^{h,g} + \delta^{h,g}$ across time h for different supply chain groups $g \in \{3, 2, 1\}$ in the remaining panels (b)–(d) of Figure 5, with 90% confidence interval obtained by the bootstrap procedure described in footnote 19. The estimate $\beta^{h,g}$, shown in the solid lines, represents the impulse response functions at horizon h for upstream group g's value-added when they situate along the supply chain for a downstream sector j with a relatively small hump-shaped AR(2) component (i.e., with the dummy P_j turned off). The estimate $\beta^{h,g} + \delta^{h,g}$, shown in the dashed lines, represents the impulse response functions at horizon h for upstream group g's value-added when they situate along the supply chain for a downstream sector j with a relatively large hump-shaped AR(2) component (i.e., with the dummy P_j turned on).

The heterogeneity in the estimated impulse responses provides a sharp test of the theory, and the estimates are consistent with all four predictions. First, downstream sectors where AR(2) shocks dominate (group $g = 4$, dashed line in panel a) exhibit a distinctly hump-shaped response of their own value-added. Second, the responses along the supply chain are also hump-shaped and initially amplified as shocks travel upstream: the initial response grows from the downstream sectors (panel a) to group $g = 3$ (panel b), and further to group $g = 2$ (panel c). Third, the amplification diminishes further upstream (group $g = 1$, panel d). Fourth—and most importantly—both the downstream hump shape and the immediate upstream amplification are markedly stronger where AR(2) shocks dominate and weak or absent where AR(1) shocks dominate. The gap between the dashed and solid lines widens and then narrows both over the response horizon and

[20] As a robustness check, in Appendix D.9 we replace the two lags of upstream value-added in the control variables by $\Delta VA_{i,t-1}^{up,g}$. This is a long-difference specification that could suppress small-sample bias (Jordà and Taylor, 2025).

across positions along the chain—opening from the downstream sectors toward intermediate suppliers and closing for the most upstream group—as the theory predicts. Because P_i is built from the lag-2 dynamics of the same series, the own-response contrast in panel (a) partly reflects the sorting itself; the discriminating evidence is the upstream pattern in panels (b)–(d), which the sorting statistic does not mechanically generate.

Alternative channels—congestion, capacity strain, backlog accumulation, cyclical inventory adjustment—can also generate delayed upstream responses. As commonly modeled, however, these channels respond to the persistence of demand rather than to the shape of its expected path, and so provide no reason for amplification to appear only after shocks to AR(2)-dominant downstream sectors while remaining muted after equally persistent monotone-decay shocks. The differential pattern documented above is therefore the discriminating evidence for the time-profile channel—for its presence, not for the absence of the alternatives.

4.4 Quantitative Implications

In this section we estimate the demand shock process from the data, and we demonstrate the quantitative importance of hump-shaped shocks in a time-to-build environment for explaining sectoral fluctuations along the supply chain.

As we do not separately observe the two types of shocks, we estimate the one-innovation representation of Section 3.3 directly. By Proposition 3, this representation nests the two-shock model when agents forecast demand by Kalman filtering: observed demand in that economy follows exactly this process, with the Kalman gain as the weight matrix $\boldsymbol{D}$. The observables follow the linear state-space system

$$\widetilde{\boldsymbol{\theta}}_t = \rho\widetilde{\boldsymbol{\theta}}_{t-1} + \boldsymbol{x}_t + (\boldsymbol{I} - \boldsymbol{D})\,\boldsymbol{v}_t, \qquad \boldsymbol{x}_t = \rho\boldsymbol{x}_{t-1} + \boldsymbol{D}\boldsymbol{v}_t, \qquad \boldsymbol{v}_t \sim N\left(\boldsymbol{0}, \boldsymbol{\Sigma}_v\right), \tag{32}$$

and we estimate $\{\rho, \boldsymbol{D}, \boldsymbol{\Sigma}_v\}$ by Gaussian maximum likelihood on the panel of realized downstream value added, via the prediction-error decomposition.

For parsimony and ease of interpretation, in the baseline specification we restrict the weight matrix to be diagonal, $\boldsymbol{D} = \mathrm{diag}\,(\delta_1, \ldots, \delta_N)$, so that a single weight $\delta_i \in [0,1]$ splits sector i's innovation between the two time profiles. The restriction does not shut down cross-sector dependence, which remains through the covariance matrix of the innovations. For the innovation covariance we impose a one-factor structure $\boldsymbol{\Sigma}_v = \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \mathrm{diag}\,(\boldsymbol{\chi}^2)$, a common demand factor plus sector-specific components. These restrictions are for parsimony; as we show below, the results are qualitatively insensitive to them.

Identification Because the Gaussian likelihood depends on the data only through the second moments of observed demand, identification is equivalent to invertibility of the mapping from

$\{\rho, \boldsymbol{D}, \boldsymbol{\Sigma}_v\}$ to the autocovariance function of (32). Proposition A.2 in the Appendix establishes invertibility for a general weight matrix and a general innovation covariance; we next describe which moments are informative about each parameter.

Under the diagonal-$\boldsymbol{D}$ specification that we impose in the baseline, the mapping from moments to parameters is especially transparent: each sector's observable demand is an ARMA(2,1) whose autoregressive polynomial is common across sectors while the moving-average root carries the sector-specific weight. Accordingly: (i) ρ is pinned down by the pooled geometric decay of autocovariance at longer lags; (ii) given ρ, the weight δ_i is identified from the short-lag autocorrelations—most sharply the lag-2 partial autocorrelation, which is monotone in δ_i;[21] and (iii) the innovation covariance $\boldsymbol{\Sigma}_v$ is identified by the contemporaneous cross-sector covariance of one-step-ahead forecast errors, which carry no additional information about the dynamic parameters.

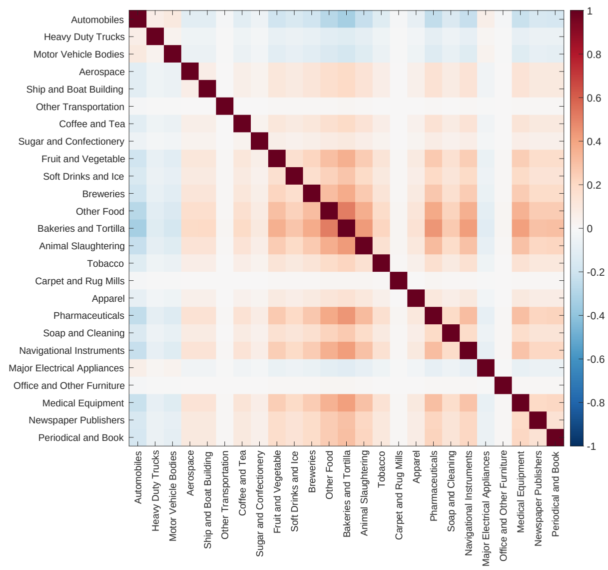
Estimation Results The details of MLE estimation are documented in Appendix C. We find $\rho = 0.829$ (standard error 0.007): demand is highly persistent. Figure 6 visualizes the correlation of cross-sector v shocks; we plot correlations rather than the covariance matrix $\boldsymbol{\Sigma}_v$ because differences in variance across sectors would otherwise dominate the scale of the figure. The figure shows that the demand shocks are not too correlated across sectors. Table 2 reports the estimated hump-share weights δ_i. A higher δ_i indicates that hump-shaped AR(2) shocks are the more important driver of a downstream sector's demand variation. The weights display wide variation across the 25 downstream sectors, from 0.11 for "Sugar and Confectionery Products" and 0.16 for "Major Electrical Household Appliances" to essentially one for "Navigational/Measuring/Electromedical/Control Instruments". The sectors most exposed to hump-shaped shocks are "Navigational/Measuring/Electromedical/Control Instruments", "Apparel", "Medical Equipment and Supplies", "Aerospace Products and Parts", "Periodical, Book, and Other Publishers", and "Soft Drinks and Ice", with δ_i between 0.82 and 1.00. Individual weights carry standard errors with median approximately 0.09; the analysis, however, uses only cross-sectional comparisons—the ranking of sectors and the median split P_i—which are far less sensitive to noise in any single weight.

Our baseline specification assumes diagonal $\boldsymbol{D}$ with a one-factor $\boldsymbol{\Sigma}_v$. We have explored alternative specifications to relax these restrictions, at a substantial cost in parameter counts: with

[21] In fact, each sector's first three autocovariances alone exactly determine $(\rho, \delta_i, [\boldsymbol{\Sigma}_v]_{ii})$: with $b_i \equiv \frac{\delta_i(1-\rho^2)\left[(1-\rho^2)+\delta_i\rho^2\right]}{(1-\rho^2)^2+2\delta_i\rho^2(1-\rho^2)+\delta_i^2\rho^2(1+\rho^2)}$ as in footnote 18, the autocorrelations of sector i's demand at lag k satisfy $r_{i,k} = \rho^k(1+kb_i)$ for $k \geq 1$. Eliminating ρ from $r_{i,1} = \rho(1+b_i)$ and $r_{i,2} = \rho^2(1+2b_i)$ yields a quadratic equation in b_i: $r_{i,2}(1+b_i)^2 = r_{i,1}^2(1+2b_i)$. Its two roots have opposite signs, so the requirement $b_i \geq 0$—which holds for every $\rho, \delta_i \in (0,1)$—selects a unique solution, and $\rho = r_{i,1}/(1+b_i)$ follows. Given ρ, the weight δ_i follows from b_i, which is monotone in δ_i, and $[\boldsymbol{\Sigma}_v]_{ii}$ from the variance $\Gamma_{i,0}$.

25 downstream sectors, the baseline has 76 parameters; estimating an unrestricted Cholesky $\mathbf{\Sigma}_v$ raises the parameter count by 275, and estimating an unrestricted weight matrix $\boldsymbol{D}$ raises the parameter count by 600. When both $\mathbf{\Sigma}_v$ and $\boldsymbol{D}$ are unrestricted, there are 951 parameters, nearly one for every five sector-quarter observations. Both Akaike and Bayesian information criteria (AIC and BIC), which trade fit against these parameter counts, favor the baseline; more importantly, the objects entering the quantitative exercises below are insensitive to these choices.[22] We therefore adopt the diagonal-$\boldsymbol{D}$, one-factor specification as the baseline on parsimony grounds.

Figure 6. Correlation matrices for shocks



Validation of the Structural Estimates against the Reduced-Form PACF Coefficients Our MLE procedure processes all second moments at once, so its estimates alone do not reveal which moments pin down which parameter. We perform the following validation check to illustrate the moment-to-parameter assignment directly. Specifically, the model implies a precise, decreasing relationship linking a sector's weight δ_i on the AR(2) component to the sector's lag-2 partial autocorrelation. Figure 7 scatterplots the lag-2 partial autocorrelation against our estimated $\hat{\delta}_i$ for each sector, with the blue curve showing the model-implied relationship based on the estimated

[22]The baseline has AIC and BIC scores of -409 and -162, respectively, versus -363 and 781 under the unrestricted covariance $\mathbf{\Sigma}_v$. The unrestricted weight matrix $\boldsymbol{D}$ paired with the one-factor covariance attains a lower AIC (-509) than the baseline but is overwhelmingly rejected by BIC (1693). The specification with both matrices unrestricted also yields worse AIC and BIC scores (-400 and 2698) compared to the baseline. All specifications yield similar estimates for the persistent parameter ρ as well as sectoral loadings on monotone and hump-shaped shocks: under the unrestricted covariance, $\rho = 0.817$ instead of 0.829 in the baseline, and the sector weights (δ_i) correlate 0.98 across specifications. Under the unrestricted weight matrix $\boldsymbol{D}$ specification, $\rho = 0.796$ with a one-factor covariance and 0.778 with an unrestricted covariance, and the diagonal of the estimated $\boldsymbol{D}$ correlates 0.99 and 0.97 with the baseline δ_i estimates across the two variants.

Table 2. Estimation results

$\rho = 0.829$ (s.e. 0.007)		δ_i
Transportation	Automobiles and Light Duty Motor Vehicles	0.30 (0.09)
	Heavy Duty Trucks	0.36 (0.10)
	Ship and Boat Building	0.39 (0.09)
	Motor Vehicle Bodies and Trailers	0.51 (0.07)
	Other Transportation Equipment	0.81 (0.10)
	Aerospace Products and Parts	0.83 (0.09)
Food Products	Sugar and Confectionery Products	0.11 (0.06)
	Coffee and Tea	0.17 (0.06)
	Fruit and Vegetable Preserving and Specialty Foods	0.28 (0.08)
	Breweries	0.61 (0.10)
	Other Food Except Coffee and Tea	0.64 (0.09)
	Bakeries and Tortilla	0.69 (0.10)
	Animal Slaughtering and Processing	0.79 (0.09)
	Soft Drinks and Ice	0.82 (0.10)
Other Final Goods	Major Electrical Household Appliances	0.16 (0.07)
	Office And Other Furniture	0.37 (0.10)
	Carpet and Rug Mills	0.39 (0.09)
	Tobacco	0.44 (0.07)
	Soap, Cleaning Compounds, and Toilet Preparation	0.66 (0.11)
	Newspaper Publishers	0.70 (0.11)
	Pharmaceuticals and Medicines	0.73 (0.09)
	Periodical, Book, and Other Publishers	0.82 (0.11)
	Medical Equipment and Supplies	0.85 (0.09)
	Apparel	0.92 (0.08)
	Navigational/Measuring/Electromedical/Control Instruments	1.00 (0.10)

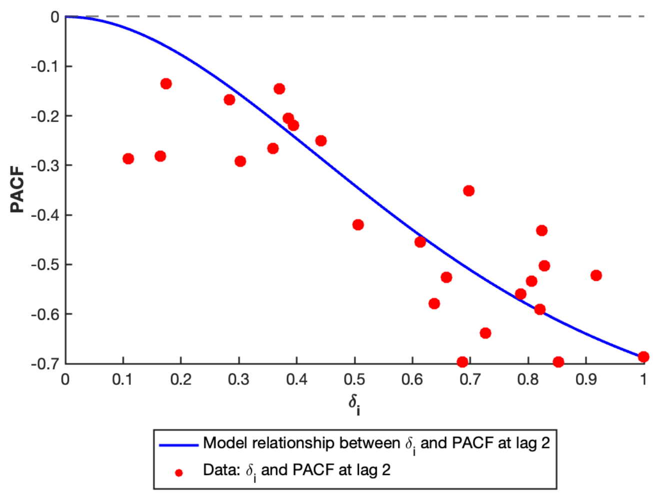
persistence $\hat{\rho}$. The scatterplot shows a strong, decreasing relationship that aligns very well with the model-implied curve: sectors that load more on AR(2) shocks have more negative lag-2 partial autocorrelation.

Inverting the curve at each sector's sample partial autocorrelation makes the comparison quantitative: the resulting one-moment estimates of δ_i, which use no information about the innovation covariance, broadly reproduce the likelihood estimates (correlation 0.84; the implied classification of sectors into AR(2)- versus AR(1)-dominant agrees for 92% of sectors). The likelihood weights all lags and pools across sectors; MLE thus adds efficiency, not information from another source.

The impulse-response exercises in Section 4.3 are performed by sorting downstream sectors into those with above- and below-median magnitudes of the PACF lag 2 coefficients. In Appendix D.4, we show the estimated impulse response functions are robust to sorting sectors by δ_i.

Variance Decompositions The source of the bullwhip effect—that downstream demand shocks create amplified volatility in upstream sectors—is the hump-shaped AR(2) component of consumer demand. How quantitatively important are hump-shaped shocks for supply chain volatility as a whole? Under the one-innovation representation, the demand innovation of downstream good i splits into a hump-shaped component, with weight δ_i, and a monotone-decay component,

Figure 7. Estimated weights (δ_i's) on hump-shaped component and lag-2 partial autocorrelations

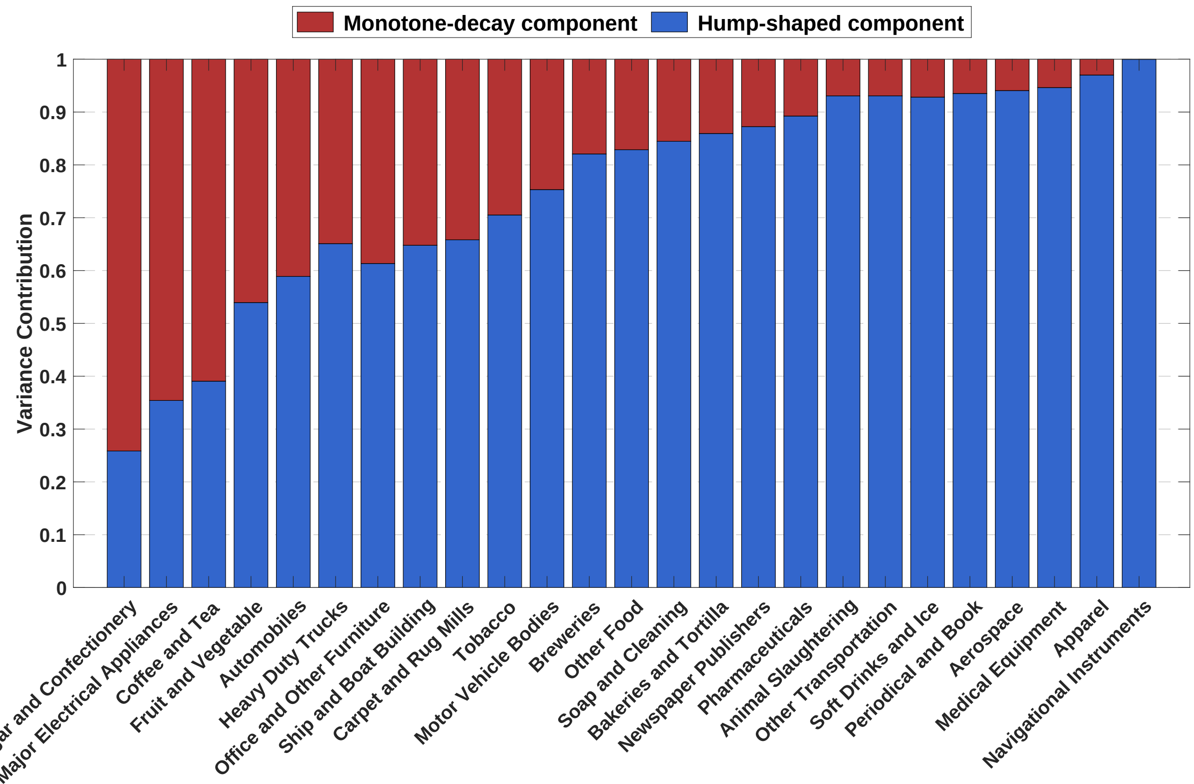


with weight $1 - \delta_i$. For each upstream sector j along the supply chain of downstream good i, we compute the share of sector j's volatility attributable to each component of demand volatility in good i, $\frac{\left[\boldsymbol{G}^x_\infty \boldsymbol{D} \boldsymbol{\Sigma}^{(i)}_v [\boldsymbol{G}^\epsilon_\infty (\boldsymbol{I}-\boldsymbol{D})+\boldsymbol{G}^x_\infty \boldsymbol{D}]'\right]_{jj}}{\left[[\boldsymbol{G}^\epsilon_\infty (\boldsymbol{I}-\boldsymbol{D})+\boldsymbol{G}^x_\infty \boldsymbol{D}]\boldsymbol{\Sigma}^{(i)}_v [\boldsymbol{G}^\epsilon_\infty (\boldsymbol{I}-\boldsymbol{D})+\boldsymbol{G}^x_\infty \boldsymbol{D}]'\right]_{jj}}$ and $\frac{\left[\boldsymbol{G}^\epsilon_\infty (\boldsymbol{I}-\boldsymbol{D}) \boldsymbol{\Sigma}^{(i)}_v [\boldsymbol{G}^\epsilon_\infty (\boldsymbol{I}-\boldsymbol{D})+\boldsymbol{G}^x_\infty \boldsymbol{D}]'\right]_{jj}}{\left[[\boldsymbol{G}^\epsilon_\infty (\boldsymbol{I}-\boldsymbol{D})+\boldsymbol{G}^x_\infty \boldsymbol{D}]\boldsymbol{\Sigma}^{(i)}_v [\boldsymbol{G}^\epsilon_\infty (\boldsymbol{I}-\boldsymbol{D})+\boldsymbol{G}^x_\infty \boldsymbol{D}]'\right]_{jj}}$, where $\boldsymbol{\Sigma}^{(i)}_v$ denotes $Var_t\left(\widetilde{\boldsymbol{\theta}}_{t+1}\right)$ with all entries other than the (i,i)-th element set to zero. The former (latter) is the share of sector j's volatility attributable to the hump-shaped (monotone-decay) component of good i's demand innovation, where each component is credited with its own variance plus half of their shared cross-term, so the two shares sum to one. Then, for each downstream sector i, we calculate the average share of volatility across sectors along i's supply chain that is attributable to each of the two components.

Figure 8 shows the decomposition results. Each bar represents a downstream sector and has two portions that add up to one. The blue portion captures the supply chain volatility associated with the downstream sector's hump-shaped shocks. The red portion captures the supply chain volatility associated with the downstream sector's monotone-decay shocks. The downstream sectors are arranged in increasing order of their $\hat{\delta}_i$, meaning that hump-shaped shocks tend to play a larger role in sectors towards the right-hand side.

Across supply chains in US manufacturing, hump-shaped AR(2) shocks account for 75.44% of volatility on average. The fraction ranges from 25.85% on the lower end (for the supply chain associated with "Sugar and Confectionery Products") to 99.98% on the higher end (for the supply chain associated with "Navigational/Measuring/Electromedical/Control Instruments").

Figure 8. Variance decompositions along supply chains



Counterfactuals We now compute several counterfactuals based on the estimated model to demonstrate the importance of hump-shaped shocks and time-to-build dynamics in generating supply chain volatility. Table 3 reports the results by downstream group. The delay columns report route-weighted average supply chain delays based on ξ_{ij} of equation (25), evaluated at the estimated persistence $\hat{\rho} = 0.829$.

Our first two counterfactuals ask how much of the volatility that supply chains inherit from downstream demand is due to the composition of demand shocks rather than to their level. For these we hold the level of demand volatility fixed and vary the composition, leaving the network and the delays unchanged.

Specifically, for each group we first isolate the volatility a supply chain inherits from it: setting the entries of $\mathrm{Var}_t\left(\widetilde{\boldsymbol{\theta}}_{t+1}\right)$ outside the group to zero and applying equation (28) at the estimated weights $\boldsymbol{D} = \mathrm{diag}\left(\hat{\delta}_i\right)$ gives $\widehat{\mathrm{Var}}_t\left(\widetilde{\gamma}_{j,t+1}\right)$, the volatility of sector j attributable to demand volatility in that group. We then recompute it with $\boldsymbol{D} = \boldsymbol{I}$, so that the group's innovations are entirely hump-shaped, and with $\boldsymbol{D} = \boldsymbol{0}$, so that they are entirely monotone-decay. In both exercises, only $\boldsymbol{D}$ changes: the network and the delays enter through $\boldsymbol{G}^{\epsilon}_{\infty}$ and $\boldsymbol{G}^{x}_{\infty}$, and the level of demand volatility through $\mathrm{Var}_t\left(\widetilde{\boldsymbol{\theta}}_{t+1}\right)$. Columns (1) and (2) report the resulting percentage changes relative to $\widehat{\mathrm{Var}}_t\left(\widetilde{\gamma}_{j,t+1}\right)$, averaged over the sectors j along each group's supply chains.

Row (d) of the table activates all twenty-five innovations jointly, including their cross-group covariances; thus it is not an average of rows (a)–(c) and can fall outside their range, as in column (2).

Hump-shaped shocks amplify supply chain volatility substantially. Along the supply chains of the transportation group, volatility rises by 241.6%; when all downstream demand shocks become hump-shaped, overall supply chain volatility rises by 155.5% (row d). Making the demand innovations entirely monotone-decay instead lowers volatility by 62 to 70% across groups. The ordering across groups is likely driven by delays rather than by composition: amplification is larger for transportation than for food products (98.4%) or other final goods (75.5%), and columns (4) and (5) show that transportation has the longest supply chain delays, both on average and among the five longest links.

A third exercise changes the model rather than the shocks. Column (3) compares the dynamic model with a static benchmark, again holding demand volatility constant. Specifically, we use the formula $\mathrm{Var}_t^{Static}\left(\widetilde{\boldsymbol{\gamma}}_{t+1}\right) = \left(\boldsymbol{I}-\boldsymbol{\Omega}\right)^{-1}\mathrm{Var}_t\left(\widetilde{\boldsymbol{\theta}}_{t+1}\right)\left(\boldsymbol{I}-\boldsymbol{\Omega}'\right)^{-1}$, which follows from the relationship between sectoral revenue $\boldsymbol{\gamma}$ and demand $\boldsymbol{\theta}$ in Acemoglu et al. (2012), given by $\boldsymbol{\gamma} = \left(\boldsymbol{I}-\boldsymbol{\Omega}\right)^{-1}\boldsymbol{\theta}$. We find that a static model without time-to-build dynamics significantly understates the supply chain volatility attributable to downstream demand shocks on average—by 22.2% overall, and by close to 44% for other final goods.[23]

In Appendix Table A.3, we report the outcomes for the same set of counterfactual exercises, except that we examine each downstream industry individually rather than in groups.

Table 3. Counterfactual change in supply chain volatility based on the nature of downstream demand shocks

	% Change in supply chain volatility			Delays along supply chain	
	Only AR(2) shocks	Only AR(1) shocks	Static model	Average	Top 5
Downstream Group	(1)	(2)	(3)	(4)	(5)
(a) Transportation	241.62	-63.47	-6.95	2.33	3.61
(b) Food Products	98.40	-66.11	-35.83	2.11	3.07
(c) Other Final Goods	75.53	-70.49	-44.18	2.12	2.74
(d) All Sectors	155.49	-61.87	-22.19	2.17	3.05

[23] The group comprises "Apparel", "Tobacco", "Carpet and Rug Mills", "Newspaper Publishers", "Pharmaceuticals and Medicines", "Soap, Cleaning Compounds, and Toilet Preparation", "Navigational/Measuring/Electromedical/Control Instruments", "Major Electrical Household Appliances", "Office And Other Furniture", "Medical Equipment and Supplies", and "Periodical, Book, and Other Publishers".

5 Conclusion

We develop a framework to analyze dynamic interactions over production networks, with two main ingredients: sectoral demand shocks and heterogeneous time-to-build delays. Each producer's input choices depend on expectations of future demand at all time horizons, arising from both exogenous consumer demand shocks and endogenous responses from other producers along the supply chain, each potentially facing distinct time-to-build delays. The central result is a mapping from the temporal profile of downstream demand into the cross-section of upstream responses: shocks that decay monotonically are dampened as they propagate upstream, while hump-shaped shocks—under which higher demand today signals higher demand ahead—are amplified, generating the bullwhip effect. We characterize the sectoral impulse response in closed form over primitives and show how equilibrium input decisions depend on the entire temporal structure of shocks and supply-chain linkages. Empirically, hump-shaped, amplified upstream responses appear in the supply chains of downstream sectors whose demand carries momentum, and not in those with monotone-decay demand; quantitatively, hump-shaped shocks account for 75 percent of supply chain volatility in U.S. manufacturing.

References

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

Alfaro, Laura, Pol Antràs, Davin Chor, and Paola Conconi, "Internalizing Global Value Chains: A Firm-Level Analysis," *Journal of Political Economy*, 2019, *127* (2), 508–559.

Antràs, Pol and Vitalii Tubdenov, "Measuring the Average Period of Production," *Working Paper*, 2025.

__ , Davin Chor, Thibault Fally, and Russell Hillberry, "Measuring the Upstreamness of Production and Trade Flows," *American Economic Review*, 2012, *102* (3), 412–416.

Baqaee, David Rezza and Emmanuel Farhi, "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem," *Econometrica*, 2019, *87*, 1155–1203.

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

Bayraktar, Erhan and Christian Keller, "Path-dependent Hamilton–Jacobi equations in infinite dimensions," *Journal of Functional Analysis*, 2018, *275* (8), 2096–2161.

Beaudry, Paul and Franck Portier, "Stock Prices, News, and Economic Fluctuations," *American Economic Review*, 2006, *96* (4), 1293–1307.

Bigio, Saki and Jennifer La'O, "Distortions in Production Networks," *The Quarterly Journal of Economics*, November 2020, *135* (4), 2187–2253.

Bizzarri, Matteo, Marco Pangallo, and Francisco Queiros, "Production Networks, Time to Build, and Endogenous Oscillations," *Working paper*, 2025.

Blanchard, Olivier J., Jean-Paul L'Huillier, and Guido Lorenzoni, "News, Noise, and Fluctuations: An Empirical Exploration," *American Economic Review*, 2013.

Blinder, Alan S and Louis J Maccini, "Taking Stock: A Critical Assessment of Recent Research on Inventories," *Journal of Economic Perspectives*, 1991, *5* (1), 73–96.

Boerma, Job, Georgii Riabov, and Aleh Tsyvinski, "Making Memories: Memory Effects and Their Dynamics," *Working paper*, 2024.

Box, George EP, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung, *Time Series Analysis: Forecasting and Control*, John Wiley & Sons, 2015.

Carvalho, Vasco, "Aggregate Fluctuations and the Network Structure of Intersectoral Trade," *Working paper*, 2010.

Chen, Frank, Zvi Drezner, Jennifer K Ryan, and David Simchi-Levi, "Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information," *Management science*, 2000, *46* (3), 436–443.

Chen, Li and Hau L Lee, "Bullwhip effect measurement and its implications," *Operations research*, 2012, *60* (4), 771–784.

Elliott, Matthew, Benjamin Golub, and Matthew V Leduc, "Supply Network Formation and Fragility," *American Economic Review*, 2022, *112* (8), 2701–2747.

Farmer, Leland E, "The Discretization Filter: A Simple Way to Estimate Nonlinear State Space Models," *Quantitative Economics*, 2021, *12* (1), 41–76.

Ferrari, Alessandro, "Inventories, Demand Shocks Propagation, and Amplification in Supply Chains," *Working Paper*, 2025.

Foerster, Andrew T, Pierre-Daniel G Sarte, and Mark W Watson, "Sectoral Versus Aggregate Shocks: A Structural Factor Analysis of Industrial Production," *Journal of Political Economy*, 2011, *119* (1), 1–38.

Forrester, Jay W, "Industrial Dynamics: A Major Breakthrough for Decision Makers," *Harvard business review*, 1958, *36* (4), 37–66.

Gabaix, Xavier, "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 2011, *79* (3), 733–772.

Hamilton, James D, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica: Journal of the econometric society*, 1989, pp. 357–384.

Hannan, Edward J. and Manfred Deistler, *The Statistical Theory of Linear Systems*, New York: Wiley, 1988.

Humpherys, Jeffrey, Preston Redd, and Jeremy West, "A Fresh Look at the Kalman Filter," *SIAM review*, 2012, *54* (4), 801–823.

Ito, Kazufumi and Kaiqi Xiong, "Gaussian Filters for Nonlinear Filtering Problems," *IEEE transactions on automatic control*, 2000, *45* (5), 910–927.

Jones, Charles I., "Intermediate Goods and Weak Links in the Theory of Economic Development," *American Economic Journal: Macroeconomics*, April 2011, *3* (2), 1–28.

— , "Misallocation, Economic Growth, and Input-Output Economics," *Advances in Economics and Econometrics*, 2013, *2.*

Jordà, Òscar and Alan M Taylor, "Local Projections," *Journal of Economic Literature*, 2025, *63* (1), 59–110.

Kahn, James A, "Why Is Production More Volatile Than Sales? Theory and Evidence on the Stockout-Avoidance Motive for Inventory-Holding," *The Quarterly Journal of Economics*, 1992, *107* (2), 481–510.

Kalman, Rudolph E and Richard S Bucy, "New Results in Linear Filtering and Prediction Theory," *Journal of Basic Engineering*, 1961, *83* (1), 95–108.

Kalman, Rudolph Emil, "A New Approach to Linear Filtering and Prediction Problems," 1960.

Khan, Aubhik and Julia K Thomas, "Inventories and the Business Cycle: An Equilibrium Analysis of (S, s) Policies," *American Economic Review*, 2007, *97* (4), 1165–1188.

— **and** — , "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment dynamics," *Econometrica*, 2008, *76* (2), 395–436.

— **and** — , "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, 2013, *121* (6), 1055–1107.

Kopytov, Alexandr, Bineet Mishra, Kristoffer Nimark, and Mathieu Taschereau-Dumouchel, "Endogenous Production Networks under Supply Chain Uncertainty," *Econometrica*, 2024, *92* (5), 1621–1659.

Kydland, Finn E and Edward C Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica: Journal of the Econometric Society*, 1982, pp. 1345–1370.

Lee, Hau L, Venkata Padmanabhan, and Seungjin Whang, "Information Distortion in A Supply Chain: The Bullwhip Effect," *Management science*, 1997, *43* (4), 546–558.

— **, Vineet Padmanabhan, and Seungjin Whang**, "Comments on "Information distortion in a supply chain: The bullwhip effect"," *Management science*, 2004, *50* (12_supplement), 1887–1893.

Liu, Ernest, "Industrial Policies in Production Networks*," *The Quarterly Journal of Economics*, November 2019, *134* (4), 1883–1948.

— **and Aleh Tsyvinski**, "A Dynamic Model of Input–Output Networks," *Review of Economic Studies*, 2024, p. rdae012.

— **and Song Ma**, "Innovation Networks and R&D Allocation," *Working Paper*, 2024.

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

Lorenzoni, Guido, "A Theory of Demand Shocks," *American Economic Review*, 2009, *99* (5), 2050–2084.

— , "News and Aggregate Demand Shocks," *Annual Review of Economics*, 2011.

Nikolakoudis, George, "Incomplete Information in Production Networks," *Working Paper*, 2024.

Oberfield, Ezra, "A Theory of Input-Output Architecture," *Econometrica*, March 2018, *86*, 559–589.

Pellet, Thomas and Alireza Tahbaz-Salehi, "Rigid Production Networks," *Journal of Monetary Economics*, 2023, *137*, 86–102.

Ramey, Valerie A, "Macroeconomic shocks and their propagation," *Handbook of macroeconomics*, 2016, *2*, 71–162.

Schaal, Edouard and Mathieu Taschereau-Dumouchel, "Echoes and Delays: Time-to-Build in Production Networks," *Working Paper*, 2025.

Sterman, John D, "Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment," *Management science*, 1989, *35* (3), 321–339.

Taschereau-Dumouchel, Mathieu, "Cascades and Fluctuations in an Economy with an Endogenous Production Network," *Working Paper*, 2020.

Tian, Can, "Input-output linkages in Pigouvian industrial fluctuations," *Journal of Monetary Economics*, 2021, *117*, 1078–1095.

vom Lehn, Christian and Thomas Winberry, "The Investment Network, Sectoral Comovement, and the Changing US Business Cycle," *The Quarterly Journal of Economics*, 2022, *137* (1), 387–433.

Wen, Yi, "Input and Output Inventory Dynamics," *American Economic Journal: Macroeconomics*, 2011, *3* (4), 181–212.

Appendix

Table of Contents

A Proofs

A.1 Proof of Theorem 1

Proof. We iteratively substitute the left-hand side of (9) into the right-hand side to obtain

$$\begin{aligned}
\boldsymbol{\gamma}_t &= \boldsymbol{\theta}_t + \sum_{d=1}^{\infty} \beta^d \boldsymbol{\Omega}_d \mathbb{E}_t \left[\boldsymbol{\gamma}_{t+d}\right] \\
&= \boldsymbol{\theta}_t + \mathbb{E}_t \left[\sum_{d=1}^{\infty} \beta^d \boldsymbol{\Omega}_d \left(\boldsymbol{\theta}_{t+d} + \sum_{d'=1}^{\infty} \beta^{d'} \boldsymbol{\Omega}_{d'} \boldsymbol{\gamma}_{t+d+d'} \right) \right] \\
&\vdots \\
&= \boldsymbol{\theta}_t + \beta \boldsymbol{\Omega}_1 \mathbb{E}_t \left[\boldsymbol{\theta}_{t+1}\right] + \beta^2 \left(\boldsymbol{\Omega}_1^2 + \boldsymbol{\Omega}_2\right) \mathbb{E}_t \left[\boldsymbol{\theta}_{t+2}\right] \\
&\quad + \beta^3 \left(\boldsymbol{\Omega}_1^3 + \boldsymbol{\Omega}_2 \boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_1 \boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_3\right) \mathbb{E}_t \left[\boldsymbol{\theta}_{t+3}\right] + \cdots \\
&= \boldsymbol{\theta}_t + \sum_{s=1}^{\infty} \boldsymbol{G}_s \mathbb{E}_t \left[\boldsymbol{\theta}_{t+s}\right], \qquad \text{(A.1)}
\end{aligned}$$

where $\boldsymbol{G}_s \equiv \beta^s \sum_{\phi \in \Phi_s} \prod_{\phi_j \in \phi} \boldsymbol{\Omega}_{\phi_j}$. Note that the proof of Corollary 1 below implies that $\boldsymbol{I} + \sum_{s=1}^{\infty} \boldsymbol{G}_s = \left(\boldsymbol{I} - \sum_{d=1}^{\infty} \beta^d \boldsymbol{\Omega}_d\right)^{-1}$. Both this invertibility and the convergence of the series rest on the spectral radius of $\boldsymbol{\Omega}$ being strictly below one: each column of $\boldsymbol{\Omega}$ sums to $1 - \alpha_i \leq 1$, and by the assumption of footnote 2 every sector's production eventually involves labor along the supply chain, so for some finite K every column of $\boldsymbol{\Omega}^K$ sums to strictly less than one and $r\left(\boldsymbol{\Omega}\right) \leq \left\|\boldsymbol{\Omega}^K\right\|_1^{1/K} < 1$. Since $0 \leq \left[\sum_{d=1}^{\infty} \beta^d \boldsymbol{\Omega}_d\right]_{ji} \leq \boldsymbol{\Omega}_{ji}$ entrywise and the spectral radius of a nonnegative matrix is monotone under entrywise domination, $r\left(\sum_{d=1}^{\infty} \beta^d \boldsymbol{\Omega}_d\right) < 1$. Together with the fact that demand is mean-reverting in the long-run (i.e., $\lim_{s \to \infty} \mathbb{E}_t \left[\boldsymbol{\theta}_{t+s}\right] = \bar{\boldsymbol{\theta}}$ exists), this implies that the infinite series in (A.1) always converges. Finally, after n rounds the substitution leaves the remainder $\sum_{(d_1,\ldots,d_n)} \beta^{d_1+\cdots+d_n} \boldsymbol{\Omega}_{d_1} \cdots \boldsymbol{\Omega}_{d_n} \mathbb{E}_t \left[\boldsymbol{\gamma}_{t+d_1+\cdots+d_n}\right]$; because expected revenues are bounded along the equilibrium path and the spectral radius above is below one, the remainder vanishes as $n \to \infty$, so (A.1) is the unique bounded solution. □

A.2 Proof of Corollary 1

Proof. Let $\boldsymbol{G}_\infty \equiv \boldsymbol{I} + \sum_{s=1}^{\infty} \boldsymbol{G}_s$. To show $\boldsymbol{G}_\infty = \left(\boldsymbol{I} - \sum_{d=1}^{\infty} \beta^d \boldsymbol{\Omega}_d\right)^{-1}$, note

$$\boldsymbol{G}_\infty = \boldsymbol{I} + \beta \boldsymbol{\Omega}_1 \boldsymbol{G}_\infty + \beta^2 \boldsymbol{\Omega}_2 \boldsymbol{G}_\infty + \cdots = \boldsymbol{I} + \left(\sum_{d=1}^{\infty} \beta^d \boldsymbol{\Omega}_d\right) \boldsymbol{G}_\infty.$$

Rearranging yields the result. Invertibility of $\left(\boldsymbol{I} - \sum_{d=1}^{\infty} \beta^d \boldsymbol{\Omega}_d\right)$ was established in the proof of Theorem 1: the spectral radius of the sum is bounded by $r\left(\boldsymbol{\Omega}\right) < 1$. □

A.3 Proof of Proposition 1

Proof. By the law of iterated expectations, under (14) the response of expected demand to a current innovation satisfies $\boldsymbol{H}_0 = \boldsymbol{I}$ and $\boldsymbol{H}_s = \sum_{k=1}^{s} \boldsymbol{\Delta}_k \boldsymbol{H}_{s-k}$: the innovation moves demand at $t+s$ through each lag k of the process, via $\boldsymbol{\Delta}_k$ applied to the response at horizon $s-k$. Unrolling the recursion enumerates the lag sequences $\phi \in \Phi_s$ and yields $\boldsymbol{H}_s = \sum_{\phi \in \Phi_s} \prod_{\phi_j \in \phi} \boldsymbol{\Delta}_{\phi_j}$, with each product taken in the order of the sequence. Substituting $\partial \mathbb{E}_t \left[\boldsymbol{\theta}_{t+s}\right] / \partial \boldsymbol{e}_t = \boldsymbol{H}_s$ into Theorem 1 delivers $\partial \boldsymbol{\gamma}_t / \partial \boldsymbol{e}_t = \sum_{s=0}^{\infty} \boldsymbol{G}_s \boldsymbol{H}_s$ with $\boldsymbol{G}_0 = \boldsymbol{I}$, which is the expression in the proposition. □

A.4 Proof of Proposition 2

Proof. Substituting the labor allocation $\frac{\ell_{it}}{\bar{\ell}_t} = \frac{\alpha_i \gamma_{it}}{\sum_k \alpha_k \gamma_{kt}}$ and the consumption share $\frac{c_{it}}{y_{it}} = \frac{\theta_{it}}{\gamma_{it}}$ from (8) into the production function (2) yields (15). We obtain (16) by summing (5) across j and substituting out w_t using (6). Finally, (17) rearranges the input share $\frac{m_{ijt}}{y_{jt}}$ in (8) using $p_{jt} = \gamma_{jt}/y_{jt}$. □

A.5 Proof of Theorem 2

Proof. Equation (21) and Theorem 1 jointly imply that

$$\widetilde{\boldsymbol{\gamma}}_t = \widetilde{\boldsymbol{\theta}}_t + \sum_{s=1}^{\infty} \boldsymbol{G}_s \left(\rho^s \left(\widetilde{\boldsymbol{\theta}}_t - \widetilde{\boldsymbol{x}}_t \right) + (s+1)\, \rho^s \widetilde{\boldsymbol{x}}_t \right).$$

Collect the terms and write $\widetilde{\boldsymbol{\gamma}}_t = \boldsymbol{G}_{\infty}^{\epsilon} \left(\widetilde{\boldsymbol{\theta}}_t - \widetilde{\boldsymbol{x}}_t \right) + \boldsymbol{G}_{\infty}^{x} \widetilde{\boldsymbol{x}}_t$, and we will derive $\boldsymbol{G}_{\infty}^{\epsilon}$ and $\boldsymbol{G}_{\infty}^{x}$. The matrix $\boldsymbol{G}_{\infty}^{\epsilon}$ summarizes the impact of demand's monotone-decay component on revenue, and $\boldsymbol{G}_{\infty}^{x}$ captures the impact of the hump-shaped component. We derive these two components separately.

To derive $\boldsymbol{G}_{\infty}^{\epsilon}$, we can follow the same logic as in the proof of Corollary 1 and decompose the impact of demand's monotone-decay component recursively as the direct and indirect network effects, where each round of the network effect is summarized by $\boldsymbol{G}_{\infty}^{\epsilon}$, discounted to the present by both the discount rate (β) and also the rate at which the shock dissipates (ρ). Specifically, note

$$\begin{aligned}
\boldsymbol{G}_{\infty}^{\epsilon} &= \boldsymbol{I} + \rho\beta \boldsymbol{\Omega}_1 + (\rho\beta)^2 \left(\boldsymbol{\Omega}_1^2 + \boldsymbol{\Omega}_2 \right) + (\rho\beta)^3 \left(\boldsymbol{\Omega}_1^3 + \boldsymbol{\Omega}_1 \boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_2 \boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_3 \right) + \cdots \\
&= \boldsymbol{I} + \rho\beta \boldsymbol{\Omega}_1 \boldsymbol{G}_{\infty}^{\epsilon} + (\rho\beta)^2 \boldsymbol{\Omega}_2 \boldsymbol{G}_{\infty}^{\epsilon} + \cdots \\
&= \boldsymbol{I} + \left(\sum_{d=1}^{\infty} (\rho\beta)^d \boldsymbol{\Omega}_d \right) \boldsymbol{G}_{\infty}^{\epsilon},
\end{aligned}$$

and re-arranging yields that $\boldsymbol{G}_{\infty}^{\epsilon} = \left(\boldsymbol{I} - \sum_{d=1}^{\infty} (\rho\beta)^d \boldsymbol{\Omega}_d \right)^{-1}$.

It is slightly more involved to derive $\boldsymbol{G}_{\infty}^{x}$ because the impact of a hump-shaped shock does not decay monotonically (for $\rho > 1/2$); instead, the impact has an arithmetically increasing component as well as an exponentially decaying component. We can write $\boldsymbol{G}_{\infty}^{x}$ as

$$\boldsymbol{G}_{\infty}^{x}=\boldsymbol{I}+2\rho\beta\boldsymbol{\Omega}_1+3\left(\rho\beta\right)^2\left(\boldsymbol{\Omega}_1^2+\boldsymbol{\Omega}_2\right)+4\left(\rho\beta\right)^3\left(\boldsymbol{\Omega}_1^3+\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2+\boldsymbol{\Omega}_2\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_3\right)+\cdots$$

1. The first term $\boldsymbol{I}$ captures the impact of the hump-shaped component $\widetilde{\boldsymbol{x}}$ on time-t revenue through concurrent demand.

2. The second term $2\rho\beta\boldsymbol{\Omega}_1$ captures the impact of $\widetilde{\boldsymbol{x}}_t$ on time-t revenue through expected demand at $t+1$. The impact of $\widetilde{\boldsymbol{x}}$ on $t+1$ demand is captured by 2ρ. The impact of $t+1$ demand on time-t revenue goes through input-output linkages with one-period delay and is captured by $\boldsymbol{\Omega}_1$.

3. The third term $3\left(\rho\beta\right)^2\left(\boldsymbol{\Omega}_1^2+\boldsymbol{\Omega}_2\right)$ captures the impact of $\widetilde{\boldsymbol{x}}_t$ on time-t revenue through expected demand at $t+2$. The impact of $\widetilde{\boldsymbol{x}}$ on $t+2$ demand is captured by $3\rho^2$. The $t+2$ demand on time-t revenue goes through input-output linkages with two-period delay and is captured by $\boldsymbol{\Omega}_1^2+\boldsymbol{\Omega}_2$. The term $\boldsymbol{\Omega}_1^2$ captures the linkages from inputs supplied by the focal sector to the sector experiencing the demand shock through another intermediate producer, with two walks in total and one-period delay along each walk. The term $\boldsymbol{\Omega}_2$ captures the linkages from the focal sector directly to the sector experiencing the demand shock, with one walk of two-period delay.

Each successive term in $\boldsymbol{G}_{\infty}^{x}$ summarizes the impact of $\widetilde{\boldsymbol{x}}_t$ on time-t revenue through demand at a future time $t+s$, accounting for all possible direct and indirect network linkages with a delay of s periods in total. To derive $\boldsymbol{G}_{\infty}^{x}$ in closed-form, we re-write it recursively.

Define $\boldsymbol{Z}\equiv\sum_{d=1}^{\infty}d\left(\rho\beta\right)^d\boldsymbol{\Omega}_d$. We can then write $\boldsymbol{G}_{\infty}^{x}$ as

$$\begin{aligned}
\boldsymbol{G}_{\infty}^{x} &= \boldsymbol{I}+\rho\beta\boldsymbol{\Omega}_1\left(2\boldsymbol{I}+3\rho\beta\boldsymbol{\Omega}_1+4\left(\rho\beta\right)^2\left(\boldsymbol{\Omega}_1^2+\boldsymbol{\Omega}_2\right)+\cdots\right)\\
&\quad+\left(\rho\beta\right)^2\boldsymbol{\Omega}_2\left(3\boldsymbol{I}+4\rho\beta\boldsymbol{\Omega}_1+5\left(\rho\beta\right)^2\left(\boldsymbol{\Omega}_1^2+\boldsymbol{\Omega}_2\right)+\cdots\right)\\
&\quad+\left(\rho\beta\right)^3\boldsymbol{\Omega}_3\left(4\boldsymbol{I}+5\rho\beta\boldsymbol{\Omega}_1+6\left(\rho\beta\right)^2\left(\boldsymbol{\Omega}_1^2+\boldsymbol{\Omega}_2\right)+\cdots\right)\\
&= \boldsymbol{I}+\rho\beta\boldsymbol{\Omega}_1\left(\boldsymbol{G}_{\infty}^{\epsilon}+\boldsymbol{G}_{\infty}^{x}\right)+\left(\rho\beta\right)^2\boldsymbol{\Omega}_2\left(2\boldsymbol{G}_{\infty}^{\epsilon}+\boldsymbol{G}_{\infty}^{x}\right)+\left(\rho\beta\right)^3\boldsymbol{\Omega}_3\left(3\boldsymbol{G}_{\infty}^{\epsilon}+\boldsymbol{G}_{\infty}^{x}\right)+\cdots\\
&= \boldsymbol{I}+\boldsymbol{Z}\boldsymbol{G}_{\infty}^{\epsilon}+\sum_{d=1}^{\infty}\left(\rho\beta\right)^d\boldsymbol{\Omega}_d\boldsymbol{G}_{\infty}^{x}\\
&= \left(\boldsymbol{I}-\sum_{d=1}^{\infty}\left(\rho\beta\right)^d\boldsymbol{\Omega}_d\right)^{-1}\left(\boldsymbol{I}+\boldsymbol{Z}\boldsymbol{G}_{\infty}^{\epsilon}\right)\\
&= \boldsymbol{G}_{\infty}^{\epsilon}+\boldsymbol{G}_{\infty}^{\epsilon}\boldsymbol{Z}\boldsymbol{G}_{\infty}^{\epsilon}
\end{aligned}$$

To derive the impulse-response, equation (20) implies that

$$\partial\widetilde{\boldsymbol{\theta}}_{t+s}\Big/\partial\boldsymbol{\epsilon}_t=\rho^s,\qquad \partial\widetilde{\boldsymbol{x}}_{t+s}\Big/\partial\boldsymbol{\epsilon}_t=0,$$

$$\partial\widetilde{\boldsymbol{\theta}}_{t+s}\Big/\partial\boldsymbol{u}_t=(s+1)\,\rho^s,\quad\text{and}\quad\partial\widetilde{\boldsymbol{x}}_{t+s}\Big/\partial\boldsymbol{u}_t=\rho^s.$$

Hence

$$\frac{\partial\widetilde{\gamma}_{t+s}}{\partial\boldsymbol{\epsilon}_t}=\boldsymbol{G}^{\epsilon}_{\infty}\frac{\partial\widetilde{\boldsymbol{\theta}}_{t+s}}{\partial\boldsymbol{\epsilon}_t}=\rho^s\boldsymbol{G}^{\epsilon}_{\infty},$$

$$\begin{aligned}\frac{\partial\widetilde{\gamma}_{t+s}}{\partial\boldsymbol{u}_t} &= \boldsymbol{G}^{\epsilon}_{\infty}\frac{\partial\widetilde{\boldsymbol{\theta}}_{t+s}}{\partial\boldsymbol{u}_t}+\left(\boldsymbol{G}^{x}_{\infty}-\boldsymbol{G}^{\epsilon}_{\infty}\right)\frac{\partial\widetilde{\boldsymbol{x}}_{t+s}}{\partial\boldsymbol{u}_t}\\ &= (s+1)\,\rho^s\boldsymbol{G}^{\epsilon}_{\infty}+\rho^s\boldsymbol{G}^{\epsilon}_{\infty}\left(\sum_{d=1}^{\infty}d\,(\rho\beta)^d\,\boldsymbol{\Omega}_d\right)\boldsymbol{G}^{\epsilon}_{\infty}\\ &= s\rho^s\boldsymbol{G}^{\epsilon}_{\infty}+\rho^s\boldsymbol{G}^{x}_{\infty},\qquad\text{as desired.}\end{aligned}$$

□

A.6 Proof of Corollary 2

Proof. Equation (22) implies that

$$\begin{aligned}Var_t\left(\widetilde{\gamma}_{t+1}\right) &= Var_t\left(\boldsymbol{G}^{\epsilon}_{\infty}\left(\widetilde{\boldsymbol{\theta}}_{t+1}-\boldsymbol{x}_{t+1}\right)+\boldsymbol{G}^{x}_{\infty}\boldsymbol{x}_{t+1}\right)\\ &= Var_t\left(\boldsymbol{G}^{\epsilon}_{\infty}\left(\rho\widetilde{\boldsymbol{\theta}}_t+\boldsymbol{\epsilon}_{t+1}\right)+\boldsymbol{G}^{x}_{\infty}\boldsymbol{u}_{t+1}\right)\\ &= \boldsymbol{G}^{\epsilon}_{\infty}\boldsymbol{\Sigma}^{\epsilon}_t\boldsymbol{G}^{\epsilon\prime}_{\infty}+\boldsymbol{G}^{x}_{\infty}\boldsymbol{\Sigma}^{u}_t\boldsymbol{G}^{x\prime}_{\infty},\qquad\text{as desired.}\end{aligned}$$

□

A.7 Proof of Proposition 3

Suppose demand follows the two-shock process in equations (18) and (19) with Gaussian innovations and time-invariant covariance matrices, and agents observe only the history of realized demand $\left\{\widetilde{\boldsymbol{\theta}}_{t-s}\right\}_{s\geq 0}$. The conditional expectation operator is then defined as $\mathbb{E}_t\left[\cdot\right]\equiv\mathbb{E}\left[\cdot|\left\{\widetilde{\boldsymbol{\theta}}_{t-s}\right\}_{s\geq 0}\right]$. In this setting, agents, including the social planner, need to solve a filtering problem and forecast future demand based on the history of demand shocks. Under incomplete information, only $\widetilde{\boldsymbol{\theta}}_t$ is observable; $\boldsymbol{x}_t$, the hump-shaped component of current demand, is a hidden state. There is an additional endogenous state variable that captures the agents' belief $\hat{\boldsymbol{x}}_t\equiv\mathbb{E}_t\left[\boldsymbol{x}_t|\left\{\widetilde{\boldsymbol{\theta}}_{t-s}\right\}_{s\geq 0}\right]$ of the hidden state based on the history of demand up to time t. We refer to $\hat{\boldsymbol{x}}_t$ as the nowcast (see Kalman, 1960, Kalman and Bucy, 1961).

The demand shock at time t is defined as the unexpected innovation in demand:

$$\boldsymbol{v}_t \equiv \widetilde{\boldsymbol{\theta}}_t - \mathbb{E}_{t-1}\left[\widetilde{\boldsymbol{\theta}}_t\right] = \widetilde{\boldsymbol{\theta}}_t - \rho\left(\widetilde{\boldsymbol{\theta}}_{t-1} + \hat{\boldsymbol{x}}_{t-1}\right).$$

We posit that learning follows a Kalman filter, which provides the optimal nowcast under Gaussian shocks and the best linear nowcast (that minimizes the expected squared loss) under non-Gaussian shocks (Humpherys et al., 2012). Generally, demand shocks need to be non-stationary and non-Gaussian in order for demand to be non-negative; nevertheless, our analysis under incomplete information can be interpreted as analyzing the linearized model around a steady-state with shocks approximated by Gaussian distributions. Hence, for expositional simplicity and also relevant for our empirical specification, we assume the covariance matrices $\boldsymbol{\Sigma}^u$ and $\boldsymbol{\Sigma}^\epsilon$ are time-invariant, so that the Kalman filter is stationary.[24]

In this case, the demand innovation $\boldsymbol{v}_t$ updates the nowcast according to the law of motion:

$$\hat{\boldsymbol{x}}_t = \rho\hat{\boldsymbol{x}}_{t-1} + \boldsymbol{K}\boldsymbol{v}_t \tag{A.2}$$

The Kalman gain $\boldsymbol{K}$ is an $N \times N$ matrix as derived in Appendix A.8:

$$\boldsymbol{K} = \left(\rho^2\boldsymbol{F} + \boldsymbol{\Sigma}^u\right)\left(\boldsymbol{\Sigma}^\epsilon + \rho^2\boldsymbol{F} + \boldsymbol{\Sigma}^u\right)^{-1}, \tag{A.3}$$

where $\boldsymbol{\Sigma}^u$ is the covariance matrix of hump-shaped shocks, $\boldsymbol{\Sigma}^\epsilon$ is the covariance matrix of monotone-decay shocks, and $\boldsymbol{F} \equiv Var_t\left(\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t\right)$, a measure of the accuracy of the nowcast, satisfies the following equation:

$$\boldsymbol{F} = \boldsymbol{\Sigma}^\epsilon\left(\boldsymbol{\Sigma}^\epsilon + \rho^2\boldsymbol{F} + \boldsymbol{\Sigma}^u\right)^{-1}\left(\rho^2\boldsymbol{F} + \boldsymbol{\Sigma}^u\right). \tag{A.4}$$

Substituting $\rho\hat{\boldsymbol{x}}_{t-1} = \hat{\boldsymbol{x}}_t - \boldsymbol{K}\boldsymbol{v}_t$ from the filter update,

$$\widetilde{\boldsymbol{\theta}}_t = \rho\widetilde{\boldsymbol{\theta}}_{t-1} + \hat{\boldsymbol{x}}_t + \left(\boldsymbol{I} - \boldsymbol{K}\right)\boldsymbol{v}_t, \qquad \hat{\boldsymbol{x}}_t = \rho\hat{\boldsymbol{x}}_{t-1} + \boldsymbol{K}\boldsymbol{v}_t,$$

which is the one-innovation process of Section 3.3 verbatim under the dictionary $\boldsymbol{K} \mapsto \boldsymbol{D}$, $\hat{\boldsymbol{x}}_t \mapsto \boldsymbol{x}_t$,$(\boldsymbol{u}_t, \boldsymbol{\epsilon}_t) \mapsto (\boldsymbol{K}\boldsymbol{v}_t, (\boldsymbol{I} - \boldsymbol{K})\boldsymbol{v}_t)$. Quasi-differencing twice gives the observable ARMA form

$$(1 - \rho L)^2\widetilde{\boldsymbol{\theta}}_t = \left(\boldsymbol{I} - \rho\left(\boldsymbol{I} - \boldsymbol{K}\right)L\right)\boldsymbol{v}_t.$$

Because the steady-state filter is stabilizing, the eigenvalues of $\rho\left(\boldsymbol{I} - \boldsymbol{K}\right)$ lie strictly inside the unit circle, so the moving-average polynomial is invertible and $\boldsymbol{v}_t$ is fundamental: it is recoverable from the history of observed demand. Hence, with the innovation variance matched, $\mathrm{Var}\left(\boldsymbol{v}_t\right) = \rho^2\boldsymbol{F} + \boldsymbol{\Sigma}^u + \boldsymbol{\Sigma}^\epsilon$, the two formulations generate identical autocovariances and an identical Gaussian likelihood for observed demand. The equivalence holds at the level of observables: under the dictionary, the disturbances $\left(\boldsymbol{K}\boldsymbol{v}_t, \left(\boldsymbol{I} - \boldsymbol{K}\right)\boldsymbol{v}_t\right)$ are perfectly correlated, unlike

[24] There is a large literature on nonlinear and nonstationary filtering (e.g., Hamilton, 1989; Ito and Xiong, 2000; Farmer, 2021); the extension to this case is straightforward but less expositionally clear.

the mutually independent $(\boldsymbol{u}_t, \boldsymbol{\epsilon}_t)$; the two environments are indistinguishable from demand data alone even though the decomposition into components differs.

The following proof derives the impulse-response of expected future demand $\mathbb{E}_t\left[\widetilde{\boldsymbol{\theta}}_{t+s}\right]$ and revenue $\mathbb{E}_t\left[\widetilde{\gamma}_{t+s}\right]$ to demand shocks. Equation (A.2) states that

$$\hat{\boldsymbol{x}}_t = \rho\hat{\boldsymbol{x}}_{t-1} + \boldsymbol{K}\left(\widetilde{\boldsymbol{\theta}}_t - \rho\left(\widetilde{\boldsymbol{\theta}}_{t-1} + \hat{\boldsymbol{x}}_{t-1}\right)\right)$$

Hence,

$$\mathbb{E}_t\left[\widetilde{\boldsymbol{\theta}}_{t+s}\right] = \rho^s\widetilde{\boldsymbol{\theta}}_t + s\rho^s\hat{\boldsymbol{x}}_t$$

The demand innovation at time t is defined as

$$\boldsymbol{v}_t \equiv \tilde{\boldsymbol{\theta}}_t - \mathbb{E}_{t-1}\left[\widetilde{\boldsymbol{\theta}}_t\right] = \widetilde{\boldsymbol{\theta}}_t - \rho\left(\widetilde{\boldsymbol{\theta}}_{t-1} + \hat{\boldsymbol{x}}_{t-1}\right)$$

We have

$$\mathbb{E}_{t-1}\left[\widetilde{\boldsymbol{\theta}}_{t+s}\right] = \rho^{s+1}\left(\widetilde{\boldsymbol{\theta}}_{t-1} + \hat{\boldsymbol{x}}_{t-1}\right) + s\rho^{s+1}\hat{\boldsymbol{x}}_{t-1}$$

$$\begin{aligned}
\mathbb{E}_t\left[\widetilde{\boldsymbol{\theta}}_{t+s}\right] &= \rho^s\widetilde{\boldsymbol{\theta}}_t + s\rho^s\hat{\boldsymbol{x}}_t \\
&= \rho^s\left(\mathbb{E}_{t-1}\left[\widetilde{\boldsymbol{\theta}}_t\right] + \boldsymbol{v}_t\right) + s\rho^s\left(\rho\hat{\boldsymbol{x}}_{t-1} + \boldsymbol{K}\boldsymbol{v}_t\right)
\end{aligned}$$

$$\mathbb{E}_t\left[\boldsymbol{x}_{t+s}\right] = \rho^s\hat{\boldsymbol{x}}_t = \rho^{s+1}\hat{\boldsymbol{x}}_{t-1} + \rho^s\boldsymbol{K}\boldsymbol{v}_t$$

Hence

$$\frac{\partial\mathbb{E}_t\left[\tilde{\boldsymbol{\theta}}_{t+s}\right]}{\partial\boldsymbol{v}_t} = \rho^s\boldsymbol{I} + s\rho^s\boldsymbol{K}, \qquad \frac{\partial\mathbb{E}_t\left[\boldsymbol{x}_{t+s}\right]}{\partial\boldsymbol{v}_t} = \rho^s\boldsymbol{K}.$$

Since

$$\mathbb{E}_t\left[\widetilde{\boldsymbol{\gamma}}_{t+s}\right] = \boldsymbol{G}_\infty^\epsilon\left(\mathbb{E}_t\left[\widetilde{\boldsymbol{\theta}}_{t+s}\right] - \mathbb{E}_t\left[\widetilde{\boldsymbol{x}}_{t+s}\right]\right) + \boldsymbol{G}_\infty^x\mathbb{E}_t\left[\widetilde{\boldsymbol{x}}_{t+s}\right]$$

We have

$$\begin{aligned}
\frac{\partial\mathbb{E}_t\left[\widetilde{\boldsymbol{\gamma}}_{t+s}\right]}{\partial\boldsymbol{v}_t} &= \boldsymbol{G}_\infty^\epsilon\left(\frac{\partial\mathbb{E}_t\left[\widetilde{\boldsymbol{\theta}}_{t+s}\right]}{\partial v_t} - \frac{\partial\mathbb{E}_t\left[\boldsymbol{x}_{t+s}\right]}{\partial v_t}\right) + \boldsymbol{G}_\infty^x\frac{\partial\mathbb{E}_t\left[\boldsymbol{x}_{t+s}\right]}{\partial v_t} \\
&= \boldsymbol{G}_\infty^\epsilon\left(\rho^s\boldsymbol{I} + s\rho^s\boldsymbol{K} - \rho^s\boldsymbol{K}\right) + \boldsymbol{G}_\infty^x\rho^s\boldsymbol{K} \\
&= \rho^s\boldsymbol{G}_\infty^\epsilon\left(\boldsymbol{I} + (s-1)\boldsymbol{K}\right) + \boldsymbol{G}_\infty^x\rho^s\boldsymbol{K},
\end{aligned}$$

as desired.

A.8 Derivation of Kalman Gains

We observe $\widetilde{\boldsymbol{\theta}}_t$, and the state variables are $\widetilde{\boldsymbol{\theta}}_t$ and $\boldsymbol{x}_t$. We write down the state equation and measurement equation:

$$\boldsymbol{S}_t \equiv \begin{bmatrix} \widetilde{\boldsymbol{\theta}}_t \\ \boldsymbol{x}_t \end{bmatrix} = \underbrace{\begin{bmatrix} \rho\boldsymbol{I} & \rho\boldsymbol{I} \\ \boldsymbol{0} & \rho\boldsymbol{I} \end{bmatrix}}_{\equiv \boldsymbol{M}} \boldsymbol{S}_{t-1} + \underbrace{\begin{bmatrix} \boldsymbol{I} & \boldsymbol{I} \\ 0 & \boldsymbol{I} \end{bmatrix}}_{\equiv \boldsymbol{B}} \begin{bmatrix} \boldsymbol{\epsilon}_t \\ \boldsymbol{u}_t \end{bmatrix}$$

$$\widetilde{\boldsymbol{\theta}}_t = \underbrace{\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix}}_{\equiv \boldsymbol{H}'} \boldsymbol{S}_t$$

Updating equation:

$$\boldsymbol{S}_{t|t} = \boldsymbol{M}\boldsymbol{S}_{t-1|t-1} + \boldsymbol{P}_t \left(\widetilde{\boldsymbol{\theta}}_t - \boldsymbol{H}'\boldsymbol{M}\boldsymbol{S}_{t-1|t-1}\right),$$

where

$$\boldsymbol{P}_t = \boldsymbol{\Sigma}_{t|t-1}\boldsymbol{H}\left(\boldsymbol{H}'\boldsymbol{\Sigma}_{t|t-1}\boldsymbol{H}\right)^{-1}$$

and the conditional variance terms evolve according to

$$\boldsymbol{\Sigma}_{t|t} = \left(\boldsymbol{I} - \boldsymbol{P}_t\boldsymbol{H}'\right)\boldsymbol{\Sigma}_{t|t-1}$$

$$\boldsymbol{\Sigma}_{t+1|t} = \boldsymbol{M}\boldsymbol{\Sigma}_{t|t}\boldsymbol{M}' + \boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}'$$

where $\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{\Sigma}^{\epsilon} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}^{u} \end{bmatrix}$. As $t \to \infty$, $\boldsymbol{P}_t$, $\boldsymbol{\Sigma}_{t|t}$, and $\boldsymbol{\Sigma}_{t+1|t}$ converge to:

$$\boldsymbol{P} \equiv \lim_{t\to\infty} \boldsymbol{P}_t = \hat{\boldsymbol{\Sigma}}\boldsymbol{H}\left(\boldsymbol{H}'\hat{\boldsymbol{\Sigma}}\boldsymbol{H}\right)^{-1},$$

$$\boldsymbol{\Sigma} \equiv \lim_{t\to\infty} \boldsymbol{\Sigma}_{t|t} = \left(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{H}'\right)\hat{\boldsymbol{\Sigma}},$$

$$\hat{\boldsymbol{\Sigma}} \equiv \lim_{t\to\infty} \boldsymbol{\Sigma}_{t+1|t} = \boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}' + \boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}'.$$

We know that matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{P}$ must take the form $\boldsymbol{\Sigma} = \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{F} \end{bmatrix}$ and $\boldsymbol{P} = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{K} \end{bmatrix}$, where $\boldsymbol{F}$ is the nowcast error variance and $\boldsymbol{K}$ is the Kalman gain. Substituting these forms into $\boldsymbol{P}$, $\boldsymbol{\Sigma}$, and $\hat{\boldsymbol{\Sigma}}$ enables us to derive $\boldsymbol{F}$ and $\boldsymbol{K}$:

$$\begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{F} \end{bmatrix} = \boldsymbol{\Sigma} \quad = \quad \left(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{H}'\right)\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 0 & 0 \\ -\boldsymbol{K} & \boldsymbol{I} \end{bmatrix}\hat{\boldsymbol{\Sigma}}$$

$$\begin{aligned} \hat{\boldsymbol{\Sigma}} &= \rho^2 \begin{bmatrix} \boldsymbol{I} & \boldsymbol{I} \\ 0 & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{F} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & 0 \\ \boldsymbol{I} & \boldsymbol{I} \end{bmatrix} + \begin{bmatrix} \boldsymbol{I} & \boldsymbol{I} \\ 0 & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}^{\epsilon} & 0 \\ 0 & \boldsymbol{\Sigma}^{u} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & 0 \\ \boldsymbol{I} & \boldsymbol{I} \end{bmatrix} \\ &= \rho^2 \begin{bmatrix} \boldsymbol{F} & \boldsymbol{F} \\ \boldsymbol{F} & \boldsymbol{F} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Sigma}^{\epsilon} + \boldsymbol{\Sigma}^{u} & \boldsymbol{\Sigma}^{u} \\ \boldsymbol{\Sigma}^{u} & \boldsymbol{\Sigma}^{u} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\Sigma}^{\epsilon} + \rho^2\boldsymbol{F} + \boldsymbol{\Sigma}^{u} & \rho^2\boldsymbol{F} + \boldsymbol{\Sigma}^{u} \\ \rho^2\boldsymbol{F} + \boldsymbol{\Sigma}^{u} & \rho^2\boldsymbol{F} + \boldsymbol{\Sigma}^{u} \end{bmatrix} \end{aligned}$$

Substitute out $\hat{\boldsymbol{\Sigma}}$ from the right-hand side of $\boldsymbol{\Sigma}$ to get

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{F} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ -\boldsymbol{K} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}^{\epsilon} + \rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u & \rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u \\ \rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u & \rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u \end{bmatrix}$$

which implies

$$\boldsymbol{F} = (\boldsymbol{I} - \boldsymbol{K}) \left(\rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u\right) = \boldsymbol{\Sigma}^{\epsilon} \left(\boldsymbol{\Sigma}^{\epsilon} + \rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u\right)^{-1} (\rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u).$$

The form of $\boldsymbol{P}$ enables us to obtain the Kalman gain matrix:

$$\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{K} \end{bmatrix} = \boldsymbol{P} = \hat{\boldsymbol{\Sigma}} \boldsymbol{H} \left(\boldsymbol{H}' \hat{\boldsymbol{\Sigma}} \boldsymbol{H}\right)^{-1} = \begin{bmatrix} \boldsymbol{I} \\ (\rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u) (\boldsymbol{\Sigma}^{\epsilon} + \rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u)^{-1} \end{bmatrix}$$

which implies

$$\boldsymbol{K} = (\rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u) \left(\boldsymbol{\Sigma}^{\epsilon} + \rho^2 \boldsymbol{F} + \boldsymbol{\Sigma}^u\right)^{-1},$$

as desired.

B Theoretical Results not in the Main Text

B.1 Decentralizing the Planner's Solution as the Competitive Equilibrium

We now set up the competitive equilibrium to decentralize the planner's solution, and we show that the Lagrange multipliers in the planner's solution indeed correspond to prices.

Consumer Problem At each time t, the consumer solves

$$W_t(a_t) = \max_{\{c_{it}, \bar{\ell}_t\}} \sum_i \theta_{it} \ln c_{it} - v\left(\bar{\ell}_t\right) + \beta \mathbb{E}_t \left[W_{t+1}(a_{t+1})\right]$$

subject to the budget constraint

$$\sum_i p_{it} c_{it} + a_{t+1}/R_t = w_t \bar{\ell}_t + a_t + \Pi_t,$$

where R_t is the interest rate, a_t is savings in terms of a numeraire bond in zero net supply, p_{it} is the price of good i, w_t is the wage rate, and Π_t is the firm profits. We choose the numeraire good so that the marginal utility of income in any period t (i.e., the Lagrange multiplier on the budget constraint in each period) is always equal to one. Given the normalization, the intra-temporal Euler equation is

$$\theta_{it} = p_{it} c_{it}, \tag{A.5}$$

and the inter-temporal Euler equation is

$$1 = \beta R_t.$$

Producer Problem The representative producer in sector i solves

$$V_{it}\left(\{m_{ijt-q}\}_q\right) = \max_{\ell_{it},\{m_{ijt}\}} p_{it}y_{it} - w_t\ell_{it} - \sum_j p_{jt}m_{ijt} + \mathbb{E}_t\left[\Lambda_{t,t+1}V_{it+1}\left(\{m_{ijt+1-q}\}_q\right)\right]$$

$$\max_{m_{ijt}} \mathbb{E}_t\left[\sum_s \Lambda_{t,t+s} \max_{\ell_{it+s}}\left(p_{it+s}y_{it+s} - w_{t+s}\ell_{it+s}\right) - p_{jt}m_{ijt}\right]$$

where $\Lambda_{t,t+s}$ is the stochastic discount factor. The second problem follows from the first by an envelope argument: future input and labor choices are held at their optimal values and drop out of the marginal decision, so the first-order condition with respect to m_{ijt} is the input-order rule stated in Section 2.2. Note that given log-linear utility, the stochastic discount factor is equal to $\Lambda_{t,t+s} = \beta^s$ in all states—across all realizations of the productivity or demand shocks. Specifically, $\Lambda_{t,t+s} = \beta^s \frac{\theta_{i,t+s}c_{it}p_{it}}{\theta_{it}c_{i,t+s}p_{i,t+s}}$ for all i, which simplifies to β^s given (A.5).

Planner's Lagrange Multipliers By comparing the equilibrium conditions with the planner's solution, it is evident that the planner's Lagrange multipliers coincide with equilibrium prices.

Remark on the Choice of Numeraire and Our Characterization The object $\boldsymbol{\gamma}_t$ in the planner's solution coincides with the vector of sectoral revenues in the decentralized equilibrium with the marginal utility of income normalized to one in each period. This normalization is equivalent to choosing a numeraire in each period and is therefore inconsequential for analyzing the relative revenue across sectors at each time (i.e., γ_{it}/γ_{jt}). The nominal interest rate is $1/\beta$ under our choice of numeraire. Theorem 1 further implies that the relative sectoral revenue is invariant to the labor supply function $v(\cdot)$ and sectoral productivities. The labor supply function and sectoral productivities affect quantities—as analyzed in section 2.2—but they do not affect relative revenue across sectors. The property also holds in static production network models with log-linear preferences and production functions (Acemoglu et al. 2012) as well as in dynamic production network models with log-linear adjustment costs (Liu and Tsyvinski 2024). Our analysis here establishes this property in a dynamic production network model with heterogeneous time-to-build.

B.2 Heterogeneous Rates of Convergence for Demand Shock Components

Our analysis thus far has focused on the case where both the monotone-decay and hump-shaped components of demand shocks have the same parameter ρ capturing the rate of convergence. In this section, we generalize our results in section 3 to the case where the two demand shock components have heterogeneous rates of convergence.

Specifically, we extend the demand shock process described by equations (18) and (19) as follows:

$$\theta_{it} - \bar{\theta}_i = \rho_\epsilon \left(\theta_{it-1} - \bar{\theta}_i\right) + x_{it} + \epsilon_{it}, \tag{A.6}$$

$$x_{it} = \rho_x x_{i,t-1} + u_{it}. \tag{A.7}$$

where ρ_ϵ defines the rate at which demand reverts back to the steady-state under monotone-decay shocks and ρ_x defines the rate of convergence under hump-shaped shocks. The previous case is recovered in the limit $\rho_x \to \rho_\epsilon$. The response to a u-shock is hump-shaped when $\rho_\epsilon + \rho_x > 1$, which holds at the estimates reported in Appendix D.13.

The expected future demand can be then determined as

$$\mathbb{E}_t\left[\widetilde{\boldsymbol{\theta}}_{t+s}\right] \quad = \quad \rho_\epsilon^s\left(\widetilde{\boldsymbol{\theta}}_t - \widetilde{\boldsymbol{x}}_t\right) + \frac{\rho_\epsilon^{s+1} - \rho_x^{s+1}}{\rho_\epsilon - \rho_x}\widetilde{\boldsymbol{x}}_t$$

which generalizes equation (21) of the baseline model to the case of heterogeneous persistence. We now generalize Theorem 2 and Corollary 2 to this setting.

Proposition A.1. *Under complete information, with demand characterized by equations (A.6) and (A.7) and with $\rho_\epsilon \neq \rho_x$, sectoral revenue at time t is given by*

$$\widetilde{\gamma}_t = \boldsymbol{G}_\infty^\epsilon\left(\widetilde{\boldsymbol{\theta}}_t - \widetilde{\boldsymbol{x}}_t\right) + \boldsymbol{G}_\infty^x \widetilde{\boldsymbol{x}}_t, \tag{A.8}$$

where $\boldsymbol{G}_\infty^\epsilon$ determines the impact of demand's monotone-decay component on revenue, and $\boldsymbol{G}_\infty^x$ determines the impact of demand's hump-shaped component, and

$$\boldsymbol{G}_\infty^\epsilon \equiv \left(\boldsymbol{I} - \sum_{d=1}^{\infty}\left(\rho_\epsilon\beta\right)^d \boldsymbol{\Omega}_d\right)^{-1}, \qquad \boldsymbol{G}_\infty^x \equiv \frac{\rho_\epsilon}{\rho_\epsilon - \rho_x}\boldsymbol{G}_\infty^\epsilon - \frac{\rho_x}{\rho_\epsilon - \rho_x}\left(\boldsymbol{I} - \sum_{d=1}^{\infty}\left(\rho_x\beta\right)^d \boldsymbol{\Omega}_d\right)^{-1}.$$

The impulse-response functions are:

$$\frac{\partial \mathbb{E}_t\left[\widetilde{\gamma}_{t+s}\right]}{\partial \boldsymbol{\epsilon}_t} = \rho_\epsilon^s \boldsymbol{G}_\infty^\epsilon, \qquad \frac{\partial \mathbb{E}_t\left[\widetilde{\gamma}_{t+s}\right]}{\partial \boldsymbol{u}_t} = \boldsymbol{G}_\infty^\epsilon \frac{\rho_\epsilon^{s+1} - \rho_x^{s+1}}{\rho_\epsilon - \rho_x} + \left(\boldsymbol{G}_\infty^x - \boldsymbol{G}_\infty^\epsilon\right)\rho_x^s.$$

The revenue volatility under complete information is $Var_t\left(\widetilde{\gamma}_{t+1}\right) = \boldsymbol{G}_\infty^\epsilon \boldsymbol{\Sigma}_t^\epsilon \boldsymbol{G}_\infty^{\epsilon\prime} + \boldsymbol{G}_\infty^x \boldsymbol{\Sigma}_t^u \boldsymbol{G}_\infty^{x\prime}$, and, when demand instead follows the one-innovation process (*26*) with heterogeneous persistence (per the Remark below), sectoral revenue volatility is:

$$Var_t\left(\widetilde{\boldsymbol{\gamma}}_{t+1}\right) = \left(\boldsymbol{G}^{\epsilon}_{\infty} - \boldsymbol{G}^{\epsilon}_{\infty}\boldsymbol{D} + \boldsymbol{G}^{x}_{\infty}\boldsymbol{D}\right)\boldsymbol{\Sigma}_v\left(\boldsymbol{G}^{\epsilon}_{\infty} - \boldsymbol{G}^{\epsilon}_{\infty}\boldsymbol{D} + \boldsymbol{G}^{x}_{\infty}\boldsymbol{D}\right)'. \tag{A.9}$$

Proof. The expected future demand can be determined as

$$\mathbb{E}_t\left[\widetilde{\boldsymbol{\theta}}_{t+s}\right] = \rho_\epsilon^s\left(\widetilde{\boldsymbol{\theta}}_t - \widetilde{\boldsymbol{x}}_t\right) + \frac{\rho_\epsilon^{s+1} - \rho_x^{s+1}}{\rho_\epsilon - \rho_x}\widetilde{\boldsymbol{x}}_t \tag{A.10}$$

To derive $\boldsymbol{G}^{\epsilon}_{\infty}$, we can follow the same logic as in the proof of Corollary 1 and decompose the impact of demand's monotone-decay component recursively as the direct and indirect network effects, where each round of the network effect is summarized by $\boldsymbol{G}^{\epsilon}_{\infty}$, discounted to the present by both the discount rate (β) and also the rate at which the shock dissipates (ρ_ϵ). Specifically, note

$$\begin{aligned}
\boldsymbol{G}^{\epsilon}_{\infty} &= \boldsymbol{I} + \rho_\epsilon\beta\boldsymbol{\Omega}_1 + (\rho_\epsilon\beta)^2\left(\boldsymbol{\Omega}_1^2 + \boldsymbol{\Omega}_2\right) + (\rho_\epsilon\beta)^3\left(\boldsymbol{\Omega}_1^3 + \boldsymbol{\Omega}_1\boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_2\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_3\right) + \cdots \\
&= \boldsymbol{I} + \rho_\epsilon\beta\boldsymbol{\Omega}_1\boldsymbol{G}^{\epsilon}_{\infty} + (\rho_\epsilon\beta)^2\boldsymbol{\Omega}_2\boldsymbol{G}^{\epsilon}_{\infty} + \cdots \\
&= \boldsymbol{I} + \left(\sum_{d=1}^{\infty}(\rho_\epsilon\beta)^d\boldsymbol{\Omega}_d\right)\boldsymbol{G}^{\epsilon}_{\infty},
\end{aligned}$$

To derive $\boldsymbol{G}^{x}_{\infty}$ in closed-form, we re-write it recursively.

$$\begin{aligned}
\boldsymbol{G}^{x}_{\infty} &= \boldsymbol{I} + \frac{\rho_\epsilon^2 - \rho_x^2}{\rho_\epsilon - \rho_x}\beta\boldsymbol{\Omega}_1 + \frac{\rho_\epsilon^3 - \rho_x^3}{\rho_\epsilon - \rho_x}\beta^2\left(\boldsymbol{\Omega}_1^2 + \boldsymbol{\Omega}_2\right) + \frac{\rho_\epsilon^4 - \rho_x^4}{\rho_\epsilon - \rho_x}\beta^3\left(\boldsymbol{\Omega}_1^3 + \boldsymbol{\Omega}_1\boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_2\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_3\right) + \cdots \\
&= \frac{\rho_\epsilon}{\rho_\epsilon - \rho_x}\left[\boldsymbol{I} + \rho_\epsilon\beta\boldsymbol{\Omega}_1 + (\rho_\epsilon\beta)^2\left(\boldsymbol{\Omega}_1^2 + \boldsymbol{\Omega}_2\right) + (\rho_\epsilon\beta)^3\left(\boldsymbol{\Omega}_1^3 + \boldsymbol{\Omega}_1\boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_2\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_3\right) + \cdots\right] \\
&\quad - \frac{\rho_x}{\rho_\epsilon - \rho_x}\left[\boldsymbol{I} + \rho_x\beta\boldsymbol{\Omega}_1 + (\rho_x\beta)^2\left(\boldsymbol{\Omega}_1^2 + \boldsymbol{\Omega}_2\right) + (\rho_x\beta)^3\left(\boldsymbol{\Omega}_1^3 + \boldsymbol{\Omega}_1\boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_2\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_3\right) + \cdots\right] \\
&= \frac{\rho_\epsilon}{\rho_\epsilon - \rho_x}\left(\boldsymbol{I} - \sum_{d=1}^{\infty}(\rho_\epsilon\beta)^d\boldsymbol{\Omega}_d\right)^{-1} - \frac{\rho_x}{\rho_\epsilon - \rho_x}\left(\boldsymbol{I} - \sum_{d=1}^{\infty}(\rho_x\beta)^d\boldsymbol{\Omega}_d\right)^{-1}
\end{aligned}$$

To derive the impulse-response, equation A.10 implies that

$$\frac{\partial\mathbb{E}_t\left[\theta_{i,t+s}\right]}{\partial\epsilon_{it}} = \rho_\epsilon^s$$

$$\frac{\partial\mathbb{E}_t\left[\theta_{i,t+s}\right]}{\partial u_{it}} = \frac{\rho_\epsilon^{s+1} - \rho_x^{s+1}}{\rho_\epsilon - \rho_x}$$

Hence

$$\frac{\partial\mathbb{E}_t\left[\widetilde{\gamma}_{t+s}\right]}{\partial\boldsymbol{\epsilon}_t} = \rho_\epsilon^s\left(\boldsymbol{I} - \sum_{d=1}^{\infty}(\rho_\epsilon\beta)^d\boldsymbol{\Omega}_d\right)^{-1}$$

$$\frac{\partial \mathbb{E}_t\left[\widetilde{\gamma}_{t+s}\right]}{\partial \boldsymbol{u}_t} = \left(\boldsymbol{I}-\sum_{d=1}^{\infty}\left(\rho_\epsilon\beta\right)^d\boldsymbol{\Omega}_d\right)^{-1}\frac{\rho_\epsilon^{s+1}-\rho_x^{s+1}}{\rho_\epsilon-\rho_x}$$
$$+\frac{\rho_x}{\rho_\epsilon-\rho_x}\left(\left(\boldsymbol{I}-\sum_{d=1}^{\infty}\left(\rho_\epsilon\beta\right)^d\boldsymbol{\Omega}_d\right)^{-1}-\left(\boldsymbol{I}-\sum_{d=1}^{\infty}\left(\rho_x\beta\right)^d\boldsymbol{\Omega}_d\right)^{-1}\right)\rho_x^s \qquad \text{as desired.}$$

Hence, under complete information,

$$Var_t\left(\widetilde{\gamma}_{t+1}\right)=\boldsymbol{G}_\infty^\epsilon\boldsymbol{\Sigma}_t^\epsilon\boldsymbol{G}_\infty^{\epsilon\prime}+\boldsymbol{G}_\infty^x\boldsymbol{\Sigma}_t^u\boldsymbol{G}_\infty^{x\prime}$$

Under one demand innovation that loads on both time profiles,

$$Var_t\left(\widetilde{\gamma}_{t+1}\right)=\left(\boldsymbol{G}_\infty^\epsilon-\boldsymbol{G}_\infty^\epsilon\boldsymbol{D}+\boldsymbol{G}_\infty^x\boldsymbol{D}\right)\boldsymbol{\Sigma}_v\left(\boldsymbol{G}_\infty^\epsilon-\boldsymbol{G}_\infty^\epsilon\boldsymbol{D}+\boldsymbol{G}_\infty^x\boldsymbol{D}\right)', \qquad \text{as desired.}$$

□

Remark: the innovations representation of Proposition 3 extends to heterogeneous persistence: when agents observe only the history of realized demand, the steady-state filter delivers $\hat{\boldsymbol{x}}_t=\rho_x\hat{\boldsymbol{x}}_{t-1}+\boldsymbol{K}\boldsymbol{v}_t$ and observed demand follows $\widetilde{\boldsymbol{\theta}}_t=\rho_\epsilon\widetilde{\boldsymbol{\theta}}_{t-1}+\hat{\boldsymbol{x}}_t+\left(\boldsymbol{I}-\boldsymbol{K}\right)\boldsymbol{v}_t$ with $\boldsymbol{v}_t$ the one-step-ahead forecast error, so the process estimated in Appendix D.13 again nests the two-shock economy with Kalman-filtering agents.

B.3 Identification from Autocovariances

Proposition A.2. *(Identification from autocovariances.) Suppose demand follows the one-innovation process (26) with $0<\rho<1$, full-rank innovation covariance $\boldsymbol{\Sigma}_v$, and weight matrix $\boldsymbol{D}$ with $\det\left(\boldsymbol{D}\right)\neq 0$ such that all eigenvalues of $\rho\left(\boldsymbol{I}-\boldsymbol{D}\right)$ lie strictly inside the unit circle. Then $\left(\rho,\boldsymbol{D},\boldsymbol{\Sigma}_v\right)$ is identified from the autocovariance structure of observed demand.*

Proof. Quasi-differencing gives $\left(1-\rho L\right)\widetilde{\boldsymbol{\theta}}_t=\boldsymbol{x}_t+\left(\boldsymbol{I}-\boldsymbol{D}\right)\boldsymbol{v}_t$ and $\left(1-\rho L\right)\boldsymbol{x}_t=\boldsymbol{D}\boldsymbol{v}_t$, hence

$$\boldsymbol{w}_t\equiv\left(1-\rho L\right)^2\widetilde{\boldsymbol{\theta}}_t=\boldsymbol{v}_t+\boldsymbol{\Theta}\boldsymbol{v}_{t-1}, \qquad \boldsymbol{\Theta}\equiv-\rho\left(\boldsymbol{I}-\boldsymbol{D}\right).$$

Observed demand is therefore a vector ARMA(2,1) process with scalar autoregressive polynomial $\left(1-\rho L\right)^2$. Since $\det\left(\boldsymbol{D}\right)\neq 0$, the moving-average polynomial $\boldsymbol{I}+\boldsymbol{\Theta}L$ does not cancel the autoregressive root (at $L=1/\rho$ it equals $\boldsymbol{I}+\rho^{-1}\boldsymbol{\Theta}=\boldsymbol{D}$, which is invertible), and ρ is identified from the autoregressive polynomial of demand. The quasi-differenced series is a first-order vector moving average with autocovariances $\boldsymbol{\Gamma}_0=\boldsymbol{\Sigma}_v+\boldsymbol{\Theta}\boldsymbol{\Sigma}_v\boldsymbol{\Theta}'$, $\boldsymbol{\Gamma}_1=\boldsymbol{\Theta}\boldsymbol{\Sigma}_v$, and $\boldsymbol{\Gamma}_k=\boldsymbol{0}$ for $k\geq 2$. The eigenvalue condition implies $\det\left(\boldsymbol{I}+\boldsymbol{\Theta}z\right)\neq 0$ for $|z|\leq 1$, so $\boldsymbol{v}_t+\boldsymbol{\Theta}\boldsymbol{v}_{t-1}$ is the invertible moving-average representation of $\boldsymbol{w}_t$. By uniqueness of the invertible spectral factorization of a full-rank vector moving-average process (Hannan and Deistler, 1988), $\left(\boldsymbol{\Theta},\boldsymbol{\Sigma}_v\right)$ is the unique pair consistent with these autocovariances, and $\boldsymbol{D}=\boldsymbol{I}+\rho^{-1}\boldsymbol{\Theta}$ recovers the weight matrix. Given ρ, which the previous step identifies from the autoregressive polynomial, the parameter count matches the moment count exactly: $\mathbf{\Gamma}_1$ has N^2 elements and $\mathbf{\Gamma}_0$ has $N(N+1)/2$, against N^2 parameters in $\boldsymbol{D}$ and $N(N+1)/2$ in $\mathbf{\Sigma}_v$. Identification weakens as $\boldsymbol{D} \to \mathbf{0}$, where the moving-average root approaches the autoregressive root and the two cancel in the limit. □

C MLE Algorithm Details

We estimate the demand process by Gaussian maximum likelihood, casting the one-innovation representation directly in state-space form. Substituting the law of motion of $\boldsymbol{x}_t$ into the demand equation gives $\widetilde{\boldsymbol{\theta}}_t = \rho\widetilde{\boldsymbol{\theta}}_{t-1} + \rho\boldsymbol{x}_{t-1} + \boldsymbol{v}_t$, so we stack the state vector $\boldsymbol{\alpha}_t = \left(\widetilde{\boldsymbol{\theta}}_t, \boldsymbol{x}_t\right)$ with

$$\boldsymbol{\alpha}_t = \boldsymbol{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{R}\boldsymbol{v}_t, \qquad \boldsymbol{T} = \begin{pmatrix} \rho\boldsymbol{I} & \rho\boldsymbol{I} \\ \mathbf{0} & \rho\boldsymbol{I} \end{pmatrix}, \qquad \boldsymbol{R} = \begin{pmatrix} \boldsymbol{I} \\ \boldsymbol{D} \end{pmatrix},$$

and the observation equation selects the first block, $\widetilde{\boldsymbol{\theta}}_t = [\boldsymbol{I}\ \mathbf{0}]\,\boldsymbol{\alpha}_t$, with no measurement error. Both blocks load on the same innovation $\boldsymbol{v}_t$, so the perfect correlation between the demand and state disturbances—the defining feature of the one-innovation process—is built into the selection matrix.

For the heterogeneous-persistence specification of Appendix D.13, the transition matrix is replaced by $\begin{bmatrix} \rho_\epsilon\boldsymbol{I} & \rho_x\boldsymbol{I} \\ \mathbf{0} & \rho_x\boldsymbol{I} \end{bmatrix}$, with ρ_ϵ governing the demand recursion and ρ_x the hidden-state recursions.

The parameters are ρ, the weights $\boldsymbol{\delta}$ with $\boldsymbol{D} = \operatorname{diag}(\boldsymbol{\delta})$, and the innovation covariance $\mathbf{\Sigma}_v$. In the baseline one-factor specification $\mathbf{\Sigma}_v = \mathbf{\Lambda}\mathbf{\Lambda}' + \operatorname{diag}(\boldsymbol{\chi}^2)$, for $1+3N$ parameters; the unrestricted specification parametrizes $\mathbf{\Sigma}_v = \boldsymbol{L}\boldsymbol{L}^\top$ with $\boldsymbol{L}$ lower triangular, for $1+N+N(N+1)/2$ parameters, and is warm-started from the one-factor estimates. ρ and each δ_i are constrained to $(0,1)$ by logistic transformations; χ_i and the Cholesky diagonals are constrained to positive values by exponential transformations.

Let $\hat{\boldsymbol{x}}_t \equiv \mathbb{E}\left[\boldsymbol{x}_t \mid \widetilde{\boldsymbol{\theta}}_t, \ldots, \widetilde{\boldsymbol{\theta}}_1\right]$ denote the filtered estimate of the hidden component and $\boldsymbol{\mathcal{Q}}_t \equiv \operatorname{Var}(\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t)$ its error covariance. Since $\widetilde{\boldsymbol{\theta}}_{t-1}$ is observed, the one-step-ahead forecast of demand is $\rho\widetilde{\boldsymbol{\theta}}_{t-1} + \rho\hat{\boldsymbol{x}}_{t-1}$, with forecast error and variance

$$\boldsymbol{e}_t \equiv \widetilde{\boldsymbol{\theta}}_t - \rho\widetilde{\boldsymbol{\theta}}_{t-1} - \rho\hat{\boldsymbol{x}}_{t-1} = \rho(\boldsymbol{x}_{t-1} - \hat{\boldsymbol{x}}_{t-1}) + \boldsymbol{v}_t, \qquad \boldsymbol{W}_t = \rho^2\boldsymbol{\mathcal{Q}}_{t-1} + \mathbf{\Sigma}_v.$$

The Kalman recursions update the estimate and its covariance as

$$\hat{\boldsymbol{x}}_t = \rho\hat{\boldsymbol{x}}_{t-1} + \boldsymbol{\mathcal{G}}_t\boldsymbol{e}_t, \qquad \boldsymbol{\mathcal{G}}_t = \boldsymbol{C}_t\boldsymbol{W}_t^{-1}, \qquad \boldsymbol{C}_t \equiv \operatorname{Cov}(\boldsymbol{x}_t - \rho\hat{\boldsymbol{x}}_{t-1}, \boldsymbol{e}_t) = \rho^2\boldsymbol{\mathcal{Q}}_{t-1} + \boldsymbol{D}\mathbf{\Sigma}_v,$$

$$\boldsymbol{\mathcal{Q}}_t = \rho^2\boldsymbol{\mathcal{Q}}_{t-1} + \boldsymbol{D}\mathbf{\Sigma}_v\boldsymbol{D}' - \boldsymbol{C}_t\boldsymbol{W}_t^{-1}\boldsymbol{C}_t',$$

initialized at the stationary distribution of the state, $\hat{\boldsymbol{x}}_0 = \mathbf{0}$ and $\boldsymbol{\mathcal{Q}}_0 = (1-\rho^2)^{-1}\boldsymbol{D}\mathbf{\Sigma}_v\boldsymbol{D}'$, which is well defined since $|\rho| < 1$. The Gaussian log-likelihood follows from the prediction-error decomposition:

$$\mathcal{L}\left(\{\widetilde{\boldsymbol{\theta}}_t\}_{t=1}^T; \rho, \boldsymbol{\delta}, \boldsymbol{\Sigma}_v\right) = -\frac{1}{2}\sum_{t=1}^{T}\left(N\log 2\pi + \log|\boldsymbol{W}_t| + \boldsymbol{e}_t'\boldsymbol{W}_t^{-1}\boldsymbol{e}_t\right).$$

Because the state and the observation share the same innovation, the filter degenerates after a short transient: $\boldsymbol{\mathcal{Q}}_t \to \mathbf{0}$, $\boldsymbol{\mathcal{G}}_t \to \boldsymbol{D}$, $\boldsymbol{e}_t \to \boldsymbol{v}_t$, and $\boldsymbol{W}_t \to \boldsymbol{\Sigma}_v$, and the recursion reduces to inverting the one-innovation process. These recursions are the econometrician's filter for the one-innovation process, distinct from the agents' filter for the two-shock economy derived in Appendix A.8. (When the one-innovation process arises from the two-shock economy, Proposition 3 shows that $\boldsymbol{D}$ equals the agents' steady-state gain $\boldsymbol{K}$.) The filter therefore serves a single purpose: it integrates out the unobserved initial state $\boldsymbol{x}_0$ over its stationary distribution, rather than conditioning on a presample initialization. Results are insensitive to the initialization: under an approximate-diffuse prior the estimates are nearly identical (the $\boldsymbol{\delta}$ estimates correlate 0.99 and the median-split classification is unchanged).

Practically, we implement the algorithm with the "MLEModel" class of the Python package "statsmodels", with the initialization option set to "stationary", and minimize the negative log-likelihood by "L-BFGS-B" with the projected-gradient tolerance set to 1e-5 and function tolerance to the default 2.22e-09.

D Alternative Empirical Specifications and Robustness Checks

In this section, we conduct several alternative specifications and robustness checks to complement our empirical results in the main text.

D.1 Higher Volatility in Upstream Sectors Across Supply Chains

Figure 4 in the main text shows the relationship between upstreamness, as measured by supply chain delays, and volatility, along the sector-specific supply chain for 9 of the 25 downstream sectors. Appendix Figure A.1 shows the relationship for the remaining 16 downstream sectors. Appendix Figure A.2 reproduces the main text figure after dropping the four commodity sectors ("Oil and Gas Extraction", "Coal Mining", "Copper, Nickel, Lead, and Zinc Mining", and "Iron, Gold, Silver and Other Metal Ore Mining").

Figure A.1. Along supply chains, upstream sectors exhibit higher volatility (supply chains for downstream sectors not shown in the main text)

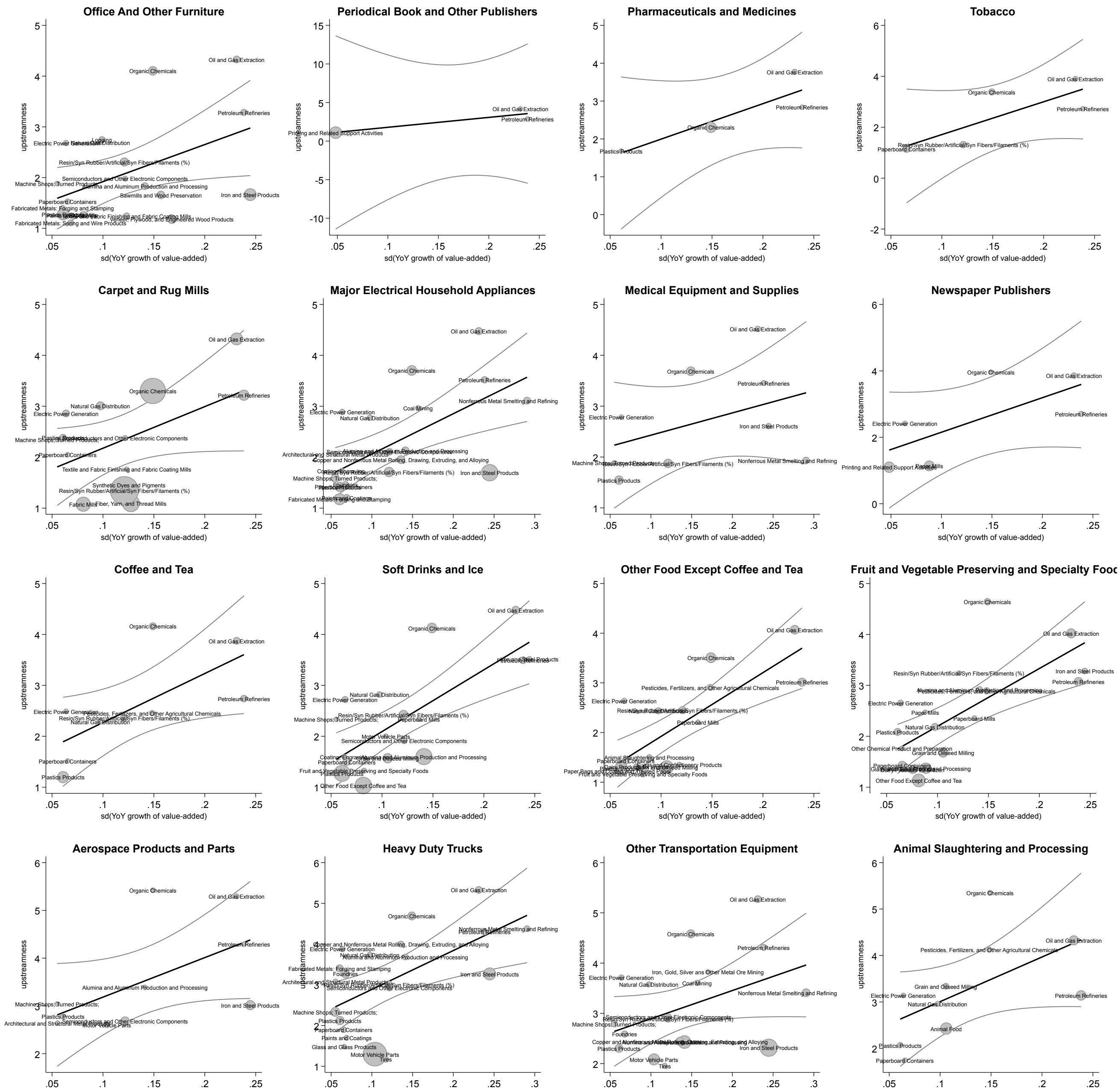
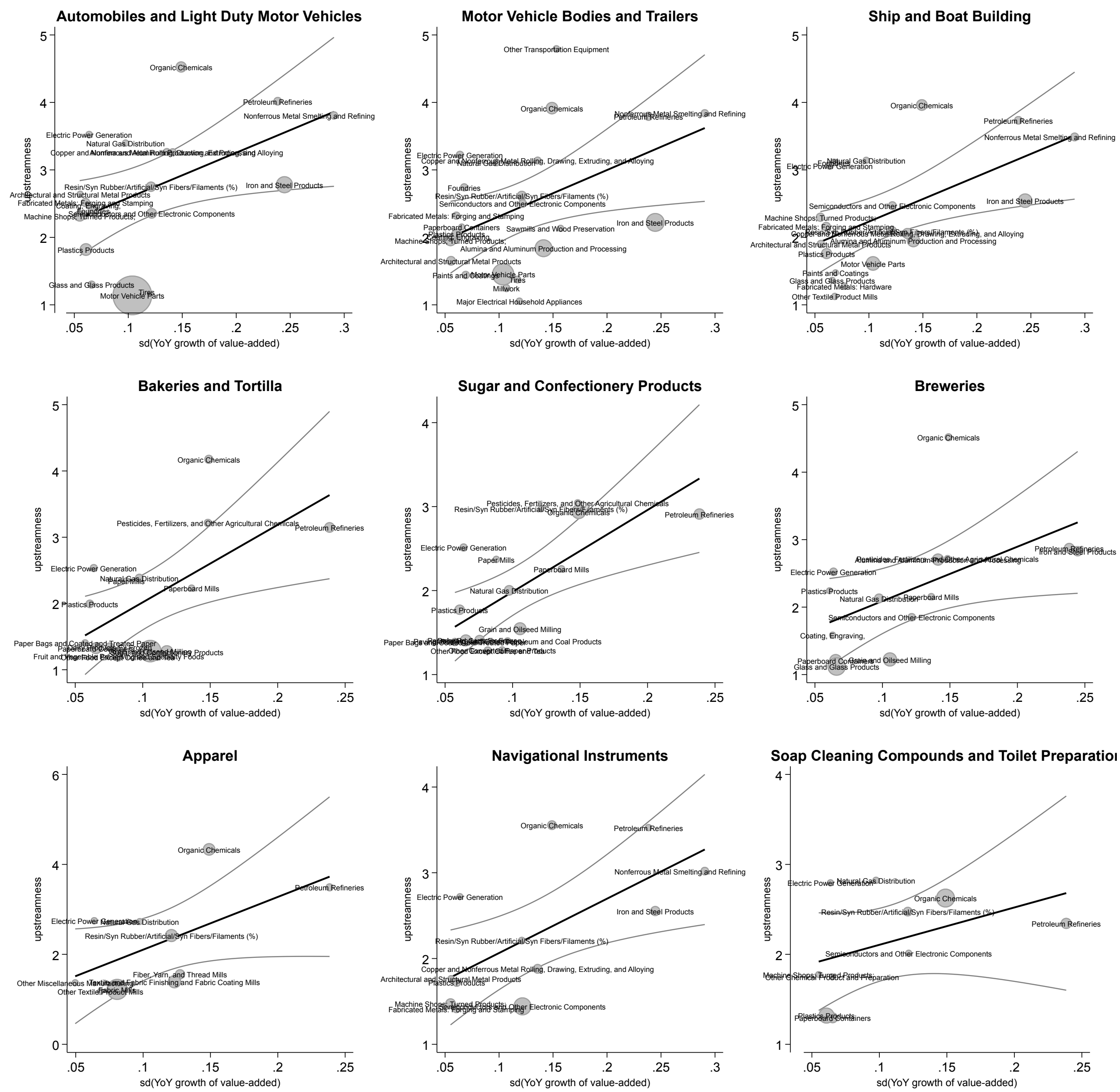


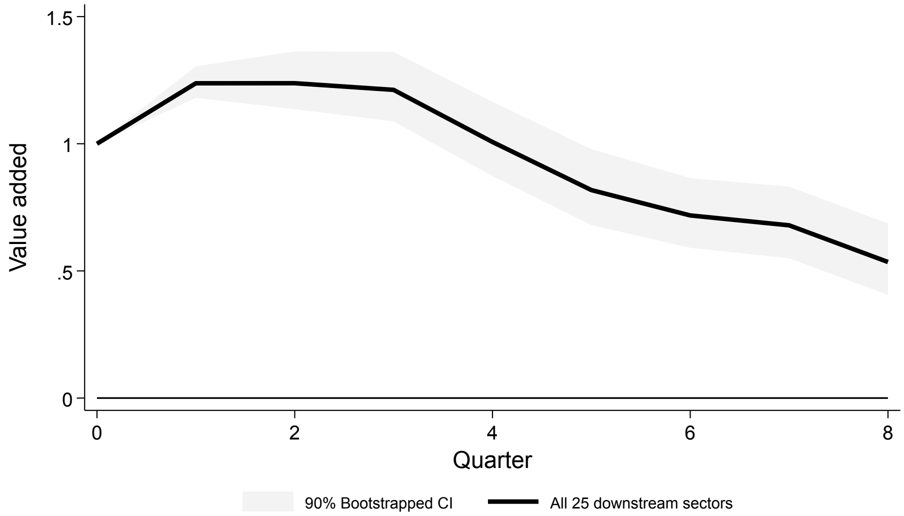
Figure A.2. Along supply chains, upstream sectors exhibit higher volatility (dropping commodity sectors)



D.2 Downstream Sector's Own Impulse-Response: Pooled Specification

Figure 5 panel (a) in the main text shows the impulse-response of downstream value-added to its own innovations. Downstream sectors in that figure are classified into two groups according to whether a sector has a dominant AR(2) component. Appendix Figure A.3 reproduces the impulse-response by pooling all 25 downstream sectors together. The figure demonstrates a hump-shaped pattern.

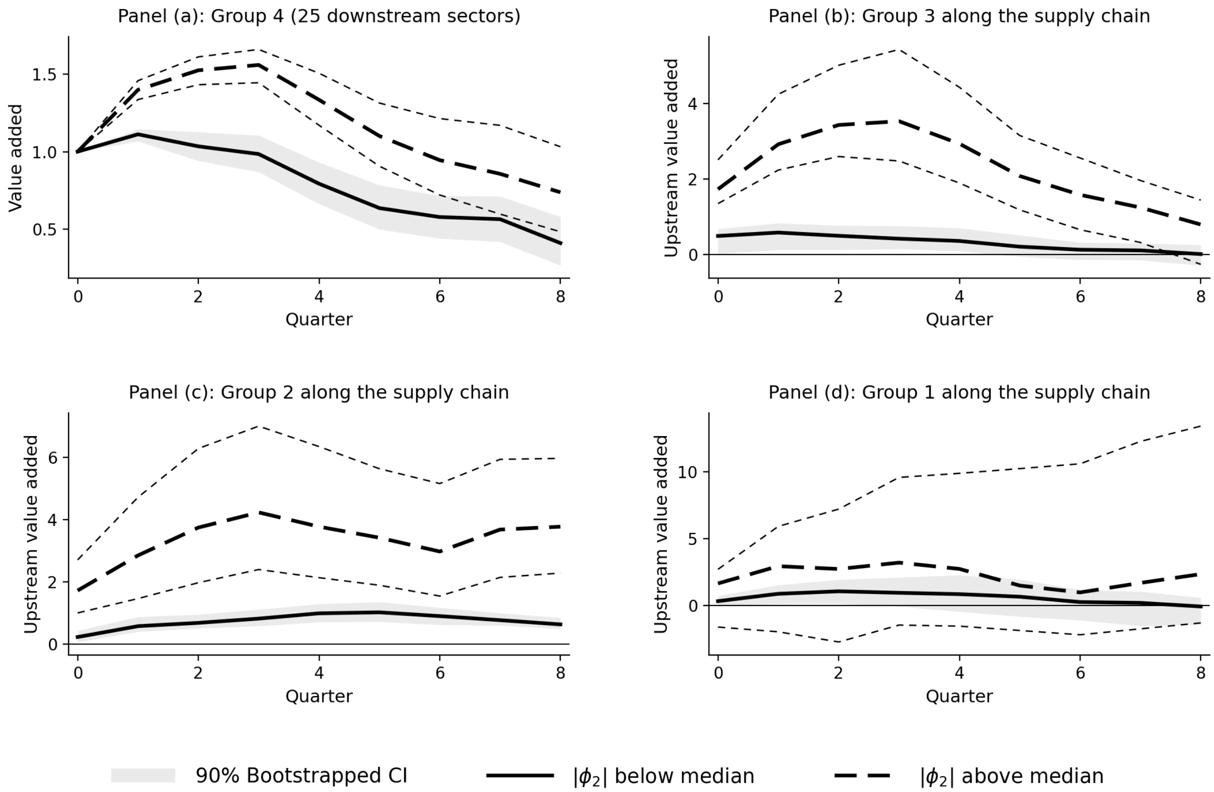
Figure A.3. Downstream sector's own impulse-response: pooled specification



D.3 Picking Downstream Sectors According to U_i^{delay}

Figure 5 in the main text shows the estimated impulse response functions when identifying a set $\mathcal{D}$ of 25 downstream sectors that are important in the consumption bundle following Antràs et al. (2012). Appendix Figure A.4 shows that our results are robust to identifying a set $\mathcal{D}$ of 25 downstream sectors that are important in the consumption bundle according to U_i^{delay}. The two rankings share 22 of the 25 sectors.

Figure A.4. Cumulative impulse response along the supply chain (picking downstream sectors according to U_i^{delay})



D.4 Impulse-Response Using Estimated δ_i

Figure 5 in the main text shows the estimated impulse response functions when downstream sectors (and the associated supply chains) are classified by their PACF lag 2 coefficients. Figure A.5 shows that our results are robust to classifying downstream sectors according to the estimated δ_i. Specifically, we estimate the following local projection regressions:

$$VA^{down}_{j,t+h} - VA^{down}_{j,t-1} = \alpha_0^h r_{jt} + \alpha_1^h r_{jt} \times D_j + X_{jt} + \zeta_{jt}, \tag{A.11}$$

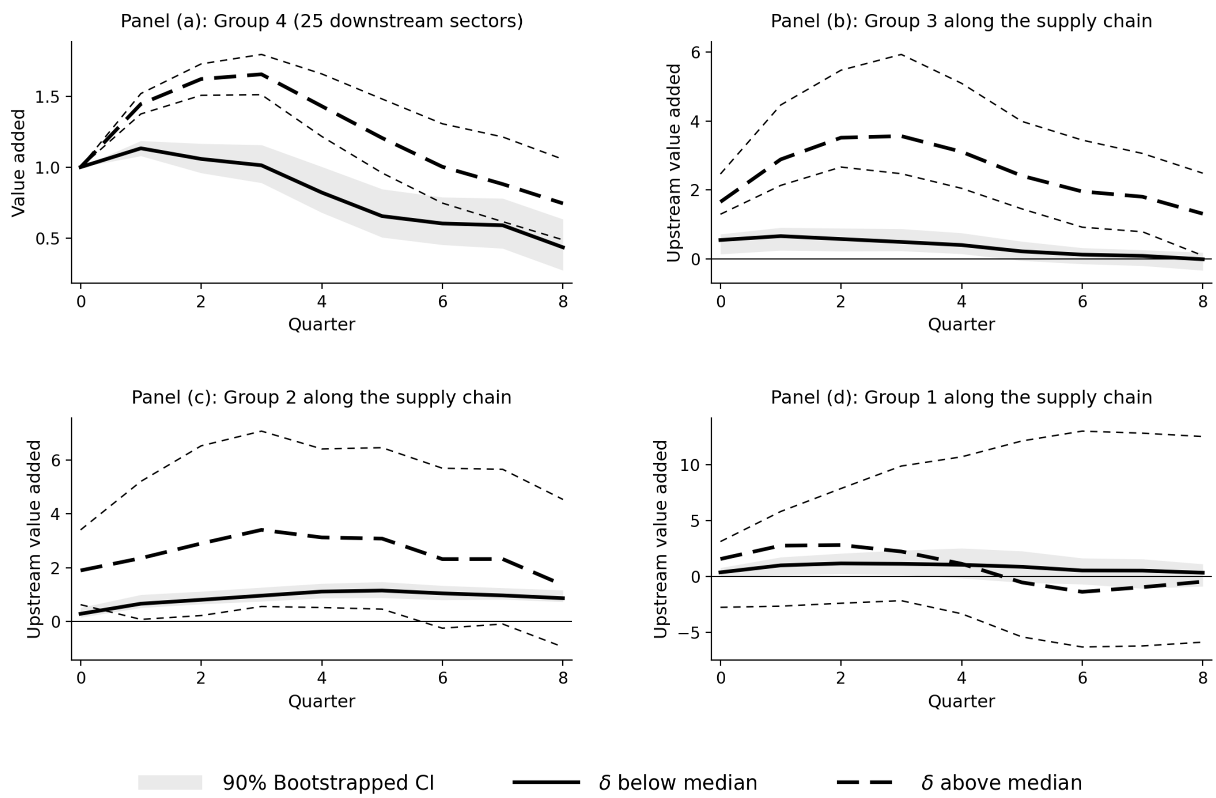
$$\text{with controls} \quad X_{jt} \in \left\{ \left\{ VA^{down}_{j,t-s} \right\}_{s=1}^{p}, \left\{ D_j \cdot VA^{down}_{j,t-s} \right\}_{s=1}^{p}, \mu_j, \mu_t \right\}.$$

$$VA^{up,g}_{i,t+h} - VA^{up,g}_{i,t-1} = \beta^{h,g} \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} r_{jt} + \delta^{h,g} \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} r_{jt} \times D_j + Z^g_{it} + \epsilon_{it}, \tag{A.12}$$

$$\text{with controls} \quad Z^g_{it} = \left\{ \left\{ \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} VA^{down}_{j,t-s}, \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} VA^{down}_{j,t-s} D_j, VA^{up,g}_{i,t-s} \right\}_{s=1}^{p}, \mu_i, \mu_t \right\}.$$

where D_j is a dummy variable for whether the j-th estimated δ_j is above the median among 25 downstream sectors.

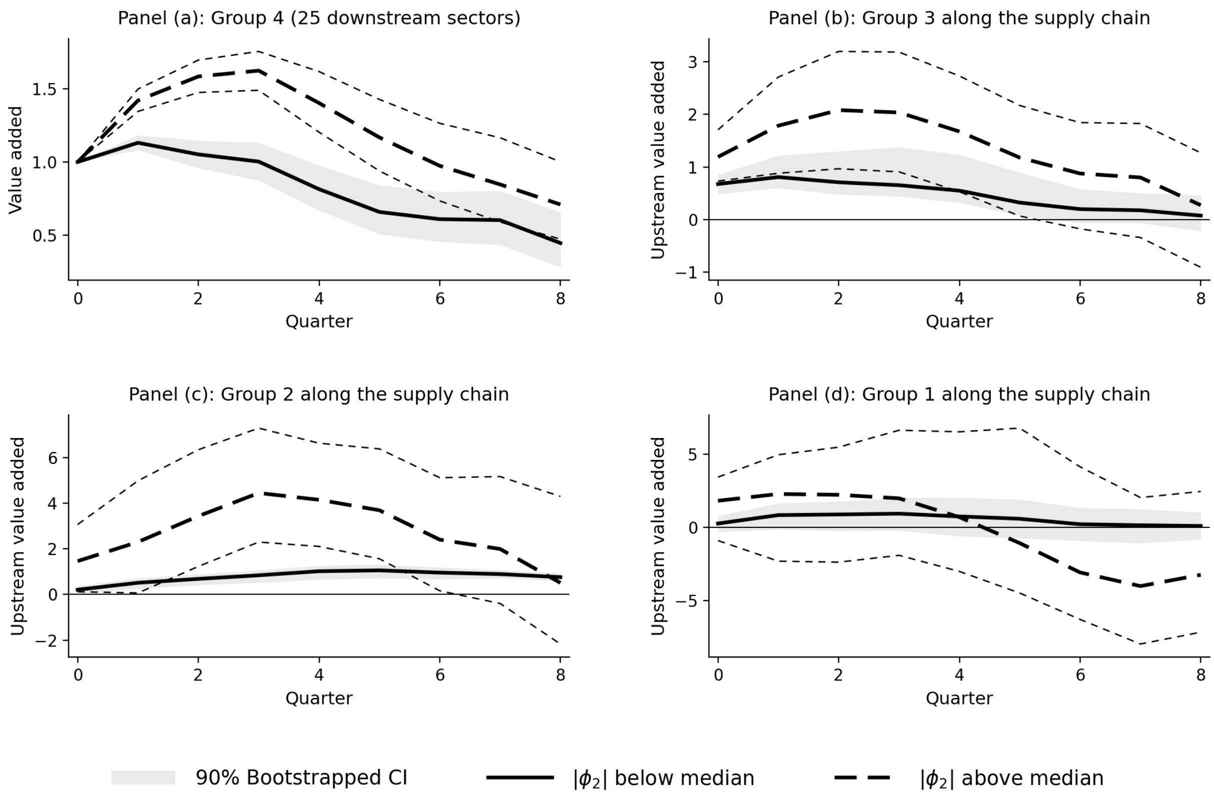
Figure A.5. Cumulative impulse response along the supply chain (classifying downstream sectors by their estimated δ_i)



D.5 Alternative Threshold

Figure 5 in the main text shows the estimated impulse response functions when suppliers enter a chain if L_{ij} is greater than 1 percent. Appendix Figure A.6 shows that our qualitative results are robust to lowering the threshold to 0.5 percent.

Figure A.6. Cumulative impulse response along the supply chain
(L_{ij} is greater than 0.5 percent)



D.6 Alternative Measures of Time-to-Build

Figure 5 in the main text shows the estimated impulse response functions when supply chain delays are based on the backlog-ratio for each input, i.e., the ratio between the stock value of unfilled orders and the flow value of goods delivered (Liu and Tsyvinski, 2024). The strategy yields a supplier-specific measure of delay.

In this Appendix we re-estimate the impulse-response functions using two alternative strategies to measure delays. The first alternative strategy follows Antràs and Tubdenov (2025) and infers an input-buyer-specific measure of time-to-build based on the ratio between the value of inventories and cost of goods sold (COGS) using the COMPUSTAT data. Because this measure captures the average production period but not shipping delays, we round up the delay measures when converting from months to quarters. Appendix Figure A.7 shows the results.

The second alternative strategy specifies the total time-to-build delay from an input producer j to input user i is the sum of the backlog ratio of good j and the inventory-to-COGS ratio of good i. Appendix Figure A.8 shows the results.

Our main results remain robust across these alternative measures of time-to-build.

Figure A.7. Cumulative impulse response along the supply chain
(average production period as the time-to-build measure)

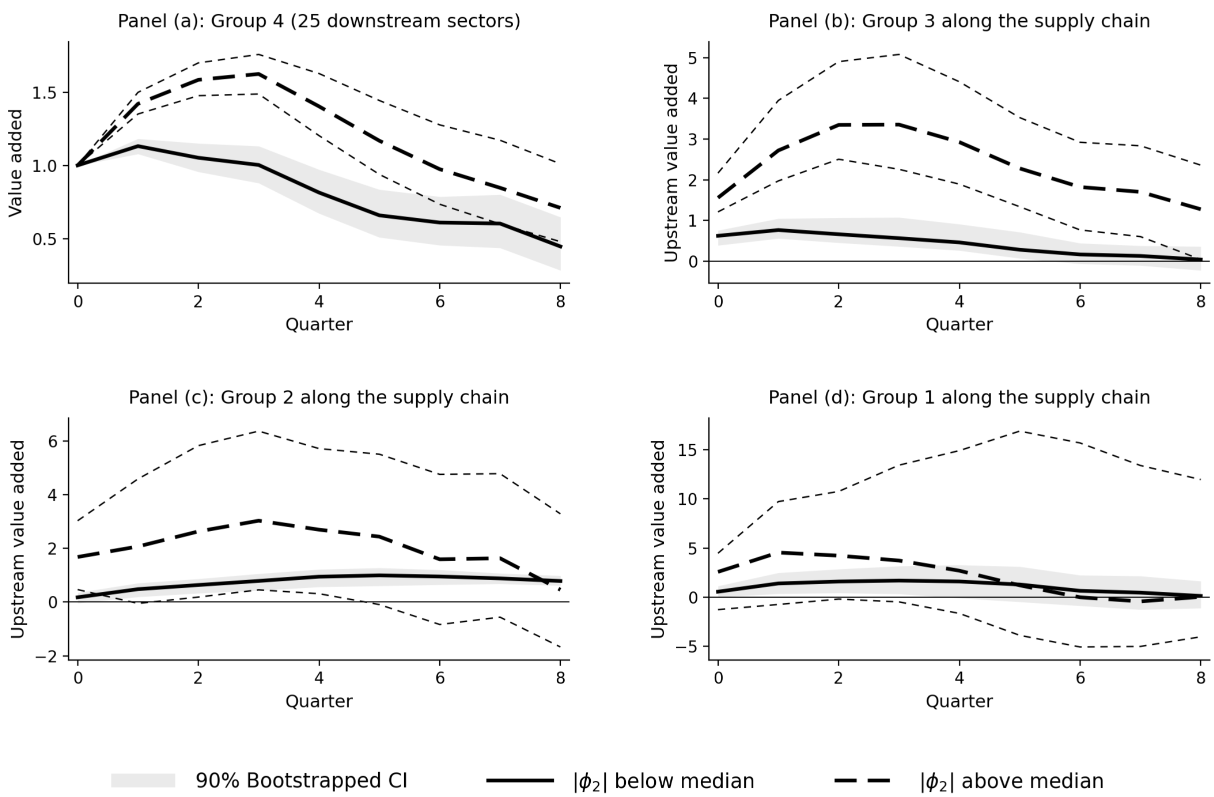
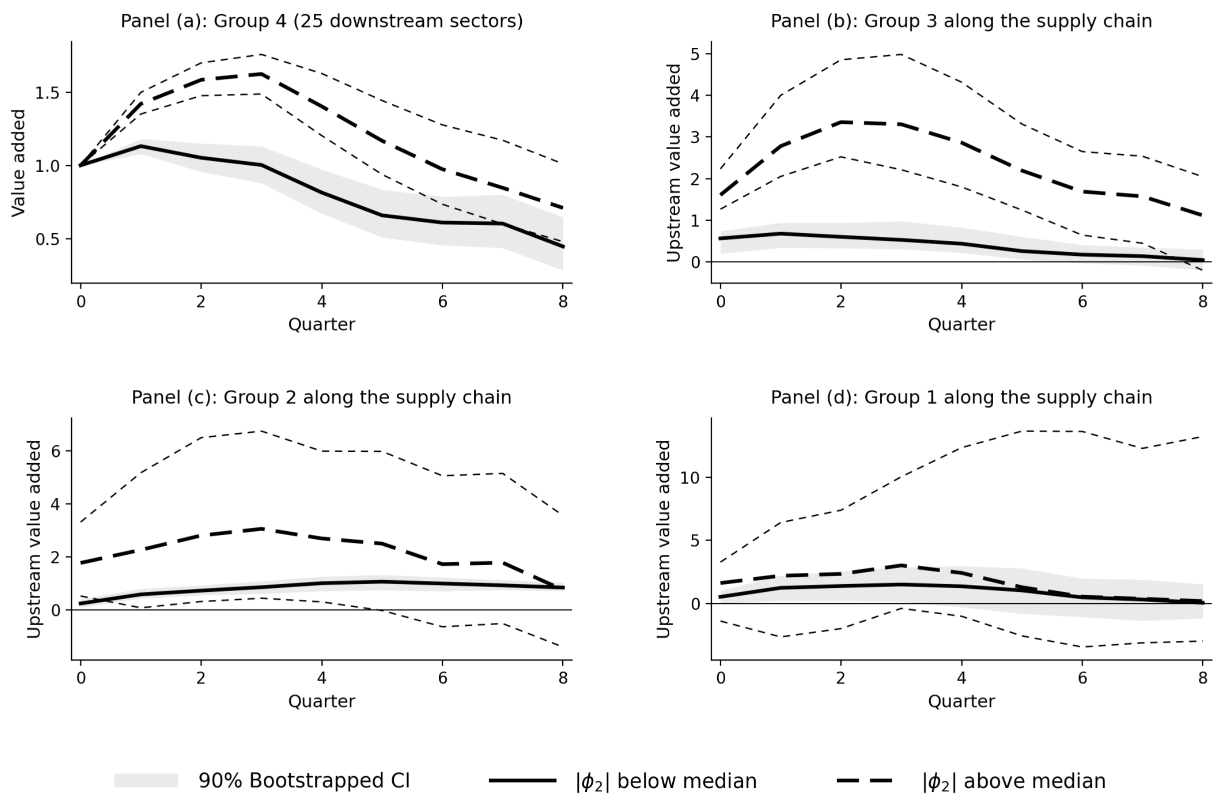


Figure A.8. Cumulative impulse response along the supply chain
(sum of the backlog-ratio and average production period as the time-to-build measure)



D.7 Using Alternative Parameters to Construct Supply Chain Distance

Figure 5 in the main text shows the estimated impulse response functions when the supply chain distance ξ_{ij} is calculated according to (25) while setting $\rho = 1$. Appendix Figure A.9 reproduces the result using the rate of decay parameter $\rho = 0.829$ estimated from the data to construct our delay measure. Our results are robust to this change.

Appendix Figure A.10 further shows the correlation (Y-axis) between ξ_{ij} calculated using specific values of ρ (shown on the X-axis) and that calculated using $\rho = 1$. The figure shows that the bilateral delay measure ξ_{ij} remains highly correlated across a wide range of ρ.

Figure A.9. Cumulative impulse response along the supply chain ($\rho = 0.829$)

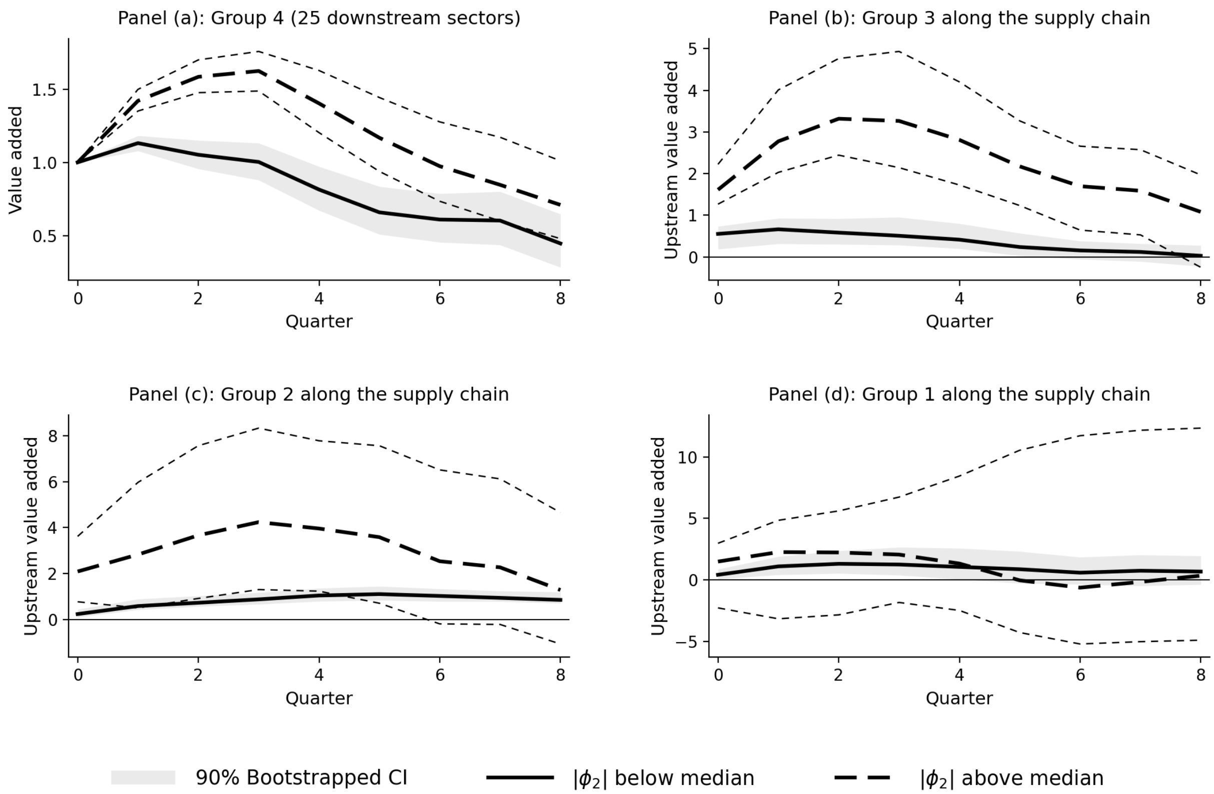
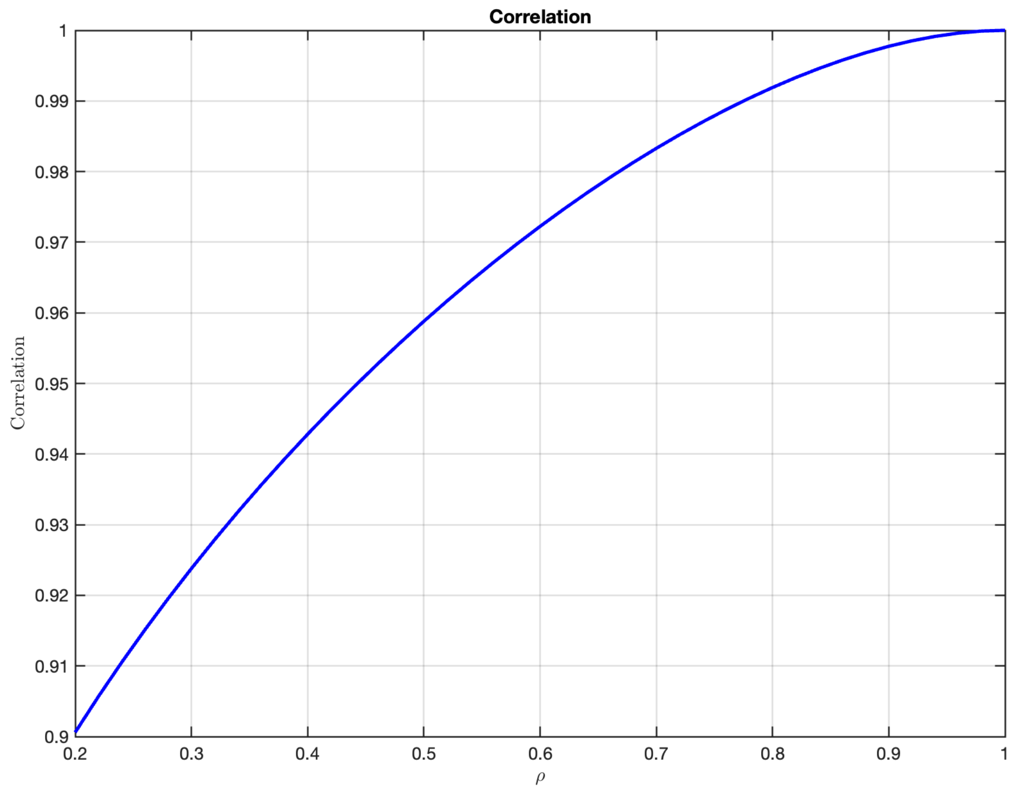


Figure A.10. Correlation between ξ_{ij} under different values of ρ and the baseline measure ($\rho = 1$)



D.8 Using Alternative Numbers of Lags When Estimating Impulse-Response Functions

Figure 5 in the main text shows the estimated impulse response functions when we use $p = 2$ lags in the regression that extracts downstream value-added innovations (equation 29). In this Appendix we reproduce the impulse-response functions when we use alternative numbers of lags. For each specification, the number of lags p in the control variables for regression equations (30) and (31) is adjusted accordingly.

Appendix Figure A.11 shows the results for $p = 1$. Appendix Figure A.12 shows the results for $p = 3$.

Figure A.11. Cumulative impulse response along the supply chain ($p = 1$ lags)

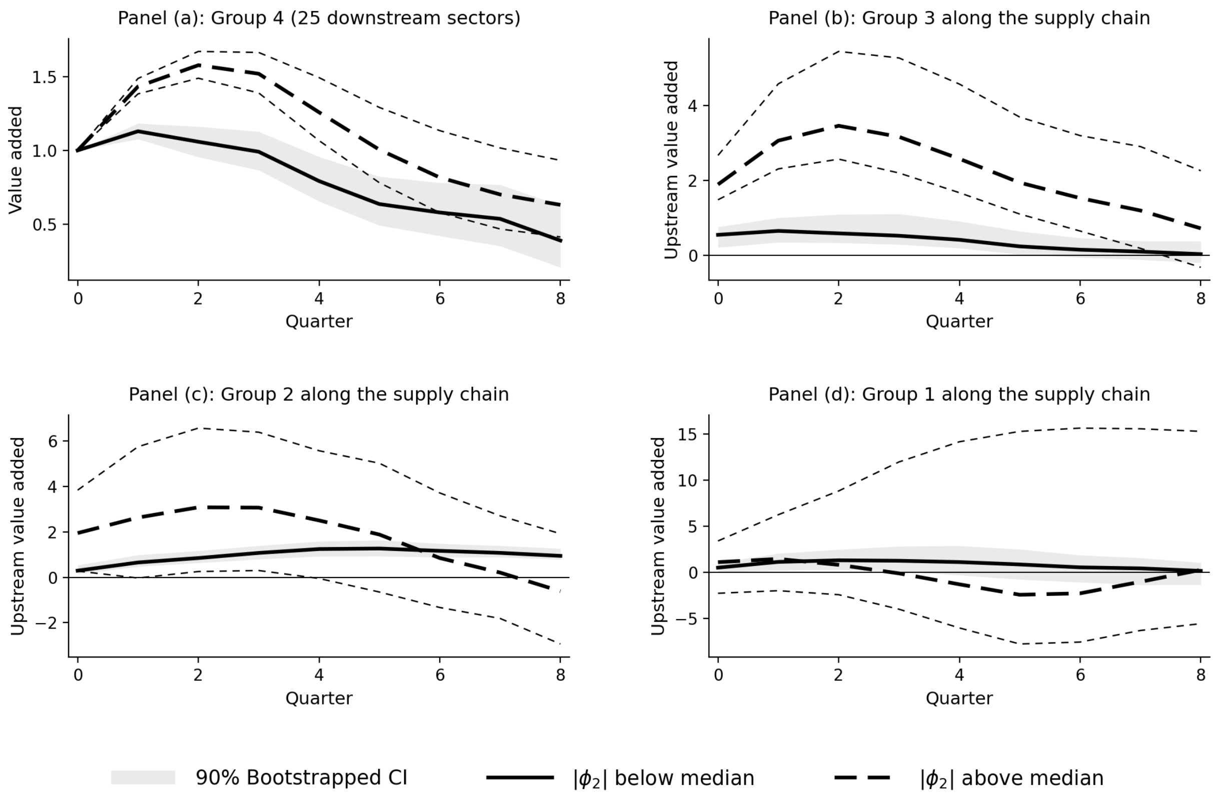
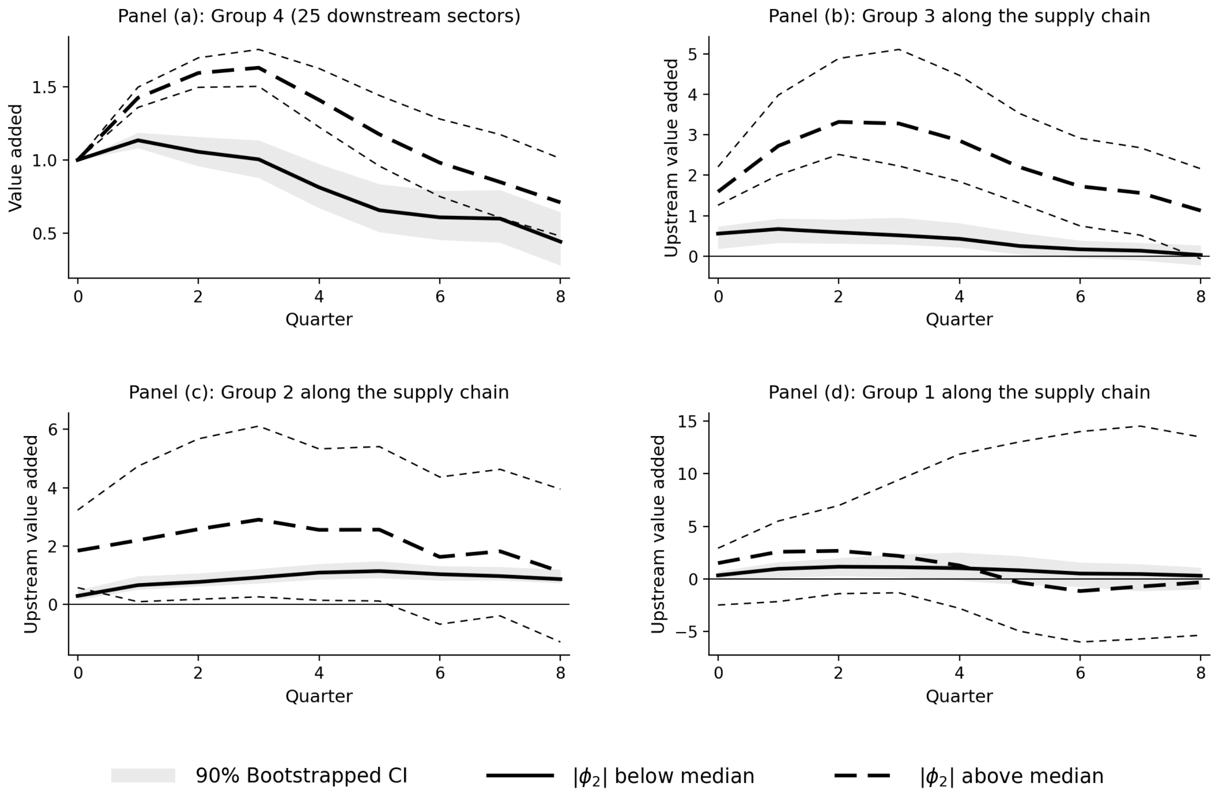


Figure A.12. Cumulative impulse response along the supply chain ($p = 3$ lags)



D.9 Impulse-Response Using Long Difference Specification

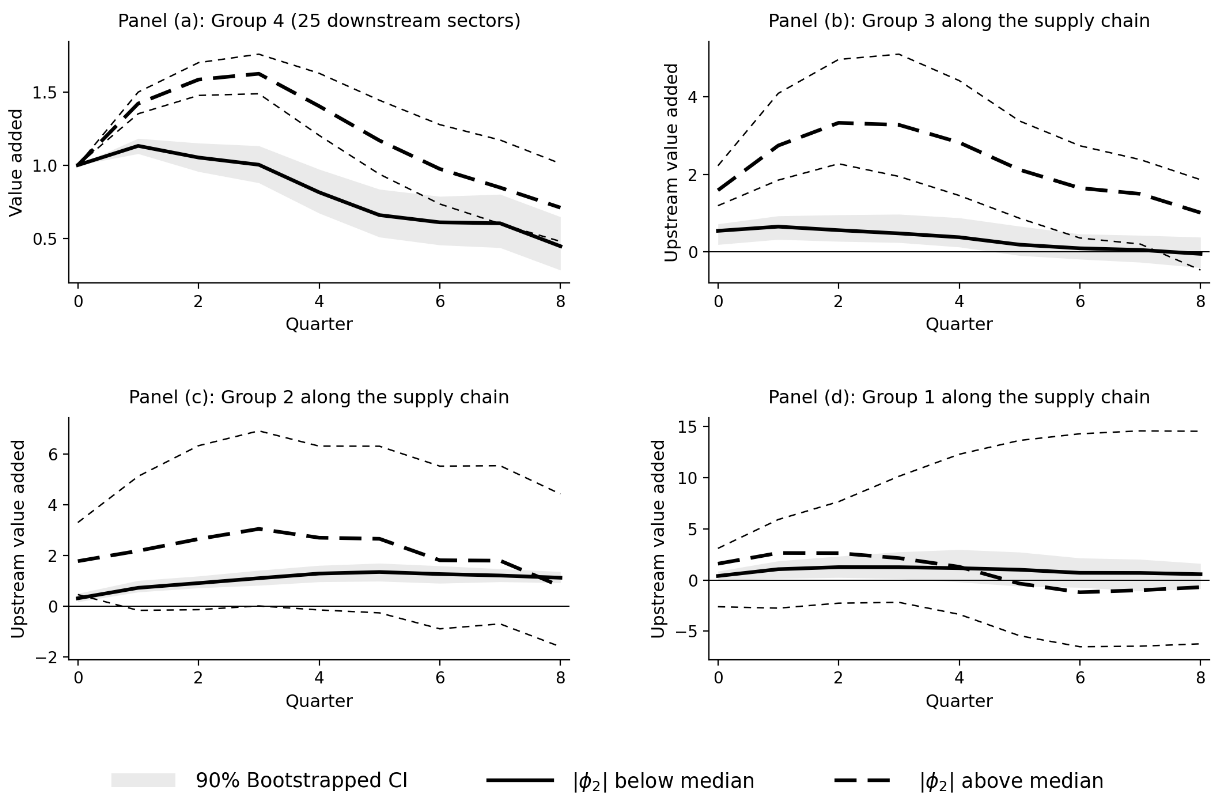
Appendix Figure A.13 reproduces the estimated impulse-response functions when using the following specification:

$$VA^{up,g}_{i,t+h} - VA^{up,g}_{i,t-1} = \beta^{h,g} \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} r_{jt} + \delta^{h,g} \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} r_{jt} \times P_j + Z^g_{it} + \epsilon_{it}, \tag{A.13}$$

$$\text{with controls} \quad Z^g_{it} \in \left\{ \left\{ \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} VA^{down}_{j,t-s}, \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} VA^{down}_{j,t-s} \times P_j \right\}_{s=1}^{p}, \Delta VA^{up,g}_{i,t-1}, \mu_i, \mu_t \right\}.$$

We include $\Delta VA^{up,g}_{i,t-1}$ in our control instead of two lags of upstream value added. This is a long-difference specification that could suppress small-sample bias (Jordà and Taylor, 2025).

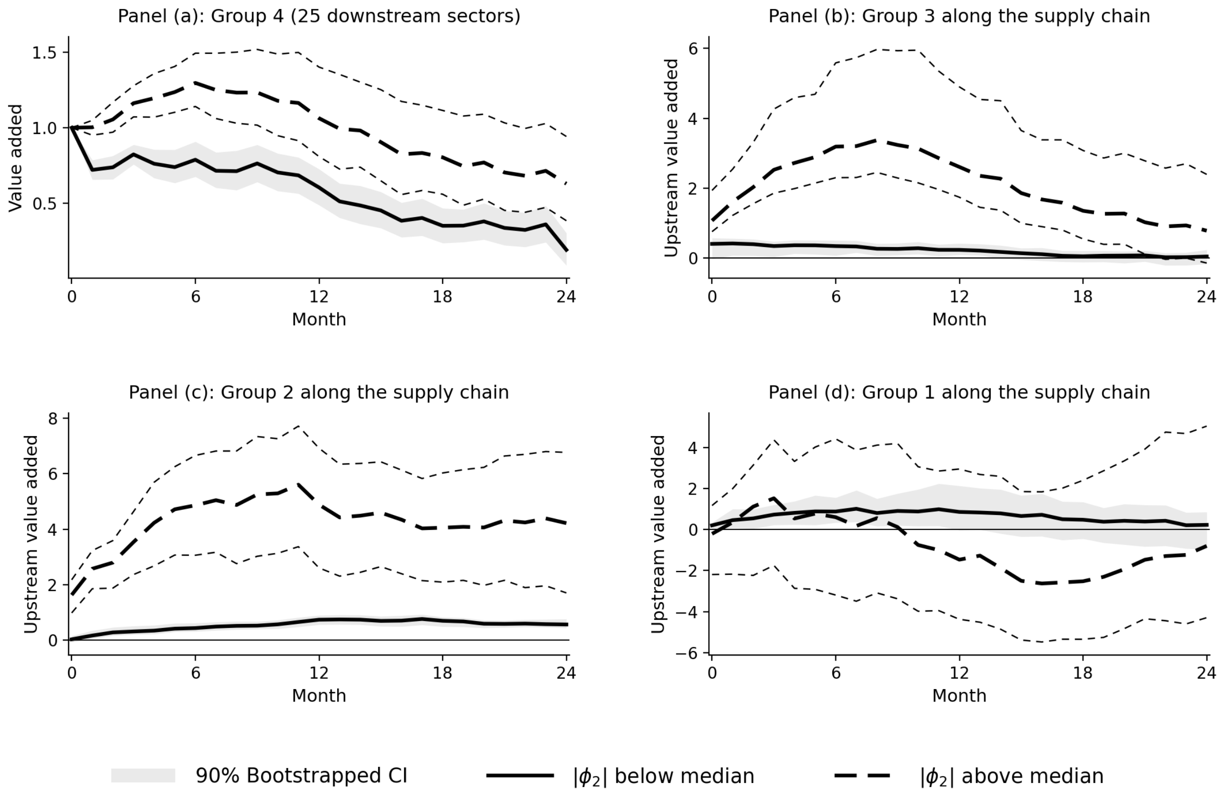
Figure A.13. Cumulative impulse response along the supply chain (long difference specification)



D.10 Impulse-Response with Monthly Data

Figure 5 in the main text shows the estimated impulse response functions using quarterly data. Appendix Figure A.14 reproduces the results using monthly data.

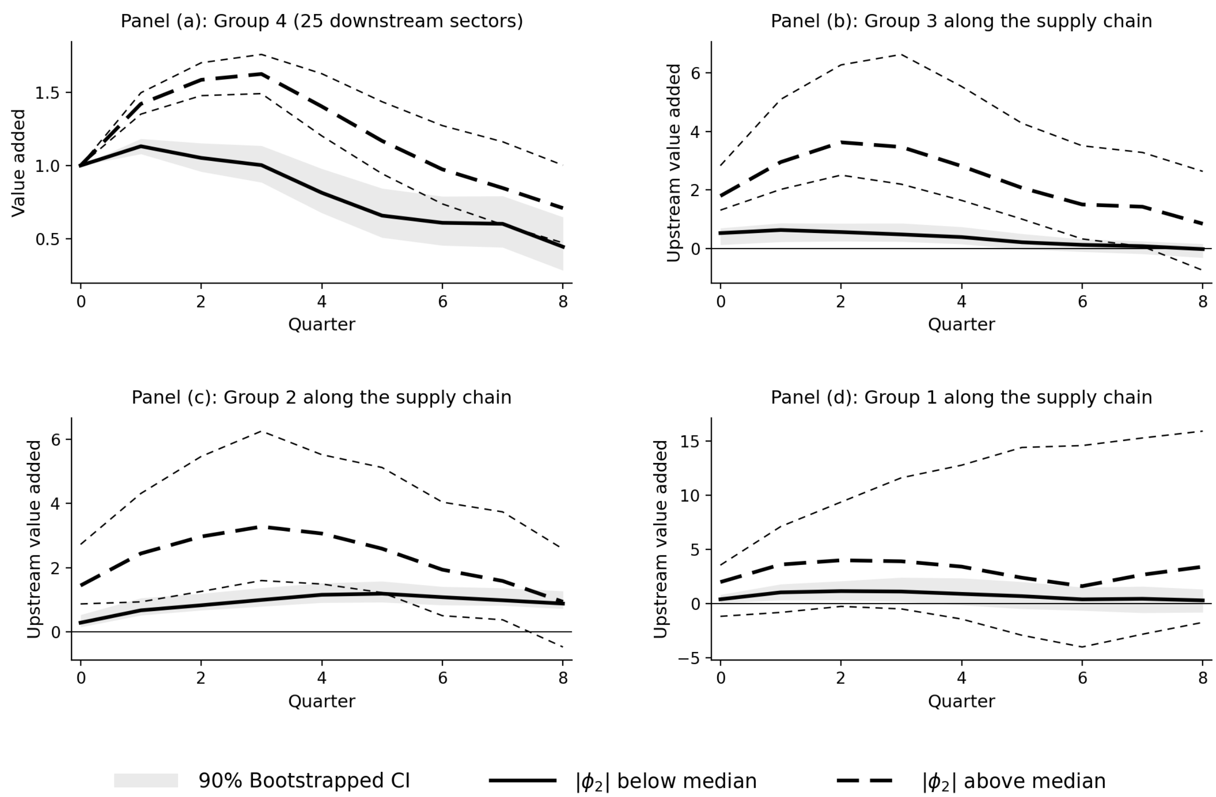
Figure A.14. Cumulative impulse response along the supply chain (monthly data)



D.11 Impulse-Response when Dropping Commodity Sectors

Appendix Figure A.15 reproduces the main text figure after dropping the four commodity sectors ("Oil and Gas Extraction", "Coal Mining", "Copper, Nickel, Lead, and Zinc Mining", and "Iron, Gold, Silver and Other Metal Ore Mining").

Figure A.15. Cumulative impulse response along the supply chain (dropping commodity sectors)



D.12 Controlling for Productivity Shocks

The impulse-response functions estimated in the main text are based on extracted innovations to downstream value-added. Although in our theoretical framework, fluctuations in value-added (i.e., revenue multiplied by sectoral value-added intensity) originate exclusively from demand shocks—since productivity shocks have offsetting effects on prices and quantities—our empirically testable predictions directly concern fluctuations in sectoral value-added.

In this Appendix section, we conduct the following robustness check to control for sector-specific cost shocks. Specifically, we solve a generalized model with constant-elasticity-of-substitution (CES) production functions with one-period delays, linearize it around a steady-state, and derive separate implications of cost shocks and demand shocks.

Consider consumer preferences

$$V\left(\mathcal{S}_t\right) = \mathbb{E}\left[\sum_{s=t}^{\infty} \beta^{s-t}\left(\ln\left(\sum_i \theta_{is}^{1/\eta} c_{is}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} - v\left(\bar{\ell}_s\right)\right) |\mathcal{S}_t\right]$$

with production functions

$$y_{jt}=z_{jt}\ell_{jt}^{\alpha_j}\left(\sum_{k=1}^{N}\lambda_{jk}^{\frac{1}{\xi}}m_{jk,t-d_{jk}}^{\frac{\xi-1}{\xi}}\right)^{\frac{\xi}{\xi-1}(1-\alpha_j)}.$$

Here, η is the elasticity of substitution across goods for the consumer, ξ is the elasticity of substitution across intermediate inputs for producers, λ_{jk} parametrizes input-output relationships (with $\sum_k \lambda_{jk}=1$), and θ_{it} again parametrizes consumer demand.

For simplicity, we solve this CES model assuming all inputs take one period to build, $d_{jk}=1$ for all j,k.

Following our derivations in the main text, we can show that

$$p_{kt}m_{jkt}=\beta\left(1-\alpha_j\right)\mathbb{E}_t\left[\gamma_{jt+1}\right]\frac{\lambda_{jk}p_{kt}^{1-\xi}}{\sum_i \lambda_{ji}p_{it}^{1-\xi}},$$

where $\gamma_{jt}\equiv p_{jt}y_{jt}$ is sectoral revenue as in the main text. That is, $\beta\left(1-\alpha_j\right)\mathbb{E}_t\left[\gamma_{jt+1}\right]$ is the total expenditure by producer j on intermediate inputs at time t, and $\frac{\lambda_{jk}p_{kt}^{1-\xi}}{\sum_i \lambda_{ji}p_{it}^{1-\xi}}$ is the fraction of intermediate input expenditure spent on input k.

Substitute m_{jkt} into the production function and then log-linearize around the steady-state (using $\widetilde{\cdot}$ to denote log-deviation from the steady-state) we obtain:

$$\widetilde{p}_{jt}=\left(1-\alpha_j\right)\left(\widetilde{\gamma}_{jt}-\mathbb{E}_{t-1}\left[\widetilde{\gamma}_{jt}\right]\right)-\widetilde{z}_{jt}+\alpha_j\widetilde{w}_t+\left(1-\alpha_j\right)\sum_k\hat{\omega}_{jk}\widetilde{p}_{kt-1},\tag{A.14}$$

where $\hat{\omega}_{jk}$ is the steady-state cost expenditure share by sector j on intermediate input k and $\sum_k\hat{\omega}_{jk}=1$. Equation (A.14) enables us to extract productivity shocks from the data on prices. To get data on prices, we first divide value-added (IP data) by value-added intensity (IO data) to recover revenue, which is then divided by sectoral quantities (IP data).

We then use the following procedure based on equation (A.14) to extract productivity shocks.

1. We residualize $\ln\gamma_{jt}$ with respect to information available up to time $t-1$. Specifically, we regress $\ln\gamma_{jt}$ on the first two principal components of all sectoral value-added at time $\{t-s\}_{s=1,\dots 4}$ and the first two principal components of all sectoral prices at time $\{t-s\}_{s=1,\dots 4}$. The residual from this regression is our measure of the innovation, $\widetilde{\gamma}_{jt}-\mathbb{E}_{t-1}\left[\widetilde{\gamma}_{jt}\right]$.

2. We multiply the innovation obtained in the previous step by $\left(1-\alpha_j\right)$ and subtract the resulting product from $\ln p_{jt}$.

3. Finally, we residualize the outcome of the previous step with respect to information available up to time $t-1$ (using the same conditioning set as step 1), to obtain residual $S_{jt}=-\widetilde{z}_{jt}+\alpha_j\widetilde{w}_t-\mathbb{E}_{t-1}\left[-\widetilde{z}_{jt}+\alpha_j\widetilde{w}_t\right]$.

S_{jt} captures innovations in sectoral unit costs arising from productivity and labor, purged of the dominant sources of common input-price variation.

We re-estimate our local projections as follows:

$$VA^{down}_{i,t+h} - VA^{down}_{i,t-1} = \alpha^h_0 r_{it} + \alpha^h_1 r_{it} \times P_i + X_{it} + \zeta_{it}, \tag{A.15}$$

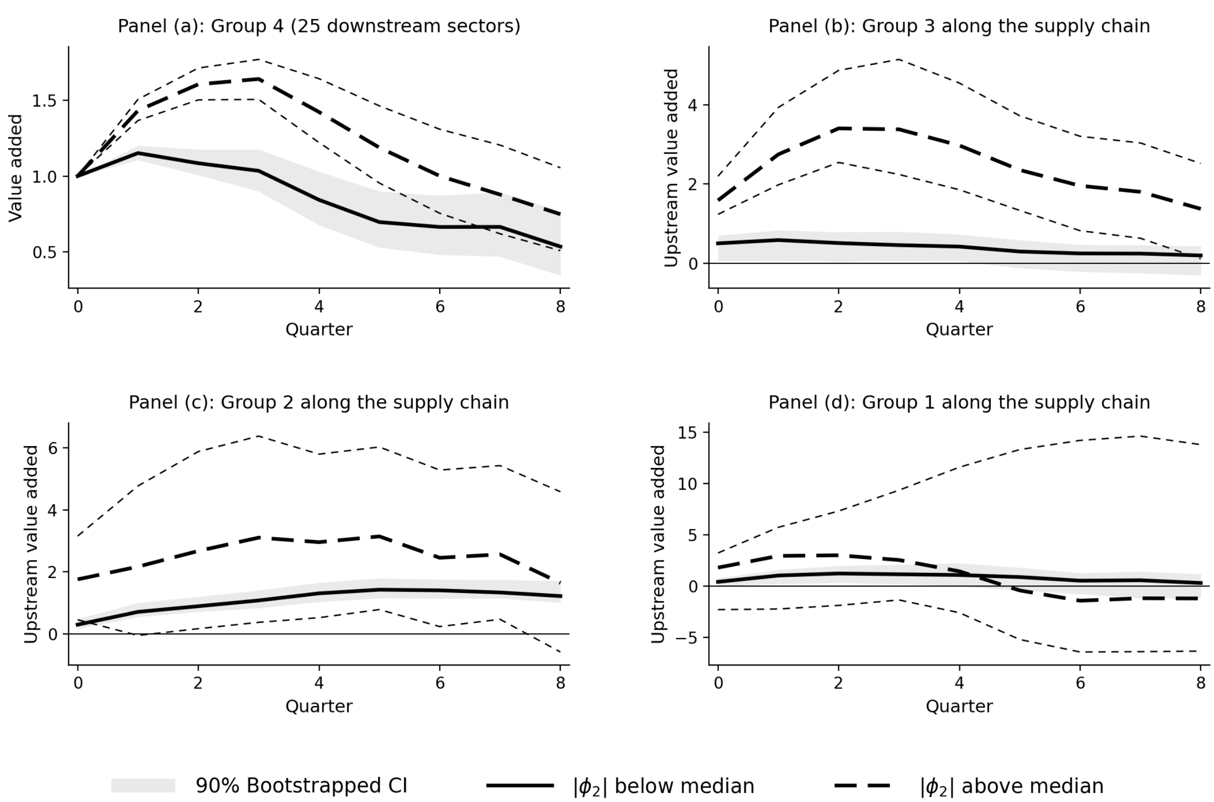
$$\text{with controls} \quad X_{it} \in \left\{ \left\{ VA^{down}_{i,t-s} \right\}^p_{s=1}, \left\{ P_i \cdot VA^{down}_{i,t-s} \right\}^p_{s=1}, S_{it}, \mu_i, \mu_t \right\}.$$

$$VA^{up,g}_{i,t+h} - VA^{up,g}_{i,t-1} = \beta^{h,g} \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} r_{jt} + \delta^{h,g} \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} r_{jt} \times P_j + Z^g_{it} + \epsilon_{it}, \tag{A.16}$$

$$\text{with controls} \quad Z^g_{it} \in \left\{ \left\{ \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} VA^{down}_{j,t-s}, \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} VA^{down}_{j,t-s} \times P_j, VA^{up,g}_{i,t-s} \right\}^p_{s=1}, \sum_{j\,:\,i\in\mathcal{U}_j} w_{ij} S_{jt}, S_{it}, \mu_i, \mu_t \right\}.$$

Figure A.16 shows the estimated impulse response functions under this specification.

Figure A.16. Cumulative impulse response along the supply chain (controlling for cost shocks)



D.13 Quantitative Implications Given Heterogeneous Rates of Convergence

Section 4.4 in the main text conducts estimation and quantification based on the baseline model where both hump-shaped and monotone-decay shocks share a common persistence parameter ρ. In this Appendix we reproduce the analysis but allow for the two types of shocks to have different persistence parameters, based on our derivations in Section B.2. We apply the MLE procedure described in Appendix C. The likelihood is invariant to exchanging (ρ_ϵ, ρ_x) together with a corresponding transformation of $\boldsymbol{D}$, so we resolve this ambiguity with the normalization $\rho_\epsilon > \rho_x$: the monotone-decay component retains its interpretation as the conventional persistent demand shock, while the hump-shaped component captures transitory surges, such as stockpiling and pent-up demand, that build and unwind at business-cycle frequencies. This labeling also preserves continuity with the baseline: under $\rho_\epsilon > \rho_x$, the estimated weights $\hat{\delta}_i$ are close to their common-persistence counterparts. The reported estimates below correspond to this normalization. The likelihood is invariant to the exchange, as are G^x_∞, the fitted revenue volatility, and the $\boldsymbol{D} = \boldsymbol{I}$ counterfactual; the $\boldsymbol{D} = \boldsymbol{0}$ counterfactual and the route-weighted delays, which depend on ρ_ϵ alone through G^ϵ_∞, are reported under the $\rho_\epsilon > \rho_x$ labeling, as are the component shares of the variance decomposition in Figure A.17. We find $\hat{\rho}_x = 0.573$ (s.e. 0.038) and $\hat{\rho}_\epsilon = 0.945$ (s.e. 0.010). Table A.1 reports the estimated δ_i. This set of $\left\{\hat{\delta}_i\right\}$ exhibits a correlation of 0.96 with the ones presented in Table 2 in the main text. Figure A.17 reproduces the variance decomposition results. Table A.2 reproduces the counterfactual exercises, where delays are route-weighted average delays computed from $\xi^\epsilon_{ij} = \frac{\left[\boldsymbol{G}^\epsilon_\infty\left(\sum_d d(\rho_\epsilon\beta)^d \boldsymbol{\Omega}_d\right)\boldsymbol{G}^\epsilon_\infty\right]_{ij}}{\left[\boldsymbol{G}^\epsilon_\infty\right]_{ij}}$, $\quad \boldsymbol{G}^\epsilon_\infty = \left(\boldsymbol{I} - \sum_d \left(\rho_\epsilon\beta\right)^d \boldsymbol{\Omega}_d\right)^{-1}$. Hump-shaped shocks continue to amplify supply chain volatility substantially—by 75.0% overall—though by less than the baseline's 155.5%: at $\hat{\rho}_x = 0.573$ the hump-shaped component's response is uniformly lower over the relevant delivery horizons than under the common estimate of 0.829 (its peak timing is essentially unchanged), while the higher $\hat{\rho}_\epsilon = 0.945$ lengthens the route-weighted delays in columns (4) and (5).

Table A.1. Estimation results: $\hat{\delta}_i$ across downstream sectors (shock-specific persistence)

		δ_i
Transportation	Ship and Boat Building	0.28 (0.13)
	Automobiles and Light Duty Motor Vehicles	0.34 (0.15)
	Heavy Duty Trucks	0.39 (0.16)
	Motor Vehicle Bodies and Trailers	0.64 (0.11)
	Aerospace Products and Parts	1.00 (0.14)
	Other Transportation Equipment	1.00 (0.17)
Food Products	Coffee and Tea	0.09 (0.09)
	Fruit and Vegetable Preserving and Specialty Foods	0.17 (0.13)
	Sugar and Confectionery Products	0.30 (0.09)
	Breweries	0.66 (0.14)
	Other Food Except Coffee and Tea	0.69 (0.14)
	Bakeries and Tortilla	0.79 (0.15)
	Animal Slaughtering and Processing	0.90 (0.14)
	Soft Drinks and Ice	0.99 (0.16)
Other Final Goods	Major Electrical Household Appliances	0.21 (0.13)
	Office And Other Furniture	0.28 (0.14)
	Carpet and Rug Mills	0.38 (0.14)
	Tobacco	0.40 (0.11)
	Newspaper Publishers	0.68 (0.15)
	Soap, Cleaning Compounds, and Toilet Preparation	0.75 (0.17)
	Pharmaceuticals and Medicines	0.81 (0.14)
	Periodical, Book, and Other Publishers	0.85 (0.16)
	Medical Equipment and Supplies	0.99 (0.14)
	Apparel	1.00 (0.12)
	Navigational/Measuring/Electromedical/Control Instruments	1.00 (0.15)

Figure A.17. Variance decomposition along supply chains (shock-specific persistence)

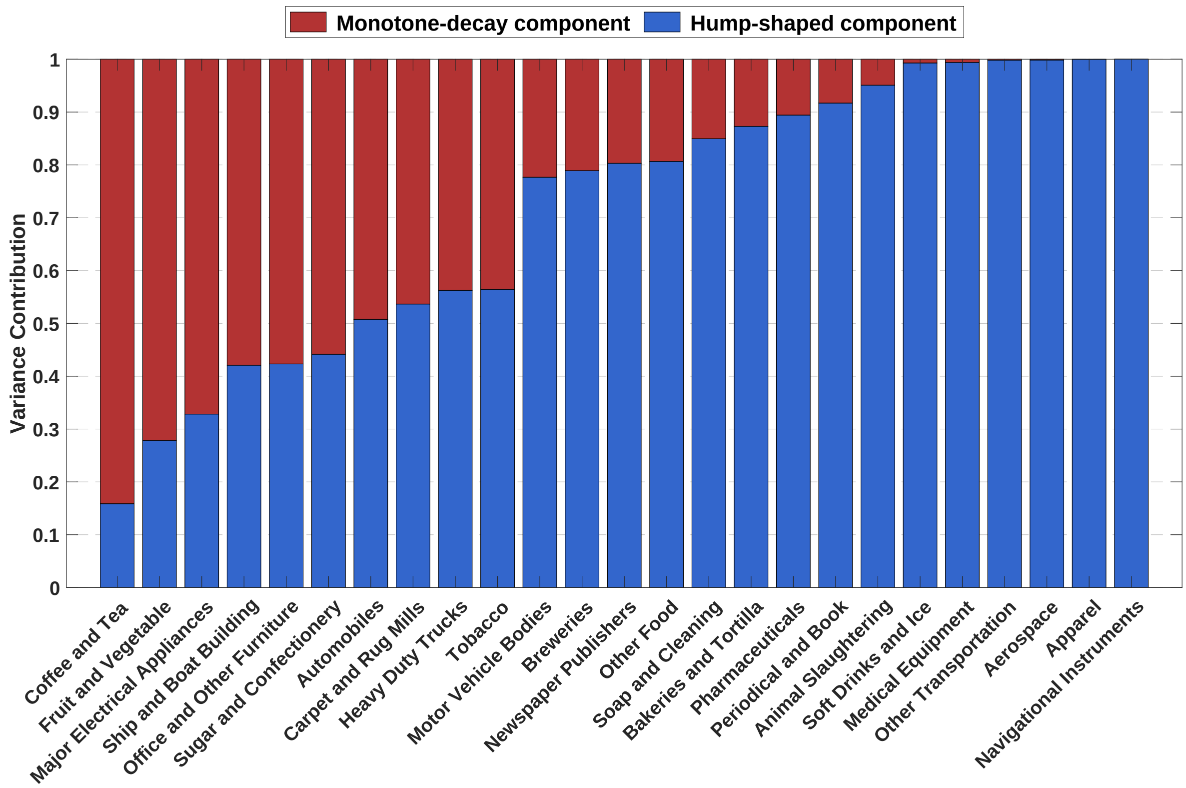


Table A.2. Counterfactual change in supply chain volatility based on the nature of downstream demand shocks (shock-specific persistence)

		% Change in supply chain volatility			Delays along supply chain	
		Only AR(2) shocks	Only AR(1) shocks	Static model	Average	Top 5
Downstream Group		(1)	(2)	(3)	(4)	(5)
(a)	Transportation	106.94	-44.74	-19.55	2.65	4.04
(b)	Food Products	55.04	-49.35	-33.40	2.28	3.32
(c)	Other Final Goods	45.10	-53.79	-39.98	2.31	3.00
(d)	All Sectors	75.05	-45.36	-26.57	2.38	3.35

D.14 Counterfactuals for Each Downstream Industry Individually

In Appendix Table A.3, we report the outcomes for the same set of counterfactual exercises as in Table 3 of the main text, except that we examine each downstream industry individually rather than in groups. Column (3) is positive for the four sectors with the smallest estimated $\hat{\delta}_i$: for AR(1)-dominated demand, the static model tends to overstate volatility, while it understates volatility wherever hump-shaped shocks matter, as the time-to-build mechanism implies. The delay columns report route-weighted average delays based on ξ_{ij} of equation (25), evaluated at $\hat{\rho} = 0.829$, as in Table 3 of the main text.

Table A.3. Counterfactual change in supply chain volatility based on the nature of downstream demand shocks, by downstream industry

	% Change in supply chain volatility			Delays along supply chain	
	Only AR(2) shocks	Only AR(1) shocks	Static model	Average	Top 5
Downstream Sector	(1)	(2)	(3)	(4)	(5)
Sugar and Confectionery Products	477.23	-30.53	61.04	1.89	2.75
Major Electrical Household Appliances	375.64	-39.81	35.81	1.85	3.17
Coffee and Tea	412.57	-44.94	38.45	2.11	2.85
Fruit and Vegetable Preserving and Specialty Foods	266.45	-57.89	1.11	2.07	3.22
Automobiles and Light Duty Motor Vehicles	282.27	-64.37	-0.44	2.43	3.74
Heavy Duty Trucks	231.47	-69.38	-10.72	2.49	3.83
Office And Other Furniture	177.47	-61.52	-18.20	1.82	3.03
Ship and Boat Building	184.53	-66.44	-18.71	2.05	3.34
Carpet and Rug Mills	179.98	-67.30	-21.51	2.07	2.98
Tobacco	157.31	-70.87	-29.04	2.23	2.23
Motor Vehicle Bodies and Trailers	122.49	-74.03	-37.35	2.16	3.89
Breweries	79.44	-77.61	-49.23	2.05	2.99
Other Food Except Coffee and Tea	69.24	-76.80	-50.54	1.90	2.84
Soap, Cleaning Compounds, and Toilet Preparation	64.51	-78.79	-52.49	1.90	2.45
Bakeries and Tortilla	57.01	-79.00	-54.42	1.97	3.02
Newspaper Publishers	56.74	-80.96	-55.36	2.22	2.44
Pharmaceuticals and Medicines	51.11	-84.10	-58.60	2.24	2.24
Animal Slaughtering and Processing	40.06	-88.89	-64.20	2.86	3.65
Other Transportation Equipment	33.43	-86.77	-61.75	2.39	3.60
Soft Drinks and Ice	28.22	-83.33	-62.98	2.02	3.21
Periodical, Book, and Other Publishers	29.01	-85.05	-64.03	2.44	2.44
Aerospace Products and Parts	29.23	-87.75	-61.71	2.46	3.25
Medical Equipment and Supplies	23.44	-86.27	-65.05	2.24	2.83
Apparel	11.77	-85.82	-67.89	2.18	3.08
Navigational/Measuring/Electromedical/Control Instruments	0.08	-87.46	-71.29	2.10	3.22