Dynamic Public Learning in Networks of Strategic Agents: The Role of Inter/Intra-community Ties

Amir Ajorlou† Ali Jadbabaie† Munther A. Dahleh†

Working Draft

Abstract

We study the quality of information aggregation in a dynamic quadratic game of incomplete information. Agents in each generation have access to a public history of noisy aggregate actions in previous generations. Each agent also belongs to an information community, where they share a private noisy signal about the state. We quantify the quality of information aggregation as the asymptotic precision of the publicly learned signal about the state, and study how the interactions between/within information communities affect the quality of learning. By characterizing the precision of aggregate learning for slow walks, we show the inefficiency of learning from the public history: while for a static state, public history fully reveals the state, a small perturbation in the state from generation to generation, significantly degrades the quality of public learning. As for the effect of information-action structure, we show that the inter-community interactions exhibit a double-edged effect: while the inter-community interaction of local influencers and influencées improves the quality of learning for slow dynamics, the inter-community interaction of global influencers and influencées defects learning for fast dynamics. As a result, we may observe phase transition in quality of learning between different information-action structures.

I. Introduction

Many complex systems involve large-scale interconnection of agents with heterogeneous information who interact strategically. Such systems are prevalent in diverse domains ranging from cyber-physical systems, energy markets, transportation networks, water networks, financial markets, consumer networks, and more broadly, complex social and economic networks Savla

†Institute for Data, Systems, and Society, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. E-mail: {ajorlou,jadbabai,dahleh}@mit.edu.

This work was supported by ARO MURI W911NF-12-1-0509.

et al. (2014); Thrampoulidis et al. (2015); Perelman et al. (2016); Tavafoghi and Teneketzis (2016); Bramoulle and Kranton (2007); Battiston et al. (2012); August et al. (2014). The strategic interactions of the agents affect both the direct and indirect (inferred) flow of information, hence introducing new challenges to modeling, analysis, and control Ajorlou and Jadbabaie (2016); Ajorlou et al. (2016).

Recent advances in information technology and Internet have provided a wide range of channels to aggregate information, from blogs to wikis to the open source movement to prediction markets Servan-Schreiber et al. (2004). There are abundant sources of information (e.g., news channels, social networks, online forums, surveys, and data bases), with every person having access to a personalized set of the sources. Each individual's decision making is not only influenced by her own sources of information, but also by the action of her friends who themselves are influenced by other information sources and friends. The aggregate (or average) action of the whole population thus partially accumulates the information shared within different information communities. An interesting question here is how the quality of information aggregation is affected by the influences between and within different information communities, specially in a dynamic setting?

Learning from the observations of the aggregate action (called *public learning* hereafter) is a twin of systemic volatility; the less volatile the aggregate action of the whole system, the higher the quality of public learning. There is a large body of literature on systemic risk and volatility, highlighting the importance of the interaction structure in prorogation and amplification of shocks and disruptive events hitting any single agent Acemoglu et al. (2012); Bergemanna et al. (2015); Allen and Gale (2000); Bidkhori et al. (2016); Bimpikis et al. (2016); Blume et al. (2011). Many such works model the problem as a one-shot static game of incomplete information, where they use the equilibrium strategies of the agents to assess the resilience of the network of interconnections. They in turn come up with a (sometimes contradicting[1]) ranking of network structures in terms of their performance against shocks. For example, authors in Acemoglu et al. (2012) show that the volatility of the economy's aggregate output is increasing with the heterogeneity in the role of different firms as input-suppliers. Such approaches neglect the fact that in many practical applications, network entities can observe and learn from the

[1]For instance, while some work suggest that denser interconnections reduce the systemic risk (e.g., Allen and Gale (2000)), others find denser networks destabilizing (e.g., Blume et al. (2011)).

system responses to shocks and revise their actions accordingly.

There is also similar research investigating the effect of local disturbances on the steady state systemic volatility in networked dynamical systems, with a special attention to the consensus problem Young et al. (2010); Bamieh et al. (2012); Siami and Motee (2016); Jadbabaie and Olshevsky (2016); Siami et al. (2017). While taking into account the opinion dynamics, no underlying dynamic unknown state is present in these models. In a closely related work (Amador and Weill (2012)), authors consider learning from both private and public observations of the aggregate action in a static setting, and show that initial public revelation of information slows down information diffusion in the long run.

To this end, we consider a model where in each period there is a generation of agents with linear quadratic payoffs, each having access to a public history of noisy aggregate actions of previous generations. Each agent also belongs to an information community, where they share a private noisy signal about the state. We quantify the quality of information aggregation as the asymptotic precision of the publicly learned signal about the state, and study how the interactions between/within information communities affect the quality of aggregate learning. As the first finding, we show the inefficiency of learning from the public history: while for a static state, public history eventually reveals the true state, a small perturbation in the state from generation to generation significantly degrades the quality of aggregate learning. As a rough measure, improving the precision of the learned signal by a factor of 10 requires the noise in the state transition to be 10^4 times less volatile. As another major contribution, we show that the inter-community interactions exhibit a double-edged effect: while the inter-community interaction of local influencers and influencées improves the quality of information aggregation for slow dynamics, the inter-community interaction of global influencers and influencées defects the aggregate learning for fast dynamics. This may result in a phase transition in quality of learning between different information-action structures.

II. Problem Description

Consider a population of agents with linear quadratic payoffs, residing in a network whose structure is represented by a matrix G. The payoff of each agent is a linear quadratic function of a common unknown state of the world, her own action, and actions of her neighbors in G, with the entries of G representing the payoff externalities. Each agent also belongs to an

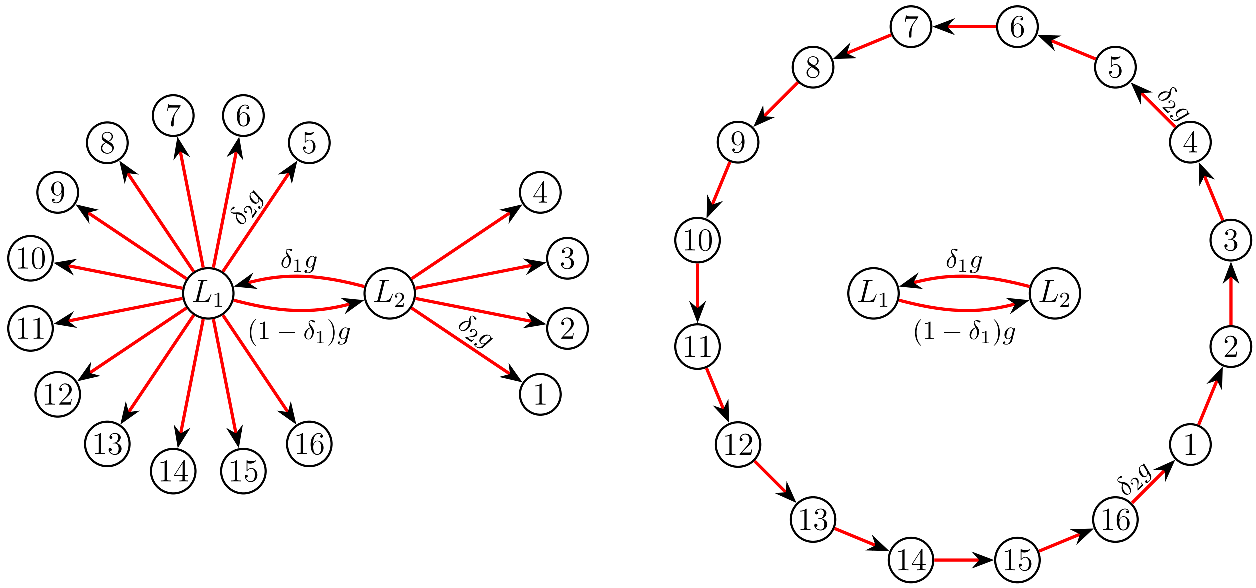

Fig. 1. G^1 and G^2 differ in their inter-community structures.In G^1 the out-of-community influence comes from either L_1 or L_2. In G^2, the out-of-community influences for $1, \ldots, 16$ form a ring graph. Note that weight of all edges entering $1, \ldots, 16$ is $\delta_2 g$ in both structures.

information community, where they share a private noisy observation from the state. Upon making the observations, agents simultaneously take actions maximizing their expected payoffs. One common macro statistics of interest here is the aggregate (or average) action of the agents. The uncertainty in the observations results in volatility in the aggregate action, the extent of which depends on the structure of the network and information communities among many other factors.

Now consider a scenario where the above setup is a single episode of an ongoing sequence of generations, except that a noisy version of the aggregate action in each generation is available to all generations to come as a common public history. There may also be a change in the state in transition between successive generations.

History results in a common public prior on the state in each generation. Our objective in this manuscript is to study the effect of the information-action structure on the quality of this publicly learned signal. To make our objective more clear, consider the two different inter-community wirings of the same information communities depicted in Figure 1. Agents within each community receive the same out-of-community influence from the average action of the agents in some other community. The difference between the two structures is that L_1 and L_2

are highly influential in G^1 influencing 12 and 4 communities each, respectively. In G^2, on the other hand, the distribution of inter-community influences is more even with each community influencing another community forming a ring, and L_1 and L_2 influencing each other. Everything else (including the realization of the signals) are exactly identical. Then, can we say that generations residing in one structure are more informed about the state than their concurrent generation in the other? If yes, how can we identify that superior structure. Or is it possible that the order of the qualities get flipped over time? What are the main network characteristics affecting the quality of aggregate learning and whether/how they depend on the state dynamics? These are the type of the questions that we tackle in this draft.

III. Model

In each period $t = 0, 1, \ldots$ there is a generation of size n of short-lived agents living in a social network whose structure is represented by a matrix $G = [g_{ij}]$ with $g_{ii} = 0$ and $\sum_{j=1}^n g_{ij} = g$ for some $0 \leq g < 1$, where $g_{ij} \geq 0$ quantifies the influence agent i receives from agent j. We refer to the set of agents $N_i = \{j | g_{ij} > 0\}$ as the friends of agent i in her generation.

The payoff of agent i in generation t is a quadratic function that depends on the state of the world θ_t and the vector of actions $a_t = (a_{1t}, \ldots, a_{nt})$:

$$u_{it} = \theta_t a_{it} - \frac{1}{2} a_{it}^2 + \sum_{j=1}^{n} g_{ij} a_{it} a_{jt}. \tag{1}$$

The state of the world evolves on a random walk $\theta_{t+1} = \theta_t + \nu_t$, where $\nu_t \sim N(0, \tau_v^{-1})$ is a white noise normal process.

Each agent has access to a private signal $x_{it} = \theta_t + \epsilon_{it}$, where $\epsilon_{it} \sim N(0, \tau_\epsilon^{-1})$. These signals are provided by a set of m information sources which partition the agents in each generation into m *information communities* $\mathcal{I} = \{I_1, \ldots, I_m\}$. Agents within each community share the same private signal while the noises in the signals $\{\epsilon_{it}\}$ are independent across communities and generations. A noisy version of the aggregate action in each generation $s_t = \bar{a}_t + \eta_t$, with $\eta_t \sim N(0, \tau_\eta^{-1})$, where $\bar{a}_t = \frac{1}{n} \sum_{i=1}^n a_{it}$ is also available to all generations to come. Agent i in generation t is thus endowed with the information vector $S_{it} = (s_0, \ldots, s_{t-1}, x_{it})$.

IV. Bayes Nash Equilibrium

Given the quadratic form of the payoff in (1), the best response of agent i in period t is given by the linear condition

$$a_{it} = \mathbb{E}[\theta_t | S_{it}] + \sum_{j=1}^{n} g_{ij} \mathbb{E}[a_{jt} | S_{it}]. \tag{2}$$

We define the Bayes Nash equilibrium of this game below.

Definition 1. *A Bayes Nash equilibrium for generation t is defined by strategies $a^*_{it} : \mathbb{R}^{t+1} \to \mathbb{R}$ such that*

$$a^*_{it}(S_{it}) = \mathbb{E}[\theta_t | S_{it}] + \sum_{j=1}^{n} g_{ij} \mathbb{E}[a^*_{jt}(S_{jt}) | S_{it}], \ \forall S_{it} \in \mathbb{R}^{t+1}, \ \forall i \in \mathbb{N}_n. \tag{3}$$

Using the linear form in (3), we can show that this game admits a unique linear Bayes Nash equilibrium. We can characterize this linear equilibrium using the precision of the public prior on the state in each generation which is formed based on the common history, the precision of the private signals, and *Bonachic centrality* of the agents in a modified graph constructed by scaling the weights of the inter-community edges as we will elaborate below.

Denote with $N(\omega_t, \tau_t)$ the common prior belief on θ_t in generation t formed based on the public history $\{s_0, \ldots, s_{t-1}\}$, where $\omega_t = \mathbb{E}[\theta_t | s_0, \ldots, s_{t-1}]$ is the conditional expectation of θ_t given the history and $\tau_t = \text{Var}^{-1}[\theta_t | s_0, \ldots, s_{t-1}]$ is the precision of the publicly learned signal about the state θ_t. Given the normality of random variables, the conditional expectation ω_t is a linear function of $\{s_0, \ldots, s_{t-1}\}$ and is a sufficient statistics for the history in estimating θ_t. Using this we can show that the linear equilibrium strategies in each period can be represented as a linear function of ω_t and the private signals x_{it}.

We next construct a new graph $G(\rho_t)$ by scaling the weights of inter-community edges in G by ρ_t, where $\rho_t = \frac{\tau_\epsilon}{\tau_t + \tau_e}$. More precisely, $G(\rho_t) = [g_{ij}(\rho_t)]$, where

$$g_{ij}(\rho_t) = \begin{cases} g_{ij}, & i \text{ and } j \text{ are in the same community} \\ \rho_t g_{ij}, & \text{otherwise.} \end{cases} \tag{4}$$

We characterize the equilibrium strategies of the agents in terms of the Bonachic centrality of the agents in $G(\rho_t)$ in the following lemma.

Lemma 1. *The game described in Section III admits a unique linear Bayes Nash equilibrium.*

The equilibrium strategies of the agents in each generation are of the form

$$a^*_{it} = c_{it}x_{it} + (\frac{1}{1-g} - c_{it})\omega_t, \tag{5}$$

where $c_{it} = \rho_t k_{it}$*, and* $k_t = [k_{1t}, \ldots, k_{nt}]^T$ *is the Bonachic centrality of the agents in graph* $G(\rho_t)$ *given by* $k_t = (I - G(\rho_t))^{-1}\mathbf{1}$.

Proof. See the appendix. ■

V. QUALITY OF THE AGGREGATE LEARNING

The quality of information aggregation is determined by the precision of the publicly learned signal about the state θ_t, i.e., $\tau_t = \text{Var}^{-1}[\theta_t|s_0, \ldots, s_{t-1}]$. We can use the equilibrium strategies of Lemma 1 to derive the update rule for τ_t.

For the equilibrium strategies given by (5), the aggregate action is

$$\bar{a}^*_t = \frac{\rho_t \sum_{i=1}^n k_{it}x_{it}}{n} + (\frac{1}{1-g} - \frac{\rho_t \sum_{i=1}^n k_{it}}{n})\omega_t. \tag{6}$$

Define

$$\hat{s}_t = \frac{\rho_t \sum_{i=1}^n k_{it}x_{it}}{n} + \eta_t, \tag{7}$$

where we also recall that $s_t = \bar{a}^*_t + \eta_t$ was the noisy aggregate action in period t which is available to generations thereafter. $\hat{s}_t$ is the new information contained in s_t, as s_t can be inferred from $\hat{s}_t$ given ω_t and vice versa. This especially yields

$$\text{Var}^{-1}[\theta_t|s_0, \ldots, s_{t-1}, s_t] = \text{Var}^{-1}[\theta_t|\omega_t, \hat{s}_t] = \tau_t + \text{Var}^{-1}[\theta_t|\hat{s}_t], \tag{8}$$

using the independence of ω_t and $\hat{s}_t$ conditional on θ_t. We define the information centrality of source j in $G(\rho_t)$ denoted with K_{jt}, as the sum of the centralities k_{it} within community j, that is $K_{jt} = \sum_{i \in I_j} k_{it}$. Let also x^j_t denote the private signal provided by the source j to the community members in period t. Noting that $x_{it} = x^j_t$ for $i \in I_j$, we can rewrite (7) as

$$\hat{s}_t = \frac{\rho_t \sum_{j=1}^m K_{jt}x^j_t}{n} + \eta_t. \tag{9}$$

Using this, we can obtain

$$\mathrm{Var}^{-1}[\theta_t|\hat{s}_t] = \frac{1}{\frac{\sum_{j=1}^m K_{jt}^2}{\tau_\epsilon(\sum_{j=1}^m K_{jt})^2} + \frac{1}{\tau_\eta \rho_t^2 \bar{k}_t^2}} = \frac{\tau_\epsilon \bar{k}_t^2}{\frac{\sum_{j=1}^m K_{jt}^2}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho_t^{-2}}, \tag{10}$$

where $\bar{k}_t = \frac{1}{n}\sum_{j=1}^m K_{jt} = \frac{1}{n}\sum_{i=1}^n k_{it}$ is the average centrality in graph $G(\rho_t)$. Given the evolution of the state on the random walk $\theta_{t+1} = \theta_t + \nu_t$, the update rule for the precision of the publicly learned signal $\tau_{t+1} = \mathrm{Var}^{-1}[\theta_{t+1}|s_0, \ldots, s_t]$ can then be obtained as

$$\tau_{t+1} = \frac{\tau_\nu \hat{\tau}_t}{\tau_\nu + \hat{\tau}_t}, \tag{11}$$

where

$$\hat{\tau}_t = \tau_t + \frac{\tau_\epsilon \bar{k}_t^2}{\frac{\sum_{j=1}^m K_{jt}^2}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho_t^{-2}}, \tag{12}$$

from (8) and (10). As the precision of the publicly learned signal (i.e., τ_t) increases, agents put more weight on the history hence revealing less of what they learn from their private information sources. This in turn results in a less precise public signal about the aggregate action. This may reduce the precision of the public prior in the next period, depending on the public and private precisions and the volatility of the random walk, resulting in fluctuations in the precision of the publicly learned signal. However, we can ensure the convergence of the sequence $\{\tau_t\}$ as $t \to \infty$ by imposing a constraint on the relative precision of the public and private signals.

Assumption 1. $\frac{\tau_\eta}{\tau_\epsilon} \le (1-g)^3$.

The rough idea behind this assumption is to make sure that the update rule for τ_t is contracting. (11) and (12) can be viewed as an update rule $\tau_{t+1} = f(\tau_t)$ for some appropriate choice of $f : [0, \tau_v] \to [0, \tau_v]$. As we show in the proof of Theorem 1, Assumption 1 yields $|\frac{\partial f}{\partial \tau}(\tau)| < 1$ for all $0 \le \tau \le \tau_v$, hence making f a contraction[2], which in turn implies the convergence of τ_t to a fixed point of f as $t \to \infty$.

Given a partition of the agents into m information communities $\cup_{j=1}^m I_j$, we quantify the quality of information aggregation for a network structure G as the asymptotic precision of the publicly learned signal. We characterize this in the following theorem.

[2]Since $[0, \tau_v]$ is compact, having $|\frac{\partial f}{\partial \tau}(\tau)| < 1$ for all $0 \le \tau \le \tau_v$ implies $|\frac{\partial f}{\partial \tau}(\tau)| \le \beta$ for some $\beta < 1$, making f a contraction with module β.

Theorem 1. *Let $\tau_t = Var^{-1}[\theta_t | s_0, \ldots, s_{t-1}]$ be the precision of the publicly learned signal about the state in generation t. Assume also that the precisions of the public and private signals satisfy Assumption 1. Then, the sequence $\{\tau_t\}_{t=1}^{\infty}$ is convergent, where the limit $\tau = \lim_{t\to\infty} \tau_t$ is the unique solution of*

$$\frac{\tau^2}{\tau_\nu - \tau} = \frac{\tau_\epsilon \bar{k}^2}{\frac{\sum_{j=1}^{m} K_j^2}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho^{-2}}, \tag{13}$$

where $\rho = \frac{\tau_\epsilon}{\tau + \tau_\epsilon}$, $k = [k_1, \ldots, k_n]^T = (I - G(\rho))^{-1}\mathbf{1}$, $\bar{k} = \frac{1}{n}\sum_{i=1}^{n} k_i$, and $K_j = \sum_{i \in I_j} k_i$.

Proof. See the appendix. ∎

We can use the above theorem to derive several interesting insights about the interplay of the information-action structure and the volatility of the random walk on the quality of aggregate learning. We are specially interested in characterizing such effect for the extreme cases of slow and fast walks.

Proposition 1. *Denote with $\tau(G, \mathcal{I})$ the asymptotic quality of the publicly learned signal about the state (i.e., the asymptotic precision of the common prior in each generation) for information-action structure $(G, \mathcal{I})$. Then,*

$$\lim_{\tau_\nu \to 0} \frac{\tau_\nu}{\tau(G, \mathcal{I})} = 1, \tag{14}$$

for fast walks. For the slow walks, we have

$$\lim_{\tau_\nu \to \infty} \frac{\sqrt[4]{\tau_\eta \tau_\epsilon^2 \bar{\kappa}^2(G, \mathcal{I}) \tau_\nu}}{\tau(G, \mathcal{I})} = 1, \tag{15}$$

where $\bar{\kappa}(G, \mathcal{I}) = \frac{1}{n}\mathbf{1}^T (I - G(0))^{-1}\mathbf{1}$.

Proof. See the appendix. ∎

A non-trivial consequence of this result is the inefficiency of learning from public history. For the static case $\theta_t = \theta_0$ (i.e., $\tau_\nu = \infty$), agents can asymptotically learn the state from the history. However, a very small perturbation in the state in each generation significantly degrades the quality of the publicly learned signal according to (15). For instance, when τ is large, in order to improve the quality of the publicly learned signal by a factor of 10, the changes in the state from generation to generation has to become 10^4 times less volatile.

Proposition 1 also reveals the first order effect of the information-action structure on the quality of aggregate learning. This, however, provides no useful information on the qualitative effect of different structures for fast walks. We study the higher order effects of the information-action structure on the quality of aggregate learning in more details in the following subsections.

A. The effect of the information-action structure on aggregate learning for slow walks

In this subsection, we study the effect of the information-action structure on the quality of aggregate learning for slow walks. Proposition 1 already captures the effect of average in-community centralities $\bar{\kappa}(G, \mathcal{I})$ on the quality of aggregate learning. However, it provides no further insight for comparative statics of $\tau(G, \mathcal{I})$ for structures with similar in-community average centralities.

Given an information-interaction structure $(G, \mathcal{I})$, the in-community interactions are determined by $G(0)$ (defined in (4)), while $G - G(0)$ characterizes the inter-community interactions. Define,

$$\bar{\xi}(G, \mathcal{I}) = \frac{1}{n}\mathbf{1}^T(I - G(0))^{-1}(G - G(0))(I - G(0))^{-1}\mathbf{1}. \tag{16}$$

Then, $\bar{\xi}(G, \mathcal{I})$ captures the inter-community interaction of local (in-community) influencers and influencées. We can show the following result.

Proposition 2. *For any two information-action structures $(G^1, \mathcal{I}^1)$ and $(G^2, \mathcal{I}^2)$, there exists $\bar{\tau}_\nu > 0$ such that for $\tau_\nu \geq \bar{\tau}_\nu$,*
i) if $\bar{\kappa}(G^1, \mathcal{I}^1) > \bar{\kappa}(G^2, \mathcal{I}^2)$, then $\tau(G^1, \mathcal{I}^1) > \tau(G^2, \mathcal{I}^2)$.
ii) if $\bar{\kappa}(G^1, \mathcal{I}^1) = \bar{\kappa}(G^2, \mathcal{I}^2)$, then $\tau(G^1, \mathcal{I}^1) > \tau(G^2, \mathcal{I}^2)$ if $\bar{\xi}(G^1, \mathcal{I}^1) > \bar{\xi}(G^2, \mathcal{I}^2)$.

Proof. See the appendix. ■

B. The effect of the information-action structure on aggregate learning for fast walks

When $\tau_\nu \to 0$, the average in-community centrality approaches $\frac{1}{1-g}$ since $\frac{1}{n}\mathbf{1}^T(I - G)^{-1}\mathbf{1} = \frac{1}{1-g}$ for all information-action structures with externality coefficient g. In this case, the quality of aggregate learning is mainly determined by the dispersion of centrality across communities (or size of the communities given that all the centralities approach $\frac{1}{1-g}$ in limit), as well as the effect of the out-of-community influencées via inter-community interactions, as we elaborate in what follows.

Define,

$$\psi(G,\mathcal{I}) = (I-G)^{-1}(G-G(0))(I-G)^{-1}\mathbf{1}. \tag{17}$$

$\psi(G,\mathcal{I})$ captures the effect of the out-of-community influencées via direct interaction or interactions of other community members. Let

$$\bar{\psi}(G,\mathcal{I}) = \frac{1}{n}\mathbf{1}^T(I-G)^{-1}(G-G(0))(I-G)^{-1}\mathbf{1}, \tag{18}$$

and $\Psi_j = \sum_{i\in I_j}\psi_i$, where $\psi = [\psi_1,\ldots,\psi_n]^T$. The next proposition characterizes the quality of aggregate learning when state moves on a fast walk.

Proposition 3. *For any two information-action structures* $(G^1,\mathcal{I}^1)$ *and* $(G^2,\mathcal{I}^2)$*, with* $\mathcal{I}^1 = \{I_1^1,\ldots,I_{m_1}^1\}$ *and* $\mathcal{I}^2 = \{I_1^2,\ldots,I_{m_2}^2\}$*, there exists* $\underline{\tau}_\nu > 0$ *such that for* $\tau_\nu \le \underline{\tau}_\nu$*,*
i) $\tau(G^1,\mathcal{I}^1) > \tau(G^2,\mathcal{I}^2)$ *if*

$$\sum_{j=1}^{m_1}|I_j^1|^2 < \sum_{j=1}^{m_2}|I_j^2|^2. \tag{19}$$

ii) if $\sum_{j=1}^{m_1}|I_j^1|^2 = \sum_{j=1}^{m_2}|I_j^2|^2$*, then* $\tau(G^1,\mathcal{I}^1) > \tau(G^2,\mathcal{I}^2)$ *if*

$$\sum_{j=1}^{m_1}\alpha_j^1\Psi_j(G^1,\mathcal{I}^1) < \sum_{j=1}^{m_2}\alpha_j^2\Psi_j(G^2,\mathcal{I}^2), \tag{20}$$

where $\alpha_j^1 = \frac{\tau_\epsilon}{\tau_\eta}(1-g)^2 - \frac{|I_j^1|}{n} + \frac{\sum_{l=1}^{m_1}|I_l^1|^2}{n^2}$ *and* $\alpha_j^2 = \frac{\tau_\epsilon}{\tau_\eta}(1-g)^2 - \frac{|I_j^2|}{n} + \frac{\sum_{l=1}^{m_2}|I_l^2|^2}{n^2}$*.*

Proof. See the appendix. ■

For the case of equal-sized communities, α_j simplifies to $\frac{\tau_\epsilon}{\tau_\eta}(1-g)^2$ and we have the following result.

Corollary 1. *For any two information-action structures* $(G^1,\mathcal{I}^1)$ *and* $(G^2,\mathcal{I}^2)$*, where* $\mathcal{I}^1$ *and* $\mathcal{I}^2$ *consist of* m_1 *and* m_2 *equal-sized communities, there exists* $\underline{\tau}_\nu > 0$ *such that for* $\tau_\nu \le \underline{\tau}_\nu$*,*
i) if $m_1 > m_2$*, then* $\tau(G^1,\mathcal{I}^1) > \tau(G^2,\mathcal{I}^2)$*.*
ii) if $m_1 = m_2$*, then* $\tau(G^1,\mathcal{I}^1) > \tau(G^2,\mathcal{I}^2)$ *if* $\bar{\psi}(G^1,\mathcal{I}^1) < \bar{\psi}(G^2,\mathcal{I}^2)$*.*

Similar result holds for large population of communities where community sizes have finite mean and variance, that is, $n\to\infty$ with $\mathbb{E}[|I_j|], \text{Var}[|I_j|] < \infty$. Comparing the results for the slow and fast dynamics, we can easily see that inter-community interactions play a more crucial role for slow dynamics as the dispersion of centralities in fast dynamics case dominates the effect of inter-community interactions. More importantly, while the inter-community interaction of local influencers and influencées improves the quality of information aggregation for slow dynamics, the inter-community interaction of global influencers and influencées defects the aggregate learning for fast dynamics. We can use this result to show a phase transition in the quality of information aggregation for the two graphs depicted in Figure 1.

Example 1. *Consider the network structures G^1 and G^2 depicted in Figure 1 over 18 equal-sized information communities. Agents in community j, with $j = 1, \ldots, 16$, receive a fraction δ_2 of their externality from another community, and receive the rest of their externality from other agents within the same community. As for the leader communities, agents in L_1 (L_2) receive a fraction δ_1 ($1 - \delta_1$) of their externality from agents in L_2 (L_1), and receive the rest of their externality from other agents within their own community. For the structures shown in Figure 1, $\delta_1 = 0.2$ and $\delta_2 = 0.3$. We can show that the inter-community effect of both local and global influences quantified by $\bar{\xi}$ and $\bar{\psi}$ is higher in G^1 compared with G^2, i.e. $\bar{\xi}(G^1, \mathcal{I}) > \bar{\xi}(G^2, \mathcal{I})$ and $\bar{\psi}(G^1, \mathcal{I}) > \bar{\psi}(G^2, \mathcal{I})$, for $g \geq 0.6$. This implies a phase transition in the quality of information aggregation between G^1 and G^2. While G^1 has a better performance in terms of aggregating the private information across communities for slow walks, it gets outperformed by G^2 over sufficiently fast walks.*

VI. Conclusions

We investigate the problem of learning from public noisy observations of the history of aggregate actions in a network of agents with linear quadratic payoffs. Each agent also belongs to an information community, where they share a private noisy signal about an underlying dynamic state. We quantify the quality of public learning as the asymptotic precision of the publicly learned signal about the state, and study how the interactions between/within information communities affect the quality of aggregate learning. We first show the inefficiency of learning from the public history: while for a static state, public history asymptotically reveals the true state, a small perturbation in the state from generation to generation significantly degrades learning. As another major contribution, we show that the inter-community interactions exhibit a double-edged effect: while the inter-community interaction of local influencers and influencées improves the quality of learning for slow dynamics, the inter-community interaction of global influencers and influencées defects learning for fast dynamics. We may thus observe a phase transition in quality of learning between different information-action structures.

REFERENCES

Daron Acemoglu, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. The network origins of aggregate fluctuations. *Econometrica*, 80:1977–2016, 2012.

Amir Ajorlou and Ali Jadbabaie. Slopey quantizers are locally optimal for Witsenhausen's counterexample. In *Proceedings of the 55th IEEE Conferene on Decision and Control*, pages 3566–3571, 2016.

Amir Ajorlou, Ali Jadbabaie, and Ali Kakhbod. Dynamic pricing in social networks: The word-of-mouth effect. *Management Science*, 2016.

Franklin Allen and Douglas Gale. Financial contagion. *Journal of Political Economic*, 108:1–33, 2000.

Manuel Amador and Pierre-Olivier Weill. Learning from private and public observations of others' actions. *Journal of Economic Theory*, 147:910–940, 2012.

Terrence August, Robert August, and Hyoduk Shin. Designing user incentives for cybersecurity. *Communications of the ACM*, 57:43–46, 2014.

B. Bamieh, M. Jovanović, P. Mitra, and S. Patterson. Coherence in large-scale networks: Dimension-dependent limitations of local feedback. *IEEE Transactions on Automatic Control*, 57:2235–2249, 2012.

Stefano Battiston, Michelangelo Puliga, Rahul Kaushik, Paolo Tasca, and Guido Caldarelli. Debtrank: Too central to fail? financial networks, the FED and systemic risk. *Scientific Reports*, 2:541, 2012.

Dirk Bergemanna, Tibor Heumanna, and Stephen Morrisb. Information and volatility. *Journal of Economic Theory*, 158:427–465, 2015.

Hoda Bidkhori, David Simchi-Levi, and Yehua Wei. Analyzing process flexibility: A distribution-free approach with partial expectations. 44:291–296, 2016.

Kostas Bimpikis, Ozan Candogan, and Shayan Ehsani. Supply disruptions and optimal network structures. *submitted*, 2016.

Lawrence Blume, David Easley, Jon Kleinberg, Robert Kleinberg, and Éva Tardos. Which networks are least susceptible to cascading failures? In *52nd IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 393–402, 2011.

Yann Bramoulle and Rachel Kranton. Public goods in networks. *Journal of Economic Theory*, 135:478–494, 2007.

Ali Jadbabaie and Alex Olshevsky. On performance of consensus protocols subject to noise: Role of hitting times and network structure. In *Proceedings of the 55th IEEE Conferene on Decision and Control*, pages 179–184, 2016.

Lina Sela Perelman, Waseem Abbas, Xenofon Koutsoukos, and Saurabh Amin. Sensor placement for fault location identification in water networks. *Automatica*, 72:166–176, 2016.

K. Savla, Munther A. Dahleh, and G. Como. Robust network routing under cascading failures. *IEEE Transactions on Network Science and Engineering*, 1:53–66, 2014.

Emile Servan-Schreiber, Justin Wolfers, David M. Pennock, and Brian Galebach. Prediction markets: Does money matter? *Electronic Markets*, 14:243–251, 2004.

Milad Siami and Nader Motee. Fundamental limits and tradeoffs on disturbance propagation in linear dynamical networks. *IEEE Transactions on Automatic Control*, 61:4055–4062, 2016.

Milad Siami, Sadegh Bolouki, Bassam Bamieh, and Nader Motee. Centrality measures in linear consensus networks with structured network uncertainties. *IEEE Transactions on Control of Network Systems*, 2017.

H. Tavafoghi and D. Teneketzis. Multi-dimensional forward contracts under uncertainty for electricity markets. *IEEE Transactions on Control of Networked Systems*, 2016.

Christos Thrampoulidis, Subhonmesh Bose, and Babak Hassibi. Optimal placement of distributed energy storage in power networks. *IEEE Transactions on Automatic Control*, 61:416–429, 2015.

George Forrest Young, Luca Scardovi, and Naomi Ehrich Leonard. Robustness of noisy consensus dynamics with directed communication. In *Proceedings of the American Control Conference*, pages 6312–6317, 2010.

APPENDIX

Proof of Lemma 1. Noting that the conditional expectation $\omega_t = \mathbb{E}[\theta_t|s_0,\dots,s_{t-1}]$ is a sufficient statistics for the history, we consider the linear equilibrium strategies to be of the form $a^*_{it} = c_{it}x_{it} + d_{it}\omega_t$ and aim to find the appropriate coefficients c_{it} and d_{it}. It follows from the independence of ω_t and x_{it} conditional on θ_t that

$$\mathbb{E}[\theta_t|S_{it}] = \mathbb{E}[\theta_t|\omega_t, x_{it}] = \rho_t x_{it} + (1-\rho_t)\omega_t, \tag{21}$$

where we recall that $\rho_t = \frac{\tau_\epsilon}{\tau_t+\tau_\epsilon}$. On the other hand,

$$\mathbb{E}[a^*_{jt}|S_{it}] = c_{jt}\mathbb{E}[x_{jt}|S_{it}] + d_{jt}\omega_t = \begin{cases} c_{jt}x_{it} + d_{jt}\omega_t, & j \in I_i^{-1} \\ c_{jt}\mathbb{E}[\theta_t|S_{it}] + d_{jt}\omega_t, & j \notin I_i^{-1}, \end{cases} \tag{22}$$

where I_i^{-1} denotes the information community to which agent i belongs. Substituting (21) and (22) in (3) and comparing the coefficients of ω_t and x_{it} in both sides of the equality, we arrive at

$$\begin{aligned} c_{it} &= \rho_t + \sum_{j\in I_i^{-1}} g_{ij}c_{jt} + \sum_{j\notin I_i^{-1}} \rho_t g_{ij} c_{jt}, \\ d_{it} &= 1-\rho_t + \sum_{j\in I_i^{-1}} g_{ij}d_{jt} + \sum_{j\notin I_i^{-1}} g_{ij}(d_{jt} + (1-\rho_t)c_{jt}). \end{aligned} \tag{23}$$

It follows from the above that $c_{it}+d_{it} = 1+\sum_{j=1}^n g_{ij}(c_{jt}+d_{jt})$. This, together with $\sum_{j=1}^n g_{ij} = g$ implies that $c_{it}+d_{it} = \frac{1}{1-g}$. The equilibrium strategies can hence be written as

$$a^*_{it} = c_{it}x_{it} + (\frac{1}{1-g} - c_{it})\omega_t, \tag{24}$$

where c_{it} satisfies (23). With a change of variables $c_{it} = \rho_t k_{it}$, this then becomes

$$k_{it} = 1 + \sum_{j\in I_i^{-1}} g_{ij}k_{jt} + \sum_{j\notin I_i^{-1}} \rho_t g_{ij} k_{jt}, \tag{25}$$

whose solution is $k_t = [k_{1t},\dots,k_{nt}]^T = (I - G(\rho_t))^{-1}\mathbf{1}$, which completes the proof. ■

Proof of Theorem 1. We prove the convergence of the sequence $\{\tau_t\}$ by showing that $\tau_{t+1} = f(\tau_t)$, where f is uniquely determined from (11) and (12), is a contraction mapping. Noting the compactness of the support of f (i.e., $\tau_{t+1} \in [0,\tau_v]$), it suffices to show that $|\frac{\partial \tau_{t+1}}{\partial \tau_t}| < 1$. From

(11), we have $0 \leq \frac{\partial \tau_{t+1}}{\partial \hat{\tau}_t} < 1$. Therefore, we aim to show that $|\frac{\partial \hat{\tau}_t}{\partial \tau_t}| < 1$ using Assumption 1. Define

$$\Delta_t^\tau = \frac{\tau_\epsilon \bar{k}_t^2}{\frac{\sum_{j=1}^m K_{jt}^2}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho_t^{-2}}. \tag{26}$$

Then, noting (12) we need to show that $-2 < \frac{\partial \Delta_t^\tau}{\partial \tau_t} < 0$. Using $\frac{\partial \rho_t}{\partial \tau_t} = -\frac{\tau_\epsilon}{(\tau_t+\tau_\epsilon)^2}$ and that $0 < \rho_t \leq 1$, it thus suffices to show that $0 < \frac{\partial \Delta_t^\tau}{\partial \rho_t} < 2\tau_\epsilon$.

Define $\hat{k}_t = [\hat{k}_{1t}, \ldots, \hat{k}_{nt}]^T = \frac{\partial k_t}{\partial \rho_t}$ (similarly, $\hat{K}_{jt} = \sum_{i \in I_j} \hat{k}_{it}$). Then, it follows from $k_t = (I - G(\rho_t))^{-1}\mathbf{1}$ that

$$\hat{k}_t = (I - G(\rho_t))^{-1}\frac{\partial G(\rho_t)}{\partial \rho_t} k_t \leq \frac{g}{1-g} k_t, \tag{27}$$

where we have used $k_t \leq \frac{\mathbf{1}}{1-g}$ and $\frac{\partial G(\rho_t)}{\partial \rho_t}\mathbf{1} \leq g\mathbf{1}$. As a result, we have $\frac{\sum_{j=1}^m K_{jt}}{n} = \frac{\sum_{i=1}^n k_{it}}{n} \leq \frac{1}{1-g}$ and $\frac{\sum_{j=1}^m \hat{K}_{jt}}{n} = \frac{\sum_{i=1}^n \hat{k}_{it}}{n} \leq \frac{g}{(1-g)^2}$. We now get back to proving that $0 < \frac{\partial \Delta_t^\tau}{\partial \rho_t} < 2\tau_\epsilon$. We start by showing that the denumerator of Δ_t^τ is decreasing with ρ_t. We write,

$$\begin{aligned}
\frac{\partial \frac{\sum_{j=1}^m K_{jt}^2}{n^2}}{\partial \rho_t} + \frac{\tau_\epsilon}{\tau_\eta}\frac{\partial \rho_t^{-2}}{\partial \rho_t} &= 2\frac{\sum_{j=1}^m K_{jt}\hat{K}_{jt}}{n^2} - 2\frac{\tau_\epsilon}{\tau_\eta}\rho_t^{-3} \\
&\leq 2\frac{\sum_{j=1}^m K_{jt}}{n}\frac{\sum_{j=1}^m \hat{K}_{jt}}{n} - \frac{2}{(1-g)^3} \\
&\leq \frac{2g}{(1-g)^3} - \frac{2}{(1-g)^3} < 0,
\end{aligned} \tag{28}$$

using Assumption 1. On the other hand, $\frac{\partial \bar{k}_t^2}{\partial \rho_t} = 2\bar{k}_t\frac{\sum_{i=1}^n \hat{k}_{it}}{n} \geq 0$. These together imply that $\frac{\partial \Delta_t^\tau}{\partial \rho_t} > 0$. Similarly, we can show that $\frac{\partial \Delta_t^\tau}{\partial \rho_t} < 2\tau_\epsilon$ is equivalent to

$$\bar{k}_t\frac{\sum_{i=1}^n \hat{k}_{it}}{n}(\frac{\sum_{j=1}^m K_{jt}^2}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho_t^{-2}) - \bar{k}_t^2(\frac{\sum_{j=1}^m K_{jt}\hat{K}_{jt}}{n^2} - \frac{\tau_\epsilon}{\tau_\eta}\rho_t^{-3}) < (\frac{\sum_{j=1}^m K_{jt}^2}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho_t^{-2})^2. \tag{29}$$

To prove this, we write

$$\begin{aligned}
\text{LHS} &\leq \frac{g}{(1-g)^3}\frac{\sum_{j=1}^m K_{jt}^2}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho_t^{-3}(\frac{g}{(1-g)^3} + \frac{1}{(1-g)^2}) \\
&< \frac{\tau_\epsilon}{\tau_\eta}\frac{\sum_{j=1}^m K_{jt}^2}{n^2} + \frac{\tau_\epsilon^2}{\tau_\eta^2}\rho_t^{-3} \\
&< (\frac{\sum_{j=1}^m K_{jt}^2}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho_t^{-2})^2.
\end{aligned} \tag{30}$$

This completes the proof of the convergence of $\{\tau_t\}$ to the fixed point $\tau = \tau_{t+1} = f(\tau_t) = \tau_t$.

It follows from (11) and (12) that

$$\frac{\tau^2}{\tau_\upsilon - \tau} = \frac{\tau_\epsilon \bar{k}^2}{\frac{\sum_{j=1}^m K_j^2}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho^{-2}}. \tag{31}$$

The uniqueness of τ easily follows from the fact that RHS is increasing in ρ (hence decreasing in τ) while LHS is increasing in τ. This completes the proof. ■

Proof of Proposition 1. We rewrite (13) as

$$\frac{(\tau+\tau_\epsilon)^2\tau^2}{\tau_\eta\tau_\epsilon} + \frac{\tau^2\sum_{j=1}^m K_j^2}{n^2} + \tau_\epsilon\bar{k}^2\tau = \tau_\epsilon\bar{k}^2\tau_\nu. \tag{32}$$

Then (14) follows by dividing both sides by τ and shifting $\tau \to 0$. Similarly, (15) follows by dividing both sides by τ^4 and then shifting $\tau \to \infty$. ■

Proof of Proposition 2. Let us starting by finding some useful but straightforward properties for τ. From (13), we have

$$\tau^2 \geq \frac{\tau_\epsilon(\tau_\nu - \tau)}{\frac{1}{(1-g)^2} + \frac{(\tau+\tau_\epsilon)^2}{\tau_\epsilon\tau_\eta}}. \tag{33}$$

This yields $\tau \to \infty$ as $\tau_\nu \to \infty$. For any two information-action structures $(G^1, \mathcal{I}^1)$ and $(G^2, \mathcal{I}^2)$ and any $\bar{\tau} > 0$, this implies the existence of $\bar{\tau}_\nu > 0$ such that $\tau(G^1, \mathcal{I}^1), \tau(G^2, \mathcal{I}^2) > \bar{\tau}$ for $\tau_\nu > \bar{\tau}_\nu$. Define,

$$h(G, \mathcal{I}, \rho) = \frac{\tau_\epsilon\rho^2\bar{k}^2(G(\rho))}{\frac{\rho^2\sum_{j=1}^m K_j^2(G(\rho),\mathcal{I})}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}}. \tag{34}$$

Then, $\frac{\tau^2}{\tau_\nu - \tau} = h(G, \mathcal{I}, \rho)$. it is easy to verify that $h(G, \mathcal{I}, 0) = \frac{\partial h}{\partial \rho}(G, \mathcal{I}, 0) = 0$. Also,

$$\begin{aligned} \frac{\partial^2 h}{\partial \rho^2}(G, \mathcal{I}, 0) &= 2\tau_\eta\bar{k}^2(G(0)) = 2\tau_\eta\bar{\kappa}^2(G, \mathcal{I}) \\ \frac{\partial^3 h}{\partial \rho^3}(G, \mathcal{I}, 0) &= 4\tau_\eta\bar{k}(G(0))\frac{\partial\bar{k}(G(0))}{\partial\rho} = 4\tau_\eta\bar{\kappa}(G, \mathcal{I})\bar{\xi}(G, \mathcal{I}), \end{aligned} \tag{35}$$

where we have used $k(G(\rho)) = (I - G(\rho))^{-1}\mathbf{1}$. Now, for both cases (i) and (ii) (i.e., $\bar{\kappa}(G^1, \mathcal{I}^1) > \bar{\kappa}(G^2, \mathcal{I}^2)$, or $\bar{\kappa}(G^1, \mathcal{I}^1) = \bar{\kappa}(G^2, \mathcal{I}^2)$ and $\bar{\xi}(G^1, \mathcal{I}^1) > \bar{\xi}(G^2, \mathcal{I}^2)$), (35) yields $h(G^1, \mathcal{I}^1, \rho) > h(G^2, \mathcal{I}^2, \rho)$ in a vicinity of zero. Since $\rho = \frac{\tau_\epsilon}{\tau+\tau_\epsilon}$, there exists $\bar{\tau} > 0$ such that $h(G^1, \mathcal{I}^1, \rho) > h(G^2, \mathcal{I}^2, \rho)$ for $\tau > \bar{\tau}$. As we discussed before, by choosing τ_ν sufficiently large we can ensure $\tau(G^1, \mathcal{I}^1), \tau(G^2, \mathcal{I}^2) > \bar{\tau}$. Therefore, $h(G^1, \mathcal{I}^1, \rho^1) > h(G^2, \mathcal{I}^2, \rho^1)$, where $\rho^1 = \frac{\tau_\epsilon}{\tau(G^1,\mathcal{I}^1)+\tau_\epsilon}$. Now, if $\rho^1 \geq \rho^2$ ($\rho^2 = \frac{\tau_\epsilon}{\tau(G^2,\mathcal{I}^2)+\tau_\epsilon}$), then $h(G^1, \mathcal{I}^1, \rho^1) > h(G^2, \mathcal{I}^2, \rho^1) \geq h(G^2, \mathcal{I}^2, \rho^2)$, noting that h is increasing in ρ as shown in the proof of Theorem 1. This in turn implies $\tau(G^1, \mathcal{I}^1) > \tau(G^2, \mathcal{I}^2)$ since $\frac{\tau^2}{\tau_\nu - \tau} = h(G, \mathcal{I}, \rho)$, which clearly contradicts $\rho^1 \geq \rho^2$. Therefore, we should have $\rho^1 < \rho^2$ or $\tau(G^1, \mathcal{I}^1) > \tau(G^2, \mathcal{I}^2)$ which completes the proof. ■

Proof of Proposition 3. Part (i) easily follows from the fact that if $h(G^1, \mathcal{I}^1, 1) > h(G^2, \mathcal{I}^2, 1)$, then $\tau(G^1, \mathcal{I}^1) > \tau(G^2, \mathcal{I}^2)$, where $h(G, \mathcal{I}, \rho) = \frac{\tau_\epsilon \bar{k}^2(G(\rho))}{\frac{\sum_{j=1}^m K_j^2(G(\rho), \mathcal{I})}{n^2} + \frac{\tau_\epsilon}{\tau_\eta}\rho^{-2}}$. For the second part, similar to the proof of Proposition 2, we can show that if $\frac{\partial h}{\partial \rho}(G^1, \mathcal{I}^1, 1) < \frac{\partial h}{\partial \rho}(G^2, \mathcal{I}^2, 1)$, then $\tau(G^1, \mathcal{I}^1) > \tau(G^2, \mathcal{I}^2)$. The proof then follows from simple algebra and noting that $K_j(G(1), \mathcal{I}) = \frac{|I_j|}{1-g}$ and $\frac{\partial k(G(1))}{\partial \rho} = \psi(G, \mathcal{I})$. ■