DISPERSED BEHAVIOR AND PERCEPTIONS IN ASSORTATIVE SOCIETIES

By

Mira Frick, Ryota Iijima, and Yuhta Ishii

April 2018

Revised October 2021

COWLES FOUNDATION DISCUSSION PAPER NO. 2128R2

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Dispersed Behavior and Perceptions in Assortative Societies*

Mira Frick Ryota Iijima Yuhta Ishii

October 26, 2021

Abstract

We formulate a model of social interactions and misinferences by agents who *neglect assortativity* in their society, mistakenly believing that they interact with a representative sample of the population. A key component of our approach is the interplay between this bias and agents' strategic incentives. We highlight a mechanism through which assortativity neglect, combined with strategic complementarities in agents' behavior, drives up action dispersion in society (e.g., socioeconomic disparities in education investment). We also show how the combination of assortativity neglect and strategic incentives may help to explain empirically documented misperceptions of income inequality and political attitude polarization.

*First posted version: April 2018. Frick: Yale University (mira.frick@yale.edu); Iijima: Yale University (ryota.iijima@yale.edu); Ishii: Pennsylvania State University (yxi5014@psu.edu). This research was supported by NSF grant SES-1824324. We are extremely grateful to co-editor Roland Bénabou and four anonymous referees for insightful suggestions that greatly improved the paper. We also thank Nageeb Ali, Attila Ambrus, Dirk Bergemann, Ron Berman, Phoebe Cai, Kyle Chauvin, Benjamin Enke, Drew Fudenberg, Simone Galperti, Wayne Yuan Gao, Helena Garcia, Ben Golub, Andrei Gomberg, Sanjeev Goyal, Marina Halac, Johannes Hörner, Matt Jackson, Philippe Jehiel, Navin Kartik, Daria Khromenkova, David Miller, Weicheng Min, Stephen Morris, Pietro Ortoleva, Romans Pancs, Jacopo Perego, Larry Samuelson, Joel Sobel, Rani Spiegler, Philipp Strack, Linh To, Leeat Yariv, as well as audiences at ASU Economic Theory Conference, Berkeley, Bielefeld, Bonn, BRIC (Columbia), Cambridge, Cornell, Drexel, Duke, East Coast Behavioral and Experimental Workshop (Maryland), Glasgow, ITAM, Kyoto, Marketing-IO Conference (Yale), Michigan, Michigan State, NSF Conference on Network Science in Economics (Vanderbilt), Penn State, Princeton, Rochester, St. Andrews, Stanford, Tokyo, TSE, UCLA, UT Austin, WE_ARE online seminar series, and Yale.

1 Introduction

A central channel through which people learn about their societies is by interacting with and observing the behavior among their peers (e.g., neighbors, coworkers, online acquaintances). Peers' behavior (e.g., their consumption choices or political activities) may provide information about behavior in society as a whole, as well as about key population characteristics (e.g., income or political attitude distributions). However, many social interactions are *assortative*, in the sense that people interact more with others with similar characteristics: Richer people are more likely than poorer people to have rich friends, and conservatives are more likely than liberals to know other conservatives; indeed, evidence suggests that societies may be growing increasingly assortative.[1] As a result, the behavior that individuals observe among their peers need not be representative of society. At the same time, there is ample evidence from psychology and behavioral economics that people are prone to misinferences from non-representative data (see Section 3.1).

In this paper, we formulate a model of social interactions and misinferences by agents who suffer from *assortativity neglect*, i.e., mistakenly believe that they interact with a representative (or more representative than actual) sample of the population. A key component of our approach is the interplay between this bias and agents' strategic incentives. In particular, we highlight a mechanism through which assortativity neglect, combined with strategic complementarities in agents' behavior, drives up action dispersion in society (e.g., socioeconomic disparities in education investment). We also show how the combination of assortativity neglect and strategic incentives may help to explain some central empirical findings about people's misperceptions of the income and political attitude distributions in their societies.

Our main contributions are twofold. First, we introduce an equilibrium concept that allows us to analyze the interplay between agents' strategic behavior and the misinferences they draw from their peers' behavior under assortativity neglect. We consider population games, where agents with ordered types are matched in an assortative manner, and each agent's optimal action may depend on her type, the *global* action distribution in society as a whole, as well as the *local* action distribution among her matches. In an *assortativity neglect equilibrium (ANE)*, each agent correctly observes her local action distribution, but misperceives this to coincide with the global action distribution and best-responds based on this misperception. Our main results focus on linear-best response games with strategic complementarities. Here, we show that ANE amplifies action dispersion relative to Nash equilibrium, by generating a gap between high and low types' perceptions of global behavior. This both increases the difference between high and low types' actions in any fixed society, and exacerbates the effect of social changes,

[1] E.g., Jargowsky (1996); Reardon and Bischoff (2011) find increased residential segregation by income, and Bishop (2009); Brown and Enos (2021) document growing segregation by political attitudes in the US.

such as increased assortativity.

Second, we provide a theory of how agents form misperceptions about the type distribution in society. Specifically, suppose that agents seek to explain their observed local action distributions in a *coherent* manner, that is, by assuming that their peers are best-responding to their incentives in the population game. Then we show that assortativity neglect leads agents to systematically misperceive the type distribution, and we characterize how the nature of social interactions shapes these misperceptions. Importantly, in our model, agents draw inferences from their peers' *behavior*, which is subject to strategic motivations; this gives rise to different predictions than if agents directly observed peers' types and made the purely *statistical error* of projecting these onto society. For example, in the latter case, one would expect agents to underestimate type dispersion, as peers are on average less diverse than the overall population. In contrast, we show that our model additionally generates an *attribution error* in agents' reasoning about their peers' incentives; under strategic complementarities, this pushes in the opposite direction of the statistical error and can lead to overestimation of type dispersion.

Misperceptions of income and political attitude distributions have received much attention in recent empirical work, in part due to their potential to affect voters' choices on important policy issues, such as redistribution. As we discuss, the interplay between the statistical and attribution errors that we identify might help shed light on some key findings in this literature, in particular, the fact that both under- and overestimation of income inequality are common and evidence of widespread overestimation of political attitude polarization (Section 4.4).

The paper proceeds as follows. To illustrate the model and some of our main findings, Section 1.1 presents a simple parametric example in the context of education investment and income-based residential sorting. Section 2 introduces general assortative societies and population games. Sections 3 and 4 consider, respectively, agents' equilibrium behavior and formation of coherent perceptions under assortativity neglect. Beyond the Gaussian societies in Section 1.1, ANE strategies and perceptions need not admit closed-form solutions. Instead, a key observation facilitating our analysis is that every assortative society can be recast as a monotone Markov process over its space of types. In linear best-response games with strategic complementarities, this allows us to analyze ANE behavior and perceptions in arbitrary societies by considering the higher-order expectations of this process. Section 5 extends our analysis to weaker forms of assortativity neglect and more general best-response functions (including strategic substitutes) and discusses related literature. Section 6 is a conclusion.

1.1 Illustrative Example

Consider a continuum population of agents, each of whom is identified with an income level $\theta \in \mathbb{R}$. The income distribution F in the population is Gaussian, with mean μ and variance

$\sigma^2 > 0$. Each agent knows her own income, but does not directly observe other agents' incomes. Due to neighborhood sorting by income, the richer an agent the more likely she is to interact with other high-income agents. Specifically, pairwise interaction probabilities between any agents θ and θ' are summarized by a symmetric bivariate Gaussian distribution P with marginal distribution F and correlation coefficient $\rho \in (0,1)$,

$$(\theta, \theta') \sim_P \mathcal{N}\left(\begin{pmatrix} \mu \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix}\right).$$

The higher ρ the greater the degree of assortativity.

Each agent θ chooses a level $s(\theta) \in \mathbb{R}$ of education investment.[2] Assume that θ's best-response against strategy profile s takes the form

$$\mathrm{BR}_\theta(s) = \theta + \beta \underbrace{\mathbb{E}_F[s(\theta')]}_{:=\int s(\theta')\,dF(\theta')} + \gamma \underbrace{\mathbb{E}_P[s(\theta')|\theta]}_{:=\int s(\theta')\,dP(\theta'|\theta)},$$

with $\beta, \gamma \geq 0$, $\beta + \gamma < 1$. Thus, richer agents have an intrinsic tendency to invest more in education; additionally, θ's optimal education investment is increasing in *global* average investment $\mathbb{E}_F[s(\theta')]$ in society as a whole (e.g., due to anticipated competition for college admissions/jobs), and in *local* average investment $\mathbb{E}_P[s(\theta')|\theta]$ among the agents she interacts with (e.g., due to peer effects in learning).[3] In the unique Nash equillibrium, $s^{NE}(\theta) = \frac{\theta-\mu}{1-\gamma\rho} + \frac{\mu}{1-\beta-\gamma}$ for each θ.

Nash equilibrium assumes that agents best-respond to correct perceptions of behavior in society. In contrast, we assume that each agent correctly observes the local distribution of education investments among the agents she interacts with; however, she mistakenly believes this to coincide with the global distribution of education investments in society (which she does not directly observe), because she neglects that society is assortative. We formalize this using the concept of *assortativity neglect equilibrium (ANE)*, where

$$s^{AN}(\theta) = \theta + (\beta + \gamma)\mathbb{E}_P[s^{AN}(\theta')|\theta],$$

i.e., each agent θ's action is a best-response to the correct local average investment $\mathbb{E}_P[s^{AN}(\theta')|\theta]$ but the misperception that this is the same as global average investment. Equilibrium actions are uniquely given by $s^{AN}(\theta) = \frac{\theta-\mu}{1-(\beta+\gamma)\rho} + \frac{\mu}{1-\beta-\gamma}$ for each θ (see Example 1 for details).

[2] We consider education investment to include decisions such as expenditures on educational materials or tutors or the amount of effort exerted at school, but assume school choice (and other decisions that might endogenously affect sorting) to be exogenous (e.g., because everyone enrolls in their neighborhood school).

[3] Bénabou (1993, 1996a,b); Fernández and Rogerson (1996, 2001); Durlauf (1996) consider related models (without misperception) of education investment with sorting and global and/or local complementarities.

Increased action dispersion. Our first main finding is that, under strategic complementarities, assortativity neglect increases action dispersion relative to Nash, through two channels. First, in any given society, ANE leads to greater socioeconomic differences in education investment than Nash: Average education investment in society is the same under ANE and Nash, but the variance is higher under assortativity neglect,

$$\mathrm{Var}_F\left[s^{AN}(\theta)\right] = \left(\frac{1}{1-(\beta+\gamma)\rho}\right)^2 \sigma^2 \geq \mathrm{Var}_F\left[s^{NE}(\theta)\right] = \left(\frac{1}{1-\gamma\rho}\right)^2 \sigma^2. \tag{1}$$

The intuition is simple and reflects a mutually reinforcing interplay between agents' misperceptions and behavior: Since richer agents are more likely than poorer ones to interact with other rich agents, they tend to observe higher education investment among their peers. Under assortativity neglect, this gives rise to a "false consensus effect:" Perceptions of global average education investment in society, and hence of returns to education, are increasing in agents' income and thus in their own investment. Relative to correct perceptions, this increases education investment differences between the rich and poor, which in turn, through observation of their peers' investment, feeds into the false consensus effect.

Second, assortativity neglect acts as a *multiplier* of social changes that increase action dispersion. For instance, the effect of an increase in the degree ρ of neighborhood sorting is

$$\frac{\partial}{\partial\rho}\mathrm{Var}_F\left[s^{AN}(\theta)\right] \geq \frac{\partial}{\partial\rho}\mathrm{Var}_F\left[s^{NE}(\theta)\right] \geq 0.$$

Thus, socioeconomic education differences rise under Nash, but even more so under ANE. Intuitively, greater sorting has a direct effect on the education gap under Nash, by increasing differences in local peer effects between richer and poorer agents. However, under ANE, this additionally magnifies the false consensus effect, because both richer and poorer agents mistakenly attribute these new local education investment levels to a global trend in society, further polarizing their responses. An increase in income inequality σ^2 has an analogous effect.

Coherent perceptions under assortativity neglect. Next, we ask whether and how agents can "make sense" of their observed local action distributions through the lens of assortativity neglect. That is, when an agent θ suffers from assortativity neglect, can she explain the distribution of education investments she observes among her peers by assuming that they are behaving optimally? Our second main finding is that the answer is yes, but that to explain her observations, θ must misperceive the income distribution in society in a particular way.

Specifically, under ANE, the local distribution of education investments that θ observes has mean $\mathbb{E}_P[s^{AN}(\theta')|\theta] =: \overline{a}_\theta$ and variance $\mathrm{Var}_P[s^{AN}(\theta') \mid \theta]$. Since θ neglects assortativity, she believes that this local distribution is representative of the investment distributions in all other neighborhoods and in society as a whole. Thus, from θ's perspective, each agent θ''s optimal

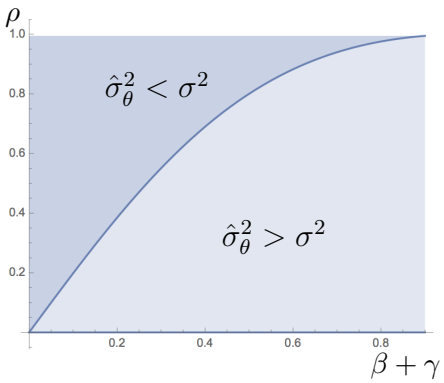

Figure 1: Under-/overestimation of income inequality. (Perceptions are correct if $\rho = 0$ or $\rho = \frac{2(\beta+\gamma)}{1+(\beta+\gamma)^2}$).

education investment choice is $\theta' + (\beta + \gamma)\overline{a}_\theta$. As Example 2 verifies, the only way that θ can explain her observed local investment mean and variance as arising from such optimal choices is if θ perceives the income mean and variance among her peers, and hence in society, to be

$$\hat{\mu}_\theta = (1 - \beta - \gamma)\overline{a}_\theta, \qquad \hat{\sigma}_\theta^2 = \mathrm{Var}_P[s^{AN}(\theta') \mid \theta] = \frac{\sigma^2(1-\rho^2)}{(1-(\beta+\gamma)\rho)^2}. \tag{2}$$

Misperceptions of income inequality. Based on this, we can examine how agents' misperceptions of the income distribution are influenced by the nature of their social interactions. For example, consider θ's perceived income inequality $\hat{\sigma}_\theta^2$ in (2). This is increasing in $\beta+\gamma$, and exceeds the true income inequality σ^2 if and only if $\frac{2(\beta+\gamma)}{1+(\beta+\gamma)^2} > \rho$. Thus, as Figure 1 shows, θ underestimates (resp. overestimates) income inequality when complementarities $\beta+\gamma$ are small (resp. large) relative to the degree of assortativity ρ.

This finding reflects two opposing errors in θ's reasoning under assortativity neglect. On the one hand, a purely *statistical error*: Income inequality $\sigma^2(1-\rho^2)$ among θ's peers is lower than in the overall population. Thus, viewing her peers as representative of society pushes θ to underestimate income inequality. On the other hand, an *attribution error*: Rather than directly observing her peers' incomes, θ must infer these from their observed investment decisions. However, due to her assortativity neglect, θ fails to take into account that the rich and poor are subject to different peer effects, because she mistakenly believes that everyone faces the *same* local average investment $\overline{a}_\theta$. As a result, she misattributes all observed investment differences to variation in income. This creates a force for overestimating income inequality. Under larger complementarities, the second channel is stronger and can dominate the first one.

2 Setting

There is a continuum of agents with mass normalized to 1. Each agent is endowed with a type $\theta \in \mathbb{R}$. An agent's type is her private information. Agents interact according to a random

matching technology. A society P specifies the probability with which any pair of types θ and θ' are matched:[4]

Definition 1. A ***society*** is a joint cdf P over $\mathbb{R} \times \mathbb{R}$ that is:

1. *symmetric*: $P(\theta, \theta') = P(\theta', \theta)$ for all θ, θ'
2. *assortative*: $P(\cdot|\theta)$ first-order stochastically dominates $P(\cdot|\theta')$ if $\theta \geq \theta'$.

Symmetry is a consistency condition required to describe a random matching in a population. Assortativity captures the idea that higher types are (weakly) more likely than lower types to interact with other high types. Note that a society P jointly summarizes an underlying ***type distribution***, described by the marginal distribution $F := \text{marg}_\Theta P$, and a ***matching technology***, which for every type θ specifies the conditional distribution $P(\cdot|\theta)$ of θ's matches. We assume that the type distribution F is absolutely continuous with $\int |\theta| dF(\theta) < \infty$ and has a connected support, denoted by Θ. Let $\mathcal{F}$ denote the set of all cdfs with these properties. We call society P ***non-assortative*** if $P = F \times F$ is the independent product of its marginals, so that each type θ's match distribution $P(\cdot|\theta) = F$ coincides with the type distribution in society as a whole. In the Gaussian parametrization in Section 1.1, P is non-assortative if and only if the correlation coefficient $\rho = 0$.

Society P is engaged in the following population game. Agents have a common action set, given by a measurable $A \subseteq \mathbb{R}$. A strategy profile is a measurable function $s : \Theta \to A$ that specifies an action $s(\theta)$ for each type θ and satisfies $\int |s(\theta)| dF(\theta) < \infty$. Each strategy profile s induces a ***global action distribution*** $G^{s,P}$, i.e., the cdf over actions when types are drawn according to F and behave according to s:

$$G^{s,P}(a) := \int_\Theta \mathbb{1}_{\{s(\theta') \leq a\}} \, dF(\theta') \text{ for all } a \in A.$$

For each type θ, s also induces a ***local action distribution*** $L_\theta^{s,P}$. This is the distribution of actions among θ's matches, i.e., the cdf over actions when types are drawn according to $P(\cdot|\theta)$:

$$L_\theta^{s,P}(a) := \int_\Theta \mathbb{1}_{\{s(\theta') \leq a\}} \, dP(\theta'|\theta) \text{ for all } a \in A.$$

Note that when $P = F \times F$ is non-assortative, each type's local action distribution $L_\theta^{s,P}$ coincides with the global action distribution $G^{s,P}$ for all s. However, $L_\theta^{s,P}$ generally differs from $G^{s,P}$ when P is assortative. Finally, for each type θ, the game specifies a ***best-response correspondence*** $\text{BR}_\theta : \Delta(A) \times \Delta(A) \to A$. For any strategy profile s, the set of optimal actions for θ is

[4]Throughout, we treat society P as exogenous. See Pin and Rogers (2016) for a survey of potential sources of assortativity, including institutional constraints or socio-psychological factors.

given by $\text{BR}_\theta(G^{s,P}, L_\theta^{s,P})$, which depends on s only through the induced global and local action distributions $G^{s,P}$ and $L_\theta^{s,P}$.

The next section defines an equilibrium concept that applies to any population game of the above form. However, our main results will focus on ***linear best-response functions with strategic complementarities*** (Section 5.2 discusses how to extend the analysis beyond this case, including to settings with strategic substitutes): Here, $A = \mathbb{R}$ and there exist coefficients $\beta, \gamma \geq 0$ with $\beta + \gamma < 1$ such that each type θ's best-response against strategy profile s is the unique action given by

$$\text{BR}_\theta(G^{s,P}, L_\theta^{s,P}) := \theta + \beta \int a \, dG^{s,P}(a) + \gamma \int a \, dL_\theta^{s,P}(a) = \theta + \beta \mathbb{E}_F[s(\theta')] + \gamma \mathbb{E}_P[s(\theta')|\theta]. \quad (3)$$

The first term captures that higher types have an intrinsic tendency to take higher actions. The second term captures society-wide strategic complementarities, whereby each type θ's best-response is increasing in the global average action. Finally, by the third term, θ's best-response is also increasing in the average behavior among her matches, reflecting local strategic complementarities.[5]

While stylized, best-response functions of this form are widely used in the literature on network games (for a survey, see Jackson and Zenou, 2013) and can capture a rich class of economic applications. In addition to education investment (Section 1.1), other examples include many consumption decisions that depend on income positively, but may also exhibit both direct peer-to-peer consumption complementarities and material or socio-psychological global payoff complementarities (e.g., economy-wide technological spillovers or a desire to adhere to a social norm). Likewise, types might represent political attitudes on a left-right spectrum and actions the extent to which agents manifest support for particular positions, where related local and global complementarity/conformity motives may be at play.

3 Behavior under Assortativity Neglect

3.1 Assortativity Neglect Equilibrium

The standard solution concept of Nash equilibrium assumes that all agents best-respond to correct perceptions about others' behavior. That is, in our population game, strategy profile s is a ***Nash equilibrium*** if each type θ's action $s(\theta) \in \text{BR}_\theta(G^{s,P}, L_\theta^{s,P})$ is a best-response to the true global and local action distributions under s.

[5] While our analysis takes best-response functions as its primitive, one possible utility function that induces (3) is $U(a, \theta, G, L_\theta) = -\int\int \left(a - \theta - \beta a' - \gamma a''\right)^2 dG(a') \, dL_\theta(a'')$; that is, θ faces a quadratic miscoordination cost that reflects the gap between her action a and a weighted sum of her type and the realized actions a' in society and a'' among her matches.

In this paper, we introduce an alternative solution concept, assortativity neglect equilibrium. Here, each agent θ correctly perceives her local action distribution $L_\theta^{s,P}$, but mistakenly perceives the global action distribution to coincide with this local action distribution, and best-responds based on this misperception:

Definition 2. A strategy profile s is an ***assortativity neglect equilibrium (ANE)*** if $s(\theta) \in \mathrm{BR}_\theta(L_\theta^{s,P}, L_\theta^{s,P})$ for each θ.

One interpretation of ANE is as the steady state of a hypothetical dynamic setting, where successive generations of agents choose actions based on (i) only observing the behavior in their local neighborhoods and (ii) the misperception that society is non-assortative, which we refer to as ***assortativity neglect***. Concretely, suppose that before a t-th generation agent θ chooses her action $s_t(\theta)$, the only information available to her is the local action distribution $L_\theta^{s_{t-1},P}$ among the previous generation's types from her match distribution $P(\cdot|\theta)$. She does not observe the global action distribution $G^{s_{t-1},P}$ (nor her matches' types or payoffs). However, since she suffers from assortativity neglect, she mistakenly believes $G^{s_{t-1},P}$ to coincide with her observed local action distribution, because she perceives her matches to be a representative sample of society. Given this, she best-responds by choosing $s_t(\theta) \in \mathrm{BR}_\theta(L_\theta^{s_{t-1},P}, L_\theta^{s_{t-1},P})$. Steady states of this setting correspond to the fixed-point condition in Definition 2.[6]

Agents' misperception that their observed local action distributions match behavior in society can be viewed as a manifestation of the idea in cognitive psychology that people are "naive intuitive statisticians" (Fiedler and Juslin, 2006; Juslin, Winman, and Hansson, 2007), who take for granted that their observed samples are representative. More broadly, this relates to the "What You See Is All There Is" bias (Kahneman, 2011) and the experimental literature on "selection neglect" (Esponda and Vespa, 2014, 2018; Barron, Huck, and Jehiel, 2019; Enke, 2020), which documents various learning settings where subjects fail to recognize that the information they see is subject to selection effects. Misperceiving society to be non-assortative also gives rise to a form of "projection bias," where each agent projects her own local action observations onto everyone else in society.[7]

Finally, a basic idea motivating ANE is that agents are better able to observe their local action distributions than the global action distribution. This seems natural in many applications, such as the setting in Section 1.1: Education investment levels among one's peers are readily observable based on day-to-day interactions, but learning about education investment levels in society as a whole might require one to research other sources or to draw inferences

[6] Analogous learning motivations, but under different forms of misinference from observed feedback, underlie solution concepts such as analogy-based expectation and cursed equilibrium; see Section 5.3.

[7] Related projections of local observations are documented by the literature on "network cognition" (for a survey, see Brands, 2013): E.g., Dessi, Gallo, and Goyal (2016) elicit subjects' assessments of degree distributions on a network and find that subjects project their own number of neighbors onto others in the network.

from peers' payoffs (e.g., labor market outcomes) that may not yet be known at the time of choosing one's own education investment.[8] At the same time, the assumption underlying ANE that agents have no information about the global action distribution and view their local action distributions as fully representative of it is, of course, extreme. In Section 5.1, we consider less extreme formulations, where some fraction of agents know the true global action distribution or agents only partially project their local action distributions onto the global distribution.

Remark 1. ANE clearly coincides with Nash equilibrium in non-assortative societies P or in population games where best-responses $\mathrm{BR}_\theta(G^{s,P}, L_\theta^{s,P}) = \mathrm{BR}_\theta(L_\theta^{s,P})$ depend only on local action distributions. Beyond these special cases, the two solution concepts generally differ, even in the widely studied setting where best-responses $\mathrm{BR}_\theta(G^{s,P}, L_\theta^{s,P}) = \mathrm{BR}_\theta(G^{s,P})$ depend only on global action distributions (i.e., $\gamma = 0$ in the case of linear best-responses). In the latter environment, local action distributions are not directly relevant for *incentives* and hence play no role under Nash, but they nevertheless affect ANE through agents' *misinference* that their local action distributions match the global action distribution. We allow best-responses to depend on both global and local action distributions as this is natural in many economic applications, but our main insights will apply even under purely global strategic externalities. ▲

3.2 Action Dispersion under Assortativity Neglect

To analyze and contrast behavior under Nash and ANE in detail, we focus on the linear best-response functions with strategic complementarities given by (3).

A key observation facilitating our analysis is the following. Even though our model is static, we can think of any society P as inducing a discrete-time Markov process over its space of types Θ: The initial distribution is given by the type distribution $F = \mathrm{marg}_\Theta P$ and the transition kernel is the matching technology $(P(\cdot|\theta))_{\theta\in\Theta}$. That is, this process first draws an initial type $\theta_0 \in \Theta$ according to F, then draws type θ_0's match θ_1 according to $P(\cdot|\theta_0)$, type θ_1's match θ_2 according to $P(\cdot|\theta_1)$, and so on. We refer to this Markov process as the process of ***t*-step ahead matches** in society and also denote it by P. Assortativity of P corresponds precisely to this process being ***monotone*** (Daley, 1968); this feature will play an essential role throughout our analysis.

The process of t-step ahead matches yields a simple description of the Nash equilibrium of our game. By the best-response condition (3), s is a Nash equilibrium if and only if, for all θ,

$$s(\theta) = \mathrm{BR}_\theta(G^{s,P}, L_\theta^{s,P}) = \theta + \beta\mathbb{E}_F[s(\theta')] + \gamma\mathbb{E}_P[s(\theta')|\theta]. \tag{4}$$

By iterating this fixed-point condition under the Markov process P and exploiting the linearity

[8] Indeed, Section 3.4 discusses empirical evidence on misperceptions of returns to education.

of the best-response function, we obtain that, for all θ and $\tau \in \mathbb{N}$,

$$\begin{aligned} s(\theta) &= \theta + \beta \mathbb{E}_F[s(\theta')] + \gamma \mathbb{E}_P\left[s(\theta_1) \mid \theta_0 = \theta\right] \\ &= \theta + \beta \mathbb{E}_F[s(\theta')] + \gamma \Big(\mathbb{E}_P\left[\theta_1 \mid \theta_0 = \theta\right] + \beta \mathbb{E}_F[s(\theta')] + \gamma \mathbb{E}_P\left[s(\theta_2) \mid \theta_0 = \theta\right] \Big) = \dots \\ &= \sum_{t=0}^{\tau} \gamma^t \Big(\mathbb{E}_P\left[\theta_t \mid \theta_0 = \theta\right] + \beta \mathbb{E}_F[s(\theta')] \Big) + \gamma^{\tau+1} \mathbb{E}_P[s(\theta_{\tau+1}) \mid \theta_0 = \theta], \end{aligned}$$

where each step applies the law of iterated expectations. In Appendix B.1, we verify that the higher-order term $\gamma^{\tau+1}\mathbb{E}_P[s(\theta_{\tau+1})|\theta_0 = \theta]$ vanishes as $\tau \to \infty$ and obtain the following result:

Lemma 1 (Nash equilibrium)**.** *For any (P, β, γ), there exists a unique Nash equilibrium. Nash strategies are strictly increasing in types, with*

$$s^{NE}(\theta) = \sum_{t=0}^{\infty} \gamma^t \mathbb{E}_P[\theta_t|\theta_0 = \theta] + \frac{\beta \mathbb{E}_F[\theta']}{(1-\gamma)(1-\beta-\gamma)} \quad \textit{for all } \theta. \tag{5}$$

Thus, θ's Nash equilibrium action is a weighted average of a γ-discounted sum of her expected t-step ahead matches and a constant that depends only on β, γ, and the type mean $\mathbb{E}_F[\theta']$ in society. Her behavior is increasing in her type for two reasons: First, higher types prefer higher actions. Second, due to local complementarities, θ's behavior is increasing in her matches' behavior, which in turn is increasing in their matches' behavior, etc., and higher types are more likely to meet other high types. This is reflected by the fact that θ's action depends on all the t-step ahead expectations $\mathbb{E}_P[\theta_t|\theta_0 = \theta]$, which are (weakly) increasing in θ due to the monotonicity of the Markov process P.

By contrast, any ANE s must satisfy the fixed-point condition

$$s(\theta) = \mathrm{BR}_\theta(L_\theta^{s,P}, L_\theta^{s,P}) = \theta + (\beta + \gamma)\mathbb{E}_P[s(\theta')|\theta]. \tag{6}$$

Analogous to the derivation of Nash equilibrium, iterating (6) implies that the ANE is uniquely given by

$$s^{AN}(\theta) = \sum_{t=0}^{\infty} (\beta+\gamma)^t \mathbb{E}_P[\theta_t|\theta_0 = \theta] \text{ for all } \theta. \tag{7}$$

That is, θ's action is a $(\beta+\gamma)$-discounted sum of her expected t-step ahead matches under P.[9]

An important implication is that, under global complementarities, assortativity neglect increases action dispersion relative to Nash equilibrium. Formally, comparing (5) and (7), the

[9] In line with the learning motivation of ANE in Section 3.1, (7) can be seen as stable steady state of the following adjustment process: Starting with any monotone strategy profile s_0, if successive generations of agents best-respond to previous period behavior according to $s_t(\theta) \in \mathrm{BR}_\theta(L_\theta^{s_{t-1},P}, L_\theta^{s_{t-1},P})$, play converges to (7).

fact that t-step ahead expectations $\mathbb{E}_P[\theta_t|\theta_0=\theta]$ are increasing in θ and that $\beta \geq 0$ implies

$$s^{AN}(\theta) - s^{AN}(\theta') \geq s^{NE}(\theta) - s^{NE}(\theta') \text{ for all } \theta > \theta'.$$

Equivalently, the global action distribution G^{AN} under ANE is ***more dispersive*** than the Nash global action distribution G^{NE}; that is, $G^{AN^{-1}}(x) - G^{AN^{-1}}(y) \geq G^{NE^{-1}}(x) - G^{NE^{-1}}(y)$ for all $0 < y \leq x < 1$. Since the average actions under (5) and (7) are the same (namely, $\frac{\mathbb{E}_F[\theta]}{1-\beta-\gamma}$), this implies that G^{AN} is a mean-preserving spread of G^{NE}.[10]

Proposition 1 (Assortativity neglect equilibrium)**.** *For any (P, β, γ), the unique ANE s^{AN} is given by (7). The global action distribution G^{AN} under ANE is more dispersive than the Nash action distribution G^{NE}.*

As illustrated in Section 1.1, Proposition 1 reflects the following intuition: Under any monotonic strategy profile, higher types face higher local action distributions, which under assortativity neglect, they view as representative of the global action distribution. This is reminiscent of the widely documented "false consensus effect" in social psychology (Ross, Greene, and House, 1977; Marks and Miller, 1987), whereby people's perceptions of others' behaviors (or attributes) tend to be positively correlated with their own behaviors and attributes. Proposition 1 highlights that, when combined with global strategic complementarities, this effect drives up the gap between higher and lower types' best-responses, further reinforcing the differences in their local action distributions and hence in their perceived global action distributions.[11]

Example 1. In the Gaussian parametrization from Section 1.1, each type θ's match distribution $P(\cdot|\theta)$ is also normal with mean $\mathbb{E}_P[\theta_1|\theta_0=\theta] = (1-\rho)\mu + \rho\theta$; inductively, $\mathbb{E}_P[\theta_t|\theta_0=\theta] = (1-\rho^t)\mu + \rho^t\theta$ for all t. Thus, applying (5) and (7) yields the strategy profiles $s^{NE}(\theta) = \frac{\theta-\mu}{1-\gamma\rho} + \frac{\mu}{1-\beta-\gamma}$ and $s^{AN}(\theta) = \frac{\theta-\mu}{1-(\beta+\gamma)\rho} + \frac{\mu}{1-\beta-\gamma}$ from Section 1.1. As noted in (1), ANE features a higher action variance than Nash, in line with Proposition 1. ▲

3.3 Multiplier Effect of Assortativity Neglect

Proposition 1 shows that, under strategic complementarities, assortativity neglect increases action dispersion relative to Nash in any fixed environment. Continuing to focus on linear best-response games with strategic complementarities, we now highlight a second channel through

[10] Recall that cdf G_1 is a mean-preserving spread of G_2 if $\int \phi(a)\, dG_1(a) \geq \int \phi(a)\, dG_2(a)$ for any convex function $\phi : \mathbb{R} \to \mathbb{R}$ for which the integrals are well-defined. For cdfs that share the same mean, the dispersive order is stronger than the mean-preserving spread order (e.g., Shaked and Shanthikumar, 2007).

[11] The welfare implications of Proposition 1 depend on whether ANE payoffs are interpreted objectively or subjectively (i.e., based on the true global action distribution G^{AN} or agents' perceived global action distributions $L_\theta^{s^{AN},P}$). E.g., under the quadratic miscoordination utilities in footnote 5, objective ANE welfare is always Pareto-worse than Nash due to the higher action dispersion, but subjective ANE welfare can be lower or higher than Nash depending on parameters (see Appendix E.3 of the previous version, Frick, Iijima, and Ishii, 2019).

which assortativity neglect can lead to more dispersed behavior: by amplifying the effect of several key social changes.

We first consider increases in assortativity or type heterogeneity. To disentangle these two changes, we use an equivalent representation of societies that re-expresses who interacts with whom in terms of type quantiles $x \in [0,1]$. For any society P, let $C(x,y)$ be the probability that two types with quantiles below x and y are matched:

$$C(x,y) := P\left(F^{-1}(x), F^{-1}(y)\right) \text{ for all } x,y \in (0,1), \tag{8}$$

and $C(x,0) = C(0,x) = 0$, $C(x,1) = C(1,x) = x$ for all x. Note that C is a copula (i.e., a cdf over $[0,1]^2$ with uniform marginals) and inherits symmetry and assortativity from P. We refer to symmetric and assortative copulas as ***interaction structures***. Any society induces an interaction structure via (8). Conversely, any interaction structure C and type distribution F yield a society by setting $P(\theta,\theta') := C(F(\theta),F(\theta'))$. Under this decomposition of societies into pairs (F,C), one can vary each component freely while holding the other fixed.

We call interaction structure C_1 ***more assortative*** than C_2 if $C_1(x,y) \geq C_2(x,y)$ for any $x,y \in (0,1)$; that is, C_1 assigns higher probability than C_2 to "low-low" matches between agents with quantiles below any cutoffs x and y.[12] The least assortative interaction structure is given by $C_I(x,y) = xy$ for all x,y; for any F, this induces the non-assortative society $(F,C_I) = F \times F$. The most assortative C matches each quantile only with types of the same quantile. In Gaussian societies, the interaction structure depends only on the correlation coefficient ρ, where higher ρ corresponds to more assortativity. For simplicity, we henceforth focus on the class $\mathcal{C}$ of interaction structures that admit positive and absolutely continuous densities on $(0,1)^2$.

To consider the effect of increased assortativity, we compare Nash and ANE global action distributions and strategies G_i^{NE}, s_i^{NE}, G_i^{AN}, s_i^{AN} across environments (F,C_i,β,γ) that differ only in their interaction structures C_i:

Proposition 2 (Effect of assortativity)**.** *For any $C_1, C_2 \in \mathcal{C}$, the following are equivalent:*

1. *C_1 is more assortative than C_2.*
2. *G_1^{NE} is a mean-preserving spread of G_2^{NE} for all (F,β,γ).*
3. *G_1^{AN} is a mean-preserving spread of G_2^{AN} for all (F,β,γ).*
4. *For all $\theta^* \in \Theta$ and (F,β,γ),*

$$\mathbb{E}_F[s_1^{AN}(\theta) - s_2^{AN}(\theta) \mid \theta \geq \theta^*] \geq \mathbb{E}_F[s_1^{NE}(\theta) - s_2^{NE}(\theta) \mid \theta \geq \theta^*] \geq 0.$$

[12] Equivalently, C_1 assigns higher probability to "high-high" matches between quantiles above any two cutoffs. This corresponds to the PQD order over bivariate cdfs used in statistics (e.g., Shaked and Shanthikumar, 2007).

Proposition 2 captures a tight connection between increased assortativity and action dispersion under strategic complementarities: Not only does more assortativity lead to greater action dispersion (by "$1 \Rightarrow 2, 3$"), but greater action dispersion is indeed a defining feature of more assortative societies (by "$2, 3 \Rightarrow 1$"). However, while this is true under both Nash and ANE, part 4 highlights that any given rise in assortativity has a *stronger* effect on action dispersion under assortativity neglect: High types' actions increase more on average (equivalently, low types' actions decrease more) than under Nash.[13] Indeed, under purely global complementarities ($\gamma = 0, \beta > 0$), a rise in assortativity has *no* effect on Nash action dispersion, which always equals type dispersion; however, it still increases ANE action dispersion.

Section 1.1 discussed the intuition for the forward direction: A rise in assortativity increases differences in *local* complementarity incentives across types (under Nash and ANE), but under ANE it additionally increases differences in *perceived global* complementarity incentives by magnifying the false consensus effect. To prove the formal equivalence, we exploit the Markov process representations of Nash and ANE in (5) and (7), which reduces the problem to a comparison of expected t-step ahead matches across different societies. A key step is to establish an equivalence between the more-assortative order over interaction structures C_i and the mean-preserving spread order over the distributions of all t-step ahead expectations of the Markov processes $P_i = (F, C_i)$; this relies crucially on the monotonicity of these processes.

We obtain an analogous result by comparing G_i^{NE} and G_i^{AN} across environments (F_i, C, β, γ) that differ only in their type distributions: Increased type dispersion corresponds to increased action dispersion under both Nash and ANE, but the effect is stronger under ANE.

Proposition 3 (Effect of type dispersion)**.** *For any $F_1, F_2 \in \mathcal{F}$, the following are equivalent:*

1. *F_1 is more dispersive than F_2.*
2. *G_1^{NE} is more dispersive than G_2^{NE} for all (C, β, γ).*
3. *G_1^{AN} is more dispersive than G_2^{AN} for all (C, β, γ).*
4. *For all (C, β, γ) and $x, y \in (0, 1)$ with $x > y$,*

$$\Delta_{x,y} G_1^{AN} - \Delta_{x,y} G_2^{AN} \geq \Delta_{x,y} G_1^{NE} - \Delta_{x,y} G_2^{NE} \geq 0,$$

where $\Delta_{x,y} G := G^{-1}(x) - G^{-1}(y)$ for any cdf G.

Finally, we show that changes in complementarity motives can have qualitatively distinct effects under Nash and ANE:

[13] For $\bullet \in \{\text{NE}, \text{AN}\}$, $G_1^\bullet$ is a mean-preserving spread of $G_2^\bullet$ iff $\mathbb{E}_F[s_1^\bullet(\theta) - s_2^\bullet(\theta) \mid \theta \geq \theta^*] \geq 0$ for all θ^*. Thus, part 4 captures that the mean-preserving spread increase from $G_2^\bullet$ to $G_1^\bullet$ is greater under ANE than Nash.

Proposition 4 (Effect of complementarity motives)**.** *Consider G_i^{NE} and G_i^{AN} in environments $(F, C, \beta_i, \gamma_i)$ $(i = 1, 2)$. We have:*

1. $\gamma_1 \geq \gamma_2 \iff G_1^{NE}$ *is more dispersive than* G_2^{NE} *for all* (F, C) *and all* β_1, β_2.

2. $\beta_1 + \gamma_1 \geq \beta_2 + \gamma_2 \iff G_1^{AN}$ *is more dispersive than* G_2^{AN} *for all* (F, C).

Thus, Nash action dispersion is increasing in local complementarity motives γ, but is unaffected by changes in global complementarity motives β. In contrast, ANE action dispersion additionally increases with β, as this amplifies the role of the false consensus effect. In particular, if $\gamma_1 > \gamma_2$, but $\beta_1 + \gamma_1 < \beta_2 + \gamma_2$, then Nash action dispersion is greater in the first environment than the second, but ANE action dispersion is greater in the second than the first.

3.4 Discussion

We briefly interpret the preceding analysis in the context of the education investment example from Section 1.1. A large literature studies the effect of income-based residential sorting on socioeconomic education inequality (for a survey, see Fernández, 2003). Much theoretical and empirical work highlights the role of local complementarities, such as peer effects or local provision of educational facilities. Under Nash, we saw that such local complementarities are the only channel through which sorting affects the socioeconomic education gap in our setting.

In contrast, the ANE analysis captures an additional *inferential* channel that is absent under Nash: If marginal returns to education are increasing in global education investment (e.g., due to labor market competition) and if people project their peers' education choices onto society, then sorting also affects education inequality by leading to a socioeconomic *perception gap* about the returns to education. In line with this channel, more recent empirical work documents the role of perceived returns to education in shaping individuals' education investment and finds that disadvantaged individuals often substantially underestimate these returns.[14][15]

Taking into account this inferential channel may be important for two reasons. First, as formalized by the multiplier effect in Proposition 2, this further strengthens the rationale for programs, such as "Moving to Opportunity," that aim to reduce income-based residential segregation. Second, the inferential channel suggests that additional reductions in education inequal-

[14] E.g., Jensen (2010) finds that predominantly poor students in the Dominican Republic underestimate the returns of secondary vs. primary school (resp. college vs. secondary school) by on average 78% (resp. 70%), with underestimation strongest among the poorest students; moreover, perceived returns significantly predict subsequent actual years of schooling. See also Nguyen (2008) and Attanasio and Kaufmann (2014).

[15] Streufert (2000) proposes a different model of misperceptions about returns to education: For each education level s, there is an exogenous true earnings distribution F_s, but poor agents observe a truncation F_s^α that omits earnings above an exogenous cap α (e.g., because successful agents leave poor neighborhoods). He assumes that poor agents misperceive F_s^α to represent true earnings, but highlights that this has an *ambiguous* effect: their perceived marginal returns $\mathbb{E}[F_s^\alpha] - \mathbb{E}[F_{s-1}^\alpha]$ to education may under- or overstate the true returns $\mathbb{E}[F_s] - \mathbb{E}[F_{s-1}]$.

ity might be achieved through *informational* interventions aimed at correcting misperceptions. Recent empirical work has begun to study such interventions and has found significant effects.[16]

4 Perceptions under Assortativity Neglect

4.1 Coherent Perceptions

So far, we have focused on agents' behavior under assortativity neglect, by defining a solution concept, ANE, where agents only observe their local action distributions and best-respond based on the misperception that these local action distributions match the global action distribution.

In this section, we take the perspective that agents not only use their local action observations to decide on a best-response. Rather, as emphasized, for instance, by the social psychology literature, we assume that agents also seek to build "coherent stories" that can *explain* this observed behavior within its social context.[17] Through these stories, agents' assortativity neglect not only shapes their equilibrium behavior, but also their perceptions of the type distribution in society and of other types' behavior.

Specifically, in any population game, we introduce the following simple formalization of how agents form coherent perceptions under assortativity neglect:

Definition 3. Given any society P and ANE s^{AN}, a ***coherent assortativity neglect perception*** for type θ consists of a perceived non-assortative society $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$ and a perceived strategy profile $\hat{s}_\theta$ such that:

1. *Observational consistency:* $L_\theta^{s^{AN},P} = L_\theta^{\hat{s}_\theta,\hat{P}_\theta}$.
2. *Perceived best-response:* for each θ', $\hat{s}_\theta(\theta') \in \mathrm{BR}_{\theta'}(G^{\hat{s}_\theta,\hat{P}_\theta}, L_{\theta'}^{\hat{s}_\theta,\hat{P}_\theta})$.

That is, given an ANE s^{AN}, we continue to assume that the only information that agent θ observes is her local action distribution $L_\theta^{s^{AN},P}$. Based on this, she forms a perception $\hat{F}_\theta$ of the type distribution (which, since she neglects assortativity, means that her perceived society is $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$) and a perception $\hat{s}_\theta$ of the strategy profile in society. These perceptions satisfy two requirements:

First, they are consistent with θ's observed local action distribution. That is, θ's perceived local action distribution $L_\theta^{\hat{s}_\theta,\hat{P}_\theta}$ matches her true local action distribution $L_\theta^{s^{AN},P}$.

[16] The aforementioned Jensen (2010) finds that providing one-time information about true returns to education generates lasting increases in students' perceived returns and increases completed schooling by 0.2-0.35 years over 4 years. In Nguyen (2008), similar interventions raise students' test scores by 0.2-0.37 standard deviations.

[17] The idea that people are "naive psychologists" who do not simply take note of the behavior of those around them, but try to explain this behavior based on a combination of internal and environmental causes underlies the field of attribution theory in social psychology (e.g., Heider, 1958; Kelley and Michela, 1980). Kahneman (2011) (p. 85) also emphasizes "the coherency of the story it manages to create" as a central measure of success of any heuristic (e.g., "What You See is All There Is") that the "intuitive" System 1 employs.

Second, they allow θ to explain her observed local action distribution within its social context, by assuming that other agents behave optimally in the population game. That is, the action $\hat{s}_\theta(\theta')$ that θ attributes to any other type θ' is a best-response for θ' to the global and local action distributions $G^{\hat{s}_\theta,\hat{P}_\theta}$ and $L_{\theta'}^{\hat{s}_\theta,\hat{P}_\theta}$ that θ perceives θ' to face.

To interpret the second requirement, note that, since θ perceives society to be non-assortative, she believes that her own local action distribution coincides with the global action distribution and all other agents' local action distributions. Thus, combined with the first assumption that θ is correct about her local action distribution, the second requirement simplifies to

$$\hat{s}_\theta(\theta') \in \mathrm{BR}_{\theta'}(L_\theta^{s^{AN},P}, L_\theta^{s^{AN},P}). \tag{9}$$

This way of reasoning about others' behavior is a natural continuation of the projection bias underlying assortativity neglect: θ's own action $s^{AN}(\theta)$ is a best-response to the perception that the global and local action distributions throughout society are $L_\theta^{s^{AN},P}$, and she naively believes that the same is true of other agents' actions.

Remark 2. For parsimony, Definition 3 does not explicitly model θ's perceptions about other agents' perceptions, but the perceived best-response condition implicitly suggests a self-centered view: θ perceives other agents to best-respond based on her *own* perceptions $\hat{P}_\theta$ and $\hat{s}_\theta$. While such naively self-centered perceptions can again be viewed as a natural continuation of the projection bias underlying assortativity neglect,[18] they might seem restrictive. However, we note that they are not essential for our analysis.

Indeed, Appendix C.1 extends Definition 3 to allow θ to perceive that other agents' perceived societies and strategy profiles disagree with her own perceptions. We show that θ's own perceived society and strategy profile $\hat{P}_\theta$ and $\hat{s}_\theta$ remain unchanged relative to Definition 3, as long as θ perceives that (i) other agents also perceive society to be non-assortative, (ii) other agents behave optimally given their perceptions, and (iii) other agents' perceptions are consistent with their observed local action distributions. As we discuss, this can be seen as a misspecified version of Esponda's (2013) (level-1) rationalizable conjectural equilibrium. ▲

4.2 Existence and Uniqueness of Coherent Perceptions

To analyze agents' perceptions, we return to linear best-response games with strategic complementarities. Our first result is that agents are always able to form coherent assortativity neglect perceptions. Thus, no matter how assortative the actual society P is, agents who suffer from assortativity neglect are still able to explain the behavior they observe in a coherent manner.[19]

[18] Indeed, this reasoning is reminiscent of "information projection," where agents mistakenly believe that others have similar beliefs (Madarász, 2012).

[19] This might capture a sense in which assortativity neglect can be an especially "persistent" misperception.

Moreover, for any (P, β, γ), each agent's coherent perceptions are unique:

Proposition 5 (Coherent perceptions)**.** *Fix any (P, β, γ) with corresponding ANE s^{AN}. For each type θ, there exist unique coherent assortativity neglect perceptions $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$ and $\hat{s}_\theta$.*

To see the idea, recall that in forming coherent perceptions, θ believes other agents to play best-responses. Thus, any difference in the actions that θ attributes to two agents θ_1 and θ_2 might in principle be due to two channels—the difference in their types and the difference in their local complementarity motives, which reflects their differing local action distributions:

$$\underbrace{\hat{s}_\theta(\theta_1) - \hat{s}_\theta(\theta_2)}_{\Delta(\text{perceived actions})} = \underbrace{\theta_1 - \theta_2}_{\Delta(\text{types})} + \underbrace{\gamma \left(\int a \, dL_{\theta_1}^{\hat{s}_\theta, \hat{P}_\theta}(a) - \int a \, dL_{\theta_2}^{\hat{s}_\theta, \hat{P}_\theta}(a) \right)}_{\Delta(\text{perceived local complementarity motives})}.$$

However, the fact that θ's perceived society $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$ is non-assortative has the following key implication: θ does not perceive any difference in θ_1 and θ_2's complementarity motives, as she believes all agents to face the *same* local action distribution. Thus, in explaining her observed local action distribution, θ attributes all action dispersion to the type dispersion among her matches. In other words, assortativity neglect leads to a form of misattribution that is reminiscent of the "fundamental attribution error" documented in social psychology (Ross, 1977): θ attributes any variation in others' behavior entirely to their intrinsic characteristics (types), neglecting that external factors (e.g., differences in peer effects) might also be at play.[20]

Since, under assortativity neglect, θ's perceived match distribution $\hat{P}_\theta(\cdot|\theta)$ coincides with her perceived type distribution $\hat{F}_\theta$ in the overall population, the fact that she misattributes all local action dispersion to type dispersion among her matches implies that

$$\left(L_\theta^{s^{AN},P}\right)^{-1}(x) - \left(L_\theta^{s^{AN},P}\right)^{-1}(y) = \hat{F}_\theta^{-1}(x) - \hat{F}_\theta^{-1}(y), \qquad \text{for all } x, y \in (0,1). \tag{10}$$

In Appendix B.6, we show that there is a unique $\hat{F}_\theta$ that achieves (10) while also ensuring that θ is correct about the local action mean. The perceived strategy profile $\hat{s}_\theta$ is uniquely pinned down by the perceived best-response condition.

Example 2. In Gaussian societies (Section 1.1), each agent θ's perceived type distribution $\hat{F}_\theta$ is also normal. In particular, (10) implies that θ's perceived type variance equals her local action variance: $\hat{\sigma}_\theta^2 = \text{Var}_P[s^{AN}(\theta')|\theta] = \frac{\sigma^2(1-\rho^2)}{(1-(\beta+\gamma)\rho)^2}$. Letting $\overline{a}_\theta := \mathbb{E}_P[s^{AN}(\theta')|\theta]$, the perceived best-response condition (9) implies that $\hat{s}_\theta(\theta') = \theta' + (\beta+\gamma)\overline{a}_\theta$ for all θ'. Finally, in order for θ's perceived local action mean $\mathbb{E}_{\hat{F}_\theta}[\hat{s}_\theta(\theta')]$ to match the true local action mean $\overline{a}_\theta$, θ must perceive the type mean to be $\hat{\mu}_\theta = (1-\beta-\gamma)\overline{a}_\theta = \mu + \frac{(1-\beta-\gamma)\rho(\theta-\mu)}{1-(\beta+\gamma)\rho}$. ▲

[20] Another source of the difference in θ_1 and θ_2's actual ANE actions is the false consensus effect (i.e., their different perceived global complementarity motives). This effect is also neglected by θ, who perceives everyone to best-respond to the same global action distribution. Hence, an attribution error arises even if $\gamma = 0$.

4.3 Misperceptions about Type Distributions

Since Proposition 5 uniquely pins down each agent's perceived type distribution, it provides a lens through which to study how, under assortativity neglect, agents' misperceptions of population characteristics are shaped by the nature of their social interactions. In particular, we highlight the importance of strategic considerations: In our model, agents form perceptions based on their matches' behavior, which is subject to strategic motivations; as we show, this leads to different predictions than under purely statistical misinference, where agents directly observe their matches' types and project them onto society.

In the following, for any environment (P, β, γ), we refer to agent θ's coherent assortativity neglect perception $\hat{F}_\theta$ simply as θ's perceived type distribution.

Perceived type dispersion. We first analyze agents' perceptions of type dispersion in society (e.g., income inequality, political attitude polarization). The following result shows that strategic complementarities drive up perceived type dispersion:

Proposition 6 (Perceived type dispersion)**.** *Fix any society P and type θ. If $\beta_1+\gamma_1 \geq \beta_2+\gamma_2$, then θ's perceived type distribution $\hat{F}^1_\theta$ under (P, β_1, γ_1) is more dispersive than θ's perceived type distribution $\hat{F}^2_\theta$ under (P, β_2, γ_2).*

Moreover, depending on the strength of $\beta + \gamma$, agents may under- or overestimate type dispersion, as illustrated by the Gaussian example in Section 1.1. To see the idea, note that our model generates two opposing errors in agents' reasoning about type dispersion:

- First, a purely *statistical error*: Agents' matches are on average less diverse than the overall population.[21] However, under assortativity neglect, agents believe their matches to be representative of the overall population. This error pushes agents to underestimate type dispersion.
- Second, an *attribution error*: Agents do not directly observe their matches' types; instead, they draw inferences about the type distribution from their local action distributions. However, as discussed following Proposition 5, assortativity neglect leads agents to misattribute all observed action dispersion to type dispersion, ignoring the fact that different types are also subject to different (local and/or perceived global) complementarity motives. This pushes agents to overestimate type dispersion.

Without strategic complementarities (i.e., if $\beta = \gamma = 0$), only the statistical error channel is relevant: In this case, agents' actions $s^{AN}(\theta) = \theta$ match their types, so this setting is equivalent to one where agents directly observe their local type distributions and project them onto society.

[21] Indeed, by the law of total variance, any society P satisfies $\mathbb{E}_F[\mathrm{Var}_P[\theta'|\theta]] \leq \mathrm{Var}_F[\theta']$. More strongly, for many parametric classes of societies (e.g., Gaussian), $\mathrm{Var}_P[\theta'|\theta] \leq \mathrm{Var}_F[\theta']$ holds for all θ.

However, the stronger strategic complementarities, the more important is the attribution error channel: By Proposition 4, an increase in $\beta + \gamma$ increases ANE action dispersion (due to stronger local complementarities and/or a stronger false consensus effect). Because of the attribution error, this leads agents to perceive more type dispersion. When strategic complementarities are strong enough (relative to other parameters, such as the extent of assortativity), then the attribution error channel can dominate, as illustrated by the Gaussian example.

Perceived type means. Second, we consider agents' perceptions $\hat{\mu}_\theta := \mathbb{E}_{\hat{F}_\theta}[\theta']$ of the type mean. The following result first notes that higher types θ perceive higher type means $\hat{\mu}_\theta$. That is, the false consensus effect we observed for perceived behavior following Proposition 1 extends to agents' perceptions of the type mean. However, strategic complementarities counteract this effect, by reducing the sensitivity of $\hat{\mu}_\theta$ to θ: The population distribution of perceived type means $\hat{\mu}_\theta$ (i.e., when θ is distributed according to F) is less dispersed the greater $\beta + \gamma$.

Proposition 7 (Perceived type means)**.** *For any (P, β, γ), agents' perceived type means $\hat{\mu}_\theta$ are increasing in their types θ. If $\beta_1 + \gamma_1 \geq \beta_2 + \gamma_2$, then the population distribution M^2 of perceived type means under (P, β_2, γ_2) is a mean-preserving spread of the distribution M^1 of perceived type means under (P, β_1, γ_1).*

The false consensus effect reflects the statistical error underlying assortativity neglect: If agents directly observe their local match distributions and project them onto society, the effect is immediate, as $\hat{\mu}_\theta$ coincides with θ's expected match $\mathbb{E}_P[\theta_1|\theta_0 = \theta]$, which is increasing in θ.

To see why strategic complementarities counteract this effect, recall that θ's perceived type mean satisfies $\hat{\mu}_\theta = (1-\beta-\gamma)\mathbb{E}_P[s^{AN}(\theta')|\theta]$ (see Example 2). Thus, combined with the Markov process representation of s^{AN} in (7), we have

$$\hat{\mu}_\theta = (1 - \beta - \gamma) \sum_{t=0}^{\infty} (\beta + \gamma)^t \mathbb{E}_P[\theta_{t+1}|\theta_0 = \theta]. \tag{11}$$

That is, when θ draws inferences about the type distribution from her matches' strategic behavior, then her perceived type mean $\hat{\mu}_\theta$ depends not only on her immediate expected match $\mathbb{E}_P[\theta_1|\theta_0 = \theta]$, but also on her more distant t-step ahead expected matches, as the latter affect her immediate matches' incentives. Moreover, the greater strategic complementarities, the stronger is this dependence: increasing $\beta + \gamma$ increases the weight on more distant expected matches in (11). Importantly, as we show, θ's more distant matches are less sensitive to her own type than her immediate matches. Thus, stronger strategic complementarities lead perceived type means $\hat{\mu}_\theta$ to differ less across different agents θ.

4.4 Discussion

Misperceptions of income distributions. Several papers in economics and psychology have put forward the aforementioned statistical error channel—individuals observe the incomes of their social contacts and naively project these onto society—as a potential source of underestimation of income inequality (e.g., Cruces, Perez-Truglia, and Tetaz, 2013; Windsteiger, 2018; Dawtry, Sutton, and Sibley, 2019).

However, underestimation of income inequality is not a universal empirical finding. Indeed, the survey by Hauser and Norton (2017) notes: "Overall, the bulk of evidence suggests that people around the world hold incorrect perceptions of inequality in their country—but with variation. In the U.S. and United Kingdom, for example, underestimation of inequality is relatively common, while overestimation occurs in other countries, such as France and Germany. Moreover, there are a few exceptions of high accuracy: respondents in Norway, for instance, were relatively accurate in estimating their country's income inequality."[22]

The preceding analysis points to a novel channel that might contribute to such more mixed findings: Rather than directly observing their social contacts' incomes, individuals may need to partly infer these from their consumption choices (e.g., education investment, or homes, cars, and attire), which are subject to well-documented peer effects. If individuals neglect assortativity, we saw that this additionally generates an attribution error that counteracts the statistical error. Depending on the relative strength of these errors, one may find underestimation, fairly accurate perceptions, or overestimation of income inequality.[23]

Importantly, people's misperceptions of income distributions can have *material* consequences, by influencing demand for redistribution: For instance, both empirical and theoretical work suggests that agents demand more redistribution if they perceive greater income inequality or a lower own position $\theta - \hat{\mu}_\theta$ relative to the mean.[24] Thus, by showing how these misperceptions are shaped by agents' social interactions, our findings in this section can also shed light on ways in

[22]For example, Niehues (2014) compares perceived income distributions across 23 European countries and the US and finds overestimation in most of continental Europe, relatively more accurate perceptions in several Scandinavian countries, and underestimation in the US (e.g., her imputed subjective vs. actual Gini coefficients include: Germany (0.35 vs. 0.29), France (0.36 vs. 0.3), Hungary (0.43 vs. 0.24), Czech Republic (0.38 vs. 0.25); Norway (0.26 vs. 0.23); US (0.34 vs. 0.42)). Bavetta, Li Donni, and Marino (2019) similarly find overestimation in Germany, France, Italy, Sweden, Finland, and South Korea, but underestimation in the US and UK.

[23]Establishing a conclusive link between the attribution error channel and cross-country differences in perceived inequality is beyond the scope of this paper, and various other factors (e.g., ideological differences) are likely also at play. While cross-country data on complementarity motives is less readily available, we note that socioeconomic segregation is documented to be higher in the US/UK than in continental Europe (e.g., Musterd, 2005; Quillian and Lagrange, 2016); thus, underestimation in the former and overestimation in the latter is consistent with our finding in the Gaussian example that the statistical error dominates the attribution error when assortativity is relatively strong (see Figure 1).

[24]Empirical evidence includes Cruces, Perez-Truglia, and Tetaz (2013); Gimpelson and Treisman (2018); Hvidberg, Kreiner, and Stantcheva (2021). These findings are consistent with replacing actual with perceived income distributions in classic theoretical models of demand for redistribution.

which the nature of social interactions (e.g., complementarity motives and assortativity) might affect a society's demand for redistribution.[25]

Misperceptions of political attitude distributions. Several recent studies document significant overestimation of political attitude polarization in the US: When asked to estimate others' (privately elicited) attitudes on various political issues, respondents perceive greater than actual attitude dispersion on most issues, because they exaggerate the prevalence of extreme attitudes on both sides of the political spectrum.[26]

Again, the attribution error channel that our analysis highlights might be relevant in this setting: Others' privately held attitudes are not directly observable, so people might partly infer these from the public manifestations of support they observe (e.g., social media posts, yard signs, or bumper stickers). However, differences in such public manifestations may exceed differences in private attitudes, as they may also be driven by catering to different peer groups. Neglecting this can contribute to overestimating political attitude polarization, even when people's observed samples are more politically homogenous than the overall population.

5 Extensions and Related Literature

5.1 Weaker Forms of Assortativity Neglect

Hybrid model. Under ANE, *all* agents suffer from assortativity neglect. More realistically, some agents might be less prone to misperception than others, for example, due to having access to information about global (rather than just local) action distributions. To capture this, consider a simple hybrid model: For each type θ, only fraction $\alpha \in [0,1]$ of agents suffer from assortativity neglect; the remaining share of agents best-respond to the correct local and global action distributions. An ***α-assortativity neglect equilibrium (α-ANE)*** consists of strategy profiles s_a for assortativity neglect agents and s_c for correct agents such that, for all θ,

$$\begin{aligned} s_a(\theta) &\in \mathrm{BR}_\theta \left(\alpha L_\theta^{s_a,P} + (1-\alpha) L_\theta^{s_c,P}, \alpha L_\theta^{s_a,P} + (1-\alpha) L_\theta^{s_c,P} \right), \\ s_c(\theta) &\in \mathrm{BR}_\theta \left(\alpha G^{s_a,P} + (1-\alpha) G^{s_c,P}, \alpha L_\theta^{s_a,P} + (1-\alpha) L_\theta^{s_c,P} \right). \end{aligned} \tag{12}$$

Appendix C.2 applies α-ANE to linear best-response games with strategic complementar-

[25] A previous version of this paper (Frick, Iijima, and Ishii, 2019, Appendix E.2) illustrated this in the context of Meltzer and Richard's (1981) model of voting for redistribution.

[26] E.g., the American National Election Survey elicits both citizens' own attitudes and political affiliation and their estimates of average attitudes among typical Democrats and Republicans on a wide range of issues. On average, the actual difference between Democrats and Republicans is 1 point (on a 7 point scale), but perceived differences are almost twice as large; perceptions vary with own political affiliation, but most respondents exaggerate the extremeness of attitudes on both sides (cf. Bordalo, Tabellini, and Yang, 2021; Bordalo, Coffman, Gennaioli, and Shleifer, 2016). See also Ahler (2014); Westfall, Van Boven, Chambers, and Judd (2015).

ities. We show that the equilibrium strategy profiles of both groups of agents again admit Markov process representations that generalize those for ANE and Nash. Moreover, coherent perceptions for assortativity neglect agents can be defined analogously to Definition 3. Based on this, all our results extend.

Notably, behavior among assortativity neglect agents is more dispersed than among correct agents, but action dispersion among *both* groups is exacerbated the greater the share α of assortativity neglect agents (Proposition C.1). This captures a sense in which assortativity neglect agents can exert a negative externality on society, as they drive up miscoordination among all agents.[27]

Partial assortativity neglect. Alternatively, one can relax the assumption that agents *fully* neglect assortativity, i.e., perceive the global action distribution to exactly coincide with their local action distributions. Appendix C.3 considers a reduced-form extension of Definition 2 that can capture various forms of partial assortativity neglect: A strategy profile s is a ***partial assortativity neglect equilibrium (PANE)*** if for each θ, there exists a perceived global action distribution $\hat{G}_\theta \in \Delta(A)$ such that (i) $s(\theta) \in \text{BR}_\theta(\hat{G}_\theta, L_\theta^{s,P})$ for each θ and (ii) $\hat{G}_\theta$ is FOSD-increasing in θ. That is, each agent θ best-responds to a correct perception of her local action distribution, but misperceives the global action distribution to be $\hat{G}_\theta$, where (ii) represents the false consensus effect that higher types perceive higher global action distributions. One simple parametrization of PANE is when agents' perceived global action distributions are a convex combination of the true local and global action distributions; that is, s is monotonic such that for some $\varepsilon \in [0,1]$,

$$\hat{G}_\theta = \varepsilon L_\theta^{s,P} + (1-\varepsilon)G^{s,P}, \text{ for all } \theta. \tag{13}$$

This captures a form of "partial projection" of local action distributions onto the global action distribution that nests both ANE ($\varepsilon = 1$) and Nash ($\varepsilon = 0$).[28] Generalizing Proposition 1, in linear best-response games with strategic complementarities, any PANE is more dispersive than Nash (Proposition C.2). Moreover, for the parametrization in (13), it is straightforward to generalize the Markov process representation of equilibrium strategies;[29] based on this, all results in Section 3 extend, with action dispersion intermediate between Nash and ANE.

Appendix C.3 also extends the definition of coherent perceptions to PANE. In Gaussian societies $P = (\mu, \sigma^2, \rho)$, we show that for each $\hat{\rho} \in [0, \rho]$, there is a PANE in which agents coherently underestimate assortativity to be $\hat{\rho}$. The corresponding perceived type distributions

[27] This observation contrasts with Jehiel's (2018) model of investment under selection neglect, where the effect of misperception is weakened the greater the share of agents who suffer from selection neglect.

[28] This parametrization of partial assortativity neglect is reminiscent of Eyster and Rabin's (2005) parametrization of partial cursedness.

[29] For any (P, γ, β), the unique PANE satisfying (13) is $s^\varepsilon(\theta) = \sum_{t=0}^{\infty}(\gamma+\varepsilon\beta)^t \mathbb{E}_P[\theta_t|\theta_0 = \theta] + \frac{(1-\varepsilon)\beta\mathbb{E}_F[\theta']}{(1-\gamma-\varepsilon\beta)(1-\beta-\gamma)}$.

are in between those under full assortativity neglect ($\hat{\rho} = 0$) and correct perceptions ($\hat{\rho} = \rho$), and satisfy the same qualitative predictions as in Section 4.

5.2 More General Best-Responses

Strategic substitutes. Our results have focused on linear best-response games with strategic complementarities. However, the analysis can be extended to the case with global and/or local strategic substitutes (i.e., $\beta < 0$ and/or $\gamma < 0$). The Markov process representations (5) and (7) of Nash and ANE strategies are unchanged, as is the derivation of coherent ANE perceptions. However, the directions of some effects change depending on the sign of β, γ, and $\beta + \gamma$. The following example illustrates this for Gaussian societies; Appendix C.4 presents general results.

Example 3. Consider global substitutes, $\beta < 0$. Nash and ANE strategies in a Gaussian society $P = (\mu, \sigma^2, \rho)$ take the same form as in Example 1. In particular, global action variances are still $\mathrm{Var}(G^{NE}) = \left(\frac{1}{1-\gamma\rho}\right)^2 \sigma^2$ and $\mathrm{Var}(G^{AN}) = \left(\frac{1}{1-(\beta+\gamma)\rho}\right)^2 \sigma^2$. Note that $\mathrm{Var}(G^{AN}) \leq \mathrm{Var}(G^{NE})$. Thus, in contrast with Proposition 1, assortativity neglect now *decreases* action dispersion. This is because, under global substitutes, the false consensus effect of perceiving the global action average to coincide with one's local action average leads higher (lower) types to play lower (higher) actions than under Nash.

The effect of increased assortativity depends additionally on the local complementarity parameter γ. If $\gamma < 0$, the multiplier effect in Proposition 2 holds with a flipped sign: Both $\mathrm{Var}(G^{NE})$ and $\mathrm{Var}(G^{AN})$ are decreasing in ρ, but the derivative is more negative under ANE. In contrast, if $\gamma > 0$, then increasing ρ can have *opposite* effects on Nash and ANE: Nash action dispersion always increases, but ANE action dispersion decreases if $\beta + \gamma < 0$.

Finally, as in Example 2, agents' (coherent) perception of type variance is $\hat{\sigma}_\theta^2 = \frac{\sigma^2(1-\rho^2)}{(1-(\beta+\gamma)\rho)^2}$, which under-/overestimates the true variance σ^2 when $\beta + \gamma$ is small/large relative to ρ. If $\beta+\gamma < 0$, only underestimation is possible, as the statistical and attribution errors in Section 4.3 now push in the same direction. ▲

Non-linear best-responses. ANE and the associated coherent perceptions can also be analyzed in any other population game of the form in Section 2. Appendix C.5 shows that several of our main insights hold more generally. First, the same attribution error logic underlying Proposition 5 ensures the existence and uniqueness of coherent assortativity neglect perceptions in general population games (Proposition C.5). Second, the finding that ANE increases action dispersion relative to Nash extends straightforwardly to additively separable best-response functions with global complementarities (Proposition C.6).

At the same time, a full generalization of the preceding results is beyond the scope of this paper. One challenge is that our representation of Nash and ANE in terms of iterated expectations of the Markov process P relied on the linearity of best-responses. This representation

played a central role for our comparative statics analysis in Sections 3.3 and 4.3; in contrast, to the best of our knowledge, existing comparative statics results for general games with strategic complementarities do not apply to our setting.[30]

5.3 Related Literature

This paper contributes to a growing literature in behavioral game theory that studies the steady-state behavior of players who draw misinferences from the observational feedback generated by their game. While in our setting agents neglect selection effects arising from the assortativity of social interactions, a number of papers consider agents who neglect selection due to missing feedback about non-implemented projects/transactions, e.g., in settings of adverse selection (Esponda, 2008), voting (Esponda and Pouzo, 2017), or investment (Jehiel, 2018).[31] A related inferential bias, correlation neglect, underlies cursed equilibrium (Eyster and Rabin, 2005) and analogy-based expectation equilibrium (Jehiel, 2005; Jehiel and Koessler, 2008): Here, agents are correct about the marginal distributions of opponents' actions and types, but misperceive the correlation between these two;[32] these solution concepts reduce to Nash equilibrium in the static private-value environment of this paper. Chauvin (2018) studies an equilibrium model of discrimination: Agents belong to observable groups whose outcome distributions depend jointly on members' individual traits and on population beliefs about the group, but others' beliefs about each group are based on the misinference that observed outcomes are purely due to members' traits. This misinference is similar in spirit to the fundamental attribution error that we derive from agents' assortativity neglect in Section 4.2.[33] ANE (and most aforementioned) settings can be seen as instances of Berk-Nash equilibrium (Esponda and Pouzo, 2016), which captures the steady-state behavior of players with general misspecified models of the feedback structure of their game.

Different from the aforementioned papers, Section 4 also considers how players can "rationalize" their observed action distributions as resulting from optimal behavior, by forming

[30] A large literature conducts comparisons of equilibrium action distributions in terms of first-order stochastic dominance (for a survey, see Vives, 2005), but comparative statics results in terms of dispersion (e.g., mean-preserving spread, second-order stochastic dominance) are more limited: Jensen (2018) analyzes the effect of type dispersion on action dispersion; his approach relies on players' types being independently distributed (while our setting displays correlation). Mekonnen and Leal Vizcaíno (2021) consider a different setting, where players observe signals about an uncertain fundamental, and analyze the effect of signal precision on action dispersion.

[31] Like these papers, we consider agents whose samples are biased but infinite. In Osborne and Rubinstein (1998, 2003); Salant and Cherry (2020); Gonçalves (2020), agents observe unbiased but finite samples.

[32] Relatedly, Spiegler (2016, 2017) considers an agent who infers an incorrect joint distribution over multiple observed economic variables by misperceiving causal relations. The implications of selection/correlation neglect have also been explored in settings without equilibrium feedback (e.g., Streufert, 2000; Glaeser and Sunstein, 2009; Ortoleva and Snowberg, 2015; Levy and Razin, 2015; Ellis and Piccione, 2017).

[33] See Kaneko and Matsui (1999) for a related model of discrimination based on inductive game theory. Ettinger and Jehiel (2010) formalize a form of fundamental attribution error in a bargaining setting.

coherent misperceptions about the type distribution and strategy profile. This exercise relates to the literature on rationalizable conjectural equilibrium (e.g., Rubinstein and Wolinsky, 1994; Esponda, 2013; Fudenberg and Kamada, 2015; Lipnowski and Sadler, 2019). This refines self-confirming equilibrium (Fudenberg and Levine, 1993; Battigalli, 1987) by requiring that agents' beliefs about opponents' behavior are not only consistent with their observational feedback, but are also consistent with opponents best-responding to beliefs that are themselves observationally consistent (and similarly for higher-order beliefs). Whereas these papers consider standard agents who do not ex-ante rule out the correct observational feedback structure, we consider misspecified agents who reason based on the dogmatic misperception that society is non-assortative (see also Remark 2 and Appendix C.1).

While we analyze the equilibrium implications of assortativity neglect for population games in fixed societies, other recent papers consider the effect of related selection biases on endogenous sorting. Levy and Razin (2017) study the coevolution of sorting into different school types and beliefs about school quality: agents' beliefs are shaped by communicating with school peers while ignoring selection into schools. They characterize when polarized beliefs about school quality are sustained in the long run. Windsteiger (2018) considers steady-state sorting into social classes when agents directly observe their peers' incomes but underestimate income differences across classes; she shows that this misperception reduces demand for redistribution.

As noted, assortativity neglect can be seen as a form of projection bias. Other work has studied strategic interactions under different forms of projection bias, for example, when agents project their tastes onto others (e.g., Breitmoser, 2019; Gagnon-Bartsch, Pagnozzi, and Rosato, 2021; Gagnon-Bartsch, 2017; Bohren and Hauser, 2021, in the context of auctions and social learning) or when agents overestimate the similarity of others' signals (Madarász, 2012, 2016).

Linear best-response games are also widely studied in the literature on network games. Two recent papers relate to our focus on agents' misperceptions of interaction patterns: Battigalli, Panebianco, and Pin (2020) study self-confirming equilibrium in network games, with a focus on learning dynamics and perceived centrality. Jackson (2019) studies implications of the "friendship paradox," i.e., the fact that people's neighbors on average have higher degrees than themselves. He shows that, because of this, if agents naively behave as in the local interaction case even though utilities depend on uniform global interactions, then, under strategic complementarities, this leads to higher average behavior than Nash. Our setting does not feature degree heterogeneity, so centrality/the friendship paradox play no role; instead, we focus on misperceptions of assortativity based on type heterogeneity. Our analysis of agents' coherent misperceptions also has no counterpart in these papers.

Finally, in incomplete-information games, Samet (1998) introduced the use of Markov processes to represent players' higher-order beliefs about the uncertain fundamental. Golub and Morris (2018) study incomplete-information games on networks, in which case the corresponding

Markov process depends on both the signal structure and network; with linear best-responses, Bayes-Nash equilibria can then be written as discounted sums of the higher-order expectations of this process.[34] While we consider population games without aggregate uncertainty, our Markov process over t-step ahead matches can be seen as an analog of the Markov processes over higher-order beliefs in those papers. The key novelty is that, due to the assortativity of P, our Markov process is monotone. Monotonicity plays a central role for our analysis of action dispersion (and perceived type distributions), by allowing us to translate comparisons of interaction structures/type distributions into comparisons of the distributions of t-step ahead expected matches. Beyond games with assortative interactions, our proof methods may also be useful in incomplete-information linear best-response games (e.g., beauty contests) where the signal structure displays appropriate positive correlation to ensure monotonicity of the corresponding Markov process.

6 Conclusion

We propose a model of social interactions and misperceptions under assortativity neglect. To analyze the interplay between assortativity neglect and agents' strategic incentives, we define an equilibrium notion where agents best-respond to the misperception that the local action distributions among their peers are representative of behavior in society. We also model how agents form misperceptions about the type distribution from their local action observations, by reasoning about their peers' incentives through the lens of their assortativity neglect. Based on this, we show how, when combined with strategic complementarities, assortativity neglect increases action dispersion in society. We also find that assortativity neglect generates two countervailing mistakes in agents' inferences about the type distribution—a statistical and an attribution error. Depending on the nature of social interactions, this may lead agents to either under- or overestimate type dispersion. We discuss the relevance of our results in the context of socioeconomic disparities in education investment, as well as empirically documented misperceptions of income inequality and political attitude polarization.

Beyond the class of population games considered in this paper, future work might explore the implications of assortativity neglect for behavior and misperceptions in games with aggregate uncertainty (e.g., financial markets) or in dynamic settings (e.g., social learning; see Section 7.1 of Frick, Iijima, and Ishii, 2020).

[34] A discounted-sum expression of equilibrium actions also appears in Morris and Shin (2002), although their analysis is focused on Gaussian information structures and does not make a connection with Markov processes.

Appendix

Proofs for Appendix A and C are in Online Appendix D.1 and D.2, respectively.

A Preliminaries

A.1 Operator T_C induced by interaction structure C

Many of our proofs will make use of a particular operator T_C over the the space of inverse cdfs that is induced by any interaction structure C. Let L^1 be the space of all measurable functions $f : (0,1) \to \mathbb{R}$ such that $\int_0^1 |f(x)|dx < \infty$, endowed with the L^1 norm. Let $\mathcal{I} \subseteq L^1$ denote the subset consisting of weakly increasing and absolutely continuous functions.[35] For each cdf $F \in \mathcal{F}$, we have that F^{-1} is strictly increasing, absolutely continuous and that $\int_0^1 |F^{-1}(x)|dx = \int |\theta| dF(\theta) < \infty$, so $F^{-1} \in \mathcal{I}$. Conversely, for any strictly increasing $f \in \mathcal{I}$, we have $f^{-1} \in \mathcal{F}$.

Given any interaction structure C, define the operator T_C over L^1 by

$$T_C f(x) = \int_0^1 f(y)\, dC(y|x)$$

for all $f \in L^1$. If $C \in \mathcal{C}$ with density c, then $T_C f(x) = \int_0^1 c(y,x) f(y) dy$ for all $f \in L^1$. The following lemma records basic properties of T_C that we invoke without reference from now on:

Lemma A.1. *Fix any $C \in \mathcal{C}$. Then T_C is a continuous operator from L^1 to L^1 with the following properties:*

1. *$\|T_C f\| \leq \|f\|$ for each $f \in L^1$.*
2. *$T_C f \in \mathcal{I}$ for any $f \in \mathcal{I}$.*
3. *For any $\gamma \in (-1,1)$ and $f \in L^1$,*

$$\lim_{\tau \to \infty} \sum_{t=0}^{\tau} \gamma^t (T_C)^t f = \sum_{t=0}^{\infty} \gamma^t (T_C)^t f \in L^1,$$

where $(T_C)^t$ is defined by $(T_C)^0(f) := f$ and $(T_C)^{t+1}(f) := (T_C)^t(T_C f)$ for all f and t.

A.2 Mean-preserving spread and dispersiveness orders over $\mathcal{I}$

Define a binary relation $\succsim_m$ over $\mathcal{I}$ by setting $f \succsim_m g$ if and only if $\int_0^1 \phi(f(x))dx \geq \int_0^1 \phi(g(x))dx$ for all convex functions ϕ such that $\phi \circ f, \phi \circ g \in L^1$. Note that for F, $G \in \mathcal{F}$, F is a mean-

[35] That is, for any $x, x' \in (0,1)$, there is an integrable function f' such that $f(x) = f(x') + \int_{x'}^{x} f'(y)dy$.

preserving spread of G if and only if $F^{-1} \succsim_m G^{-1}$. The following characterization of $\succsim_m$ is standard (for the proof, see Section 3.A.1 in Shaked and Shanthikumar, 2007):

Lemma A.2. *Let $f, g \in \mathcal{I}$. Then $f \succsim_m g$, if and only if, $\int_y^1 f(x)dx \geq \int_y^1 g(x)dx$ holds for all $y \in (0,1)$ and holds with equality when $y = 0$.*

Define binary relation $\succsim_d$ over $\mathcal{I}$ by $f \succsim_d g$ if and only if $f(x) - f(x') \geq g(x) - g(x')$ for all $x, x' \in (0,1)$ with $x \geq x'$. For $F, G \in \mathcal{F}$, F is more dispersive than G if and only if $F^{-1} \succsim_d G^{-1}$.

We say that a preorder (i.e., reflexive and transitive binary relation) $\succsim$ over $\mathcal{I}$ is ***linear*** if for any $f, g, h \in \mathcal{I}$ and $\alpha_1, \alpha_2 > 0$, we have $f \succsim g$ if and only if $\alpha_1 f + \alpha_2 h \succsim \alpha_1 g + \alpha_2 h$; ***continuous*** if for any $f_n \to f \in \mathcal{I}, g_n \to g \in \mathcal{I}$ with $f_n \succsim g_n$ for each n, we have $f \succsim g$; and ***isotone*** if $f \succsim g$ implies $T_C f \succsim T_C g$ for any $C \in \mathcal{C}$. Orders $\succsim_m$ and $\succsim_d$ satisfies these properties:

Lemma A.3. *$\succsim_m$ and $\succsim_d$ are preorders over $\mathcal{I}$ that are linear, continuous, and isotone.*

Finally, we show that $(T_C)^t f$ is $\succsim_m$-decreasing in t:

Lemma A.4. *$(T_C)^t f \succsim_m (T_C)^{t+1} f$ for all $t \geq 0$, $C \in \mathcal{C}$ and $f \in \mathcal{I}$.*

B Main Proofs

B.1 Proof of Lemma 1

Write $P = (F, C)$ and $\mu := \mathbb{E}_F[\theta]$. Since $F \in \mathcal{F}$, $F^{-1} \in \mathcal{I}$ with F^{-1} strictly increasing. Define

$$h(x) := \sum_{t \geq 0} \gamma^t (T_C)^t F^{-1}(x) + \frac{\beta\mu}{(1-\gamma)(1-\beta-\gamma)}$$

for each $x \in (0,1)$. Note that, by construction, $h = F^{-1} + \beta T_{C_I} h + \gamma T_C h$, where $C_I(x,y) = xy$ denotes the non-assortative interaction structure. Moreover, h is strictly increasing, since $(T_C)^t F^{-1}$ is weakly increasing for each $t \geq 0$ and strictly increasing for $t = 0$. Note also that for each t, $(T_C)^t F^{-1} \in \mathcal{I}$ and hence there exists $(T_C^t F^{-1})' : (0,1) \to \mathbb{R}_+$ such that $(T_C)^t F^{-1}(x) - (T_C)^t F^{-1}(x') = \int_{x'}^x (T_C^t F^{-1})'(y)dy$ for all $x > x'$. Thus, h is absolutely continuous as

$$\begin{aligned} h(x) - h(x') &= \lim_{\tau \to \infty} \textstyle\sum_{t=0}^{\tau} \int_{x'}^x \gamma^t (T_C^t F^{-1})'(y)dy \\ &= \lim_{\tau \to \infty} \textstyle\int_{x'}^x \sum_{t=0}^{\tau} \gamma^t (T_C^t F^{-1})'(y)dy = \int_{x'}^x \sum_{t \geq 0} \gamma^t (T_C^t F^{-1})'(y)dy, \end{aligned}$$

where the last equality holds by the monotone convergence theorem.

Let $s(\theta) := h(F(\theta))$ for each $\theta \in \Theta$. Since $h \in L^1$, we have $\int |s(\theta)| dF(\theta) = \int |h(x)|dx < \infty$. Moreover, s inherits strict monotonicity and absolute continuity (by the change of variable theorem) from h and F. Finally, s is a Nash equilibrium because for each type θ and $x = F(\theta)$, we have $s(\theta) = h(x) = F^{-1}(x) + \beta T_{C_I} h(x) + \gamma T_C h(x) = \theta + \beta \mathbb{E}_F[s(\theta')] + \gamma \mathbb{E}_P[s(\theta') \mid \theta]$.

To show uniqueness of equilibrium, consider any Nash equilibrium $\hat{s}$. Define $\hat{h}(x) := \hat{s}(F^{-1}(x))$ for each x. By the best-response condition for $\hat{s}$, we have

$$\hat{h} = F^{-1} + \beta T_{C_I}\hat{h} + \gamma T_C \hat{h}. \tag{14}$$

Iterating (14) yields

$$\begin{aligned}\hat{h} &= F^{-1} + \beta T_{C_I}\hat{h} + \gamma T_C\left(F^{-1} + \beta T_{C_I}\hat{h}\right) + \gamma^2 (T_C)^2\hat{h} = \dots \\ &= \sum_{t=0}^{\tau} \gamma^t (T_C)^t \left(F^{-1} + \beta T_{C_I}\hat{h}\right) + \gamma^{\tau+1}(T_C)^{\tau+1}\hat{h}\end{aligned}$$

for all $\tau \in \mathbb{N}$. The analogous iteration holds for h. Thus,

$$\begin{aligned}\|\hat{h} - h\| &\leq \| \sum_{t=0}^{\tau} \gamma^t (T_C)^t \left(F^{-1} - \beta T_{C_I}\hat{h} - F^{-1} + \beta T_{C_I} h\right) \| + \gamma^{\tau+1}\|(T_C)^{\tau+1}(\hat{h} - h)\| \\ &\leq \| \sum_{t=0}^{\tau} \gamma^t (T_C)^t \left(\beta T_{C_I}(h - \hat{h})\right) \| + \gamma^{\tau+1}\|\hat{h} - h\|,\end{aligned}$$

which converges to $\| \sum_{t=0}^{\infty} \gamma^t (T_C)^t \left(\beta T_{C_I}(h - \hat{h})\right) \|$ as $\tau \to \infty$. But integrating both sides of (14) with respect to x, we obtain $\int_0^1 \hat{h}(x)dx = T_{C_I}\hat{h}(y) = \frac{\mu}{1-\beta-\gamma}$ for each y, and analogously $T_{C_I}h(y) = \frac{\mu}{1-\beta-\gamma}$ from the best-response condition for h. Thus, $\|\hat{h} - h\| = 0$, whence $\hat{s} = s$. □

B.2 Proof of Proposition 1

Fix any (P, β, γ). By the best-response condition (6), any ANE s^{AN} is the Nash equilibrium at (P, β', γ'), where $\beta' = 0$, $\gamma' = \beta + \gamma$. Thus, by Lemma 1, s^{AN} is uniquely given by (7).

To show that the ANE global action distribution G^{AN} is more dispersive than the Nash distribution G^{NE}, it suffices to show that $s^{AN}(\theta) - s^{AN}(\theta') \geq s^{NE}(\theta) - s^{NE}(\theta')$ for all $\theta > \theta'$. To show this, note that for all τ,

$$0 \leq \sum_{t=0}^{\tau} ((\gamma + \beta)^t - \gamma^t) \left(\mathbb{E}_P[\theta_t | \theta_0 = \theta] - \mathbb{E}_P[\theta_t | \theta_0 = \theta']\right),$$

as the monotonicity of process P implies $\mathbb{E}_P[\theta_t|\theta_0 = \theta] \geq \mathbb{E}_P[\theta_t|\theta_0 = \theta']$ for all t. By (5) and (7), the RHS converges to $(s^{AN}(\theta) - s^{AN}(\theta')) - (s^{NE}(\theta) - s^{NE}(\theta'))$ as $\tau \to \infty$. □

B.3 Proof of Proposition 2

Let $\succsim_{MA}$ denote the more-assortative order over $\mathcal{C}$. We first show that $\succsim_{MA}$ is the "dual order" of the mean-preserving spread order $\succsim_m$:

Lemma B.1. *Fix any $C_1, C_2 \in \mathcal{C}$. Then $C_1 \succsim_{MA} C_2$ if and only if $T_{C_1}F^{-1} \succsim_m T_{C_2}F^{-1}$ for all $F \in \mathcal{F}$.*

Proof. First, observe that $C_1 \succsim_{MA} C_2$ if and only if $C_1(\cdot|x \geq y)$ first-order stochastically dominates $C_2(\cdot \mid x \geq y)$ for any $y \in (0,1)$. This is because

$$\begin{aligned} C_1 \succsim_{MA} C_2 &\iff \int_0^z \int_0^y c_1(x',x)dxdx' \geq \int_0^z \int_0^y c_2(x',x)dxdx' \ \forall y, z \in (0,1) \\ &\iff \int_0^z \int_y^1 c_1(x',x)dxdx' \leq \int_0^z \int_y^1 c_2(x',x)dxdx' \ \forall y, z \in (0,1) \\ &\iff C_1(z \mid x \geq y) \leq C_2(z \mid x \geq y) \ \forall y, z \in (0,1), \end{aligned}$$

where the second line uses $\int_0^z \int_0^1 c_i(x',x)dxdx' = \int_0^1 \int_0^z c_i(x',x)dx'dx = z$ for each $i = 1, 2$.

Next, note that for any $F \in \mathcal{F}$, we have $T_{C_1}F^{-1} \succsim_m T_{C_2}F^{-1}$ if and only if $\int_y^1 T_{C_1}f(x)dx \geq \int_y^1 T_{C_2}f(x)dx$ for all $y \in (0,1)$ with equality if $y = 0$. But

$$\begin{aligned} &\int_y^1 T_{C_1}f(x)dx \geq \int_y^1 T_{C_2}f(x)dx, \ \ \forall y \in (0,1) \\ \iff &\int_y^1 \int_0^1 c_1(x',x)f(x')dx'dx \geq \int_y^1 \int_0^1 c_2(x',x)f(x')dx'dx, \ \ \forall y \in (0,1) \\ \iff &\int_0^1 \int_y^1 \tfrac{1}{1-y} c_1(x',x)dxf(x')dx' \geq \int_0^1 \int_y^1 \tfrac{1}{1-y} c_2(x',x)dxf(x')dx', \ \ \forall y \in (0,1). \end{aligned}$$

Since the set of all F^{-1} with $F \in \mathcal{F}$ consists of all L^1, strictly increasing and absolutely continuous functions on $(0,1)$, this implies that $T_{C_1}F^{-1} \succsim_m T_{C_2}F^{-1}$ holds for all $F \in \mathcal{F}$ if and only if $C_1(\cdot \mid x \geq y)$ first-order stochastically dominates $C_2(\cdot \mid x \geq y)$ for any $y \in (0,1)$. By the first paragraph, this is equivalent to $C_1 \succsim_{MA} C_2$. □

Proof of Proposition 2. **(1.) ⇒ (2.)**: Suppose that $C_1 \succsim_{MA} C_2$ and consider any F, β, γ. Let $f := F^{-1}$, which is in $\mathcal{I}$ since $F \in \mathcal{F}$. We first show by induction that $(T_{C_1})^t f \succsim_m (T_{C_2})^t f$ for all t. For $t = 1$, this is true by Lemma B.1. Suppose the claim holds for some $t \geq 1$. Then

$$(T_{C_1})^{t+1} f = T_{C_1} (T_{C_1})^t f \succsim_m T_{C_1} (T_{C_2})^t f \succsim_m T_{C_2} (T_{C_2})^t f = (T_{C_2})^{t+1} f,$$

where the first comparison follows from the inductive hypothesis by isotonicity of $\succsim_m$, and the second one holds by Lemma B.1. Thus, by transitivity of $\succsim_m$, we have $(T_{C_1})^{t+1} f \succsim_m (T_{C_2})^{t+1} f$.

Next, note that linearity of $\succsim_m$ and $C_1 \succsim_{MA} C_2$ implies

$$\begin{aligned} &\sum_{t=0}^{\tau} \gamma^t (T_{C_1})^t F^{-1} \succsim_m \left(\gamma^\tau (T_{C_2})^\tau + \sum_{t=0}^{\tau-1} \gamma^t (T_{C_1})^t \right) F^{-1} \\ \succsim_m &\left(\sum_{t=\tau-1}^{\tau} \gamma^t (T_{C_2})^t + \sum_{t=0}^{\tau-2} \gamma^t (T_{C_1})^t \right) F^{-1} \succsim_m \cdots \succsim_m \sum_{t=0}^{\tau} \gamma^t (T_{C_2})^t F^{-1} \end{aligned}$$

for any $\tau \in \mathbb{N}$. Moreover, by Lemma A.1, as $\tau \to \infty$, we have

$$\sum_{t=0}^{\tau} \gamma^t (T_{C_1})^t F^{-1} \to \sum_{t=0}^{\infty} \gamma^t (T_{C_1})^t F^{-1}, \quad \sum_{t=0}^{\tau} \gamma^t (T_{C_2})^t F^{-1} \to \sum_{t=0}^{\infty} \gamma^t (T_{C_2})^t F^{-1}.$$

Thus, by continuity and linearity of $\succsim_m$, we have

$$\sum_{t=0}^{\infty} \gamma^t (T_{C_1})^t F^{-1} + \frac{\beta\mu}{(1-\gamma)(1-\beta-\gamma)} \succsim_m \sum_{t=0}^{\infty} \gamma^t (T_{C_2})^t F^{-1} + \frac{\beta\mu}{(1-\gamma)(1-\beta-\gamma)},$$

where $\mu = \mathbb{E}_F[\theta]$. Thus, G_1^{NE} is a mean-preserving spread of G_2^{NE} at (F, β, γ).
(2.) $\Rightarrow$ (3.): Immediate from the fact that G^{AN} at (P, β, γ) coincides with G^{NE} at $(P, 0, \beta+\gamma)$.
(3.) $\Rightarrow$ (1.): Let $g_{F,C_i,\beta,\gamma}$ denote the inverse of the ANE global action distribution at (F, C_i, β, γ). Suppose $g_{F,C_1,\beta,\gamma} \succsim_m g_{F,C_2,\beta,\gamma}$ for all (F, β, γ). Setting $f := F^{-1}$ and $\delta := \beta + \gamma$, we have

$$\sum_{t\geq 0} \delta^t (T_{C_1})^t f = g_{F,C_1,\beta,\gamma} \succsim_m g_{F,C_2,\beta,\gamma} = \sum_{t\geq 0} \delta^t (T_{C_2})^t f.$$

By linearity of $\succsim_m$ and since $(T_{C_i})^0(f) = f$ for $i = 1, 2$, this implies

$$T_{C_1} f + \sum_{t\geq 2} \delta^{t-1} (T_{C_1})^t f \succsim_m T_{C_2} f + \sum_{t\geq 2} \delta^{t-1} (T_{C_2})^t f. \tag{15}$$

Note that for each $i = 1, 2$,

$$\|T_{C_i} f + \sum_{t\geq 2} \delta^{t-1} (T_{C_i})^t f - T_{C_i} f\| \leq \sum_{t\geq 2} \delta^{t-1} \|(T_{C_i})^t f\| \leq \sum_{t\geq 2} \delta^{t-1} \|f\|.$$

Hence, as $\delta \to 0$, $T_{C_i} f + \sum_{t\geq 2} \delta^{t-1} (T_{C_i})^t f \to T_{C_i} f$. Thus, by continuity of $\succsim_m$, (15) yields $T_{C_1} f \succsim T_{C_2} f$. As this is true for all $f = F^{-1}$, we have $C_1 \succsim_{MA} C_2$ by Lemma B.1.
(1.) $\Leftrightarrow$ (4.): We first show that (1.) implies (4.). By the proof of "(1.) $\Rightarrow$ (2.)," we have $(T_{C_1})^t F^{-1} \succsim_m (T_{C_2})^t F^{-1}$ for all t. Thus,

$$\left((\beta+\gamma)^t (T_{C_1})^t + \gamma^t (T_{C_2})^t\right) F^{-1} \succsim_m \left(\gamma^t (T_{C_1})^t + (\beta+\gamma)^t (T_{C_2})^t\right) F^{-1},$$

as $(\beta+\gamma)^t \geq \gamma^t \geq 0$ and by linearity of $\succsim_m$. Then linearity and continuity of $\succsim_m$ also imply

$$\begin{aligned} &\textstyle\sum_{t=0}^{\infty} \left((\beta+\gamma)^t (T_{C_1})^t + \gamma^t (T_{C_2})^t\right) F^{-1} + \frac{\beta \mathbb{E}_F[\theta]}{(1-\gamma)(1-\beta-\gamma)} \\ &\succsim_m \textstyle\sum_{t=0}^{\infty} \left(\gamma^t (T_{C_1})^t + (\beta+\gamma)^t (T_{C_2})^t\right) F^{-1} + \frac{\beta \mathbb{E}_F[\theta]}{(1-\gamma)(1-\beta-\gamma)}. \end{aligned}$$

By monotonicity of equilibrium strategies, this yields for all θ^* that

$$\mathbb{E}_F[s_1^{AN}(\theta) + s_2^{NE}(\theta)|\theta \geq \theta^*] \geq \mathbb{E}_F[s_1^{NE}(\theta) + s_2^{AN}(\theta)|\theta \geq \theta^*],$$

which is equivalent to the first inequality in part (4.). The second inequality follows from part (2.), which is implied by part (1.) as we showed above. Finally, to see that (4.) implies (1.), note that the second inequality in (4.) implies (2.). Thus, (1.) follows from the above proofs. □

B.4 Proof of Proposition 3

(1.) ⇒ (2.): Take any $F_1, F_2 \in \mathcal{F}$ such that F_1 is more dispersive than F_2. Then $F_1^{-1} \succsim_d F_2^{-1}$. First, we inductively show that for each t, $(T_C)^t F_1^{-1} \succsim_d (T_C)^t F_2^{-1}$. Indeed, supposing that the claim is true at t, isotonicity of $\succsim_d$ implies

$$(T_C)^{t+1} F_1^{-1} = T_C(T_C)^t F_1^{-1} \succsim_d T_C(T_C)^t F_2^{-1} = (T_C)^{t+1} F_2^{-1},$$

as required. Next, since $\succsim_d$ is linear, we have $\sum_{t=0}^{\tau} \gamma^t (T_C)^t F_1^{-1} \succsim_d \sum_{t=0}^{\tau} \gamma^t (T_C)^t F_2^{-1}$ for all $\tau \in \mathbb{N}$. Since $\lim_{\tau\to\infty} \sum_{t=0}^{\tau} \gamma^t (T_C)^t F_i^{-1} = \sum_{t=0}^{\infty} \gamma^t (T_C)^t F_i^{-1}$ for each $i = 1, 2$ and any $\gamma \in [0, 1)$, continuity and linearity of $\succsim_d$ then yields

$$(G_1^{NE})^{-1} = \sum_{t\geq 0} \gamma^t (T_C)^t F_1^{-1} + \frac{\beta \mathbb{E}_{F_1}[\theta]}{(1-\gamma)(1-\beta-\gamma)} \succsim_d \sum_{t\geq 0} \gamma^t (T_C)^t F_2^{-1} + \frac{\beta \mathbb{E}_{F_2}[\theta]}{(1-\gamma)(1-\beta-\gamma)} = (G_2^{NE})^{-1},$$

whence G_1^{NE} is more dispersive than G_2^{NE}.

(2.) ⇒ (3.): Immediate from the fact that G^{AN} at (P, β, γ) coincides with G^{NE} at $(P, 0, \beta+\gamma)$.

(3.) ⇒ (1.): Immediate from the fact that G^{AN} at $(P, 0, 0)$ coincides with F.

(1.) ⇔ (4.): To see that (1.) implies (4.), note that, for any $x > x'$,

$$\begin{aligned}
& \left((G_1^{AN})^{-1}(x) - (G_1^{AN})^{-1}(x')\right) - \left((G_1^{NE})^{-1}(x) - (G_1^{NE})^{-1}(x')\right) \\
= & \sum_{t\geq 0} \left((\gamma+\beta)^t - \gamma^t\right) \left((T_C)^t F_1^{-1}(x) - (T_C)^t F_1^{-1}(x')\right) \\
\geq & \sum_{t\geq 0} \left((\gamma+\beta)^t - \gamma^t\right) \left((T_C)^t F_2^{-1}(x) - (T_C)^t F_2^{-1}(x')\right) \\
= & \left((G_2^{AN})^{-1}(x) - (G_2^{AN})^{-1}(x')\right) - \left((G_2^{NE})^{-1}(x) - (G_2^{NE})^{-1}(x')\right).
\end{aligned}$$

Here the inequality holds since by the proof of "(1.) ⇒ (2.)," we have $(T_C)^t F_1^{-1} \succsim_d (T_C)^t F_2^{-1}$. This establishes the first inequality in (4.). The second inequality in (4.) holds by the fact that (1.) implies (2.), as shown above. Finally, to see that (4.) implies (1.), note that the second inequality in (4.) implies (2.). Thus, (1.) follows from the above proofs. □

B.5 Proof of Proposition 4

We only consider ANE; the proof for Nash is analogous. Suppose $\beta_1 + \gamma_1 \geq \beta_2 + \gamma_2$. For any (F, C), consider the inverse cdf g_i of G_i^{AN} at $(F, C, \beta_i, \gamma_i)$. Then $g_i = \sum_{t\geq 0}(\beta_i + \gamma_i)^t (T_C)^t f$, where $f = F^{-1}$. Observe that $(\beta_1 + \gamma_1)^t (T_C)^t f \succsim_d (\beta_2 + \gamma_2)^t (T_C)^t f$ for all t. Thus, $\sum_{t\geq 0}^{\tau}(\beta_1 + \gamma_1)^t (T_C)^t f \succsim_d \sum_{t\geq 0}^{\tau}(\beta_2+\gamma_2)^t (T_C)^t f$ for all $\tau \geq 0$ by linearity of $\succsim_d$. Then by continuity of $\succsim_d$, it follows that $g_1 \succsim_d g_2$, whence G_1^{AN} is more dispersive than G_2^{AN}. Conversely, if $\beta_1+\gamma_1 < \beta_2+\gamma_2$, the same argument implies $g_2 \succ_d g_1$ for any (F, C) with $C \neq C_I$. □

B.6 Proof of Proposition 5

We first verify that $L_\theta^{s^{AN},P} \in \mathcal{F}$ for each θ. By Proposition 1, $(s^{AN})^{-1}$ is strictly increasing and absolutely continuous. By monotonicity of s^{AN}, $L_\theta^{s^{AN},P}(a) = P((s^{AN})^{-1}(a)|\theta)$ for each $a \in s^{AN}(\Theta)$. Since $P(\theta'|\theta)$ is absolutely continuous and strictly increasing in θ' on Θ, $L_\theta^{s^{AN},P}(a)$ is absolutely continuous (by the change of variable theorem) and strictly increasing in $a \in s^{AN}(\Theta)$. Moreover, since s^{AN} is L^1 with respect to F, $\int\int |s^{AN}(\theta')|dP(\theta'|\theta)dF(\theta) = \int |s^{AN}(\theta)|dF(\theta) < \infty$. Thus, there exists $\Theta^* \subseteq \Theta$ such that $\Theta\backslash\Theta^*$ has Lebesgue measure zero and for every $\theta \in \Theta^*$, $\int |s^{AN}(\theta')|dP(\theta'|\theta) < \infty$. Hence, $L_\theta^{s^{AN},P}$ is L^1 for all $\theta \in \Theta^*$. As $L_\theta^{s^{AN},P}$ is FOSD-monotonic in θ, this implies that $L_\theta^{s^{AN},P}$ is L^1 for every $\theta \in \Theta$.[36]

Define a type distribution $\hat{F}_\theta$ by

$$\hat{F}_\theta^{-1}(x) = (L_\theta^{s^{AN},P})^{-1}(x) - (\beta + \gamma)\int (L_\theta^{s^{AN},P})^{-1}(z)dz \tag{16}$$

for each x. Since $L_\theta^{s^{AN},P} \in \mathcal{F}$, it follows that $\hat{F}_\theta^{-1}$ is L^1, strictly increasing, and absolutely continuous, so $\hat{F}_\theta \in \mathcal{F}$. Let $\hat{P}_\theta := \hat{F}_\theta \times \hat{F}_\theta$ and let $\hat{s}_\theta$ be the Nash equilibrium at $(\hat{P}_\theta, \beta, \gamma)$. Then $\hat{s}_\theta(\theta') = \mathrm{BR}_{\theta'}(G^{\hat{s}_\theta,\hat{P}_\theta}, L_{\theta'}^{\hat{s}_\theta,\hat{P}_\theta})$ for each $\theta' \in \mathrm{supp}\hat{F}_\theta$, so the perceived best-response condition holds. Moreover, for each x,

$$\begin{aligned}(L_\theta^{\hat{s}_\theta,\hat{P}_\theta})^{-1}(x) &= \hat{F}_\theta^{-1}(x) + (\beta+\gamma)\int (L_\theta^{\hat{s}_\theta,\hat{P}_\theta})^{-1}(z)dz \\ &= (L_\theta^{s^{AN},P})^{-1}(x) + (\beta+\gamma)\int \left((L_\theta^{\hat{s}_\theta,\hat{P}_\theta})^{-1}(z) - (L_\theta^{s^{AN},P})^{-1}(z)\right)dz,\end{aligned}$$

where the first equality uses the perceived best-response condition and $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$, and the second equality uses (16). Integrating both sides with respect to x yields $L_\theta^{\hat{s}_\theta,\hat{P}_\theta} = L_\theta^{s^{AN},P}$, verifying observational consistency. Thus, $(\hat{P}_\theta, \hat{s}_\theta)$ are coherent assortativity neglect perceptions.

[36] Indeed, take any $\theta \in \Theta\backslash\Theta^*$. If $\theta \in (\inf\Theta, \sup\Theta)$, pick $\theta', \theta'' \in \Theta^*$ with $\theta' < \theta < \theta''$. Then $\int |a|dL_\theta^{s^{AN},P}(a) = -\int_{-\infty}^0 a dL_\theta^{s^{AN},P}(a) + \int_0^\infty a dL_\theta^{s^{AN},P}(a) \leq -\int_{-\infty}^0 a dL_{\theta'}^{s^{AN},P}(a) + \int_0^\infty a dL_{\theta''}^{s^{AN},P}(a) < \infty$. If $\theta = \sup\Theta$ (the case $\theta = \inf\Theta$ is analogous), then $\mathrm{supp}L_\theta^{s^{AN},P}$ is bounded above. Thus, $\int |a|dL_\theta^{s^{AN},P}(a) = -\int_{-\infty}^0 a dL_\theta^{s^{AN},P}(a) + \int_0^\infty a dL_\theta^{s^{AN},P}(a) \leq -\int_{-\infty}^0 a dL_{\theta'}^{s^{AN},P}(a) + \int_0^\infty a dL_\theta^{s^{AN},P}(a) < \infty$ for any $\theta' \in \Theta^*$.

Conversely, for any coherent assortativity neglect perceptions ($\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta, \hat{s}_\theta$), observational consistency and perceived best-response imply (16), ensuring the uniqueness of $\hat{F}_\theta$. Moreover, by the perceived best-response condition, $\hat{s}_\theta$ is the unique Nash equilibrium at $(\hat{P}_\theta, \beta, \gamma)$. □

B.7 Proof of Proposition 6

Write $P = (F, C)$. Let s_i and $(\hat{P}^i_\theta = \hat{F}^i_\theta \times \hat{F}^i_\theta, \hat{s}^i_\theta)$ denote the ANE strategy profile and θ's coherent assortativity neglect perceptions at (P, β_i, γ_i) for $i = 1, 2$. Suppose $\beta_1 + \gamma_1 \geq \beta_2 + \gamma_2$. For any $x, y \in (0, 1)$ with $x > y$, we have

$$\begin{aligned}(\hat{F}^i_\theta)^{-1}(x) - (\hat{F}^i_\theta)^{-1}(y) &= (L_\theta^{\hat{s}^i_\theta, \hat{P}^i_\theta})^{-1}(x) - (L_\theta^{\hat{s}^i_\theta, \hat{P}^i_\theta})^{-1}(y) \\ = (L_\theta^{s^i, P})^{-1}(x) - (L_\theta^{s^i, P})^{-1}(y) &= \textstyle\sum_{t\geq 0}(\beta_i + \gamma_i)^t \left((T_C)^t F^{-1}(x') - (T_C)^t F^{-1}(y')\right),\end{aligned}$$

where $x', y' \in (0, 1)$ with $x' > y'$ are defined by $C(x' \mid F^{-1}(\theta)) = x$ and $C(y' \mid F^{-1}(\theta)) = y$. Indeed, the first equality holds by the perceived best-response condition, the second equality by observational consistency, and the final one by construction of ANE strategies. Since $(\beta_1+\gamma_1)^t \geq (\beta_2+\gamma_2)^t$ and $(T_C)^t F^{-1}(x') - (T_C)^t F^{-1}(y') \geq 0$ for all t, it follows that $(\hat{F}^1_\theta)^{-1}(x) - (\hat{F}^1_\theta)^{-1}(y) \geq (\hat{F}^2_\theta)^{-1}(x) - (\hat{F}^2_\theta)^{-1}(y)$. Thus, $\hat{F}^1_\theta$ is more dispersive than $\hat{F}^2_\theta$. □

B.8 Proof of Proposition 7

Write $P = (F, C)$. Let $f := F^{-1}$ and $\eta_i := \beta_i + \gamma_i$ for each $i = 1, 2$. Since the local action average observed by each type quantile x under the ANE at (P, β_i, γ_i) is given by $\sum_{t\geq 0} \eta_i^t (T_C)^{t+1} f(x)$, we have $(M^i)^{-1} = (1 - \eta_i) \sum_{t\geq 0} \eta_i^t (T_C)^{t+1} f$. For each $\tau \geq 0$, we show

$$\frac{1}{\sum_{t=0}^{\tau} \eta_2^t} \sum_{t=0}^{\tau} \eta_2^t (T_C)^{t+1} f \succsim_m \frac{1}{\sum_{t=0}^{\tau} \eta_1^t} \sum_{t=0}^{\tau} \eta_1^t (T_C)^{t+1} f. \tag{17}$$

For $\tau = 0$, there is nothing to prove. Supposing the claim holds for some $\tau \geq 0$, we have

$$\begin{aligned}\frac{1}{\sum_{t=0}^{\tau+1} \eta_2^t} \sum_{t=0}^{\tau+1} \eta_2^t (T_C)^{t+1} f &= \frac{\sum_{t=0}^{\tau} \eta_2^t}{\sum_{t=0}^{\tau+1} \eta_2^t} \left(\frac{1}{\sum_{t=0}^{\tau} \eta_2^t} \sum_{t=0}^{\tau} \eta_2^t (T_C)^{t+1} f \right) + \frac{\eta_2^{\tau+1}}{\sum_{t=0}^{\tau+1} \eta_2^t} (T_C)^{\tau+2} f \\ &\succsim_m \frac{\sum_{t=0}^{\tau} \eta_2^t}{\sum_{t=0}^{\tau+1} \eta_2^t} \left(\frac{1}{\sum_{t=0}^{\tau} \eta_1^t} \sum_{t=0}^{\tau} \eta_1^t (T_C)^{t+1} f \right) + \frac{\eta_2^{\tau+1}}{\sum_{t=0}^{\tau+1} \eta_2^t} (T_C)^{\tau+2} f \\ \succsim_m \frac{\sum_{t=0}^{\tau} \eta_1^t}{\sum_{t=0}^{\tau+1} \eta_1^t} \left(\frac{1}{\sum_{t=0}^{\tau} \eta_1^t} \sum_{t=0}^{\tau} \eta_1^t (T_C)^{t+1} f \right) + \frac{\eta_1^{\tau+1}}{\sum_{t=0}^{\tau+1} \eta_1^t} (T_C)^{\tau+2} f &= \frac{1}{\sum_{t=0}^{\tau+1} \eta_1^t} \sum_{t=0}^{\tau+1} \eta_1^t (T_C)^{t+1} f,\end{aligned}$$

as required. Here the first dominance holds by inductive hypothesis and the second dominance follows from linearity of $\succsim_m$ along with the fact that $\eta_1 \geq \eta_2$ (so that $\frac{\sum_{t=0}^{\tau} \eta_1^t}{\sum_{t=0}^{\tau+1} \eta_1^t} \leq \frac{\sum_{t=0}^{\tau} \eta_2^t}{\sum_{t=0}^{\tau+1} \eta_2^t}$ and

$\frac{\eta_1^{\tau+1}}{\sum_{t=0}^{\tau+1}\eta_1^t} \geq \frac{\eta_2^{\tau+1}}{\sum_{t=0}^{\tau+1}\eta_2^t}$) and that $(T_C)^{t+1}f \succsim_m (T_C)^{\tau+2}f$ for all $t \leq \tau+1$ (by Lemma A.4).

Taking $\tau \to \infty$ in (17), continuity of $\succsim_m$ then yields $(1-\eta_2)\sum_{t\geq 0}\eta_2^t(T_C)^{t+1}f \succsim_m (1-\eta_1)\sum_{t\geq 0}\eta_1^t(T_C)^{t+1}f$, i.e., $(M^2)^{-1} \succsim_m (M^1)^{-1}$, as claimed. □

C Extensions

C.1 Generalization of coherent assortativity neglect perceptions

Definition 3 assumes that θ perceives all other agents θ' to share her perceptions $\hat{P}_\theta$ and $\hat{s}_\theta$. Expanding on Remark 2, we show that this assumption is not essential for our results. Specifically, suppose that we enrich type θ's coherent assortativity neglect perceptions to consist of:

- *Own perceptions:* θ's own perception of a non-assortative society $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$ and perceived strategy profile $\hat{s}_\theta$;
- *Perceptions about others' perceptions:* for each type θ', θ's perception of θ''s perceived non-assortative society $\hat{P}_{\theta,\theta'} = \hat{F}_{\theta,\theta'} \times \hat{F}_{\theta,\theta'}$ and θ's perception of θ''s perceived strategy profile $\hat{s}_{\theta,\theta'}$;

subject to three requirements:

1. *Observational consistency:* $L_\theta^{s^{AN},P} = L_\theta^{\hat{s}_\theta,\hat{P}_\theta}$;
2. *Perceived best-response:* for each θ', $\hat{s}_\theta(\theta') \in \mathrm{BR}_{\theta'}(G^{\hat{s}_{\theta,\theta'},\hat{P}_{\theta,\theta'}}, L_{\theta'}^{\hat{s}_{\theta,\theta'},\hat{P}_{\theta,\theta'}})$;
3. *Perceived observational consistency:* for each θ', $L_{\theta'}^{\hat{s}_\theta,\hat{P}_\theta} = L_{\theta'}^{\hat{s}_{\theta,\theta'},\hat{P}_{\theta,\theta'}}$.

The first condition is the same observational consistency requirement as in Definition 3. The second condition still says that θ perceives θ' to play a best-response; however, in rationalizing θ''s behavior, θ now allows that θ''s perceived society and strategy profile might be different from her own. Finally, the third condition requires the perceptions $\hat{P}_{\theta,\theta'}$ and $\hat{s}_{\theta,\theta'}$ that θ attributes to θ' to be consistent with the local action distribution $L_{\theta'}^{\hat{s}_\theta,\hat{P}_\theta}$ that she perceives θ' to observe. Definition 3 corresponds to the special case where θ perceives all other agents to share her perceptions (i.e., $\hat{P}_{\theta,\theta'} = \hat{P}_\theta$ and $\hat{s}_{\theta,\theta'} = \hat{s}_\theta$).

While this generalization allows θ to perceive others to disagree with her perceptions $\hat{P}_\theta$ and $\hat{s}_\theta$, we note that $\hat{P}_\theta$ and $\hat{s}_\theta$ themselves are unchanged relative to Definition 3: Indeed, we have

$$G^{\hat{s}_{\theta,\theta'},\hat{P}_{\theta,\theta'}} = L_{\theta'}^{\hat{s}_{\theta,\theta'},\hat{P}_{\theta,\theta'}} = L_{\theta'}^{\hat{s}_\theta,\hat{P}_\theta} = L_\theta^{\hat{s}_\theta,\hat{P}_\theta} = L_\theta^{s^{AN},P}.$$

The first equality holds because $\hat{P}_{\theta,\theta'}$ is non-assortative, the second by perceived observational consistency, the third because $\hat{P}_\theta$ is non-assortative, and the fourth by observational consistency.

Thus, just as under Definition 3, the perceived best-response condition reduces to (9), i.e.,

$$\hat{s}_\theta(\theta') \in \mathrm{BR}_{\theta'}(L_\theta^{s^{AN},P}, L_\theta^{s^{AN},P}).$$

Based on this and observational consistency, the derivation of θ's perceived type distribution $\hat{F}_\theta$ and strategy profile $\hat{s}_\theta$ is unchanged, so all results in Section 4 remain valid. At the same time, θ's perceptions about θ''s perceptions $\hat{F}_{\theta,\theta'}$ and $\hat{s}_{\theta,\theta'}$ are flexible; for example, the definition is consistent with θ being aware that coherent assortativity neglect perceptions vary across types.

The above definition can be viewed as a misspecified version of a belief system in Esponda's (2013) level-1 rationalizable conjectural equilibrium (i.e., with first-order belief in rationality and observational consistency): We assume that agents dogmatically believe that society is non-assortative and that others share this perception (and for simplicity we do not model agents' kth-order beliefs beyond $k = 2$). Note that modeling agents' entire hierarchy of perceptions and imposing rationality and observational consistency up to higher orders would put more discipline on θ's perceptions about others' perceptions; however, as any such belief system is a special case of the above definition, θ's first-order perceptions $\hat{P}_\theta$ and $\hat{s}_\theta$ remain unchanged.

C.2 α-ANE

We apply the equilibrium concept of α-ANE defined in Section 5.1 to linear best-response games with strategic complementarities. Just as in Section 3.2, we can iterate the best-response conditions (12) under the Markov process P. This yields the following α-ANE strategy profiles s_a^α and s_c^α for assortativity neglect and correct agents:

$$s_a^\alpha(\theta) = \theta + (\beta+\gamma)\sum_{t=1}^{\infty}(\gamma+\alpha\beta)^{t-1}\mathbb{E}_P[\theta_t \mid \theta_0 = \theta] + \frac{(\beta+\gamma)(1-\alpha)\beta\mathbb{E}_F[\theta']}{(1-\gamma-\alpha\beta)(1-\beta-\gamma)} \tag{18}$$

$$s_c^\alpha(\theta) = \theta + \gamma\sum_{t=1}^{\infty}(\gamma+\alpha\beta)^{t-1}\mathbb{E}_P[\theta_t \mid \theta_0 = \theta] + \frac{(1-\alpha(\beta+\gamma))\,\beta\mathbb{E}_F[\theta']}{(1-\gamma-\alpha\beta)(1-\beta-\gamma)}. \tag{19}$$

Thus, the higher-order expectation terms take a "quasi-hyperbolic" form, with geometric discount factor $\gamma + \alpha\beta$ increasing in the share α of assortativity neglect agents. Note that when $\alpha = 0$ (resp. $\alpha = 1$), s_c^α (resp. s_a^α) reduces to the expression for Nash (resp. ANE) in Section 3.2.

Let G_a^α and G_c^α denote the global action distributions among assortativity neglect and correct agents. The following result compares action dispersion across both groups of agents, as well as across different values of α:

Proposition C.1. *Fix any (P, β, γ). For any $\alpha \in [0,1]$, there is a unique α-ANE, whose strategy profiles are given by (18) and (19). Moreover, (i) G_a^α is more dispersive than G_c^α, and (ii) both G_a^α and G_c^α are more dispersive the greater α.*

Thus, behavior among correct agents is less dispersed than among assortativity neglect agents, but action dispersion among both groups is exacerbated the greater the share of assortativity neglect agents.

Given (18)–(19), similar arguments as for the main analysis imply that both G_a^α and G_c^α are subject to analogous comparative statics and multiplier effects as G^{AN} in Propositions 2-4. Moreover, coherent perceptions $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$ and $\hat{s}_\theta$ for assortativity neglect agents can be defined analogously to Definition 3.[37] As in Proposition 5, each θ admits unique coherent assortativity neglect perceptions, and the comparative statics in Propositions 6-7 extend.[38]

C.3 Partial assortativity neglect

Consider PANE as defined in Section 5.1. Extending Proposition 1, the following result shows that, in linear best-response games with strategic complementarities, any PANE induces more action dispersion than Nash (subject to a regularity condition that always holds under ANE):

Proposition C.2. *Fix any (P, β, γ). For any PANE s in which perceived global action averages $\int a\, d\hat{G}_\theta(a)$ are absolutely continuous in θ, the global action distribution $G^{s,P}$ is more dispersive than the Nash action distribution.*

Given any society P and PANE s with perceived global action distributions $(\hat{G}_\theta)_\theta$, define a ***coherent perception*** for type θ to consist of a perceived society $\hat{P}_\theta$ and a perceived strategy profile $\hat{s}_\theta$ such that:

1. (a) $L_\theta^{s,P} = L_\theta^{\hat{s}_\theta, \hat{P}_\theta}$;
   (b) $\hat{G}_\theta = G^{\hat{s}_\theta, \hat{P}_\theta}$.
2. For each θ', $\hat{s}_\theta(\theta') \in \mathrm{BR}_{\theta'}(G^{\hat{s}_\theta, \hat{P}_\theta}, L_{\theta'}^{\hat{s}_\theta, \hat{P}_\theta})$.

Conditions 1(a) and 2 are the same as in Definition 3. Condition 1(b) requires that the perceived global action distribution $\hat{G}_\theta$ to which θ best-responds in the PANE s matches the global action distribution under her perceived society $\hat{P}_\theta$ and strategy profile $\hat{s}_\theta$. In the case of coherent ANE perceptions, the latter condition is immediate from the assumption that $\hat{P}_\theta$ is non-assortative.[39]

Obtaining general analogs of Propositions 5–7 for coherent PANE perceptions is beyond the scope of this paper. However, the following example considers the case of Gaussian societies:

[37] Specifically, let $L_\theta^{\alpha,P} = \alpha L_\theta^{s_a^\alpha, P} + (1-\alpha) L_\theta^{s_c^\alpha, P}$ denote θ's true local action distribution in the α-ANE. A coherent perception for an assortativity neglect agent θ consists of a perceived non-assortative society $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$ and perceived strategy profile $\hat{s}_\theta$ such that (i) $L_\theta^{\alpha,P} = L_\theta^{\hat{s}_\theta, \hat{P}_\theta}$; and (ii) for each θ', $\hat{s}_\theta(\theta') \in \mathrm{BR}_{\theta'}(G^{\hat{s}_\theta, \hat{P}_\theta}, L_{\theta'}^{\hat{s}_\theta, \hat{P}_\theta})$.

[38] More precisely, the comparative statics are now with respect to increasing β and γ separately (as $\beta + \gamma$ is no longer a sufficient statistic). The analog of Proposition 6 is that $\hat{F}_\theta$ is increasing in (β, γ) with respect to the dilation order (defined in Appendix C.4). The false consensus effect and comparative statics of distributions of perceived means with respect to (β, γ) are the same as in Proposition 7.

[39] The fact that $\hat{P}_\theta$ is non-assortative implies that $G^{\hat{s}_\theta, \hat{P}_\theta} = L_\theta^{\hat{s}_\theta, \hat{P}_\theta}$, which is equal to $L_\theta^{s,P}$ by 1(a). Since under ANE $\hat{G}_\theta = L_\theta^{s,P}$, this implies 1(b).

Example C.1. Fix a Gaussian society $P = (\mu, \sigma^2, \rho)$ and consider linear best-response games with strategic complementarities. For each $\hat{\rho} \in [0, \rho]$, we can construct a PANE s^* and associated coherent perceptions $(\hat{P}_\theta, \hat{s}_\theta)_\theta$ such that each type θ's perceived society $\hat{P}_\theta$ is Gaussian with correlation coefficient $\hat{\rho}$: Specifically, for each θ,

1. θ's action is $s^*(\theta) = \frac{\theta-\mu}{1-\gamma\rho-\beta\frac{\rho-\hat{\rho}}{1-\hat{\rho}}} + \frac{\mu}{1-\beta-\gamma}$;

2. θ's coherent perceived society is Gaussian with $\hat{P}_\theta = (\hat{\mu}_\theta, \hat{\sigma}^2, \hat{\rho})$, where

$$\hat{\mu}_\theta = \mu + (\theta-\mu)\frac{(1-\beta-\gamma)(\rho-\hat{\rho})(\gamma+\frac{\beta}{1-\hat{\rho}})}{(\beta+\gamma(1-\hat{\rho}))(1-\gamma\rho-\beta\frac{\rho-\hat{\rho}}{1-\hat{\rho}})}, \qquad \hat{\sigma}^2 = \sigma^2\frac{(1-\rho^2)}{(1-\hat{\rho}^2)}\left(\frac{1-\gamma\hat{\rho}}{1-\gamma\rho-\beta\frac{\rho-\hat{\rho}}{1-\hat{\rho}}}\right)^2;$$

3. θ's coherent perceived strategy profile satisfies $\hat{s}_\theta(\theta') = \frac{\theta-\mu}{1-\gamma\hat{\rho}} + \frac{\hat{\mu}_\theta}{1-\beta-\gamma}$ for all θ'.

See Online Appendix D.2.3 for the derivation; in particular, the fact that agents *underestimate* assortativity (i.e., $\hat{\rho} \in [0, \rho]$) is key in ensuring that perceived global action distributions $\hat{G}_\theta = G^{\hat{s}_\theta, \hat{P}_\theta}$ are FOSD-increasing in θ, as required by PANE. Observe that the above expressions generalize the ones under Nash ($\hat{\rho} = \rho$) and ANE ($\hat{\rho} = 0$) in Examples 1–2. Moreover, the qualitative predictions for action dispersion and perceived type variances and means are the same as under full assortativity neglect.

Finally, we note that one can show conversely that any linear PANE that admits Gaussian coherent perceptions must take the form in conditions 1–3 for some $\hat{\rho} \in [0, \rho]$. ▲

C.4 Strategic substitutes

Consider linear best-response games with global and/or local strategic substitutes (i.e., $\beta \leq 0$ and/or $\gamma \leq 0$). The following result shows that Nash and ANE strategies admit the same Markov process representations as in the complementarity case. Moreover, we provide a simple condition (satisfied, e.g., by Gaussian societies) under which these strategies are monotone:

Proposition C.3. *Fix any P and β, γ with $|\beta+\gamma|, |\gamma| < 1$. The unique Nash and ANE strategies s^{NE} and s^{AN} are given by (5) and (7). Moreover, s^{NE} and s^{AN} are strictly increasing if $\left|\mathbb{E}_P[\theta_1 \mid \theta_0 = \theta] - \mathbb{E}_P[\theta_1 \mid \theta_0 = \theta']\right| \leq |\theta - \theta'|$ for all θ, θ'.*

Analyzing equilibrium behavior is more difficult than under complementarities, because when $\gamma < 0$ (resp. $\beta + \gamma < 0$) the discounted terms in (5) (resp. (7)) alternate signs across odd and even t. To extend our comparative statics results, we impose the condition on societies from Proposition C.3 to ensure monotonicity of Nash and ANE. We also employ a weakening of the more-dispersive order: G_1 is a ***dilation*** of G_2 (Shaked and Shanthikumar, 2007) if there exists $b \in \mathbb{R}$ such that $\int \phi(a) d\, G_1(a) \geq \int \phi(a+b) d\, G_2(a)$ for any convex function ϕ

under which the integrals are well-defined; that is, G_1 is a mean-preserving spread of G_2 up to normalizing means. Finally, we use the following strenghtening of the more-assortative order from Section 3.3: C_1 is ***strongly more assortative*** than C_2, denoted $C_1 \gtrsim_{SMA} C_2$, if

$$C_1(z|y) - C_1(z|x) \geq C_2(z|y) - C_2(z|x), \qquad \text{for all } x, y, z \in (0,1) \text{ with } x \geq y.$$

To interpret, recall that assortativity of C requires the distribution of matches' quantiles to be first-order stochastically increasing in own quantile; C_1 is strongly more assortative than C_2 if this effect is globally stronger under C_1 than under C_2.

Proposition C.4. *Fix $(F_i, C_i, \beta_i, \gamma_i)$ with $|\gamma_i + \beta_i|, |\gamma_i| < 1$ and $\left|\mathbb{E}_{P_i}[\theta_1 \mid \theta_0 = \theta] - \mathbb{E}_{P_i}[\theta_1 \mid \theta_0 = \theta']\right| \leq |\theta - \theta'|$ for all θ, θ', $i = 1, 2$. Let G_i^{AN} and G_i^{NE} denote the corresponding ANE and Nash global action distributions.*

1. *Suppose $F_1 = F_2$, $\beta_1 = \beta_2$, $\gamma_1 = \gamma_2$, and C_1 is strongly more assortative than C_2. Then:*
   (a) *G_2^{AN} is a dilation of G_1^{AN} if $\beta_i + \gamma_i < 0$, and vice versa if $\beta_i + \gamma_i > 0$.*
   (b) *G_2^{NE} is a dilation of G_1^{NE} if $\gamma_i < 0$, and vice versa if $\gamma_i > 0$.*
2. *Suppose $C_1 = C_2$, $\beta_1 = \beta_2$, $\gamma_1 = \gamma_2$, and F_1 is more dispersive than F_2. Then G_1^{AN} is a dilation of G_2^{AN}, and G_1^{NE} is a dilation of G_2^{NE}.*
3. *Suppose $F_1 = F_2$, $C_1 = C_2$. If $\beta_1 + \gamma_1 \geq \beta_2 + \gamma_2$ (resp. $\gamma_1 \geq \gamma_2$), then G_1^{AN} is a dilation of G_2^{AN} (resp. G_1^{NE} is a dilation of G_2^{NE}).*

Relative to the complementarities case, the comparative statics with respect to assortativity have flipped directions under substitutes. The first part also implies that, under local complementarities but stronger global substitutes (i.e., $\gamma > 0$, $\beta + \gamma < 0$), increases in assortativity have the *opposite* effect on Nash and ANE action dispersion; this contrasts with the multiplier effect in Proposition 2, where the difference between Nash and ANE was one of magnitude. Finally, the third part implies that if $\beta \leq 0$, then the Nash action distribution is a dilation of the ANE distribution (assuming monotonicity), reversing the comparison in Proposition 1.

Each agent continues to admit unique coherent assortativity neglect perceptions (see Proposition C.5 below). Proposition 6 (on comparative statics of perceived type dispersion) also remains valid up to replacing the dispersiveness with the dilation order.

C.5 Non-linear best-response functions

Consider general population games as defined in Section 2. For simplicity, we drop the regularity assumptions on strategy profiles and type distributions (L^1, absolute continuity, connected support). The following result shows that the existence and uniqueness of coherent assortativity neglect perceptions (Proposition 5) remains valid under mild conditions on best-responses:

Proposition C.5. *Assume that* $\mathrm{BR}_\theta(G, L)$ *is single-valued, and increasing and surjective in* θ *for all* $G, L \in \Delta(A)$*. Fix any* P *and ANE* s^{AN}*.*[40] *For each type* θ*, the corresponding coherent assortativity neglect perceptions* $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$ *and* $\hat{s}_\theta$ *exist and are unique.*

We can also extend the comparison of action dispersion across ANE and Nash (Proposition 1) to additively separable best-responses with purely global complementarities. This case allows one to sidestep difficulties associated with equilibrium multiplicity, as all Nash action distributions are equally dispersive:

Proposition C.6. *Assume that* $\mathrm{BR}_\theta(G, L) = \phi(\theta) + \psi(G)$ *for some increasing function* $\phi : \Theta \to A$ *and FOSD-increasing function* $\psi : \Delta(A) \to A$*. Fix any* P*. The global action distribution under any monotone ANE* s^{AN} *is more dispersive than under any Nash equilibrium* s^{NE}*.*

References

Ahler, D. J. (2014): "Self-fulfilling misperceptions of public polarization," *The Journal of Politics*, 76(3), 607–620.

Aliprantis, C. D., and K. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer.

Attanasio, O. P., and K. M. Kaufmann (2014): "Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender," *Journal of Development Economics*, 109, 203–216.

Barron, K., S. Huck, and P. Jehiel (2019): "Everyday econometricians: Selection neglect and overoptimism when learning from others," Discussion paper, Working Paper.

Battigalli, P. (1987): "Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali," *unpublished undergraduate dissertation, Bocconi University, Milano*.

Battigalli, P., F. Panebianco, and P. Pin (2020): "Learning and Selfconfirming Equilibria in Network Games," working paper.

Bavetta, S., P. Li Donni, and M. Marino (2019): "An empirical analysis of the determinants of perceived inequality," *Review of Income and Wealth*, 65(2), 264–292.

Bénabou, R. (1993): "Workings of a city: location, education, and production," *The Quarterly Journal of Economics*, 108(3), 619–652.

——— (1996a): "Equity and efficiency in human capital investment: the local connection," *The Review of Economic Studies*, 63(2), 237–264.

——— (1996b): "Heterogeneity, stratification, and growth: macroeconomic implications of community structure and school finance," *The American Economic Review*, pp. 584–609.

Bishop, B. (2009): *The big sort: Why the clustering of like-minded America is tearing us apart*. Houghton Mifflin Harcourt.

Bohren, J. A., and D. Hauser (2021): "Learning with Heterogeneous Misspecified Models: Characterization and Robustness," *Econometrica*, forthcoming.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016): "Stereotypes," *The Quarterly Journal of Economics*, 131(4), 1753–1794.

[40] ANE exist under standard conditions that ensure a solution to the fixed-point problem $s(\theta) = \mathrm{BR}_\theta(G^{s,P}, L_\theta^{s,P})$.

Bordalo, P., M. Tabellini, and D. Y. Yang (2021): "Issue Salience and Political Stereotypes," Discussion paper, National Bureau of Economic Research.

Brands, R. A. (2013): "Cognitive social structures in social network research: A review," *Journal of Organizational Behavior*, 34(S1).

Breitmoser, Y. (2019): "Knowing me, imagining you: Projection and overbidding in auctions," *Games and Economic Behavior*, 113, 423–447.

Brown, J. R., and R. D. Enos (2021): "The measurement of partisan sorting for 180 million voters," *Nature Human Behaviour*, 5, 998–1008.

Chauvin, K. (2018): "A Misattribution Theory of Discrimination," working paper.

Cruces, G., R. Perez-Truglia, and M. Tetaz (2013): "Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment," *Journal of Public Economics*, 98, 100–112.

Daley, D. (1968): "Stochastically monotone Markov chains," *Probability Theory and Related Fields*, 10(4), 305–317.

Dawtry, R. J., R. M. Sutton, and C. G. Sibley (2019): "Social sampling, perceptions of wealth distribution, and support for redistribution," in *The Social Psychology of Inequality*, ed. by J. Jetten, and K. Peters, pp. 381–396. Springer.

Dessi, R., E. Gallo, and S. Goyal (2016): "Network cognition," *Journal of Economic Behavior & Organization*, 123, 78–96.

Durlauf, S. N. (1996): "A theory of persistent income inequality," *Journal of Economic growth*, 1(1), 75–93.

Ellis, A., and M. Piccione (2017): "Correlation misperception in choice," *American Economic Review*, 107(4), 1264–92.

Enke, B. (2020): "What you see is all there is," *The Quarterly Journal of Economics*, 135(3), 1363–1398.

Esponda, I. (2008): "Behavioral equilibrium in economies with adverse selection," *The American Economic Review*, 98(4), 1269–1291.

——— (2013): "Rationalizable conjectural equilibrium: A framework for robust predictions," *Theoretical Economics*, 8(2), 467–501.

Esponda, I., and D. Pouzo (2016): "Berk–Nash Equilibrium: A Framework for Modeling Agents With Misspecified Models," *Econometrica*, 84(3), 1093–1130.

——— (2017): "Conditional Retrospective Voting in Large Elections," *American Economic Journal: Microeconomics*, 9(2), 54–75.

Esponda, I., and E. Vespa (2014): "Hypothetical thinking and information extraction in the laboratory," *American Economic Journal: Microeconomics*, 6(4), 180–202.

——— (2018): "Endogenous sample selection: A laboratory study," *Quantitative Economics*, 9(1), 183–216.

Ettinger, D., and P. Jehiel (2010): "A theory of deception," *American Economic Journal: Microeconomics*, 2(1), 1–20.

Eyster, E., and M. Rabin (2005): "Cursed equilibrium," *Econometrica*, 73(5), 1623–1672.

Fernández, R. (2003): "Sorting, Education, and Inequality," in *Advances in Economics and Econometrics*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, vol. II. Cambridge University Press.

Fernández, R., and R. Rogerson (1996): "Income distribution, communities, and the quality of public education," *The Quarterly Journal of Economics*, 111(1), 135–164.

——— (2001): "Sorting and long-run inequality," *The Quarterly Journal of Economics*, 116(4), 1305–1341.

Fiedler, K., and P. Juslin (2006): *Information sampling and adaptive cognition*. Cambridge University Press.

Frick, M., R. Iijima, and Y. Ishii (2019): "Dispersed Behavior and Perceptions in Assortative Societies," Cowles Foundation Discussion Paper # 2128R.

——— (2020): "Misinterpreting others and the fragility of social learning," *Econometrica*, 88(6), 2281–2328.

Fudenberg, D., and Y. Kamada (2015): "Rationalizable partition-confirmed equilibrium," *Theoretical Economics*, 10(3), 775–806.

Fudenberg, D., and D. K. Levine (1993): "Self-confirming equilibrium," *Econometrica*, pp. 523–545.

Gagnon-Bartsch, T. (2017): "Taste Projection in Models of Social Learning," Discussion paper, working paper.

Gagnon-Bartsch, T., M. Pagnozzi, and A. Rosato (2021): "Projection of Private Values in Auctions," *American Economic Review*, 111(10), 3256–98.

Gimpelson, V., and D. Treisman (2018): "Misperceiving inequality," *Economics & Politics*, 30(1), 27–54.

Glaeser, E. L., and C. R. Sunstein (2009): "Extremism and social learning," *Journal of Legal Analysis*, 1(1), 263–324.

Golub, B., and S. Morris (2018): "Expectations, Networks, and Conventions," working paper.

Gonçalves, D. (2020): "Sequential sampling and equilibrium," Discussion paper.

Hauser, O. P., and M. I. Norton (2017): "(Mis) perceptions of inequality," *Current opinion in psychology*, 18, 21–25.

Heider, F. (1958): *The Psychology of Interpersonal Relations*. New York: Wiley.

Hvidberg, K. B., C. Kreiner, and S. Stantcheva (2021): "Social position and fairness views," Discussion paper, National Bureau of Economic Research.

Jackson, M. O. (2019): "The friendship paradox and systematic biases in perceptions and social norms," *Journal of Political Economy*, 127(2), 777–818.

Jackson, M. O., and Y. Zenou (2013): "Games on Networks," in *Handbook of Game Theory*, ed. by P. H. Young, and S. Zamir, vol. 4. Elsevier Science.

Jargowsky, P. A. (1996): "Take the money and run: Economic segregation in US metropolitan areas," *American sociological review*, pp. 984–998.

Jehiel, P. (2005): "Analogy-based expectation equilibrium," *Journal of Economic Theory*, 123(2), 81–104.

——— (2018): "Investment strategy and selection bias: An equilibrium perspective on overoptimism," *American Economic Review*, 108(6), 1582–97.

Jehiel, P., and F. Koessler (2008): "Revisiting games of incomplete information with analogy-based expectations," *Games and Economic Behavior*, 62(2), 533–557.

Jensen, M. K. (2018): "Distributional comparative statics," *The Review of Economic Studies*, 85(1), 581–610.

Jensen, R. (2010): "The (perceived) returns to education and the demand for schooling," *The Quarterly Journal of Economics*, 125(2), 515–548.

Juslin, P., A. Winman, and P. Hansson (2007): "The naïve intuitive statistician: A naïve sampling model of intuitive confidence intervals.," *Psychological review*, 114(3), 678.

Kahneman, D. (2011): *Thinking, fast and slow*. Macmillan.

Kaneko, M., and A. Matsui (1999): "Inductive game theory: discrimination and prejudices," *Journal of Public Economic Theory*, 1(1), 101–137.

Kelley, H. H., and J. L. Michela (1980): "Attribution theory and research," *Annual review of psychology*, 31(1), 457–501.

Levy, G., and R. Razin (2015): "Correlation neglect, voting behavior, and information aggregation," *American Economic Review*, 105(4), 1634–45.

——— (2017): "The Coevolution of Segregation, Polarized Beliefs, and Discrimination: The Case of Private versus State Education," *American Economic Journal: Microeconomics*, 9(4), 141–70.

Lipnowski, E., and E. Sadler (2019): "Peer-Confirming Equilibrium," *Econometrica*, 87(2), 567–591.

Madarász, K. (2012): "Information projection: Model and applications," *The Review of Economic Studies*, 79(3), 961–985.

——— (2016): "Projection equilibrium: Definition and applications to social investment and persuasion," .

Marks, G., and N. Miller (1987): "Ten years of research on the false-consensus effect: An empirical and theoretical review.," *Psychological bulletin*, 102(1), 72.

Mekonnen, T., and R. Leal Vizcaíno (2021): "Bayesian Comparative Statics," *Theoretical Economics*.

Meltzer, A. H., and S. F. Richard (1981): "A rational theory of the size of government," *Journal of Political Economy*, 89(5), 914–927.

Morris, S., and H. S. Shin (2002): "Social value of public information," *American Economic Review*, 92(5), 1521–1534.

Musterd, S. (2005): "Social and ethnic segregation in Europe: levels, causes, and effects," *Journal of Urban Affairs*, 27(3), 331–348.

Nguyen, T. (2008): "Information, role models and perceived returns to education: Experimental evidence from Madagascar," Working paper.

Niehues, J. (2014): "Subjective perceptions of inequality and redistributive preferences: An international comparison," working paper.

Ortoleva, P., and E. Snowberg (2015): "Overconfidence in political behavior," *American Economic Review*, 105(2), 504–35.

Osborne, M. J., and A. Rubinstein (1998): "Games with procedurally rational players," *American Economic Review*, pp. 834–847.

——— (2003): "Sampling equilibrium, with an application to strategic voting," *Games and Economic Behavior*, 45(2), 434–441.

Pin, P., and B. Rogers (2016): "Stochastic Network Formation and Homophily," in *The Oxford Handbook of the Economics of Networks*.

Quillian, L., and H. Lagrange (2016): "Socioeconomic segregation in large cities in France and the United States," *Demography*, 53(4), 1051–1084.

Reardon, S. F., and K. Bischoff (2011): "Income inequality and income segregation," *American Journal of Sociology*, 116(4), 1092–1153.

Ross, L. (1977): "The Intuitive Psychologist And His Shortcomings: Distortions in the Attribution Process1," in *Advances in experimental social psychology*, vol. 10, pp. 173–220. Elsevier.

Ross, L., D. Greene, and P. House (1977): "The "false consensus effect:" An egocentric bias in social perception and attribution processes," *Journal of experimental social psychology*, 13(3), 279–301.

Rubinstein, A., and A. Wolinsky (1994): "Rationalizable conjectural equilibrium: between Nash and rationalizability," *Games and Economic Behavior*, 6(2), 299–311.

Ryff, J. (1963): "On the representation of doubly stochastic operators," *Pacific Journal of Mathematics*, 13(4), 1379–1386.

Salant, Y., and J. Cherry (2020): "Statistical inference in games," *Econometrica*, 88(4), 1725–1752.

Samet, D. (1998): "Iterated expectations and common priors," *Games and economic Behavior*, 24(1-2), 131–141.

Shaked, M., and G. Shanthikumar (2007): *Stochastic orders*. Springer.

Spiegler, R. (2016): "Bayesian networks and boundedly rational expectations," *The Quarterly Journal of Economics*, 131(3), 1243–1290.

——— (2017): ""Data Monkeys": A Procedural Model of Extrapolation from Partial Statistics," *The Review of Economic Studies*, 84(4), 1818–1841.

Streufert, P. (2000): "The effect of underclass social isolation on schooling choice," *Journal of Public Economic Theory*, 2(4), 461–482.

Vives, X. (2005): "Complementarities and games: New developments," *Journal of Economic Literature*, 43(2), 437–479.

Westfall, J., L. Van Boven, J. R. Chambers, and C. M. Judd (2015): "Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide," *Perspectives on Psychological Science*, 10(2), 145–158.

Windsteiger, L. (2018): "The Redistributive Consequences of Segregation and Misperceptions," working paper.

Online Appendix to "Dispersed Behavior and Perceptions in Assortative Societies"

Mira Frick, Ryota Iijima, and Yuhta Ishii

D Omitted Proofs

D.1 Proofs for Appendix A

D.1.1 Proof of Lemma A.1

For the first point, note that for any $f \in L^1$,

$$\|T_C f\| = \int_0^1 |T_C f(x)| dx \leq \int_0^1 \int_0^1 c(x', x)|f(x')| dx' dx = \int_0^1 |f(x')| dx' = \|f\| < \infty.$$

Thus, $T_C : L^1 \to L^1$. Moreover, since T_C is clearly linear, the above ensures that it is also continuous.

For the second point, consider $f \in \mathcal{I}$. Since C is assortative, $T_C f(x) \geq T_C f(x')$ for all $x \geq x'$, so that $T_C f$ is weakly increasing. To show that $T_C f$ is absolutely continuous, note that for each $x, x' \in (0, 1)$,

$$\begin{aligned} T_C f(x) &= \int_0^1 c(y, x) f(y) dy = \int_0^1 \left(\int_{x'}^x c_2(y, z) dz + c(y, x') \right) f(y) dy \\ &= \int_{x'}^x \int_0^1 c_2(y, z) f(y) dy dz + T_C f(x'), \end{aligned}$$

where c_2 denotes the partial derivative of c with respect to the second argument, which exists almost everywhere by the absolute continuity assumption on c. Thus $T_C f$ is absolutely continuous with $(T_C f)'(z) = \int_0^1 c_2(y, z) f(y) dy$ for each z.

Finally, for the third point, fix any $f \in L^1$ and $\gamma \in (-1, 1)$. Then for any $\tau > \tau'$,

$$\| \sum_{t=0}^{\tau} \gamma^t (T_C)^t f - \sum_{t=0}^{\tau'} \gamma^t (T_C)^t f \| \leq \sum_{t=\tau'+1}^{\tau} |\gamma|^t \|(T_C)^t f\| \leq \sum_{t=\tau'+1}^{\tau} |\gamma|^t \|f\| \leq \frac{|\gamma|^{\tau'+1}}{1-\gamma} \|f\|,$$

which vanishes as $\tau' \to \infty$. Thus, the sequence is Cauchy. Since the space L^1 is complete, this yields the desired result. □

D.1.2 Proof of Lemma A.3

$\succsim_m$-order: It is clear from the definition that $\succsim_m$ is reflexive and transitive; moreover, by Lemma A.2, $\succsim_m$ is linear. To check that $\succsim_m$ is continuous, take sequences $f_n \to f, g_n \to g$ in

$\mathcal{I}$ such that $f_n \succsim_m g_n$ for each n. For any $y \in (0,1)$, we have

$$|\int_y^1 f(x)dx - \int_y^1 f_n(x)dx| \leq \int_y^1 |f(x) - f_n(x)|dx \leq \|f - f_n\| \to 0$$

and likewise $|\int_y^1 g(x)dx - \int_y^1 g_n(x)dx| \to 0$. Since $\int_y^1 f_n(x)dx \geq \int_y^1 g_n(x)dx$ and $\int_0^1 f_n(x)dx = \int_0^1 g_n(x)dx$ for each n, this implies $\int_y^1 f(x)dx \geq \int_y^1 g(x)dx$ and $\int_0^1 f(x)dx = \int_0^1 g(x)dx$. Thus, $f \succsim_m g$ by Lemma A.2.

To verify that $\succsim_m$ is isotone, take any $f, g \in \mathcal{I}$ such that $f \succsim_m g$ and set $h := f - g$. Note that $\int_0^1 h(x)dx = \int_0^1 T_C h(x)dx = 0$. It suffices to show that $\int_y^1 T_C h(x)dx \geq 0$ for all $y \in (0,1)$. To see this, note that $\int_y^1 T_C h(x)dx$ is given by

$$\begin{aligned}
\int_y^1 \int_0^1 h(z)c(z|x)dzdx &= \int_0^1 \int_y^1 c(z|x)dx h(z)dz = \int_0^1 (1 - C(y|z))h(z)dz \\
&= -\int_0^1 \frac{\partial 1 - C(y|z)}{\partial z} \int_0^z h(z')dz'dz + \left[(1 - C(y|z)) \int_0^z h(z')dz'\right]_0^1 \\
&= \int_0^1 \frac{\partial C(y|z)}{\partial z} \int_0^z h(z')dz'dz \geq 0,
\end{aligned}$$

where the second equality uses $\int_y^1 c(z|x)dx = \int_y^1 c(x|z)dx = 1 - C(y|z)$, the third holds by integration by parts (using absolute continuity of c), the fourth uses $\int_0^1 h(z)dz = 0$, and the final inequality uses $\int_0^z h(z')dz' \leq 0$ (by $f \succsim_m g$) and assortativity of C.

$\succsim_d$-order: It is clear from the definition that $\succsim_d$ is reflexive, transitive, and linear. To check that it is continuous, take sequences $f_n \to f$ and $g_n \to g$ in $\mathcal{I}$ such that $f_n \succsim_d g_n$ for each n. By standard results (e.g., Theorem 13.6 in Aliprantis and Border (2006)), we can find subsequences $(f_{n_k})_{k\in\mathbb{N}}, (g_{n_k})_{k\in\mathbb{N}}$ such that $f_{n_k}(x) \to f(x)$ and $g_{n_k}(x) \to g(x)$ for almost all $x \in (0,1)$. This implies $f(x) - f(x') \geq g(x) - g(x')$ for almost all $x \geq x'$, which ensures $f \succsim_d g$ since f and g are continuous.

To show that $\succsim_d$ is isotone, first consider any bounded $f, g \in \mathcal{I}$ such that $f \succsim_d g$. Since f and g are absolutely continuous, there exist integrable functions $f', g' : (0,1) \to \mathbb{R}$ such that $f(x) = f(0) + \int_0^x f'(y)\,dy$ and $g(x) = g(0) + \int_0^x g'(y)\,dy$ for all $x \in (0,1)$. Then, for any $x \geq x'$ and $C \in \mathcal{C}$, integration by parts yields

$$\begin{aligned}
T_C f(x) - T_C f(x') &= \int_0^1 f(y)(c(y|x) - c(y|x'))dy \\
&= -\int_0^1 f'(y)(C(y|x) - C(y|x'))dy + [f(y)(C(y|x) - C(y|x'))]_0^1 \\
&= -\int_0^1 f'(y)(C(y|x) - C(y|x'))dy \geq -\int_0^1 g'(y)(C(y|x) - C(y|x'))dy \\
&= -\int_0^1 g'(y)(C(y|x) - C(y|x'))dy + [g(y)(C(y|x) - C(y|x'))]_0^1 \\
&= \int_0^1 g(y)(c(y|x) - c(y|x'))dy = T_C g(x) - T_C g(x').
\end{aligned}$$

Here, the inequality holds because the fact that $f \succsim_d g$ and $f, g \in \mathcal{I}$ implies $f'(y) \geq g'(y) \geq 0$ for almost all $y \in (0,1)$.

Next, consider arbitrary $f, g \in \mathcal{I}$ such that $f \succsim_d g$. By defining bounded functions

$$f_n(x) = \begin{cases} f(\frac{1}{n}) \text{ if } x \in (0, \frac{1}{n}) \\ f(x) \text{ if } x \in [\frac{1}{n}, \frac{n-1}{n}] \\ f(\frac{n-1}{n}) \text{ if } x \in (\frac{n-1}{n}, 1) \end{cases} \qquad g_n(x) = \begin{cases} g(\frac{1}{n}) \text{ if } x \in (0, \frac{1}{n}) \\ g(x) \text{ if } x \in [\frac{1}{n}, \frac{n-1}{n}] \\ g(\frac{n-1}{n}) \text{ if } x \in (\frac{n-1}{n}, 1) \end{cases} \tag{20}$$

for each $n \in \mathbb{N}$, we obtain $f_n \succsim_d g_n$ for each n and $f_n \to f, g_n \to g$. For any $C \in \mathcal{C}$, since T_C is a continuous operator, this implies $T_C f_n \to T_C f$ and $T_C g_n \to T_C g$. Thus, $T_C f \succsim_d T_C g$ by continuity of $\succsim_d$, as we already know that $T_C f_n \succsim_d T_C g_n$ from the previous part of the proof. □

D.1.3 Proof of Lemma A.4

The base case $t = 0$ holds because of the following result by Ryff (1963): Call a linear operator $T : L^1 \to L^1$ an ***𝔖-operator*** if $f \succsim_m Tf$ for all $F \in \mathcal{I}$. The representation theorem in Ryff (1963) implies that T is an $\mathfrak{S}$-operator if there exists some measurable function $K : [0,1]^2 \to \mathbb{R}$ such that $Tf(x) = \frac{d}{dx}\int_0^1 K(x,y)f(y)dy$ for all $f \in L^1$ and almost every x and the following conditions are met: (1) $K(0,y) = 0$ for all $0 \leq y \leq 1$; (2) $\operatorname{essup}_y V(K(\cdot, y)) < \infty$, where $V(\cdot)$ denotes the total variation and essup the essential supremum; (3) $\int_0^1 K(x,y)f(y)dy$ is absolutely continuous in x for all $f \in L^1$; (4) $x = \int_0^1 K(x,y)dy$; (5) $x_1 < x_2 \Longrightarrow K(x_1, \cdot) \leq K(x_2, \cdot)$; and (6) $K(1,y) = 1$ for all $y \in [0,1]$.

Since $C \in \mathcal{C}$, it is easy to see that T_C satisfies these conditions with $K(x,y) := C(x \mid y)$ for all x, y, so that T_C is an $\mathfrak{S}$-operator. Thus, $f \succsim_m T_C f$, proving the base case. The inductive step then follows from isotonicity of $\succsim_m$ (Lemma A.3). □

D.2 Proofs for Appendix C

D.2.1 Proof of Proposition C.1

Let $\mu := \mathbb{E}_F[\theta]$. Consider strategy profiles g_a^α and g_c^α of assortativity neglect and correct agents expressed as functions of quantiles. Write $g^\alpha := \alpha g_a^\alpha + (1-\alpha) g_c^\alpha$. In an α-ANE, we must have

$$g_a^\alpha(x) = F^{-1}(x) + (\beta + \gamma) T_C g^\alpha(x), \qquad g_c^\alpha(x) = F^{-1}(x) + \gamma T_C g^\alpha(x) + \beta \int_0^1 g^\alpha(y)dy$$

for each $x \in (0,1)$. Since $g^\alpha = \alpha g_a^\alpha + (1-\alpha) g_c^\alpha$, it follows that

$$g^\alpha(x) = F^{-1}(x) + (\gamma + \alpha\beta) T_C g^\alpha(x) + (1-\alpha)\beta \int_0^1 g^\alpha(y)dy$$

for each x, which implies $\int_0^1 g^\alpha(y)dy = \frac{\mu}{1-\beta-\gamma}$ by integrating both sides over x. Moreover, iterating the above equation we obtain

$$g^\alpha(x) = \sum_{t\geq 0}(\gamma+\alpha\beta)^t(T_C)^t F^{-1}(x) + \frac{(1-\alpha)\beta\mu}{(1-\gamma-\alpha\beta)(1-\beta-\gamma)},$$

where the convergence of the RHS can be shown as in the proof of Lemma 1. Note that this uniquely determines g^α for any α. By the best-response conditions, we obtain

$$\begin{aligned}
g_a^\alpha(x) &= F^{-1}(x) + (\beta+\gamma)T_C g^\alpha(x) \\
&= F^{-1}(x) + (\beta+\gamma)\sum_{t\geq 1}(\gamma+\alpha\beta)^{t-1}(T_C)^t F^{-1}(x) + \frac{(\beta+\gamma)(1-\alpha)\beta\mu}{(1-\gamma-\alpha\beta)(1-\beta-\gamma)}, \\
g_c^\alpha(x) &= F^{-1}(x) + \gamma T_C g^\alpha(x) + \beta\int_0^1 g^\alpha(y)dy \\
&= F^{-1}(x) + \gamma\sum_{t\geq 1}(\gamma+\alpha\beta)^{t-1}(T_C)^t F^{-1}(x) + \frac{(1-\alpha(\beta+\gamma))\,\beta\mu}{(1-\gamma-\alpha\beta)(1-\beta-\gamma)}
\end{aligned}$$

for each x, yielding (18)-(19). Then the claim $g_a^\alpha \succsim_d g_c^\alpha$ and the comparative statics with respect to α can be verified using linearity and continuity of $\succsim_d$. □

D.2.2 Proof of Proposition C.2

Write $P = (F, C)$. Consider any PANE s with $\int a\, d\hat{G}_\theta(a)$ absolutely continuous in θ. Then $s(\theta) = \theta + \beta\int a\, d\hat{G}_\theta(a) + \gamma\mathbb{E}_P[s(\theta')|\theta]$ for each θ. Thus, s is the Nash equilibrium in environment $(\tilde{F}, C, \tilde{\beta}, \gamma)$, where $\tilde{\beta} = 0$ and $\tilde{F}^{-1}(x) = F^{-1}(x) + \beta\int a\, d\hat{G}_{F^{-1}(x)}(a)$ for each x (note that $\tilde{F} \in \mathcal{F}$, as $\int a d\hat{G}_\theta(a)$ is increasing and absolutely continuous in θ). Since $\tilde{F}$ is more dispersive than F (and the global complementarity parameter does not affect Nash action dispersion by Proposition 4), Proposition 3 implies that $G^{s,P}$ is more dispersive than the Nash global action distribution in environment (P, β, γ). □

D.2.3 Details for Example C.1

Fix any $\hat{\rho} \in [0, \rho]$. We verify that, for the expressions in Example C.1, s^* is a PANE and $(\hat{P}_\theta, \hat{s}_\theta)$ are associated coherent perceptions. Let $x := \frac{1}{1-\gamma\rho-\beta\frac{\rho-\hat{\rho}}{1-\hat{\rho}}}$ and $\hat{x} := \frac{1}{1-\gamma\hat{\rho}}$, so that $s^*(\theta) = x(\theta-\mu) + \frac{\mu}{1-\beta-\gamma}$ and $\hat{s}_\theta(\theta') = \hat{x}(\theta'-\hat{\mu}_\theta) + \frac{\hat{\mu}_\theta}{1-\beta-\gamma}$ for all θ, θ'. Since $P(\cdot|\theta)$ is distributed $\mathcal{N}(\rho\theta + (1-\rho)\mu, (1-\rho^2)\sigma^2)$, θ's true local action distribution $L_\theta^{s^*,P}$ is distributed $\mathcal{N}(x\rho(\theta-\mu)+\frac{\mu}{1-\beta-\gamma}, x^2(1-\rho^2)\sigma^2)$. Since $\hat{P}_\theta(\cdot|\theta)$ is distributed $\mathcal{N}(\hat{\rho}\theta+(1-\hat{\rho})\hat{\mu}_\theta, (1-\hat{\rho}^2)\hat{\sigma}^2)$, θ's perceived local action distribution $L_\theta^{\hat{s}_\theta,\hat{P}_\theta}$ is distributed $\mathcal{N}(\hat{x}\hat{\rho}(\theta-\hat{\mu}_\theta)+\frac{\hat{\mu}_\theta}{1-\beta-\gamma}, \hat{x}^2(1-\hat{\rho}^2)\hat{\sigma}^2)$. Thus, condition 1(a) of coherency can be verified by observing that, by construction, the mean and variance of $L_\theta^{s^*,P}$ and $L_\theta^{\hat{s}_\theta,\hat{P}_\theta}$ are equal.

To verify condition 2, note that, by construction, θ's perceived strategy profile $\hat{s}_\theta$ is the Nash equilibrium in society $\hat{P}_\theta = (\hat{\mu}_\theta, \hat{\sigma}^2, \hat{\rho})$ (see Example 1).

Finally, we verify that s^* is a PANE with perceived global action distributions $\hat{G}_\theta = \hat{G}^{\hat{s}_\theta,\hat{P}_\theta}$, as required by condition 1(b). Note first that, by construction, $s^*(\theta) = \hat{s}_\theta(\theta)$ for all θ. Thus,

conditions 1(a) and 2 imply that $s^*(\theta) \in \mathrm{BR}_\theta(\hat{G}^{\hat{s}_\theta,\hat{P}_\theta}, L_\theta^{s^*,P})$. It remains to check that $\hat{G}^{\hat{s}_\theta,\hat{P}_\theta}$ is FOSD-increasing in θ. This holds because $\hat{G}^{\hat{s}_\theta,\hat{P}_\theta}$ is distributed $\mathcal{N}(\frac{\hat{\mu}_\theta}{1-\beta-\gamma}, \hat{x}^2\hat{\sigma}^2)$ and because $\hat{\rho} \leq \rho$ ensures that $\hat{\mu}_\theta$ is increasing in θ.

D.2.4 Proof of Proposition C.3

We only consider Nash equilibrium, as ANE at (P, β, γ) corresponds to Nash equilibrium at $(P, 0, \beta + \gamma)$. Let $\mu := \mathbb{E}_F[\theta]$ and, for each $x \in (0, 1)$, define

$$h(x) := \sum_{t\geq 0} \gamma^t (T_C)^t F^{-1}(x) + \frac{\beta\mu}{(1-\gamma)(1-\beta-\gamma)},$$

which is a well-defined function in L^1 as $|\gamma| < 1$. Following the same argument as in the proof of Lemma 1, the strategy profile defined by $s^{NE}(\theta) = h(F^{-1}(\theta))$ for each θ is the unique Nash equilibrium and satisfies (5).

To show the "moreover" part, note that

$$h = \sum_{t\geq 0} \gamma^{2t} T_C^{2t}(F^{-1} + \gamma T_C F^{-1}) + \frac{\beta\mu}{(1-\gamma)(1-\beta-\gamma)}.$$

Since $\gamma > -1$, the additional assumption on P implies that $F^{-1} + \gamma T_C F^{-1}$ is strictly increasing. Therefore, h, and hence s^{NE}, is strictly increasing. □

D.2.5 Proof of Proposition C.4

We first show that, analogously to the relationship between $\succsim_{MA}$ and $\succsim_m$ (Lemma B.1), the strongly more-assortative order $\succsim_{SMA}$ is the "dual order" of the dispersiveness order $\succsim_d$:

Lemma D.1. *Fix any $C_1, C_2 \in \mathcal{C}$. Then $C_1 \succsim_{SMA} C_2$ if and only if $T_{C_1} f \succsim_d T_{C_2} f$ for all $f \in \mathcal{I}$.*

Proof. For the "only if" part, suppose that $C_1 \succsim_{SMA} C_2$. First consider any bounded $f \in \mathcal{I}$. Then there exists an integrable function $f' : (0, 1) \to \mathbb{R}$ that is nonnegative almost everywhere such that $f(x) = f(0) + \int_0^x f'(y)dy$ for all $x \in (0, 1)$. Thus, for any $x \geq x'$, integration by parts yields

$$\begin{aligned}
T_{C_1} f(x) - T_{C_1} f(x') &= \int_0^1 f(y)(c_1(y|x) - c_1(y|x'))dy \\
&= -\int_0^1 f'(y)(C_1(y|x) - C_1(y|x'))dy + [f(y)(C_1(y|x) - C_1(y|x'))]_0^1 \\
&= -\int_0^1 f'(y)(C_1(y|x) - C_1(y|x'))dy \geq -\int_0^1 f'(y)(C_2(y|x) - C_2(y|x'))dy \\
&= -\int_0^1 f'(y)(C_2(y|x) - C_2(y|x'))dy + [f(y)(C_2(y|x) - C_2(y|x'))]_0^1 \\
&= \int_0^1 f(y)(c_2(y|x) - c_2(y|x'))dy = T_{C_2} f(x) - T_{C_2} f(x'),
\end{aligned}$$

where the inequality holds because $f'(y) \geq 0$ for almost all y. Hence, $T_{C_1} f \succsim_d T_{C_2} f$.

Next take an arbitrary $f \in \mathcal{I}$. Define the sequence of bounded functions (f_n) as in (20), so that $f_n \to f$. By the previous observation, we have $T_{C_1} f_n \succsim_d T_{C_2} f_n$ for each n. Since $T_{C_1} f_n \to T_{C_1} f$ and $T_{C_2} f_n \to T_{C_2} f$ by continuity of T_{C_1} and T_{C_2}, continuity of $\succsim_d$ then yields $T_{C_1} f \succsim_d T_{C_2} f$.

For the "if" part, we prove the contrapositive. Suppose that C_1 is not strongly more assortative than C_2. That is, there exist y and $x > x'$ such that

$$C_2(y|x) - C_2(y|x') < C_1(y|x) - C_1(y|x') \leq 0.$$

Since C_1 and C_2 admit densities, the above inequality holds throughout some interval $(y_1, y_2) \ni y$. Define $f \in \mathcal{I}$ by $f(z) = \int_0^z f'(y')dy'$ for all z, where f' is an integrable function given by $f'(y') = 1$ for $y' \in (y_1, y_2)$ and $f'(y') = 0$ for all $y' \notin (y_1, y_2)$. Using the same integration by parts argument as above, we obtain

$$\begin{aligned} T_{C_1} f(x) - T_{C_1} f(x') &= -\int f'(y)(C_1(y|x) - C_1(y|x'))dy \\ < -\int f'(y)(C_2(y|x) - C_2(y|x'))dy &= T_{C_2} f(x) - T_{C_2} f(x'). \end{aligned}$$

Thus, $T_{C_1} f \succsim_d T_{C_2} f$ fails. □

Proof of Proposition C.4. Note that Nash and ANE strategies are monotone by the assumption on P_i (Proposition C.3). We prove each part only for Nash, as the ANE at (P, β, γ) is the Nash at $(P, 0, \beta + \gamma)$. For each $f, g \in \mathcal{I}$, write $f \succsim_{dil} g$ iff $f \succsim_m g + \alpha$ for some constant function α. This order inherits linearity, isotonicity, and continuity from $\succsim_m$. Note that for $F, G \in \mathcal{F}$, F is a dilation of G iff $F^{-1} \succsim_{dil} G^{-1}$; moreover, the $\succsim_{dil}$ order is implied by the $\succsim_d$ order.

Second part: Let $\beta := \beta_1 = \beta_2$, $\gamma := \gamma_1 = \gamma_2$, $C := C_1 = C_2$. The proof of Proposition 3 carries over to the case $\gamma \geq 0$, so we focus on the case $\gamma < 0$. Since β only shifts the action mean without affecting the dilation order, we also assume $\beta = 0$ without loss.

For each $i = 1, 2$, define an operator $\Gamma_i : \mathcal{I} \to \mathcal{I}$ by $\Gamma_i f = F_i^{-1} + \gamma T_C F_i^{-1} + \gamma^2 T_C^2 f$ for each $f \in \mathcal{I}$. Note that $\Gamma_i(\cdot)$ is increasing, as $(1 + \gamma T_C) F_i^{-1}$ is increasing by the assumption on P_i. We make two preliminary observations:

1. For $i = 1, 2$, $\Gamma_i f \succsim_{dil} \Gamma_i g$ whenever $f \succsim_{dil} g$.

   This follows from isotonicity of $\succsim_{dil}$.

2. $\Gamma_1 f \succsim_{dil} \Gamma_2 f$ for each $f \in \mathcal{I}$.

   To see this, note that $F_1^{-1} \succsim_d F_2^{-1}$ implies $F_1^{-1} - F_2^{-1} \in \mathcal{I}$. Thus,

   $$F_1^{-1} - F_2^{-1} \succsim_m T_C(F_1^{-1} - F_2^{-1}) \succsim_{dil} -\gamma T_C(F_1^{-1} - F_2^{-1}),$$

   where the first comparison uses Lemma A.4 and the second uses $-1 < \gamma \leq 0$. Therefore, $F_1^{-1} + \gamma T_C F_1^{-1} \succsim_{dil} F_2^{-1} + \gamma T_C F_2^{-1}$, and thus $\Gamma_1 f \succsim_{dil} \Gamma_2 f$ for each $f \in \mathcal{I}$.

Now, fix any $f \in \mathcal{I}$. Let

$$g_i := \sum_{t \geq 0} \gamma^t T_C^t F_i = \lim_{t \to \infty} \Gamma_i^t(f).$$

This is the inverse cdf of G_i^{NE}, as s_i^{NE} is increasing. By induction, we show that $\Gamma_1^t f \succsim_{dil} \Gamma_2^t f$ for all t. The base case $t = 1$ holds by the second observation above. Moreover, if $\Gamma_1^{t-1} f \succsim_{dil} \Gamma_2^{t-1} f$, then

$$\Gamma_1^t f \succsim_{dil} \Gamma_2 \Gamma_1^{t-1} f \succsim_{dil} \Gamma_2^t f$$

holds by observations 1-2. Given this, $g_1 \succsim_{dil} g_2$ follows by continuity of $\succsim_{dil}$.

First part: Let $F := F_1 = F_2$, $\beta := \beta_1 = \beta_2$, $\gamma := \gamma_1 = \gamma_2$. The proof of Proposition 2 carries over to the case $\gamma \leq 0$, so we focus on the case $\gamma < 0$. Since β only shifts the action mean without affecting the dilation order, we also assume $\beta = 0$ without loss. Let $g_i := \sum_{t \geq 0} \gamma^t T_{C_i}^t F^{-1}$; this is the inverse cdf of G_i^{NE} since s_i^{NE} is monotone.

For each $i = 1, 2$ and any $f \in L^1$, the linearity of the operators $T_{C_i}^t$ implies

$$(\mathbf{1} - \gamma_i T_{C_i}) \sum_{t \geq 0} \gamma_i^t T_{C_i}^t f = \sum_{t \geq 0} (\gamma_i^t T_{C_i}^t)(\mathbf{1} - \gamma_i T_{C_i}) f = f, \tag{21}$$

where $\mathbf{1}$ denotes the identity operator. Observe that

$$g_2 = \sum_{t \geq 0} \gamma^t T_{C_2}^t F^{-1} = \sum_{t \geq 0} \gamma^t T_{C_2}^t (\mathbf{1} - \gamma T_{C_1}) g_1,$$

where the second equality uses (21) with $i = 1$ and $f = F^{-1}$. Likewise,

$$g_1 = \sum_{t \geq 0} \gamma^t T_{C_2}^t (\mathbf{1} - \gamma T_{C_2}) g_1,$$

by the second equality in (21) with $i = 2$ and $f = g_1$. This shows that g_1 and g_2 correspond to the inverse cdfs of the Nash action distributions in two modified environments that share a common interaction structure C_2 and complementarity parameters $(0, \gamma)$ and have type distributions $\tilde{F}_1$ and $\tilde{F}_2$ with inverse cdfs $\tilde{F}_1^{-1} := (\mathbf{1} - \gamma T_{C_2}) g_1$ and $\tilde{F}_2^{-1} := (\mathbf{1} - \gamma T_{C_1}) g_1$, respectively. Since $g_1 \in \mathcal{I}$, $\gamma < 0$, and $C_1 \succsim_{SMA} C_2$, Lemma D.1 implies $\tilde{F}_2^{-1} \succsim_d \tilde{F}_1^{-1}$.

Given this, the arguments in part 2 above imply that $g_2 \succsim_{dil} g_1$, provided we can show that $(\mathbf{1} + \gamma T_{C_2}) \tilde{F}_i^{-1}$ is increasing for $i = 1, 2$ (which ensures that the corresponding operators $\Gamma_i(\cdot)$ in the two modified societies are increasing). For $i = 2$, note that $(\mathbf{1} + \gamma T_{C_2}) \tilde{F}_2^{-1} := (\mathbf{1} + \gamma T_{C_2})(\mathbf{1} - \gamma T_{C_1}) g_1 = (\mathbf{1} + \gamma T_{C_2}) F^{-1}$ by (21), which is increasing by the assumption on P_2 and since $\gamma > -1$. For $i = 1$, note that (i) $(\mathbf{1} - \gamma^2 T_{C_1}^2) g_1 = (\mathbf{1} + \gamma T_{C_1}) F^{-1}$ is increasing (by the assumption on P_1 and since $\gamma > -1$), and (ii) $\gamma^2 T_{C_1}^2 g_1 \succsim_d \gamma^2 T_{C_2}^2 g_1$ since $C_1 \succsim_{SMA} C_2$ (Lemma D.1). Combining (i) and (ii) yields that $(\mathbf{1} + \gamma T_{C_2}) \tilde{F}_1^{-1} := (\mathbf{1} - \gamma^2 T_{C_2}^2) g_1$ is increasing, as required.

Third part: Let $F := F_1 = F_2$, $C := C_1 = C_2$. The proof of Proposition 4 carries over to the case $\gamma_i \geq 0$ for $i = 1, 2$. Thus, by the transitivity of the dilation order, we can focus on the case $\gamma_i \leq 0$ for $i = 1, 2$. Since β only shifts the action mean without affecting the dilation order, we also assume $\beta_1 = \beta_2 = 0$ without loss. Let $g_i := \sum_{t \geq 0} \gamma_i^t T_C^t F^{-1}$; this is the inverse cdf of G_i^{NE} since s_i^{NE} is monotone. Observe that

$$g_1 = \sum_{t \geq 0} \gamma_1^t T_C^t F^{-1} = \sum_{t \geq 0} \gamma_1^t T_C^t (\mathbf{1} - \gamma_2 T_C) g_2,$$

where the second equality uses (21) with $i = 1$ and $f = F^{-1}$. Likewise,

$$g_2 = \sum_{t \geq 0} \gamma_1^t T_C^t (\mathbf{1} - \gamma_1 T_C) g_2,$$

by the second equality in (21) with $i = 1$ and $f = g_2$. This shows that g_1 and g_2 can be seen as inverse cdfs of Nash action distributions in two modified environments that share a common interaction structure C and complementarity parameters $(0, \gamma_1)$ and have type distributions $\tilde{F}_1$ and $\tilde{F}_2$ with inverse cdfs $\tilde{F}_1^{-1} := (\mathbf{1} - \gamma_2 T_C) g_2$ and $\tilde{F}_2^{-1} := (\mathbf{1} - \gamma_1 T_C) g_2$, respectively. Since $0 \geq \gamma_1 \geq \gamma_2$, we have $\tilde{F}_1^{-1} \succsim_d \tilde{F}_2^{-1}$.

Given this, the arguments in part 2 above imply that $g_1 \succsim_{dil} g_2$, provided we can show that $(\mathbf{1} + \gamma_1 T_C)\tilde{F}_i^{-1}$ is increasing for $i = 1, 2$ (which ensures that the corresponding operators $\Gamma_i(\cdot)$ in the two modified societies are increasing). For $i = 1$, note that $(\mathbf{1} + \gamma_1 T_C)\tilde{F}_2^{-1} := (\mathbf{1} + \gamma_1 T_C)(\mathbf{1} - \gamma_2 T_C) g_2 = (\mathbf{1} + \gamma_1 T_C) F^{-1}$, which is increasing by the assumption on P_i and $\gamma_1 > -1$. For $i = 2$, note that (i) $(\mathbf{1} - \gamma_2^2 T_C^2) g_2 = (\mathbf{1} + \gamma_2 T_C) F^{-1}$ is increasing (by the assumption on P_i and since $\gamma_2 > -1$), and (ii) $\gamma_2^2 T_C^2 g_2 \succsim_d \gamma_1^2 T_C^2 g_2$ as $0 \geq \gamma_1 \geq \gamma_2$. Combining (i) and (ii) yields that $(\mathbf{1} + \gamma_1 T_C)\tilde{F}_1^{-1} := (\mathbf{1} - \gamma_1^2 T_C^2) g_2$ is increasing, as required. □

D.2.6 Proof of Proposition C.5

Fix any ANE $s^{AN} =: s$ and θ. For each θ', set $\hat{s}_\theta(\theta') := \mathrm{BR}_{\theta'}(L_\theta^{s,P}, L_\theta^{s,P})$ and $\hat{F}_\theta(\theta') := L_\theta^{s,P}(\hat{s}_\theta(\theta'))$, and let $\hat{P}_\theta := \hat{F}_\theta \times \hat{F}_\theta$. To verify observational consistency, note that $L_\theta^{\hat{s}_\theta, \hat{P}_\theta}(a) = \hat{F}_\theta(\hat{s}_\theta^{-1}(a)) = L_\theta^{s,P}(a)$ for each a, where the first equality uses $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$ and the inverse $\hat{s}_\theta^{-1}$ is well-defined and increasing by the surjectivity and monotonicity assumption on best-responses. To verify the perceived best-response condition, note that, for each θ',

$$\hat{s}_\theta(\theta') = \mathrm{BR}_{\theta'}(L_\theta^{s,P}, L_\theta^{s,P}) = \mathrm{BR}_{\theta'}(L_\theta^{\hat{s}_\theta, \hat{P}_\theta}, L_\theta^{\hat{s}_\theta, \hat{P}_\theta}) = \mathrm{BR}_{\theta'}(G^{\hat{s}_\theta, \hat{P}_\theta}, L_{\theta'}^{\hat{s}_\theta, \hat{P}_\theta}),$$

where the second equality uses observational consistency and the third uses non-assortativity of $\hat{P}_\theta$. Thus, $(\hat{P}_\theta, \hat{s}_\theta)$ is a coherent assortativity neglect perception for type θ.

To show uniqueness, consider any coherent assortativity neglect perception $(\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta, \hat{s}_\theta)$ for θ. Then, for each θ', the perceived best-response condition, non-assortativity of $\hat{P}_\theta$, and observational consistency imply $\hat{s}_\theta(\theta') = \mathrm{BR}_{\theta'}(G^{\hat{s}_\theta, \hat{P}_\theta}, L_{\theta'}^{\hat{s}_\theta, \hat{P}_\theta}) = \mathrm{BR}_{\theta'}(L_\theta^{\hat{s}_\theta, \hat{P}_\theta}, L_\theta^{\hat{s}_\theta, \hat{P}_\theta}) = \mathrm{BR}_{\theta'}(L_\theta^{s,P}, L_\theta^{s,P})$. Moreover, $\hat{P}_\theta = \hat{F}_\theta \times \hat{F}_\theta$ and observational consistency imply $\hat{F}_\theta(\hat{s}_\theta^{-1}(a)) = L_\theta^{\hat{s}_\theta, \hat{P}_\theta}(a) = L_\theta^{s,P}(a)$ for each a, which yields $\hat{F}_\theta(\theta') = L_\theta^{s,P}(\hat{s}_\theta(\theta'))$ for each θ'. Thus, $(\hat{P}_\theta, \hat{s}_\theta)$ coincides with the perceptions in the first paragraph. □

D.2.7 Proof of Proposition C.6

Consider any monotone ANE s^{AN} and any Nash equilibrium s^{NE}. For any types $\theta > \theta'$, the fact that ψ and ϕ are monotone yields

$$s^{AN}(\theta) - s^{AN}(\theta') = \phi(\theta) - \phi(\theta') + \psi(L_\theta^{s^{AN},P}) - \psi(L_{\theta'}^{s^{AN},P}) \geq \phi(\theta) - \phi(\theta') = s^{NE}(\theta) - s^{NE}(\theta') > 0,$$

where the first inequality holds because $L_\theta^{s^{AN},P}$ FOSD-dominates $L_{\theta'}^{s^{AN},P}$ (by monotonicity of s^{AN} and assortativity of P). Thus, $G^{s^{AN},P}$ is more dispersive than $G^{s^{NE},P}$. □