Supplementary Material for "Incentive-Aware Models of Financial Networks"

Akhil Jalan
Department of Computer Science, University of Texas at Austin, Austin, TX 78712, akhiljalan@utexas.edu

Deepayan Chakrabarti
McCombs School of Business, University of Texas at Austin, Austin, TX 78712, deepay@utexas.edu

Purnamrita Sarkar
Department of Statistics and Data Science, University of Texas at Austin, Austin, TX 78712, purna.sarkar@austin.utexas.edu

---

Appendix A: Proofs

A.1. Example of Stable Network

To illustrate Theorem 1, consider the following example.

Example 1 (Stable points). Consider a 3-firm network where the only allowed edges are given by $F = \{(1,2),(1,3)\}$. Suppose firms share the same covariance belief matrix $\Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma$, but have different mean beliefs $M = \begin{bmatrix}\boldsymbol{\mu}_1 & \boldsymbol{\mu}_2 & \boldsymbol{\mu}_3\end{bmatrix}$ and risk aversions. The firms' beliefs are:

$$M = \begin{bmatrix} 0 & 2/3 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{bmatrix}, \gamma_1 = 1, \gamma_2 = 1/2, \gamma_3 = 1/4$$

Then the $A, B_{(i,j)}$ matrices in Theorem 1 are given as:

$$A = \begin{bmatrix} 0 & 2/3 & 3/4 \\ 1/2 & 0 & 0 \\ 1/2 & 0 & 0 \end{bmatrix}, B_{(1,2)} = \begin{bmatrix} 0 & 3/4 & -1/4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$B_{(1,3)} = \begin{bmatrix} 0 & -1/4 & 3/4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B_{(2,1)} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$B_{(3,1)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3/2 & 0 & 0 \end{bmatrix}$$

Hence

$$C_{(1,2)} = \frac{1}{4}\begin{bmatrix} 0 & 7 & -1 \\ -7 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, C_{(1,3)} = \frac{1}{4}\begin{bmatrix} 0 & -1 & 9 \\ 1 & 0 & 0 \\ -9 & 0 & 0 \end{bmatrix},$$

Therefore, $Z_F = \frac{1}{4}\begin{bmatrix} 7 & -1 \\ -1 & 9 \end{bmatrix}$ and $\text{uvec}(A - A^T)_F = (1/6, 1/4)^T$. Since Z_F is full-rank, there exists a unique stable point for this network setting.

A.2. Stable Points are Common

LEMMA 1. *Define F, Z_F and Q_i as in Theorem 1. Let Q'_i be such that*

$$(Q'_i)_{j,k} = \begin{cases} (Q_i)_{j,k} + \beta & \text{if } j = k, (i,j) \in F \\ (Q_i)_{j,k} & \text{otherwise} \end{cases}$$

Then, the corresponding Z'_F has the form $Z'_F = Z_F + \beta I$.

Proof of Lemma 1. This follows from the form of the matrices $B_{(i,j)}$ and $C_{(i,j)}$ in the statement of Theorem 1. □

Now, we consider the Σ_i's (and hence the Q_i's) to be random variables. Any distribution of $\{\Sigma_i\}_{i\in[n]}$ induces a distribution on $\{Q_i\}_{i\in[n]}$, where $Q_i \succ 0$. Define $\tilde{Q}_i := Q_i - \delta I$, where $\delta > 0$ is the minimum of union of the (nonzero) eigenvalues of all the Q_i's. A distribution over $\{Q_i\}$ corresponds to a distribution over $(\{\tilde{Q}_i\}, \delta)$.

PROPOSITION 1. *If the distribution of δ given $\{\tilde{Q}_i\}_{i\in[n]}$ is continuous, then a unique stable point exists with probability* 1.

Proof of Proposition 1. Let $\tilde{Z}_F$ be the $|F| \times |F|$ matrix generated from $\{\tilde{Q}_i\}_{i\in[n]}$, and Z_F the corresponding matrix for $\{Q_i\}_{i\in[n]}$. By Lemma 1, $Z_F = \tilde{Z}_F + \delta I$. Hence, $\sigma(Z_F) = \sigma(\tilde{Z}_F) + \delta$, where $\sigma(M)$ denote the set of eigenvalues of M. Since $\sigma(\tilde{Z}_F)$ is a function of $\{\tilde{Q}_i\}$ and δ is continuous given $\{\tilde{Q}_i\}$, the eigenvalues of Z_F are non-zero with probability 1. Hence, by Theorem 1, a unique stable point exists for $\{Q_i\}$ with probability 1. □

Note that we require no condition on the distribution of $\{\tilde{Q}_i\}$. The condition of Proposition 1 is satisfied if the joint distribution of the $\{\Sigma_i\}_{i\in[n]}$ is continuous and all edges are permitted, as shown in the following example.

EXAMPLE 2. Fix some $n \geq 2$. Suppose the joint distribution of the $\{\Sigma_i\}_{i\in[n]}$ is continuous and all edges are permitted. Then $Q_i = (2\gamma_i)^{-1}\Sigma_i^{-1}$ so the joint distribution of $\{Q_i\}_{i\in[n]}$ is continuous. By Bayes' rule, $\mathbb{P}[\delta|\tilde{Q}_1, \ldots, \tilde{Q}_n] \propto \mathbb{P}[\delta, \tilde{Q}_1, \ldots, \tilde{Q}_1] = \mathbb{P}[Q_1, \ldots, Q_n]$. Since $\mathbb{P}[Q_1, \ldots, Q_n]$ is continuous, we conclude $\mathbb{P}[\delta|\tilde{Q}_1, \ldots, \tilde{Q}_n]$ is continuous.

A.3. Proof of Theorem 2

RESTATEMENT OF THEOREM 2 . *Let (W^*, P^*) be a stable feasible point. Then there is no feasible (W, P) such that $(W, P) \succ (W^*, P*)$.*

Proof of Theorem 2. *Case 1: $P = P^*$.* First, consider a feasible (W, P) such that $P = P^*$. Then $W \neq W^*$. Since W^* is stable, by definition each agent optimizes contracts with respect to P^*, so no agent is worse off under (W^*, P^*) then (W, P^*). Hence $(W, P) \not\succ (W^*, P^*)$.

Case 2: $P \neq P^*$. Second, suppose that $P \neq P^*$. Let $\Delta_i := g_i(W, P) - g_i(W, P^*)$. It follows that $\Delta_i = (W\boldsymbol{e}_i)^T((P^* - P)\boldsymbol{e}_i)$. Let $A \in \mathbb{R}^{n\times n}$ be defined as $A_{ij} = W_{ij}(P^*_{ij} - P_{ij})$. Then $\Delta_i = \boldsymbol{e}_i^T A\mathbf{1}$.

Next, notice that $A_{ji} = -A_{ij}$. Therefore, $\sum_i \Delta_i = \mathbf{1}^T A\mathbf{1} = 0$. Hence, either $\Delta_i = 0$ for all i, or there exists k such that $\Delta_k < 0$.

Case 2(i). Suppose there exists k such that $\Delta_k < 0$. Then $g_k(W, P) < g_k(W, P^*)$. By case 1, we have $g_k(W, P^*) \leq g_k(W^*, P^*)$. Therefore agent k is strictly worse off, so $(W, P) \not\succ (W^*, P^*)$.

Case 2(ii). Suppose $\Delta_i = 0$ for all i. Then $g_i(W, P) = g_i(W, P^*)$ for all i. By case 1, we have $g_i(W, P^*) \leq g_i(W^*, P^*)$. Therefore no agent is better off, so $(W, P) \not\succ (W^*, P^*)$. □

A.4. Proof of Theorem 3

RESTATEMENT OF THEOREM 3 *Any stable point (W, P) is Higher-Order Nash Stable.*

Proof of Theorem 3. First, we argue (W, P) is a Nash equilibrium. Suppose that agent i wants to shift some of their contracts at the stable feasible point (W, P). Suppose they propose $(w'_{i,j_1}, p'_{i,j_1}), \ldots, (w'_{i,j_m}, p'_{i,j_m})$ for $j_1, \ldots, j_m \in [n]$. Let (W', P') denote the new feasible point that occurs if all changes are accepted. By Theorem 2 we know that $(W', P') \not\succ (W, P)$, so at least one agent does not prefer (W', P'). Since the only changes are to edges $\{i, j_1\}, \ldots, \{i, j_m\}$, there must exist a $j \in \{j_1, \ldots, j_m\}$ who does not prefer (W', P'). Therefore, they will reject the proposal of agent i to shift to (w'_{ij}, p'_{ij}).

Then, agent i can choose to either maintain the existing contract (w_{ij}, p_{ij}) or delete the edge $\{i, j\}$. We claim that agent i prefers to keep the edge, since they could have chosen to set $w_{ij} = 0$ during the network formation process, no matter what price was offered. But $w_{ij} \neq 0$ at equilibrium (W, P). By stability of (W, P) we know w_{ij} is the optimal choice for agent i at prices P. Therefore, after agent j rejects (w'_{ij}, p'_{ij}), it follows that the edge remains at (w_{ij}, p_{ij}).

Since (W', P') was arbitrary, we conclude that at equilibrium, agent i cannot propose any set of changes that result in a strictly better network for them. Therefore, their optimal action at (W, P) is to not deviate from the equilibrium.

Next, we show cartel robustness. Suppose $S \subset [n]$ is a strict subset and $(W', P') \neq (W, P)$ is a feasible point differing only at indices $\{i, j\}$ such that $i, j \in S$. By Theorem 2, we know (W', P') cannot dominate (W, P), so there is some agent $i \in [n]$ that does not prefer (W', P') to (W, P).

Since (W', P') only changes contracts where both members are in S, the utility of agents in $[n] \setminus S$ must be unchanged. Therefore $i \in S$, and hence not all members of the cartel have higher utility under (W', P'). $\square$

A.5. Price Update Rule for Pairwise Negotiations

We give an explicit formula for the updated price of a unit contract after a pairwise negotiation.

PROPOSITION 2 **(Price after Pairwise Negotiation)**. *Consider a network setting $(\boldsymbol{\mu}_i, \gamma_i, \Sigma_i, \Psi_i)_{i\in[n]}$. Let Q_i be as in Theorem 1. Given a price matrix $P = -P^T$ and a pair of firms (i, j) that are permitted to trade, let P' be another skew-symmetric price matrix such that (a) P' differs from P only in the cells (i, j) and (j, i), (b) i and j both maximize their utility at the same contract size under P', and (c) i and j can choose their optimal contract sizes with all other agents given these prices. Then,*

$$P'_{ij} = \frac{1}{Q_{i;j,j} + Q_{j;i,i}} \Big(\boldsymbol{e}_i^T Q_j (M - P) \boldsymbol{e}_j - \boldsymbol{e}_j^T Q_i (M - P) \boldsymbol{e}_i \Big) + P_{ij}$$

Proof. Let $A_i := \gamma_i Q_i$ for $i \in [n]$. Since $\Sigma_i \succ 0$ and $\Psi_i \Sigma_i \Psi_i^T$ is a principal submatrix, we know $\Psi_i \Sigma_i \Psi_i^T$ is real symmetric and positive definite, and hence its inverse is as well. Therefore A_i is real symmetric and PSD. (It is not full rank in general, unless $\Psi_i = I$). Since $\{i, j\}$ is a permitted edge, $\Psi_i \boldsymbol{e}_j \neq \boldsymbol{0}$ and $\Psi_j \boldsymbol{e}_i \neq \boldsymbol{0}$. Therefore $A_{i;j,j} = \boldsymbol{e}_j^T A_i \boldsymbol{e}_j = (\Psi_i \boldsymbol{e}_j)^T (2\Psi_i \Sigma_i \Psi_i^T)^{-1} (\Psi_i \boldsymbol{e}_j) > 0$ since $(2\Psi_i \Sigma_i \Psi_i^T)^{-1}$ is positive definite. So, $A_{i;j,j} > 0$ and similarly $A_{j;i,i} > 0$.

Now, the optimal contracts for agent i under prices P' are given by $\boldsymbol{w}_i = A_i (M - P') \Gamma^{-1} \boldsymbol{e}_i$. Note that $P' = P + (P'_{ij} - P_{ij})(\boldsymbol{e}_i \boldsymbol{e}_j^T - \boldsymbol{e}_j \boldsymbol{e}_i^T)$. Since both i and j maximize their utility at the same contract size, we have:

$$\begin{aligned}
\boldsymbol{w}_{i;j} &= \boldsymbol{w}_{j;i} \\
\Rightarrow \boldsymbol{e}_j^T \boldsymbol{w}_i &= \boldsymbol{e}_i^T \boldsymbol{w}_j \\
\Rightarrow \boldsymbol{e}_j^T (A_i (M - P') \Gamma^{-1}) \boldsymbol{e}_i &= \boldsymbol{e}_i^T (A_j \\
&\qquad (M - P') \Gamma^{-1}) \boldsymbol{e}_j \\
\Rightarrow \gamma_j \boldsymbol{e}_j^T A_i M \boldsymbol{e}_i - \gamma_i \boldsymbol{e}_i^T A_j M \boldsymbol{e}_j &= \gamma_j \boldsymbol{e}_j^T A_i P' \boldsymbol{e}_i \\
&\quad - \gamma_i \boldsymbol{e}_i^T A_j P' \boldsymbol{e}_j
\end{aligned}$$

The last line can written:

$$\begin{aligned}
&\gamma_j \boldsymbol{e}_j^T A_i P \boldsymbol{e}_i - \gamma_i \boldsymbol{e}_i^T A_j P \boldsymbol{e}_j \\
-&(P'_{ij} - P_{ij}) \left(\gamma_j \boldsymbol{e}_j^T A_i \boldsymbol{e}_j + \gamma_i \boldsymbol{e}_i^T A_j \boldsymbol{e}_i \right)
\end{aligned}$$

Hence, we can write $(P'_{ij} - P_{ij})$ as:

$$
\begin{aligned}
P'_{ij} - P_{ij} = &\frac{1}{\gamma_j A_{i;j,j} + \gamma_i A_{j;i,i}} \Big(\boldsymbol{e}_i^T \Gamma A_j (M-P)\boldsymbol{e}_j \\
&- \boldsymbol{e}_j^T \Gamma A_i (M-P)\boldsymbol{e}_i \Big) \\
= &\frac{1}{Q_{i;j,j} + Q_{j;i,i}} \Big(\boldsymbol{e}_i^T Q_j (M-P)\boldsymbol{e}_j \\
&- \boldsymbol{e}_j^T Q_i (M-P)\boldsymbol{e}_i \Big)
\end{aligned}
$$

□

A.6. Proof of Theorem 4

First, we characterize pairwise negotiation dynamics as linear in the price updates.

Theorem 1. *Consider a network setting* $(\boldsymbol{\mu}_i, \gamma_i, \Sigma_i, \Psi_i)_{i\in[n]}$. *Define* Q_i *as in Theorem 1. Let* $s_{ij} = 1$ *if* $\{i,j\}$ *is a permitted edge and* 0 *otherwise. Let* $L, R \in \mathbb{R}^{n^2\times n^2}$ *be diagonal matrices such that* $L_{(i-1)n+j,(i-1)n+j} = Q_{i;jj} + Q_{j;ii}$ *and* $R_{(i-1)n+j,(i-1)n+j} = s_{ij}$, *and* $L^\dagger$ *be the pseudoinverse of* L. *Let* $\Delta_{(t+1)} = P(t+1) - P(t)$, *where* $P(t)$ *is the price matrix at time step* t *of pairwise negotiations. Then,*

$$
vec(\Delta_{(t+1)}) = R\Big(I_{n^2} - \eta L^\dagger K\Big) vec(\Delta_{(t)}),
$$
$$
\text{where } K = \sum_{r=1}^{n} \big(\boldsymbol{e}_r\boldsymbol{e}_r^T \otimes Q_r + Q_r \otimes \boldsymbol{e}_r\boldsymbol{e}_r^T\big).
$$

Proof. Let $\{i,j\}$ be a permitted edge. From Proposition 2, we obtain:

$$
\begin{aligned}
(\Delta_{(t+1)})_{ij} = &\frac{\eta}{Q_{i;j,j} + Q_{j;i,i}} \Big(\\
&\boldsymbol{e}_i^T Q_j (M - P(t))\boldsymbol{e}_j \\
&- \boldsymbol{e}_j^T Q_i (M - P(t))\boldsymbol{e}_i \Big) \\
\Rightarrow (\Delta_{(t+1)})_{ij} - (\Delta_{(t)})_{ij} = &\frac{\eta}{Q_{i;j,j} + Q_{j;i,i}} \Big(\\
&\boldsymbol{e}_i^T Q_j (-\Delta_{(t)})\boldsymbol{e}_j \\
&- \boldsymbol{e}_j^T Q_i (-\Delta_{(t)})\boldsymbol{e}_i \Big) \\
= &\frac{-\eta}{Q_{i;j,j} + Q_{j;i,i}} \Big(\\
&\boldsymbol{e}_i^T Q_j \Delta_{(t)}\boldsymbol{e}_j \\
&- \boldsymbol{e}_j^T Q_i \Delta_{(t)}\boldsymbol{e}_i \Big) \\
= &\frac{-\eta}{Q_{i;j,j} + Q_{j;i,i}} \boldsymbol{e}_i^T \Big(Q_j \Delta_{(t)} \\
&- (Q_i \Delta_{(t)})^T \Big) \boldsymbol{e}_j
\end{aligned}
$$

Hence,

$$(Q_{i;j,j} + Q_{j;i,i})\Big((\Delta_{(t+1)})_{ij} - (\Delta_{(t)})_{ij}\Big) = -\eta s_{ij}$$
$$\cdot \boldsymbol{e}_i^T \Big(Q_j \Delta_{(t)} + \Delta_{(t)} Q_i\Big) \boldsymbol{e}_j.$$

We assumed that $\{i, j\}$ was a permitted edge above, but notice the identity is also true for prohibited $\{i, j\}$ since both the numerator and denominator become 0, and we can define their ratio to be 0. Defining $Y_{ij} = \boldsymbol{e}_i^T \Big(Q_j \Delta_{(t)} + \Delta_{(t)} Q_i\Big) \boldsymbol{e}_j$, and recalling the definitions of L and R from the theorem statement, the above formula becomes

$$L\text{vec}(\Delta_{(t+1)} - \Delta_{(t)}) = -\eta R\text{vec}(Y). \tag{1}$$

We show next that $\text{vec}(Y) = K\text{vec}(\Delta_{(t)})$, where K is defined in the theorem statement. Let tr denote the trace operator. Then $(\boldsymbol{e}_j^T \otimes \boldsymbol{e}_i^T)\text{vec}(Y) = Y_{ij}$. Hence,

$$\begin{aligned}
Y_{ij} &= \boldsymbol{e}_i^T \Big(Q_j \Delta_{(t)} + \Delta_{(t)} Q_i\Big) \boldsymbol{e}_j \\
&= tr\Big(\boldsymbol{e}_i^T Q_j \Delta_{(t)} \boldsymbol{e}_j\Big) + tr\Big(\boldsymbol{e}_i^T \Delta_{(t)} Q_i \boldsymbol{e}_j\Big) \\
&= tr\Big(\boldsymbol{e}_j^T \Delta_{(t)}^T Q_j^T \boldsymbol{e}_i\Big) + tr\Big(\boldsymbol{e}_i^T \Delta_{(t)} Q_i \boldsymbol{e}_j\Big) \\
&= tr\Big(\Delta_{(t)}^T Q_j^T \boldsymbol{e}_i \boldsymbol{e}_j^T\Big) + tr\Big(Q_i \boldsymbol{e}_j \boldsymbol{e}_i^T \Delta_{(t)}\Big) \\
&= \text{vec}(\Delta_{(t)})^T \text{vec}(Q_j^T \boldsymbol{e}_i \boldsymbol{e}_j^T + (Q_i \boldsymbol{e}_j \boldsymbol{e}_i^T)^T) \\
&= \text{vec}(Q_j \boldsymbol{e}_i \boldsymbol{e}_j^T + \boldsymbol{e}_i \boldsymbol{e}_j^T Q_i)^T \text{vec}(\Delta_{(t)}),
\end{aligned}$$

where we used $Q_i = Q_i^T$.

Hence we need to show $(\boldsymbol{e}_j^T \otimes \boldsymbol{e}_i^T)K = \text{vec}(Q_j \boldsymbol{e}_i \boldsymbol{e}_j^T + \boldsymbol{e}_i \boldsymbol{e}_j^T Q_i)^T$. Letting δ denote the Kronecker delta, we obtain:

$$(\boldsymbol{e}_j^T \otimes \boldsymbol{e}_i^T)K = (\boldsymbol{e}_j^T \otimes \boldsymbol{e}_i^T)\Big(\sum_{r=1}^n \boldsymbol{e}_r \boldsymbol{e}_r^T \otimes Q_r \tag{2}$$
$$+ Q_r \otimes \boldsymbol{e}_r \boldsymbol{e}_r^T\Big)$$
$$= \sum_{r=1}^n \Big(\delta_{jr}(\boldsymbol{e}_j^T \otimes \boldsymbol{e}_i^T Q_r) \tag{3}$$
$$+ \delta_{ir}(\boldsymbol{e}_j^T Q_r \otimes \boldsymbol{e}_i^T)\Big)$$
$$= (\boldsymbol{e}_j^T \otimes \boldsymbol{e}_i^T Q_j) + (\boldsymbol{e}_j^T Q_i \otimes \boldsymbol{e}_i^T)$$
$$= (\boldsymbol{e}_j \otimes Q_j \boldsymbol{e}_i + Q_i \boldsymbol{e}_j \otimes \boldsymbol{e}_i)^T. \tag{4}$$

Now, we observe that $\boldsymbol{e}_j \otimes Q_j \boldsymbol{e}_i$ is the vectorization of a matrix whose j^{th} column is $Q_j \boldsymbol{e}_i$, i.e., the matrix $Q_j \boldsymbol{e}_i \boldsymbol{e}_j^T$. Similarly, $Q_i \boldsymbol{e}_j \otimes \boldsymbol{e}_i$ is the vectorization of a matrix whose i^{th} row is $(Q_i \boldsymbol{e}_j)^T$, i.e., the matrix $\boldsymbol{e}_i \boldsymbol{e}_j^T Q_i$. Hence, $(\boldsymbol{e}_j^T \otimes \boldsymbol{e}_i^T)K = \text{vec}(Q_j \boldsymbol{e}_i \boldsymbol{e}_j^T + \boldsymbol{e}_i \boldsymbol{e}_j^T Q_i)^T$, as desired.

Plugging into Eq. (1),

$$\begin{aligned}
L\text{vec}(\Delta_{(t+1)}-\Delta_{(t)}) &= -\eta RK\text{vec}(\Delta_{(t)}) \\
\Rightarrow L\text{vec}(\Delta_{(t+1)}) &= L\text{vec}(\Delta_{(t)}) - \eta RK\text{vec}(\Delta_{(t)}) \\
\Rightarrow \text{vec}(\Delta_{(t+1)}) &= \Big(L^\dagger L - \eta L^\dagger RK\Big)\text{vec}(\Delta_{(t)}) \\
\Rightarrow \text{vec}(\Delta_{(t+1)}) &= \Big(R - \eta RL^\dagger K\Big)\text{vec}(\Delta_{(t)}) \\
&= R\Big(I_{n^2} - \eta L^\dagger K\Big)\text{vec}(\Delta_{(t)}),
\end{aligned}$$

where we used the facts that $(\Delta_t)_{ij} = (\Delta_{(t+1)})_{ij} = 0$ for disallowed edges, and $L^\dagger L = R$ and $LR = RL = L$, which can be easily confirmed by inspection of these diagonal matrices. □

We use Lyapunov theory to analyze the convergence of pairwise negotiation dynamics. In particular, we need the the discrete Lyapunov equation, also called the Stein equation.

THEOREM 2 **(Callier and Desoer (1994) 7.d)**. *For the discrete-time dynamical system $\boldsymbol{x}_{t+1} = A\boldsymbol{x}_t$, with $\boldsymbol{x}_t \in \mathbb{R}^n$, the following are equivalent:*

1. *The system is globally asymptotically stable towards* $\mathbf{0}$.
2. *For any positive definite $R \in \mathbb{R}^{n\times n}$, there exists a unique solution $X \succ 0$ to the equation*

$$AXA^T - X = -R$$

3. *For any eigenvalue λ of A, $|\lambda| < 1$.*

Pairwise negotiation dynamics can be described as a discrete-time linear system in $\text{vec}(\Delta_t)$, where Δ_t is the price difference at time t. Clearly, the system converges iff Δ_t approaches zero. Therefore, we can use the Stein equation to prove global asymptotic stability conditions.

We will also need the *commutation matrix.*

LEMMA 2 **(Horn and Johnson (1994))**. *Let $\Pi^{(n,n)} : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$ be a permutation matrix (called the (n,n)* commutation matrix*) defined as $\Pi^{(n,n)} = \sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{e}_i\boldsymbol{e}_j^T \otimes \boldsymbol{e}_j\boldsymbol{e}_i^T$. Then for any $A, B \in \mathbb{R}^{n\times n}$, we have*

$$A \otimes B = \Pi^{(n,n)}(B \otimes A)(\Pi^{(n,n)})^T$$

Recall that for a linear operator T that $\sigma(T)$ denotes the eigenvalues of T. We are ready to prove Part 1 of Theorem 4.

PROPOSITION 3 **(Part 1 of Theorem 4)**. *Let L, R, K be defined as in Theorem 1. For a matrix $X \in \mathbb{R}^{n^2\times n^2}$ let $X\,|_R$ denote the principal submatrix of X corresponding to the nonzero rows/columns of R. Define $\eta^\star = \min_{\lambda\in\sigma((L^\dagger K)|_R)} \frac{2}{\lambda}$. Then, for any $\eta \in (0, \eta^\star)$, vec$(\Delta_{(t)})$ is globally asymptotically stable towards* $\mathbf{0}$.

Proof of Proposition 3. Let $T = R(I - \eta L^\dagger K)$. By Theorem 2, the dynamics are globally asymptotically stable towards $\mathbf{0}$ iff for all $\lambda \in \sigma(T)$, we have $|\lambda| < 1$.

From Eq. (4) for a prohibited edge (i,j), we see that $(\boldsymbol{e}_j^T \otimes \boldsymbol{e}_i^T)K = \mathbf{0}^T$, since $Q_i\boldsymbol{e}_j = \mathbf{0} = Q_j\boldsymbol{e}_i$. Hence, $K = RK$. Taking transposes and noting that both K and R are symmetric, we find $KR = K$. Hence, $T = R(I - \eta L^\dagger K) = R(I - \eta L^\dagger K)R$, where we used $R^2 = R$. Thus, T is zero except for the principal submatrix corresponding to the nonzero columns of R. So, to apply Theorem 2, we only require $|\lambda| < 1$ for $\lambda \in \sigma(T\,|_R)$.

For clarity of exposition we will first consider the case where $R = I$ (no prohibited edges). Then, the eigenvalues of $T\,|_R = T$ equal $1 - \eta\lambda$, where $\lambda \in \sigma(L^{-1}K) = \sigma(L^{-1/2}KL^{-1/2})$ by a similarity transformation. Also, $K = U_1 + U_2$, where $U_1 = \sum_{r=1}^{n}\left(\boldsymbol{e}_r\boldsymbol{e}_r^T \otimes Q_r\right)$ and $U_2 := \sum_{r=1}^{n}\left(Q_r \otimes \boldsymbol{e}_r\boldsymbol{e}_r^T\right)$. The matrix U_1 is block diagonal with positive-definite blocks $Q_r \succ 0$, so $U_1 \succ 0$. By Lemma 2, U_2 is similar to U_1 via a permutation matrix, so $U_2 \succ 0$. Hence, $K \succ 0$, and $L^{-1/2}KL^{-1/2} \succ 0$. So, the eigenvalues of $L^{-1}K$ are real and positive. Hence, we have convergence iff for all $\lambda \in \sigma(L^{-1}K)$, we have $1 > (1-\eta\lambda)^2 = 1 - 2\eta\lambda + \eta^2\lambda^2$. i.e., $\lambda < 2/\eta$. Hence, $\eta^\star = 2/\|L^{-1}K\|$ as required.

Now we consider the prohibited edges setting ($R \neq I$). Here, convergence occurs iff $|1-\eta\lambda| < 1$ for all $\lambda \in \sigma((L^\dagger K)\,|_R)$. Since $RL^\dagger R = L^\dagger$ and $RKR = K$, we have $(L^\dagger K)\,|_R = L^\dagger\,|_R\, K\,|_R = (L\,|_R)^{-1}K\,|_R$. Arguing as above, it suffices to show that $K\,|_R \succ 0$. We claim $K\,|_R = V_1 + V_2$ where V_1 is a block diagonal matrix with i^{th} block equal to $(2\gamma_i\Psi_i\Sigma_i\Psi_i^T)^{-1} \succ 0$, and V_2 is similar to V_1 via Lemma 2. Hence $K\,|_R \succ 0$ and the expression for $\eta^\star$ follows. □

Proposition 4 **(Part 2 of Theorem 4)**. *We define $\eta^\star$ as in Proposition 3, and L, R, K, α as in Theorem 1. Let $\eta \in (0, \eta^\star)$. Then,*

$$\|P(t) - P^\star\|_F \leq \frac{\alpha^t}{1-\alpha}\cdot\|P(1) - P(0)\|_F$$

Here, $P^\star$ is the stable point to which the negotiation converges.

Proof. Let β denote the greatest eigenvalue in absolute value of $R(I_{n^2} - \eta L^\dagger K)$. From Theorem 1, we have $\|\Delta_{t+1}\|_F \leq |\beta|\|\Delta_t\|_F$. Recall that $\lambda_{\max}, \lambda_{\min}$ denote largest and smallest eigenvalues of the matrix $(L^\dagger K)\,|_R$ respectively. Since $\|R\| = 1$, it follows that $|\beta| = \max\{|1-\eta\lambda_{\min}|, |1-\eta\lambda_{\max}|\} = \alpha$.

Then,

$$\begin{aligned}
\|P^\star - P(t)\|_F &\leq \sum_{i>t}\|\Delta_i\|_F \\
&\leq \|\Delta_t\|_F(\alpha + \alpha^2 + \ldots) \\
&\leq \|\Delta_t\|_F\frac{\alpha}{1-\alpha} \\
&\leq (\alpha^{t-1}\|\Delta_1\|_F)\frac{\alpha}{1-\alpha} \\
&= \|\Delta_1\|_F\frac{\alpha^t}{1-\alpha}
\end{aligned}$$

Since $\|\Delta_1\|_F = \|P(1) - P(0)\|_F$ we are done. $\square$

A.7. Example of Convergence Conditions and Rate

The following example illustrates Theorem 4 in the setting of Example 1 (Appendix A.1).

EXAMPLE 3 (CONVERGENCE CONDITIONS AND RATE). In the setting of Example 1, we have

$$Q_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2/3 & -1/3 \\ 0 & -1/3 & 2/3 \end{bmatrix}, Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$
$$Q_3 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

Hence

$$K = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{3}+1 & \frac{-1}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{-1}{3} & \frac{2}{3}+2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1+\frac{2}{3} & 0 & 0 & \frac{-1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{-1}{3} & 0 & 0 & 2+\frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Also, L is the diagonal matrix with $L_{i,i} = K_{i,i}$ for $i \in [9]$. Since the permitted edges are $\{(1,2),(1,3)\}$, $R = \{2,3\}$ and so $(L^\dagger K)_R = \begin{bmatrix} 1 & \frac{-1}{5} \\ \frac{-1}{8} & 1 \end{bmatrix}$. Hence $\lambda_{min} = 1 - \frac{1}{2\sqrt{10}}, \lambda_{max} = 1 + \frac{1}{2\sqrt{10}}$, and $\eta^* = \frac{2}{1+(40)^{-1/2}} \approx 1.727$.

It follows that pairwise negotiations with $\eta \in (0, \frac{2}{1+(40)^{-1/2}})$ are globally asymptotically stable. Suppose that $\eta = 0.99$. Then $\alpha = (1 - \eta \cdot (1 - \frac{1}{2\sqrt{10}})) \approx 0.17$. Hence after t rounds, the distance of $P(t)$ to P^* shrinks by a factor of $\approx \frac{0.17^t}{0.83}$.

A.8. Proof of Theorem 5

We will use a series of Lemmas to reduce the result of Theorem 5 to a matrix concentration inequality in each of the $\hat{\Sigma}_i$.

LEMMA 3. *Let $\widehat{\eta^*}, \eta^*$ be as in Theorem 5. Suppose all edges are permitted.*

Suppose that for all $i \in [n]$, we have $\|\hat{\Sigma}_i^{-1} - \Sigma^{-1}\| = o(1)$. Then, $\hat{\eta}^\star > \eta^\star(1 - o(1))$.

Proof. Let $\hat{L}, \hat{K} \in \mathbb{R}^{n^2 \times n^2}$ be as in Theorem 5, but built using $\hat{\Sigma}_1, \ldots, \hat{\Sigma}_n$ instead of $\Sigma, \ldots, \Sigma$. Let L, K be defined similarly to $\hat{L}, \hat{K}$ but using Σ in place of all $\hat{\Sigma}_i$.

Then $\widehat{\eta^*} := \frac{2}{\max \sigma(\hat{L}^{-1}\hat{K})}$ and $\eta^* := \frac{2}{\max \sigma(L^{-1}K)}$.

Let $\epsilon_L, \epsilon_K \in \mathbb{R}^{n^2 \times n^2}$ be such that $\hat{L}^{-1} = L^{-1} + \epsilon_L$ and $\hat{K} = K + \epsilon_K$. We will bound $\|\epsilon_L\|, \|\epsilon_K\|$.

Let $Q_i, \hat{Q}_i$ be defined as in Theorem 1, so $Q_i := (2\gamma_i\Sigma)^{-1}$ and $\hat{Q}_i := (2\gamma_i\hat{\Sigma}_i)^{-1}$. Let $\alpha = \max_{i \in [n]} \|\hat{Q}_i - Q_i\|$. Notice $\|\Gamma^{-1}\| = O(1)$, so $\alpha = o(1)$.

First, since L is diagonal, $\|\epsilon_L\| \leq \max_{i,j\in[n]} \big((\hat{Q}_{i;jj} - Q_{i;jj}) + (\hat{Q}_{j;ii} - Q_{j;ii})\big) \leq 2 \max_{i,j\in[n]} \big(\hat{Q}_{i;jj} - Q_{i;jj}\big) \leq 2\max_{i\in[n]} \|\hat{Q}_i - Q_i\| = 2a$.

Second, let $\hat{K} := \hat{U}_1 + \hat{U}_2$ where $\hat{U}_1, \hat{U}_2$ are defined analogously to U_1, U_2 in the proof of Theorem 4. Letting Π be the (n,n) commutation matrix of Lemma 2, we know $\hat{U}_2 = \Pi\hat{U}_1\Pi^T$, so $\|\epsilon_K\| \leq 2\|\hat{U}_1 - U_1\|$. Since $U_1, \hat{U}_1$ are block diagonal with i^{th} blocks $Q_i, \hat{Q}_i$ respectively, it follows $\|\hat{U}_1 - U_1\| = \max_{i\in[n]} \|\hat{Q}_i - Q_i\| = \alpha$. Hence $\|\epsilon_K\| \leq 2\alpha$.

Third, notice that since $\|\Sigma\|$ and $\|\Gamma\|$ are assumed to be $O(1)$ that $\|L^{-1}\| = O(\max_i \|Q_i\|) = O(1)$ and $\|K\| = O(\max_i \|Q_i\|) = O(1)$. So,

$$
\begin{aligned}
\|\hat{L}^{-1}\hat{K} - L^{-1}K\|_2 &\leq \|\epsilon_L\|\|K\| \\
&\quad + \|L^{-1}\|\|\epsilon_K\| + \|\epsilon_L\|\|\epsilon_K\| \\
&\leq 2\alpha(\|K\| + 2\alpha) \\
&\quad + 4\alpha(\|L^{-1}\| + \alpha) \\
&= 4\alpha(\|K\| + \|L^{-1}\|) + 8\alpha^2 \\
&\leq o(1)
\end{aligned}
$$

We conclude that $\|\hat{L}^{-1}\hat{K}\|_2 \leq \|L^{-1}K\|_2 + o(1)$, so $\widehat{\eta^*} \geq \frac{\eta^*}{1+(o(1)/\|L^{-1}K\|)} \geq (1-o(1))\eta^*$. □

LEMMA 4. *Suppose for $i \in [n]$, we have $\delta_i := \|\hat{\Sigma}_i - \Sigma\| = o(1)$. Then $\|\hat{\Sigma}_i^{-1} - \Sigma_i^{-1}\| = o(1)$.*

Proof. Weyl's inequality implies that $\lambda_{min}(\hat{\Sigma}_i) \geq \lambda_{min}(\Sigma) - \|\hat{\Sigma}_i - \Sigma\|$. Therefore,

$$
\begin{aligned}
\|\hat{\Sigma}_i^{-1}\| &= \frac{1}{\lambda_{min}(\hat{\Sigma}_i)} \\
&\leq \frac{1}{\lambda_{min}(\Sigma) - \delta_i} \\
&= \frac{1}{\lambda_{min}(\Sigma)}\Big(1 + \frac{\delta_i}{\lambda_{min}(\Sigma)} \\
&\quad + O\Big(\Big(\frac{\delta_i}{\lambda_{min}(\Sigma)}\Big)^2\Big)\Big) \\
&= \|\Sigma^{-1}\|(1 + o(1)) \\
\Rightarrow \|\hat{\Sigma}_i^{-1} - \Sigma^{-1}\| &= \|\Sigma^{-1}(\Sigma_i - \hat{\Sigma}_i)\hat{\Sigma}_i^{-1}\| \\
&\leq (1 + o(1))\|\Sigma^{-1}\|^2\delta_i \\
&\leq o(1)
\end{aligned}
$$

The last step follows from the fact $\|\Sigma^{-1}\| = O(1)$. □

The hypothesis of Lemma 4 follows from a standard argument on the concentration of random covariance matrices.

Theorem 3. *Under the setting of Theorem 5, with probability at least* $1-e^{-\Omega(n)}$*, we have* $\|\hat{\Sigma}_i - \Sigma\| = o(1)$ *for all* $i \in [n]$.

Proof of Theorem 3. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0}, \Sigma)$ be the samples. Let $\hat{\boldsymbol{\mu}} = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{X}_i$, and $\tilde{\Sigma}_i := \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{X}_i \boldsymbol{X}_i^T$. Then, $\hat{\Sigma}_i = m/(m-1)\cdot(\tilde{\Sigma}_i - \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^T)$. Hence,

$$\|\hat{\Sigma}_i - \Sigma\| \le m/(m-1)\cdot\left(\|\tilde{\Sigma}_i - \Sigma\| + \|\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^T\|\right) = m/(m-1)\left(\|\tilde{\Sigma}_i - \Sigma\| + \|\hat{\boldsymbol{\mu}}\|^2\right).$$

Now, $\hat{\boldsymbol{\mu}} \sim \mathcal{N}(0, \frac{1}{m}\Sigma)$, so $\sqrt{m}\Sigma^{-1/2}\hat{\boldsymbol{\mu}} \sim \mathcal{N}(0, I_n)$. By Vershynin (2018) (4.7.3 and 2.8.3), there exist constants $c, c_2 > 0$ such that for any $u, \epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left[\|\tilde{\Sigma}_i - \Sigma\|_2 \le c\|\Sigma\|_2\left(\sqrt{\frac{n+u}{m}} + \frac{n+u}{m}\right)\right] \\
&\ge 1 - 2e^{-u}, \\
&\mathbb{P}\left[\left|\frac{1}{n}\|\sqrt{m}\Sigma^{-1/2}\hat{\boldsymbol{\mu}}\|_2^2 - 1\right| \le \epsilon\right] \\
&\ge 1 - 2e^{-c_2 n \min(\epsilon, \epsilon^2)}
\end{aligned}$$

Now we set $\epsilon > 1$ and $u = c_3 n$ for some constant $c_3 > 0$. Then, when $m = \lceil n\log n\rceil$, we have $(n+u)/m = o(1)$ Then, with probability at least $1 - 2e^{-c_3 n} - 2e^{-c_2\epsilon n}$, we have

$$\begin{aligned}
\|\tilde{\Sigma}_i - \Sigma\|_2 &\le \|\Sigma\|\cdot o(1), \\
\text{and } \|\Sigma^{-1/2}\hat{\boldsymbol{\mu}}\|_2^2 &\le \frac{(1+\epsilon)n}{m} \\
\Rightarrow \|\hat{\boldsymbol{\mu}}\|^2 &\le \frac{(1+\epsilon)n\|\Sigma\|}{m} = \|\Sigma\|\cdot o(1), \\
\Rightarrow \|\hat{\Sigma}_i - \Sigma\| &\le \|\Sigma\|\cdot o(1).
\end{aligned}$$

Choosing large enough c_3 and ϵ, this statement holds for all $i \in [n]$ with probability greater than $1 - e^{logn - c_4 n} = 1 - e^{-\Omega(n)}$. □

Theorem 5 follows easily.

Proof of Theorem 5 When all edges are permitted, the proof follows from Theorem 3, Lemma 3, and Lemma 4.

If there are prohibited edges, then we must use matrix concentration to bound $\max\sigma(\hat{L}^\dagger\hat{K})$ instead of $\max\sigma(\hat{L}^{-1}\hat{K})$. Notice that prohibited edges have the effect of simply zeroing out certain rows and columns of Q_i, so that $Q_i := \Psi_i(2\gamma_i\Psi_i^T\Sigma_i\Psi_i)^{-1}\Psi_i^T$, rather than $(2\gamma_i\Sigma_i)^{-1}$. Therefore, we can use Theorem 3 to bound $\|\Psi_i^T\hat{\Sigma}_i\Psi_i - \Psi_i^T\Sigma\Psi_i\|$ for all i, and then prove the appropriate analogue of Lemma 3. In particular, the sample size requirement remains the same. □

A.9. Proof of Proposition 1

Restatement of Proposition 1. *Finding the maximum likelihood estimator of Σ under Assumption 1 is equivalent to the following SDP:*

$$\min_{\Sigma} \sum_{t=1}^{T-1} \big\|\Sigma(W(t+1)-W(t)) + (W(t+1)-W(t))\Sigma\big\|_F^2$$
$$s.\ t.\ \Sigma \succeq 0, tr(\Sigma) = 1.$$

Recall that in Assumption 1 we assumed that $M_{ij}(t)$ varies independently according to a Brownian motion with the same parameters for all (i,j). To avoid ambiguity, we recall the definition of a standard Brownian motion as follows.

Definition 1 (Brownian Motion). For $d \geq 1$, a d-dimensional Brownian motion with scale parameter $\sigma > 0$ is a stochastic process $\{\boldsymbol{X_t} : t \geq 0\}$ such that $\boldsymbol{X_t} \in \mathbb{R}^d$ for all t, the components of $\boldsymbol{X_t}$ are independent, and for all $j \in [d]$,

i) The process $\{(\boldsymbol{X_t})_j : t \geq 0\}$ has independent increments.

ii) For $r > 0$, the increment $(\boldsymbol{X_{t+r}})_j - (\boldsymbol{X_t})_j$ is distributed as $N(0, r\sigma^2)$.

iii) With probability 1, the function $t \mapsto \boldsymbol{X_t}$ is continuous on $[0,\infty)$.

We can derive the SDP of Proposition 1 as follows.

Proposition 5. *Under Assumption 1, the maximum likelihood estimator for Σ is the unique $\Sigma \succ 0$ such that $tr\Sigma = 1$ and*

- **Consistency***: For all $t \in [T]$,*

$$W(t)\Sigma + \Sigma W(t) = \frac{1}{2}(M(t) + M(t)^T)$$
$$\textit{for some } M(1), M(2), \ldots M(t)$$

- **Minimum mean shift***: The resulting $M(1), \ldots, M(T)$ minimize the objective*

$$\sum_{t=1}^{T-1} \|M(t+1) - M(t)\|_F^2$$

Proof of Proposition 5.

$$\mathbb{P}(M(1), \ldots, M(T) \mid W(1), \ldots, W(T), \Sigma)$$
$$\propto \mathbb{P}(W(1), \ldots W(T) \mid M(1), \ldots, M(T), \Sigma)$$
$$\cdot \mathbb{P}(M(1), \ldots, M(T) \mid \Sigma)$$

$$= \left(\prod_{t=1}^{T} 1_{W(t)\Sigma + \Sigma W(t) = 0.5(M(t)+M(t)^T)} \right)$$
$$\cdot \left(\prod_{t=1}^{T-1} \mathbb{P}(M(t+1) - M(t)) \right)$$
$$= \left(\prod_{t=1}^{T} 1_{\mathrm{vec}(W(t)) = 0.5(\Sigma\otimes I + I \otimes \Sigma)\mathrm{vec}(M(t)+M(t)^T)} \right)$$
$$\cdot \left(\prod_{t=1}^{T-1} \exp(-\frac{\|\mathrm{vec}(M(t+1) - M(t))\|^2}{2\sigma^2}) \right),$$

where the first step follows from Bayes' Rule, the second step from Corollary 2, and the third from Assumption 1. The theorem follows from the observation that for any matrix X, we have $\|\mathrm{vec}(X)\|^2 = \|X\|_F^2$. □

The proof of Proposition 1 follows easily.

Proof of Proposition 1. By Proposition 5, we obtain the SDP

$$\min_{\Sigma} \sum_{t=1}^{T-1} \|M(t+1) - M(t)\|_F^2$$
$$\forall t \in [T] : W(t)\Sigma + \Sigma W(t) = \frac{1}{2}(M(t) + M(t)^T)$$

under the assumptions of $\Sigma \succ 0$ and $\mathrm{tr}(\Sigma) = 1$. Since the Frobenius norm is invariant under transposes, we have

$$\sum_{t=1}^{T-1} \|M(t+1) - M(t)\|_F^2 \propto \sum_{t=1}^{T-1} \|(M(t+1) + M(t+1)^T) - (M(t) + M(t)^T)\|_F^2.$$

We can replace $M(t) + M(t)^T$ with $2W(t)\Sigma + 2\Sigma W(t)$ for all $t \in [T]$ to obtain the equivalent objective function $\sum_{t=1}^{T-1} \|(W(t+1) - W(t))\Sigma + \Sigma(W(t+1) - W(t))\|_F^2$ (up to a constant). This substitution enforces the fixed point equation $W(t)\Sigma + \Sigma W(t) = \frac{1}{2}(M(t) + M(t)^T)$ for all $t \in [T]$, so the conclusion follows. □

Remark 1 (The prohibited edges setting.). Proposition 1 generalizes straightforwardly to the setting of prohibited edges. Let E denote the set of permitted edges. Then minimum mean shift assumption is equivalent to minimizing $\sum_{t=1}^{T-1} \sum_{\{i,j\} \in E} \left(M(t+1) + M(t+1)^T - M(t) - M(t)^T\right)_{ij}^2$. In words, the objective just zeroes out prohibited edges, since mean estimates for prohibited edges have no effect on the network. For a network setting $(\boldsymbol{\mu}_j, \Sigma, \gamma_j, \Psi_j)_{j \in [n]}$, some algebra gives $M(t)_{ij} = \boldsymbol{e}_i^T 2\gamma_j (\Psi_j^T \Psi_j)\Sigma(\Psi_j^T \Psi_j) W(t) \boldsymbol{e}_j$. Notice $\Psi_j^T \Psi_j \in \mathbb{R}^n$ is a diagonal matrix with $(\Psi_j^T \Psi_j)_{ii} = 1$ if $\{i, j\} \in E$ and zero otherwise. Therefore, it is clear that upon substitution, the objective is an SDP in Σ with the same constraints.

A.10. Proof of Theorem 6

Restatement of Theorem 6 *Suppose that for each firm $i \in [n]$, the function $F_i : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable, and there exist strictly increasing functions $f_{ji} : \mathbb{R} \to \mathbb{R}$ such that for all $\boldsymbol{x} \in \mathbb{R}^n$, $\nabla F_i(\boldsymbol{x}) = [f_{1i}(x_1), \ldots, f_{ni}(x_n)]^T$. Then there exists a unique stable point.*

Proof of Theorem 6. Note that the Hessian of $F_i(\boldsymbol{w}_i)$ is a positive diagonal matrix due to the conditions on $F_i(.)$. So, any stationary point is a local maximum. Hence, it suffices to show the existence of a unique stationary point.

Let $R(W)$ be an $n \times n$ matrix whose $(i,j)^{th}$ entry $R(W)_{ij} := f_{ij}(W_{ij})$. If a stable point (W, P) exists, it must satisfy $W = W^T$, $P = -P^T$, and

$$M - P = 2\Sigma W \Gamma + R, \tag{5}$$

following the same steps as the proof for Corollary 2. Adding this equation to its transpose, the stable point must satisfy

$$\begin{aligned}(M + M^T)/2 = {} & (\Sigma W \Gamma + \Gamma W \Sigma) \\ & + (R(W) + R(W)^T)/2.\end{aligned}$$

For a stable point, $[R(W) + R(W)^T]_{ij} = f_{ij}(W_{ij}) + f_{ji}(W_{ji}) = (f_{ij} + f_{ji})(W_{ij})$, using $W = W^T$. Define $S(W)$ to be an $n \times n$ matrix with $S(W)_{ij} = (1/2) \cdot (f_{ij} + f_{ji})(W_{ij})$. Hence, the stable point must satisfy

$$(M + M^T)/2 = S(W) + (\Sigma W \Gamma + \Gamma W \Sigma) \tag{6}$$

$$\begin{aligned}\Leftrightarrow \operatorname{vec}((M + M^T)/2) = {} & \operatorname{vec}(S(W)) \\ & + \underbrace{(\Gamma \otimes \Sigma + \Sigma \otimes \Gamma)}_{Q} \operatorname{vec}(W).\end{aligned} \tag{7}$$

Note that Q is positive-definite (from the proof of Corollary 2), and each entry of $\operatorname{vec}(S(W))$ is a function of the corresponding entry of $\operatorname{vec}(W)$. By Theorems 1 and 2 of Sandberg and Willson (1972), Eq. (6) has a unique solution if (1) for all diagonal $D \succ 0$, $\det(D + Q) > 0$ and (2) for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{n^2}$ such that $\boldsymbol{x} = Q\boldsymbol{y}$, we have $\boldsymbol{x}^T \boldsymbol{y} \geq 0$. The first condition holds because $\det(D + Q) = \det(D^{1/2}(I + D^{-1/2} Q D^{-1/2}) D^{1/2}) = \det(D) \cdot \det(I + D^{-1/2} Q D^{-1/2}) > 0$. The second condition is true because $\boldsymbol{x}^T \boldsymbol{y} = \boldsymbol{y}^T Q \boldsymbol{y} \geq 0$. Hence, Eq. (6) has a unique solution $\mathcal{W}$.

We now show that this solution satisfies the conditions of the stable point, that is, $\mathcal{W} = \mathcal{W}^T$, and there exists a skew-symmetric P which satisfies Eq. (5). Observe that

$$\begin{aligned}[S(\mathcal{W})^T]_{ij} &= S(\mathcal{W})_{ji} \\ &= (1/2) \cdot (f_{ij} + f_{ji})(\mathcal{W}_{ji}) \\ &= S(\mathcal{W}^T)_{ij},\end{aligned}$$

so $S(\mathcal{W})^T = S(\mathcal{W}^T)$. Taking the transpose of Eq. (6) and using $\Sigma = \Sigma^T$, $\Gamma = \Gamma^T$, and $S(\mathcal{W})^T = S(\mathcal{W}^T)$, we find

$$(M + M^T)/2 = (\Sigma\mathcal{W}^T\Gamma + \Gamma\mathcal{W}^T\Sigma) + S(\mathcal{W}^T).$$

But since there is only one solution to Eq. (6), we must have $\mathcal{W} = \mathcal{W}^T$.

Finally, we choose

$$\begin{aligned} P &= M - 2\Sigma\mathcal{W}\Gamma - R \\ \Rightarrow P + P^T &= (M + M^T) - 2(\Sigma\mathcal{W}\Gamma + \Gamma\mathcal{W}\Sigma) \\ &\quad - 2S(W) \\ &= 0, \end{aligned}$$

using the fact that $\mathcal{W} = \mathcal{W}^T$ is a solution for Eq. (6). Hence, this choice of P is both skew-symmetric and satisfies Eq. (5). □

A.11. Proof of Theorem 7

RESTATEMENT OF 7. *Suppose $\Sigma_i = \Sigma$ for all firms, and let M be the matrix of expected returns. Then, we have the following:*

1. **Change in beliefs about expected returns:** *Let Σ have the eigendecomposition $\Sigma = V\Lambda V^T$. Then for $i, j, k, \ell \in [n]$,*

$$\begin{aligned} \frac{\partial W_{ij}}{\partial M_{k\ell}} &= \frac{1}{2\sqrt{\gamma_i\gamma_j\gamma_k\gamma_\ell}} \\ &\quad \cdot \sum_{s,t\in[n]} \frac{V_{is}V_{ks}V_{jt}V_{\ell t} + V_{is}V_{\ell s}V_{jt}V_{kt}}{\lambda_s + \lambda_t}. \end{aligned} \tag{8}$$

In particular, W_{ij} is monotonically increasing with respect to M_{ij}.

2. **Risk scaling:** *If the covariance Σ changes to $c\Sigma$ $(c > 0)$, then W changes to $(1/c)W$.*
3. **Increase in perceived risk:** *Suppose $\gamma_i = \gamma$ for all i, and the covariance Σ increases to $\Sigma' \succ \Sigma$. Let W and W' be the stable points under Σ and Σ' respectively. Then, $tr(M^T(W' - W)) < 0$.*

Proof of Theorem 7. 1. Let $(\lambda_i, \boldsymbol{v}_i)$ denote the i^{th} eigenvalue and eigenvector of $\Gamma^{-1/2}\Sigma\Gamma^{-1/2}$, and let $V_{ij} = \boldsymbol{e}_i^T\boldsymbol{v}_j$. By Corollary 2,

$$\begin{aligned} W &= \Gamma^{-1/2}\Bigg(\sum_{s=1}^{n}\sum_{t=1}^{n} \\ &\quad \frac{\boldsymbol{v}_s^T\Gamma^{-1/2}(M + M^T)\Gamma^{-1/2}\boldsymbol{v}_t}{2(\lambda_r + \lambda_s)}\boldsymbol{v}_s\boldsymbol{v}_t^T\Bigg)\Gamma^{-1/2} \\ \Rightarrow \frac{\partial W_{ij}}{\partial M_{k\ell}} &= \boldsymbol{e}_i^T\Gamma^{-1/2}\Bigg(\sum_{s=1}^{n}\sum_{t=1}^{n} \end{aligned}$$

$$\frac{\boldsymbol{v}_s^T \Gamma^{-1/2}(\boldsymbol{e}_k \boldsymbol{e}_\ell^T + \boldsymbol{e}_\ell \boldsymbol{e}_k^T)\Gamma^{-1/2}\boldsymbol{v}_t}{2(\lambda_s+\lambda_t)}$$
$$\boldsymbol{v}_s \boldsymbol{v}_t^T\Big)\Gamma^{-1/2}\boldsymbol{e}_j$$
$$= \frac{1}{2\sqrt{\gamma_i\gamma_j\gamma_k\gamma_\ell}}\Bigg(\sum_{s=1}^{n}\sum_{t=1}^{n}$$
$$\frac{\boldsymbol{v}_s^T(\boldsymbol{e}_k \boldsymbol{e}_\ell^T + \boldsymbol{e}_\ell \boldsymbol{e}_k^T)\boldsymbol{v}_t}{(\lambda_s+\lambda_t)}(\boldsymbol{e}_i^T \boldsymbol{v}_s)(\boldsymbol{v}_t^T \boldsymbol{e}_j)\Bigg)$$
$$= \frac{1}{2\sqrt{\gamma_i\gamma_j\gamma_k\gamma_\ell}}\sum_{s=1}^{n}\sum_{t=1}^{n}\Bigg($$
$$\frac{V_{is}V_{ks}V_{jt}V_{\ell t} + V_{is}V_{\ell s}V_{jt}V_{kt}}{\lambda_s+\lambda_t}\Bigg)$$

This proves Eq. (3). If $i = k, j = \ell$, we have:

$$\begin{aligned}
\frac{\partial W_{ij}}{\partial M_{ij}} &= (2\gamma_i\gamma_j)^{-1}\left(\sum_{s=1}^{n}\sum_{t=1}^{n}\underbrace{\frac{V_{is}^2V_{jt}^2 + V_{is}V_{js}V_{jt}V_{it}}{\lambda_s+\lambda_t}}_{Z_{st}}\right) \\
&= (4\gamma_i\gamma_j)^{-1}\left(\sum_{s=1}^{n}\sum_{t=1}^{n} Z_{st} + \sum_{t=1}^{n}\sum_{s=1}^{n} Z_{ts}\right) \\
&= (4\gamma_i\gamma_j)^{-1}\sum_{s=1}^{n}\sum_{t=1}^{n}(Z_{st}+Z_{ts}) \\
&= (4\gamma_i\gamma_j)^{-1}\sum_{s=1}^{n}\sum_{t=1}^{n}\frac{(V_{is}V_{jt}+V_{js}V_{it})^2}{\lambda_r+\lambda_s} > 0.
\end{aligned}$$

Hence, W_{ij} is monotonically increasing with respect to M_{ij}.

2. This follows from Corollary 2.

3. By Corollary 2, $\text{vec}(W) = \gamma^{-1}(\Sigma \otimes I + I \otimes \Sigma)^{-1}\text{vec}(\frac{M+M^T}{2})$. Let $K = \gamma(\Sigma \otimes I + I \otimes \Sigma)$ and $K' = \gamma(\Sigma' \otimes I + I \otimes \Sigma')$. Since $\Sigma' \succ \Sigma$ it follows that $K' \succ K$. Therefore $K^{-1} \succ (K')^{-1}$.

So, since $\text{vec}(W' - W) = ((K')^{-1} - K^{-1})\text{vec}(\frac{M+M^T}{2})$, we immediately obtain $\frac{1}{2}\text{vec}(M + M^T)^T\text{vec}(W' - W) < 0$. Since W, W' are symmetric it follows that $\text{vec}(M^T)^T\text{vec}(W' - W) = \text{vec}(M)^T\text{vec}(W' - W)$. So we have $\text{vec}(M)^T\text{vec}(W' - W) < 0$.

Since $\text{vec}(M)^T\text{vec}(W' - W) = \text{tr}(M^T(W' - W))$, the conclusion follows. □

A.12. Hardness of Source Detection

We begin by defining

$$\left|\frac{\partial W_{ij}}{\partial M_{k\ell}}\right|_{approx} := \frac{|V_{in}V_{kn}V_{jn}V_{\ell n}|}{2\lambda_n}. \tag{9}$$

This approximates the right hand side of Eq. (3) when the term corresponding to the smallest eigenvalue λ_n dominates the sum. We now show that if the corresponding eigenvector $\boldsymbol{v}_n$ is random, source detection becomes difficult.

Proposition 6 **(Hardness of Source Detection)**. *Suppose $\boldsymbol{v}_n$ is Haar-distributed, that is, $\boldsymbol{v}_n$ is distributed uniformly on the unit sphere S^{n-1}. Then, if $\Sigma = V\Lambda V^T$ and $\Gamma = I$,*

$$\mathbb{P}\left[\max_{i,j\in[n]:(i,j)\neq(k,\ell)}\left|\frac{\partial W_{ij}}{M_{k\ell}}\right|_{approx} < \left|\frac{\partial W_{k\ell}}{M_{k\ell}}\right|_{approx}\right] \leq O\left(\frac{1}{n}\right).$$

Proof of Proposition 6. Without loss of generality we can set $k=1, \ell=2$ (the analysis of $k=\ell$ is identical). Notice that $\left|\frac{\partial W_{ij}}{M_{k\ell}}\right|_{approx}$ is maximized at the (i,j) that maximizes $|V_{in}V_{jn}|$.

Now, consider $(i,j)\in\{(1,2),(3,4),\ldots,(n-1,n)\}$. Notice the distribution of $\boldsymbol{v}_n$ is permutation-invariant by assumption. Hence the joint distribution of (V_{in}, V_{jn}) is the same for all such pairs (i,j). Hence the distribution of $|V_{in}V_{jn}|$ is also the same for all such (i,j). Therefore,

$$\mathbb{P}\left[\arg\max_{(i,j)\in\{(1,2),(3,4),\ldots,(n-1,n)\}}\left|\frac{\partial W_{ij}}{M_{12}}\right|_{approx} = (1,2)\right] \leq O(1/n). \qquad \square$$

A.13. Proof of Proposition 2

Restatement of Proposition 2 *Suppose M, Σ, Γ exhibit community structure (Eq. (4)), and all the error terms $(\epsilon_i)_{i\in[n]}$ and $(\epsilon'_{\theta_i,j})_{i,j\in[n]}$ are independent and identically distributed. Let $\pi:[n]\to[n]$ be any intra-community permutation, and let $\Pi:\mathbb{R}^n\to\mathbb{R}^n$ be the corresponding column-permutation matrix: $\Pi(\boldsymbol{e}_i)=\boldsymbol{e}_{\pi(i)}$. Then, W and $\Pi^TW\Pi$ are identically distributed.*

Proof. Let $H=\frac{1}{2}(M+M^T)$. The fixed point equation for W is given by Corollary 2 as $\Sigma W\Gamma + \Gamma W\Sigma = H$. Vectorization implies $(\Gamma\otimes\Sigma+\Sigma\otimes\Gamma)\text{vec}(W)=\text{vec}(H)$. Let $X\sim Y$ denote that a pair of random variables X, Y are identically distributed. We want to show $\Pi^TW\Pi\sim W$. Vectorization gives $\text{vec}(\Pi^TW\Pi)=(\Pi^T\otimes\Pi^T)\text{vec}(W)$. Let $P=(\Pi^T\otimes\Pi^T)$ and $K=(\Gamma\otimes\Sigma+\Sigma\otimes\Gamma)$ for shorthand.

In this notation, we want to show that $PK^{-1}\text{vec}(H)\sim K^{-1}\text{vec}(H)$. Since P is a permutation, we have $PK^{-1}\text{vec}(H)=PK^{-1}P^TP\text{vec}(H)=(PKP^T)^{-1}P\text{vec}(H)$. Since the collections of random variables $\{\epsilon_i\}_i$ and $\{\epsilon'_{\theta_i,j}\}_{i,j}$ are independent, we know $\text{vec}(H)$ and K are independent. So to show $(PKP^T)^{-1}P\text{vec}(H)\sim K^{-1}\text{vec}(H)$ it suffices to show that $P\text{vec}(H)\sim\text{vec}(H)$ and $PKP^T\sim K$.

Notice $P\text{vec}(H)=\text{vec}(\Pi^TH\Pi)$. Hence, we want to show $\Pi^TH\Pi\sim H$, which holds iff $\Pi^T(M+M^T)\Pi\sim M+M^T$. Notice that $\Pi^TM^T\Pi=(\Pi^TM\Pi)^T$, so if $\Pi^TM\Pi\sim M$ then we obtain $\Pi^TM^T\Pi\sim M^T$ as well. It suffices to show $\Pi^TM\Pi\sim M$.

Similarly, we can simplify $PKP^T=\Pi^T\Sigma\Pi\otimes\Pi^T\Gamma\Pi+\Pi^T\Gamma\Pi\otimes\Pi^T\Sigma\Pi$. It suffices to show $\Pi^T\Gamma\Pi\sim\Gamma$ and $\Pi^T\Sigma\Pi=\Sigma$.

We are left to show that $\Pi^T\Sigma\Pi=\Sigma$ and $\Pi^TA\Pi\sim A$ for $A\in\{\Gamma,M\}$.

Proof of $\Pi^T\Sigma\Pi=\Sigma$. Let $i,j\in[n]$. Then $(\Pi^T\Sigma\Pi)_{ij}=\Sigma_{\pi(i),\pi(j)}=g(\theta_{\pi(i)},\theta_{\pi(j)})$. Since π only commutes members within communities, $g(\theta_{\pi(i)},\theta_{\pi(j)})=g(\theta_i,\theta_j)=\Sigma_{ij}$. So $\Pi^T\Sigma\Pi=\Sigma$.

Proof of $\Pi^T\Gamma\Pi\sim\Gamma$. Notice $\Pi^T\Gamma\Pi$ and Γ are both diagonal. Let $i\in[n]$. Then $(\Pi^T\Gamma\Pi)_{ii}=\Gamma_{\pi(i),\pi(i)}=h(\theta_{\pi(i)})+\epsilon_{\pi(i)}=h(\theta_i)+\epsilon_{\pi(i)}$. Since $\theta_i=\theta_{\pi(i)}$, we know $\epsilon_i\sim\epsilon_{\pi(i)}$. The conclusion follows.

Proof of $\Pi^TM\Pi\sim M$. Let $i,j\in[n]$. Then $(\Pi^TM\Pi)_{ij}=M_{\pi(i),\pi(j)}=f(\theta_{\pi(i)},\theta_{\pi(j)})+\epsilon'_{\theta_{\pi(i)},\pi(j)}=f(\theta_i,\theta_j)+\epsilon'_{\theta_i,\pi(j)}$. Since $\theta_j=\theta_{\pi(j)}$, we know that $\epsilon'_{\theta_i,\pi(j)}\sim\epsilon'_{\theta_i,j}$, and the conclusion follows. $\square$

A.14. Proof of Theorem 8

Restatement of Theorem 8 *Consider two network settings $S=(\boldsymbol{\mu}_i,\Sigma,\gamma_i)_{i\in[n]}$ and $S'=(\boldsymbol{\mu}_i,\Sigma,\gamma'_i)_{i\in[n]}$ which differ only in the risk-aversions of firms $J=\{j\mid\gamma_j\neq\gamma'_j\}\subseteq[n]$. Then, there exists a setting $S^\dagger=(\boldsymbol{\mu}^\dagger_i,\Sigma,\gamma_i)_{i\in[n]}$ such that $\boldsymbol{\mu}_i=\boldsymbol{\mu}^\dagger_i$ for all $i\notin J$ and the stable networks under $S^\dagger$ and S' are identical.*

Proof of Theorem 8. First, consider the network settings S and S'. Let $\Gamma\in\mathbb{R}^{n\times n}$ be a diagonal matrix with $\Gamma_{i,i}=\gamma_i$; define Γ' similarly under S'. Let the corresponding networks be W and W', and let $\Delta_W=W'-W$ and $\Delta_\Gamma=\Gamma'-\Gamma$. By Corollary 2, we have

$$\begin{aligned}
\Sigma W\Gamma+\Gamma W\Sigma &= \frac{M+M^T}{2}\\
&=\Sigma W'\Gamma'+\Gamma'W'\Sigma\\
\Rightarrow \frac{M+M^T}{2} &= \Sigma(W+\Delta_W)(\Gamma+\Delta_\Gamma)\\
&\quad+(\Gamma+\Delta_\Gamma)(W+\Delta_W)\Sigma\\
&=\Sigma W\Gamma+\Gamma W\Sigma \qquad (10)\\
&\quad+\Sigma\Delta_W\Gamma+\Gamma\Delta_W\Sigma\\
&\quad+\Sigma W\Delta_\Gamma+\Delta_\Gamma W\Sigma\\
&\quad+\Sigma\Delta_W\Delta_\Gamma+\Delta_\Gamma\Delta_W\Sigma\\
\Rightarrow \Sigma\Delta_W\Gamma+\Gamma\Delta_W\Sigma &= -(\Sigma W\Delta_\Gamma+\Delta_\Gamma W\Sigma\\
&\quad+\Sigma\Delta_W\Delta_\Gamma+\Delta_\Gamma\Delta_W\Sigma)\\
&=-(\Sigma W'\Delta_\Gamma+\Delta_\Gamma W'\Sigma) \qquad (11)
\end{aligned}$$

Next, consider S versus $S^\dagger$. Suppose that $M^\dagger$ has columns $\boldsymbol{\mu}^\dagger_1,\ldots,\boldsymbol{\mu}^\dagger_n$ and let $\Delta_M=M^\dagger-M$. Let $W^\dagger$ be the fixed point network under $S^\dagger$, given by $\Sigma W^\dagger\Gamma+\Gamma W^\dagger\Sigma=\frac{M^\dagger+(M^\dagger)^T}{2}$. Let $\Delta^\dagger_W=W^\dagger-W$. Then a similar argument gives:

$$\frac{\Delta_M+\Delta_M^T}{2}=\Sigma\Delta^\dagger_W\Gamma+\Gamma\Delta^\dagger_W\Sigma \qquad (12)$$

Therefore, from Eq (11) and (12), it follows that $W' = W^\dagger$ if

$$\frac{\Delta_M + \Delta_M^T}{2} = -(\Sigma W' \Delta_\Gamma + \Delta_\Gamma W' \Sigma).$$

Hence, $W' = W^\dagger$ if we set $\Delta_M = -\Sigma W' \Delta_\Gamma$.

It remains to show that $M^\dagger$ differs from M only in columns corresponding to J. Suppose that $i \notin J$. Then $\gamma_i = \gamma_i'$, so $\Delta_\Gamma \boldsymbol{e}_i = \boldsymbol{0}$. We conclude that $\Delta_M \boldsymbol{e}_i = \boldsymbol{0}$ and hence $M\boldsymbol{e}_i = M^\dagger \boldsymbol{e}_i$. □

A.15. Additional Discussion of Theorem 6

Theorem 6 considers budget constraints or penalties of the form $F_i(\boldsymbol{w}_i)$, where $\boldsymbol{w}_i$ is the vector of contracts for agent i. Consider the more general setting of $F_i(\boldsymbol{w}_i^T P e_i)$ or $F_i(\boldsymbol{w}_i \odot P\boldsymbol{e}_i)$. Using the techniques from Sandberg and Willson (1972), we cannot prove the existence and uniqueness of stable points in the general setting, except in trivial cases.

To see this, note that we must impose conditions on the first derivative $f_{ji} = \frac{\partial F_i}{\partial W_{ij}}$ of the penalty function F_i. Specifically, we need $S_{ij} := f_{ij} + f_{ji}$ to be a function of W_{ij} alone. But if F_i were to depend on P, so would f_{ij}. Each entry of P depends on all entries of W in general, not just W_{ij}. Hence, we cannot handle general forms of $F_i(W, P)$.

In the special case where $F_i(W, P) := \boldsymbol{w}_i^T P \boldsymbol{e}_i$, we have $S_{ij} = P_{ij} + P_{ji} = 0$, and Theorem 6 still applies. However, this case is trivial since it amounts to modifying the payments matrix by a factor of 2:

$$\begin{aligned}\text{agent } i\text{'s utility } g_i(W, P) &:= \boldsymbol{w}_i^T(\boldsymbol{\mu}_i - P\boldsymbol{e}_i) \\ &- \gamma_i \cdot \boldsymbol{w}_i^T \Sigma_i \boldsymbol{w}_i - F_i(\boldsymbol{w}_i) \\ &= \boldsymbol{w}_i^T(\boldsymbol{\mu}_i - 2P\boldsymbol{e}_i) \\ &- \gamma_i \cdot \boldsymbol{w}_i^T \Sigma_i \boldsymbol{w}_i.\end{aligned}$$

If we instead have a positive penalty only when the total payment is positive (say, $F_i(W, P) := \max(0, \boldsymbol{w}_i^T P \boldsymbol{e}_i)$), the approach no longer works.

Appendix B: Experimental Details

B.1. Fama-French Stock Market Data

We use the Fama-French value-weighted asset returns dataset, for 96 assets over 625 months (Fama and French 2015).

B.2. OECD International Trade Data

We use international trade statistics from the OECD to get quarterly measurements of bilateral trade between 46 large economies, including the top 15 world nations by GDP OECD (2022). The data are available at the OECD Statistics webpage (`https://stats.oecd.org/`). The data are measured quarterly from Q1 2010 to Q2 2022. We take the sum of trade flows $i \to j$ and $j \to i$ to measure the weight of an edge $\{i, j\}$.

To obtain the corresponding Σ, we run our inference procedure (Section 2.4). Since there is no data for within-country trade, the network has no self-loops ($W_{ii} = 0$). So we modify the inference according to Remark 1 in Appendix A.9.

B.3. Outlier Detection Simulation

The experiments in Figure 6 proceed as follows. Fix a number of communities k and number of firms n. Fix a value of $\sigma > 0$. For us, $k = 2$, $n \in \{20, 100, 300\}$, and $\sigma \in \{\sigma_1, \ldots, \sigma_{10}\}$, where the σ_i are logarithmically spaced on the interval $[0.1, 1]$, so that

$$\sigma \in \{0.1, 0.12915497, 0.16681005,$$
$$0.21544347, 0.27825594, 0.35938137,$$
$$0.46415888, 0.59948425, 0.77426368, 1.0\}$$

For a setting of n, k, σ, we perform the following simulation $m = 500$ times.

Generate communities. Generate the community membership matrix $\Theta \in \{0, 1\}^{n \times k}$ with rows independently and uniformly at random from $\{\boldsymbol{e_1}, \ldots, \boldsymbol{e_k}\}$.

Generate the network setting. The deterministic functions f, g, h for M, Σ, Γ respectively are as follows. First $f(\theta_1, \theta_2) = f(\theta_2, \theta_1) = 1$ and $f = 0$ otherwise. Next, let $G \in \mathbb{R}^{k \times k}$ be the matrix $G_{ij} = g(\theta_i, \theta_j)$. Then G is generated from a normalized Wishart distribution centered at I_k and with 5 degrees of freedom. Finally, $h(\theta_i) = 1$ for all i.

The noise variables for agent beliefs are as follows. Sample i.i.d. ϵ_i according to a $N(0, \sigma^2)$ distribution truncated to $[-0.5, 0.5]$ for all i. Sample $\epsilon'_{\theta_i, j} \overset{\text{iid}}{\sim} N(0, \sigma^2)$ for all i, j.

Designate an outlier. Set the the noise parameter $\epsilon_1 = -0.5$ for firm 1 (the risk-seeker), so as $\sigma \to 0$, γ_1 gets further separated from all other γ_i.

Outlier detection simulation. Then for a random firm i such that $\theta_i \neq \theta_1$, we test whether the outlier $\hat{j} := \arg\max_{j: \theta_j = \theta_1} |W_{i,j}|$ is equal to the true outlier firm 1.

Collate results. Once the $m = 500$ runs are completed for a single setting of n, k, σ, we obtain an estimate $\hat{p}$ for the probability of successful deviator detection at this setting of parameters. We plot a confidence interval $[p - 2\sqrt{\frac{\hat{p}(1-\hat{p})}{m}}, p + 2\sqrt{\frac{\hat{p}(1-\hat{p})}{m}}]$. This is plotted on the y-axis. The x-axis quantifies how much γ_1 deviates from the mean, in terms of the number of standard deviations of the truncated normal distribution ϵ_i.

References

Callier FM, Desoer CA (1994) *Linear System Theory* (Springer Verlag).

Fama EF, French KR (2015) A five-factor asset pricing model. *Journal of financial economics* 116(1):1–22.

Horn RA, Johnson CR (1994) *Topics in Matrix Analysis* (Cambridge University Press).

OECD (2022) OECD statistics. URL `https://stats.oecd.org/`.

Sandberg I, Willson A Jr (1972) Existence and uniqueness of solutions for the equations of nonlinear DC networks. *SIAM Journal on Applied Mathematics* 22(2):173–186.

Vershynin R (2018) *High-dimensional probability: An introduction with applications in data science*, volume 47 (Cambridge University Press).