THE CENTRE FOR MARKET AND PUBLIC ORGANISATION

The Centre for Market and Public Organisation (CMPO) is a leading research centre, combining expertise in economics, geography and law. Our objective is to study the intersection between the public and private sectors of the economy, and in particular to understand the right way to organise and deliver public services. The Centre aims to develop research, contribute to the public debate and inform policy-making.

CMPO, now an ESRC Research Centre was established in 1998 with two large grants from The Leverhulme Trust. In 2004 we were awarded ESRC Research Centre status, and CMPO now combines core funding from both the ESRC and the Trust.

Centre for Market and Public Organisation
Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX
http://www.bris.ac.uk/Depts/CMPO/

Tel: (0117) 33 10799
Fax: (0117) 33 10705
E-mail: cmpo-office@bristol.ac.uk

The Costs and Benefits of "Strangers": Why mixed communities are better.

Paul A. Grout, Sebastien Mitraille
and Silvia Sonderegger

February 2008

Working Paper No. 08/191

ISSN 1473-625X

CMPO Working Paper Series No. 08/191

The Costs and Benefits of "Strangers": Why Mixed Communities Are Better

Paul A. Grout[1], Sebastien Mitraille[2],
and
Silvia Sonderegger[1]

[1]CMPO, University of Bristol
[2]Toulouse Business School

February 2008

Abstract

Much of the literature on diversity assumes that individuals have an exogenous "taste for discrimination". In contrast with this approach, we build a model where preferences over the nature of one's community are derived indirectly, and arise because the composition of the community determines the behavior of its members. This allows us to gain a far deeper understanding of the forces that underpin the desirability of diversity or homogeneity within communities. Our main contribution is to show that there are always counteracting forces (heterogeneity involves both costs and benefits), and that, although people prefer to live in communities where their type is majoritarian, they always benefit from having some heterogeneity in the composition of their community.

Keywords: heterogeneity, social interactions, value of information, complementarities.

JEL Classification: C7, D82, Z1

Electronic version: http://www.bris.ac.uk/Depts/CMPO/workingpapers/wp191.pdf

Acknowledgements

We thank Fabrizio Adriani, Steven Durlauf, Miltos Makris, Eric Maskin, Peyton Young and seminar participants at the Universities of Birmingham, Bolzano-Bozen, Bristol, Exeter, Halifax, Keele, Lausanne, Louvain-la-Neuve, Milan, Royal Holloway and Vienna for useful comments and discussions. All errors are our own.

Address for Correspondence
CMPO, Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol
BS8 1TX
P.A.Grout@bristol.ac.uk
www.bris.ac.uk/Depts/CMPO/

CMPO is jointly funded by the Leverhulme Trust and the ESRC

1 Introduction

The sociopolitical debate of our times features much discussion centered around diversity in society. The rapid rise of globalization, the softening of national boundaries, and the drop in communication and transportation costs have all contributed to make these issues increasingly compelling. This paper is focused on two specific questions that arise in this context: what happens when "different" people interact with each other, and are people better off in homogeneous communities, where they only interact with others of their own type, or in mixed communities?

These themes are clearly related to the literature on diversity and economic performance, as well as the literature on discrimination. However, much of this literature – such as Becker (1957), Arrow (1973) and Alesina and La Ferrara (2000, 2005) – assumes that individuals have direct preferences over the composition of society, characterized by a "taste for discrimination". Hence, diversity is never desirable, as it simply reduces welfare without generating any benefits.[1] However, there is no reason to suppose that individual preferences over society composition are the primitive. We normally assume that utility depends on behavior and choices. In this case, different compositions of society may well affect behavior and choices and hence utility, but then preferences over the nature of one's community are derived indirectly, and arise because the composition of the community determines the behavior of individuals and their fellow community-members.

The motivation of this paper is to model this latter approach and to see what it implies for the induced preferences on community composition. We model communities in a loose sense, as environments where people interact with each other. Interactions may emerge from geographical proximity (in which case we think of communities as cities or neighborhoods), or from other types of links (as in the case of the "academic community"). We characterize the behavior of people as a function of the make up of their community. This allows us to gain a far deeper understanding of the forces that underpin the desirability of diversity or homogeneity within communities. In particular, our main contribution is to show that there are always counteracting forces (heterogeneity involves both costs and benefits), and that within our model people always benefit from having some strangers in their community.

There is a vast literature on social interactions – such as Akerlof (1997), Glaeser and Scheinkman (2002), Bisin, Horst and Özgür (2006) and Kuran and Sandholm (2007).[2] The focus of these, and our, studies are social interactions that are not regulated by the price mechanism. These models typically feature strategic complementarities, and assume that individual utility is affected by (i) the individual's action, (ii) the average action of agents in the individual's interacting group, and (iii)

[1] Other works – such as Lazear (1999) and Hong and Page (2001) – argue that diversity may be beneficial, because it may generate productivity gains. Essentially, the rationale is that individuals belonging to different types possess complementary sets of skills, abilities, or information; as a result, heterogeneous teams benefit from a competitive advantage, stemming from greater productivity. With respect to these works, our paper provides a novel rationale of why diversity may be desirable.

[2] Other examples include Cooper and John (1986) (non-market interactions and business cycles), Glaeser, Sacerdote and Scheinkman (1996) (social interactions and crime), Bénabou (1993, 1996) and Durlauf (1996a and 1996b) (social interactions and educational choices). See also Manski (2000) for a general discussion on the economic analysis of social interactions.

personal characteristics/taste shocks. In this paper, we follow a similar approach, although the objective of our investigation differs from previous studies.[3]

More precisely, we build a model where, as in Kuran and Sandholm (2007), individual behavior reflects a compromise between following one's own personal preferences, and coordinating with the choices of others. Actions may range from style of dressing, to punctuality, to alcohol consumption, to style of conducting business affairs. Returns from social interaction are characterized by strategic complementarities, and are greater when an individual selects an action that is close to the average action of those with whom he interacts. For instance, my returns from trying to meet deadlines, or from being on time for appointments, are higher if those with whom I interact behave in a similar way.[4]

We show that the utility that an individual expects to obtain when interacting in a community depends on the precision with which he can forecast the actions of others in the community. In other words, people are better off when they face little *strategic uncertainty*. Intuitively, when the actions of others are known, one can easily adapt to coordinate with them. In contrast, uncertainty over the actions of others is unpleasant, because it makes it harder to coordinate. Preferences over community composition therefore depend on the amount of strategic uncertainty faced by an individual in different communities.

We divide individuals into "types", which may reflect cultural background, ideology, religious beliefs, ethical attitudes and so on. Individuals belonging to the same type have more in common with each other than individuals from different types. More specifically, we assume that the preferences of individuals belonging to the same type are drawn from the same distribution. This implies that, keeping everything else equal, people find it easier to predict the behavior of those of their same type, and suggests that individuals should be biased towards homogeneous communities. However, we find that this conjecture is only partially correct. Although people would always rather live in communities where their type is majoritarian (what we call "self-bias"), this bias is rather mild, and does not imply that they favor homogeneity. In fact, as indicated, heterogeneity involves both costs and benefits.

The cost of heterogeneity arises because strangers are *less predictable* than people who are similar to us. Keeping everything else equal, this increases strategic uncertainty. Intuitively, "stranger" is synonymous of unknown, and this makes strangers somehow less attractive than people we are more familiar with, and over whom we possess more information.

On the other hand, heterogeneity also generates benefits. These benefits are two-fold. First, the presence of strangers acts as a coordination device for the incumbent members of the community – what we call the *coordination- (or cohesion) enhancing effect* of strangers. Intuitively, the desire to coordinate with newcomers induces the incumbents to modify their behavior. People pay less attention

[3] The paper that is most closely related to the present one is Kuran and Sandholm (2007). However, that paper takes an evolutionary approach, and concentrates on the distributions of preferences that will prevail in the very long-term, if initially heterogeneous groups are mixed with each other. Preferences are modeled endogenously, and are allowed to very from one generation to the next. Here, in contrast, we consider a shorter horizon, so that, for our purposes, preferences are taken as exogenous.

[4] As mentioned above, the existence of strategic complementarities is a common assumption. For instance, in Akerlof (1997) the probability of mutually beneficial tradeis assumed to increase in the proximity of individual choices of location in some "social space".

to their personal preferences and instead try to make their behavior more compatible with that of the newcomers, by taking their average preferences into account. This increases coordination between the incumbents and the newcomers and, crucially, it also increases coordination *among* the incumbents, since they all modify their behaviors in the same direction – and also put less weight on their (partially idiosyncratic) personal preferences. A by-product of this is that heterogeneity makes the incumbents' behavior *more predictable*.

Second, heterogeneity introduces what we call a *cultural diversification effect*. This effect arises because, in mixed groups, a greater variety of behaviors can coexist. This decreases the probability that any individual behavior is entirely out of line with the average. For instance, my effort towards meeting deadlines may be too high compared to the majority of people of type A, but may be too low compared to the majority of people of type B. Since I only care about the overall average, however, these two effects partially cancel out. Keeping everything else equal, this decreases aggregate uncertainty.

The benefits of heterogeneity are connected with the notions of tolerance and integration. In mixed communities, the desire to coordinate with others triggers a process of integration, where incumbents modify their behavior to make it more consistent with that of the newcomers – and vice-versa, the newcomers modify their behavior to make it more consistent with that of the incumbents. Integration increases coordination between types but also *within* each type. The latter effect is the coordination-enhancing effect of strangers. The cultural diversification effect of heterogeneity arises because mixed groups support a wider range of behaviors than homogeneous groups. As a result, each individual faces a lower probability of ending up as an "outlier". In that sense, therefore, mixed communities can be seen as more tolerant.

Overall, the interaction of the costs and benefits of heterogeneity implies that the ideal community for any type t individual always includes a positive fraction of type $t' \neq t$. Hence, our results suggest that people would always welcome some strangers in their midst. Once the share of strangers has become sufficiently large, however, people will start to be less welcoming. Attitudes towards strangers are therefore not invariant, but crucially depend upon the existing mix in a community. For any given type, the ideal share of strangers in their community is always greater than zero, but smaller than one-half. The natural question is then whether mixed communities are sustainable in practice. We address this issue by asking whether it is possible that *both* types t and t' may be better off when coexisting in the same community, rather than in a homogeneous community of their own. We show that the necessary and sufficient conditions for this to be the case is that (i) people care sufficiently about coordinating with others and (ii) within the same type, preferences are sufficiently dispersed.

One implication that emerges from our analysis is that the benefits and costs of mixed communities are two faces of the same coin. Strangers generate benefits *precisely* because their preferences are uncorrelated with those of incumbents, and are therefore less predictable. Policies of assimilation – that aim at eradicating the differences between types – will therefore reduce the costs of heterogeneity, but also its benefits.

Finally, although the paper's contribution is concerned with diversity, discrimination and social interactions, it is also connected to the literature on the value of information, such as Morris and

Shin (2002, 2005) and Angeletos and Pavan (2004, 2007(a) and 2007(b)). This literature focuses on environments where agents are both concerned with adapting their actions to the realization of an exogenous "fundamental", and also with coordinating with the choices of others.[5] Since information is (at least partially) dispersed, different individuals may hold different beliefs about the fundamental, and therefore also about other agents' actions. The models analyzed by this literature share similarities with our model, and in this context an interesting novel feature of our analysis is to show that, by altering the shares of different types in the population, we also affect the precision with which individuals can make predictions on the actions of others – namely, the amount of strategic uncertainty faced by individuals

The remainder of the paper is organized as follows. In Section 2, we present the model, while in Section 3 we derive our results. Section 4 discusses implications and possible extensions of our analysis. Section 5 concludes. All the proofs can be found in the Appendix.

2 Model

Background and Utility Functions A community contains a continuum of interacting individuals. In each social interaction, an individual is matched with a representative sample of people from his community, to form an interacting group. Matching is therefore exogenous. Each interacting group contains a continuum of atomless individuals of unit mass, indexed by the unit interval $[0, 1]$: the mass of interacting individuals is therefore equal to 1. The actions that people select affect their returns from social interaction.

Social interactions exhibit strategic complementarities. That is, the return from social exchange for an individual is greater the closer his action to the actions selected by the other individuals he is matched with. As mentioned in the introduction, this is a common assumption in models of social interactions . Moreover, individuals possess personal preferences, that also affect their utility of undertaking a certain action. The utility of an individual i who selects action $a_i \in \mathbb{R}$ is

$$u_i\left(a_i, \overline{a}_j, \alpha_i\right) = -\left(a_i - \theta\overline{a}_j - (1-\theta)\alpha_i\right)^2 \tag{1}$$

where $\overline{a}_j = \int_0^1 a_j dj$ is the average action of the individuals j with whom i is matched, $\theta \in (0, 1)$ is a constant that takes the same value for all individuals, and $\alpha_i \in \mathbb{R}$ is a taste parameter that reflects i's preferences.[6] The parameter θ measures the weight given to social coordination or individual preferences in the individual utility function: when θ is high (close to one) individuals are almost exclusively concerned with selecting an action that is close to the average action selected in their group. When θ is low (close to zero), individuals care almost exclusively about the extent to which their actions match their preferences.

[5] A typical example is that of an entrepreneur investing in a new sector: the returns to his investment depend both on the sector's inherent productivity (the "fundamental"), and on the amount of aggregate investment flowing in from other entrepreneurs.

[6] In what follows, the term "preferences" will always be utilized to indicate α_i. Note that our model could also accommodate alternative interpretations of α_i. For instance, α_i could capture individual i's identity or self-image, as in Akerlof and Kranton (2000).

Note that utility here only depends on the distance between an individual's action and the weighted average of the actions of others and his preferences, *not* on the action itself. As in Kuran and Sandholm (2007), actions are therefore "economically neutral", so that payoffs do not depend on the nature of the coordination achieved. This allows us to abstract from situations where one type has, on average, inherently "superior" preferences. The actions we have in mind may range from punctuality, to style of dressing, to the way of conducting business affairs – e.g., more or less formally. In each case, we argue, there is no "objective" ranking – e.g., there isn't a way of conducting business affairs which is objectively better than others. Rather, the optimal action for an individual is a weighted average between his personal preferences and the actions of those with whom he interacts.

Types Individuals are divided into mutually exclusive categories, or types. Types may be thought of as reflecting cultural background ("City People" versus "Country Folk"), age (Young versus Old), religious beliefs (Christians versus Muslims), ethical attitudes (Puritans versus Libertarians) and so on. We take a somehow crude – but, we believe, illustrative – position, and categorize individuals into two types, type A and type B. The share of individuals of each type in a community is common knowledge: we assume that the proportion of individuals of type $t = A, B$ is equal to λ^t, with $\lambda^A + \lambda^B = 1$. Individuals belonging to the same type possess, on average, the same preferences. However, there exists some within-type variation, in that, within the same type, different preferences may coexist. We assume that individual preferences are given by the sum of two components: a type-specific component equal to $\mu^t + e^t$, and an idiosyncratic component, equal to ε_i. The sum $\mu^t + e^t$ corresponds to the sum of an average preference μ^t and a random element e^t, which captures the common effect that specific circumstances have on all individuals of type t. The taste parameter α_i^t of an individual i belonging to type $t = A, B$ is therefore equal to

$$\alpha_i^t = \mu^t + e^t + \varepsilon_i \tag{2}$$

where for any type $t = A, B$, μ^t represents the average preferences of type t, e^t drawn as $N(0,1)$ represents a type-specific shock to preferences (arising from specific circumstances), with $e^A \perp e^B$, and ε_i, the idiosyncratic shock to preferences, is drawn as $N(0, \sigma)$ for a positive constant σ, with $\varepsilon_i \perp \varepsilon_j$ for $i \neq j$. Furthermore, we assume that the idiosyncratic shock to preferences is independent to the type-specific shock, $\varepsilon_i \perp e^t$ for any i and t, and that average preferences μ^A and $\mu^B \neq \mu^A$ are common knowledge.[7]

Finally, remark that, as the variance of type-specific shocks is normalized to 1, the parameter σ measures the variance of the idiosyncratic shocks *relative* to the type-specific shocks.[8]

Information Individual preferences are private information. Each individual observes his own α_i^t, but he is unable to discriminate between e^t and ε_i, the type-specific and the idiosyncratic shocks affecting his preferences. What we have in mind is that the specific circumstances surrounding social

[7] In Section 4 we discuss the robustness of our results to relaxing this assumption.

[8] Note that σ (as well as θ) is the same for both types. In section 4 we discuss the case where σ and/or θ differ across types.

interactions may vary, and attitudes or preferences may vary with them. An individual may be unsure of the attitude that people of a given type will tend to adopt in any given precise situation, although he may have a good idea of the attitudes that they adopt on average, when specific circumstances are averaged out.

Notice that the estimate that an individual i of type t can make of the preferences of a type-t' individual can only be based on prior information, i.e., $\mu^{t'}$. In contrast, when predicting the preferences of another person of type t, individual i possesses additional information, since he knows *both* μ^t *and* α_i^t – which, from (2), is correlated with the preferences of others of the same type.

This captures the intuitively–appealing feature that the predictions that people can make of the preferences of individuals of their same type are *more precise* than those they can make of the preferences of a person of a different type. However, the predictions that people make of the preferences of others also differ from individual to individual, as they are based on private information – namely, an individual's personal preferences.[9] As will become clear in the main body of the paper, this feature plays an important role in our analysis.

Timing

Social exchanges occur at fixed intervals of time within communities. In each social exchange:

$t = 0$ Individuals are matched in interacting groups.

$t = 1$ For each group, nature selects the (group-specific) realizations e^A, e^B and individual idiosyncratic shocks.

$t = 2$ Each individual observes his personal preferences and selects his action.

$t = 3$ Payoffs are realized.

Notice that the realization of type–specific shocks is group–dependent. This reflects the notion that type–specific shocks arise from the specific circumstances surrounding an interaction. These circumstances are common for all individuals in the same group, but vary across different interacting groups.[10]

3 Results

Having described our model, we are now in the position of introducing our results. We start off by characterizing the equilibrium of the game. First, utility maximization yields

$$a_i = \theta E\left(\overline{a}_j \mid \alpha_i\right) + (1-\theta)\alpha_i \tag{3}$$

Individuals select actions that are a weighted average of their expectations over the average action in their interacting group, and their preferences. Suppose that the proportion of individuals of type

[9] In a dynamic context, an individual's private information may also reflect his personal past experiences. See section 4 for a fuller discussion on this possibility.

[10] Strictly speaking, we should therefore denote this element of the type-specific shocks as e_z^t, where the subscript z indicates that the realization is specific to group z. This would however make the notation heavier, and is therefore omitted. Note that this feature only matters for the results of Section 3.1 (community cohesion) and is irrelevant for those in Section 3.2 (individual welfare).

$t = A, B$ in a given community is λ^t, and that of type $t' = \{A, B\} \backslash t$ is $\lambda^{t'}$, where $\lambda^t + \lambda^{t'} = 1$. Using the linear-normal from of the game, we can establish the following result:

Lemma 1 (Description of the Equilibrium) *In the unique equilibrium of the game, the action of an individual i of type t is equal to*

$$a_i^t = k^t \alpha_i^t + \lambda^{t'} \theta \mu^{t'} + (1 - \lambda^{t'} \theta - k^t) \mu^t \quad \textit{where} \quad k^t = \frac{(1-\theta)(\sigma+1)}{\sigma - \theta \left(1 - \lambda^{t'}\right) + 1}.$$

Proof: See Appendix.

Lemma 1 describes the equilibrium strategy followed by individuals, as a function of the composition of the community in which they interact. Each individual selects an action that is a weighted average of his preferences, the average preferences of individuals of his same type across interactions, and those of individuals of the other type. Notice that, substituting for $\alpha_i^t = \mu^t + e^t + \varepsilon_i$, the equilibrium action a_i^t played by individual i of type t can also be rewritten as

$$a_i^t = k^t \left(e^t + \varepsilon_i\right) + \lambda^{t'} \theta \mu^{t'} + (1 - \lambda^{t'} \theta) \mu^t \tag{4}$$

Expression (4) clarifies the effect that the presence of people of type t' in the community – namely, a positive value of $\lambda^{t'}$ – has on the behavior of people of type t. First, $\lambda^{t'} > 0$ induces all type t individuals to move their actions away from μ^t and towards $\mu^{t'}$ – namely, their best estimate of the preferences of type t'. This phenomenon works both ways: when their types mix, type A move their actions towards μ^B, and type B move their actions towards μ^A. As a result of this process of integration, behaviors across the two types become more similar than if the two types had been kept apart. In other words, when different types mix, each type loses some of its peculiarity, to adopt part of the other type's behavior.[11]

Second, a the presence of type t' also affects k^t, namely the weight people of type t put on their personal preferences when selecting their actions. The key feature here is that k^t is strictly *decreasing* in $\lambda^{t'}$, the proportion of individuals of type t' in the community.[12] Intuitively, the correlation between the preferences of an individual of type t and those of an individual of type t' is zero. Hence, the presence of type t' in the community reduces the usefulness for a type t of utilizing his private preferences for making predictions about the actions of others. Each type t individual therefore decreases the weight he places on his preferences when selecting his action. By inspection of (4), it is clear that a smaller k^t implies that actions are less affected by the realizations of both type-specific and idiosyncratic shocks to preferences. The following Lemma makes clear the consequences of this effect in terms of within–type variance of actions (or within–type cohesion).

[11] This shares similarities with Kuran and Sandholm (2007).

[12] The derivative of k^t with respect to $\lambda^{t'}$ is equal to $\frac{-\theta(1-\theta)(1+\sigma)}{(1+\sigma-\theta(1-\lambda^{t'}))^2}$ which is strictly negative given our assumptions on the model parameters.

Lemma 2 (Within–Type Variance of Actions) *The variance of the actions of individuals of type $t = A, B$ around the mean type t –action in a community is strictly decreasing in $\lambda^{t'}$, the proportion of type $t' \neq t$ in the community.*

Proof: See Appendix.

Lemma 2 illustrates how the presence of type t' acts as a *coordination (or cohesion-enhancing) device* on individuals of type t. Note that type t' generate a *direct* and an *indirect* effect on type t's actions. The *direct effect* arises because type t individuals forecast that they will now interact with individuals of type t', and accordingly adjust their actions, by putting less weight on their individual preferences. The *indirect (or multiplier) effect* arises because each individual realizes that, through the direct effect, all individuals of type t will now put less weight on their private preferences when selecting their actions. In turn, this decreases the weight put by each individual on his private preferences even further, and so on. This multiplier effect shares similarities with the social multiplier identified by the literature on social interactions (see for instance Glaeser et al. 2003). As a result of this effect, small variations in community composition may result in relatively large changes in individual behavior. For instance, the introduction of even a very small share of type t' in a community of type t will sensibly affect the actions of the incumbent members of the community.[13]

The following corollary shows that the cohesion-enhancing effect of strangers is greater in communities where the preferences of incumbents are rather similar (i.e., σ is small) and people are concerned with coordination, but not excessively so (i.e., θ takes intermediate values).

Corollary 1 *The value of $\mid dk^t(\lambda^{t'})/d\lambda^{t'} \mid$ is: (i) strictly decreasing in σ; (ii) concave in θ, reaching a maximum at $\theta = \frac{\sigma+1}{2\sigma+\lambda^{t'}+1} \in [1/2, 1)$.*

Proof: See Appendix.

Corollary 1 highlights how changes in the values of σ (the variance of idiosyncratic shocks to preferences) and θ (the weight of coordination in the utility function) affect the coordinating role of strangers on the incumbent members of a community.[14] Consider first a high σ. A high value of σ implies that, even within the same type, personal preferences are not very informative of the preferences of others. As a result, when predicting the preferences of others, people pay little attention to their personal preferences. This decreases the usefulness of strangers for increasing coordination. Hence, the cohesion-enhancing effect of strangers is inversely related to the value of σ.

Now consider a high value of θ. When selecting their actions, people put very little weight on their taste parameter, since (i) they care little about satisfying their preferences and (ii) (as a result of (i))

[13] This can be seen by noticing that $\lim_{\lambda^{t'} \to 0} \frac{dk^t(\lambda^{t'})}{d\lambda^{t'}} = -\theta(1-\theta)\frac{\sigma+1}{(\sigma-\theta+1)^2}$.

[14] Note that, if we allowed for θ and σ to differ across types, the value of $\mid dk^t(\lambda^{t'})/d\lambda^{t'} \mid$ would depend on θ^t and σ^t, but not on $\theta^{t'}$ and $\sigma^{t'}$. (This can be easily seen by looking at the derivation of k^t in the Appendix). As a result, the values of θ and σ affect $\mid dk^t(\lambda^{t'})/d\lambda^{t'} \mid$ through their impact on the behavior of the incumbent population *only*.

personal preferences are of little use for forecasting the actions of others. This reduces the cohesion-enhancing effect of strangers, since people put little weight on personal preferences anyway. What about a low θ? A low value of θ implies that, when deciding their actions, people are chiefly concerned with satisfying their personal preferences. As a result, coordination in homogeneous communities is low. However, the cohesion-enhancing effect of strangers is also low. This is because, since the incumbents are not particularly desirous of coordinating with them, the presence of "different" individuals has only a minimal effect on the incumbents' choice of actions.

Overall, therefore, strangers are most effective in enhancing coordination among incumbents when θ is neither too high, nor too low.

3.1 Community Cohesion

In recent times, governments have been increasingly promoting the idea that "Community Cohesion" should be fostered and encouraged. For instance, in the United Kingdom a Community Cohesion Unit has recently been created within the Home Office, to oversee initiatives on building community cohesion. In addition to the ministerial group, there are also several advisory groups and taskforces such as the Community Cohesion Review Team. As part of this initiative, the Local Government Association (LGA), together with the Home Office, has recently issued a draft guidance to local authorities on mainstreaming and promoting community cohesion.[15] Similar initiatives – such as the project for "Building the New American Community", sponsored by the National Conference of State Legislatures[16] – have also been on the rise in the United States. It is therefore important to investigate whether we should expect mixed or homogeneous communities to be more cohesive. This is what we do in this section. Since the concept of community cohesion is in itself quite vague, it is first necessary to make clear how we define it in the present context.

Definition: *We define community cohesion as the variance of the actions of individuals within a community around the community mean action.*

As seen in Lemma 2, introducing a fraction of type t' in a community increases the within–type cohesion of individuals of type t. However, it also introduces a new source of variance, arising from between–type variations in the choice of actions. Proposition 1 spells out the sufficient conditions for the latter effect to be weaker than the former.

Proposition 1 (Community Cohesion) *The sufficient condition for community cohesion to be greater in mixed communities is that*

$$\left(\mu^A - \mu^B\right)^2 < \theta \frac{\sigma + 1}{(\sigma - \theta + 1)^3} \left[(\sigma + 1)\left(4(\sigma + 1) - 3\theta\right) + \theta^2\right]$$

Proof: See Appendix.

[15] "Community Cohesion: An Action Guide", available at www.lga.gov.uk.

[16] Details are available at http://www.ncls.org/programs/immig/community_orr.htm.

Proposition 1 shows that, for some parameter values, cohesion may be higher in mixed as opposed to homogeneous communities. Consider for instance $\theta = 3/4$, $\sigma = 1$. Then the condition for mixed communities to be more cohesive than homogeneous communities is that $| \mu^A - \mu^B | < 3$. More generally, keeping everything else equal, the righthandside of the condition lain out in proposition 1 is increasing in θ, and decreasing in σ. First, consider σ. From corollary 1, a large σ decreases the usefulness of utilizing one's preferences for predicting the preferences of others. This decreases the cohesion-enhancing effect of strangers, and makes the result harder to obtain. Now consider θ. A large value of θ implies that, when selecting their actions, people pay less attention to their personal preferences, and are instead more concerned about coordinating with others. This decreases the coordination-enhancing effect of strangers (as seen in corollary 1), but it also decreases the variance of the actions of the newcomers. In this case, the latter effect is dominant.

Proposition 1 contravenes the often–heard argument that heterogeneity results in less cohesion.[17] We show that this argument may be incorrect. Although differences between types are indeed a source of greater variance in actions, in mixed communities individuals of the same type select actions that are closer to one another. For certain parameter values, this second effect may dominate, and result in greater community cohesion. Empirically, therefore, our model predicts that mixed communities may indeed exhibit lower variance of actions than homogeneous ones. This more likely in environments where (i) people care sufficiently about coordinating with others and (ii) idiosyncratic shocks are not very important, so that the preferences of people of the same type are never too dispersed. To the extent to which differences in behavior generate frictions, or even conflict, our result therefore suggests that mixed communities may actually be more harmonious and peaceful than homogeneous ones.

3.2 Preferences over Community Composition

We now investigate the preferences of individuals with respect to community composition. We do so by taking an ex-ante perspective, evaluating the expected utility of an individual as a function of the composition of his community. The idea is to gain a better understanding of the welfare characteristics of different types of communities. Are individuals always better off in homogeneous communities? Or are they willing to introduce some "different" people in their communities? In this latter case, how large is the share of strangers that people are willing to introduce? These are the types of questions that we wish to address in this section.

Substituting for the optimal action (3) into the utility function, we see that an individual's expected utility can be written as

$$E\left[u_i\left(a_i, \overline{a}_j, \alpha_i\right)\right] = -\theta^2 E\left[E\left(\left(\overline{a}_j - E\left(\overline{a}_j \mid \alpha_i\right)\right)^2 \mid \alpha_i\right)\right] = -\theta^2 E\left(Var\left(\overline{a}_j \mid \alpha_i\right)\right) \tag{5}$$

where $Var\left(\overline{a}_j \mid \alpha_i\right)$ indicates the conditional variance of $\overline{a}_j$ around $E\left(\overline{a}_j \mid \alpha_i\right)$, i's expectation of $\overline{a}_j$

[17] The British media and political discourse on the eroding effects of diversity on social cohesion is an example. See D. Goodhart, "Too Diverse" *Prospect*, February 2004 and "The Kindness of Strangers. A report on Multiculturalism", *The Economist*, 28 February 2004, pp. 3-4 for illustrations of the current debate on diversity in British society.

conditional on α_i.

In words, $Var(\overline{a}_j \mid \alpha_i)$ denotes the accuracy with which i is able to forecast the average action of others in his group. In this sense, (5) captures the amount of *strategic uncertainty* present in the game. An individual's expected utility is therefore inversely proportional to the amount of strategic uncertainty he faces.

The effect of community composition on strategic uncertainty is ambiguous. Consider a type t individual (individual i), whose community changes from being entirely composed of type t to also including a fraction of type t'. Because the preferences of t' individuals are entirely uncorrelated to his own, the change in the composition of i's community introduces people, over whom i possesses little information. In particular, the precision with which i can forecast the preferences of type t' is lower than the precision with which he can forecast those of type t. Whether this translates into greater strategic uncertainty depends on the extent to which type t' follow their preferences when selecting their actions. If θ is small – so that individuals care mostly about selecting actions that match their preferences – then predicting the actions of type t' in a mixed community is indeed *always* harder for i than predicting the actions of individuals of his same type t in a homogeneous community.[18] If on the other hand θ is large – so that people are sufficiently concerned about coordination – this happens only when the share of type t' in the mixed community is above a certain threshold, namely, $\lambda^{t'} > \widehat{\lambda}$ for some $\widehat{\lambda} \in (0, 1)$.

Overall, therefore, heterogeneity may impose some costs, arising from the fact that the actions of strangers may be harder to predict than those of individuals of our same type. This is summarized in the following lemma.

Lemma 3 (Cost of Heterogeneity) *Suppose that $\theta < 1 + \sigma - \sqrt{\sigma + \sigma^2}$. Then, the precision with which, in a mixed community, an individual of type $t = A, B$ can forecast the actions of individuals of type $t' \neq t$ is always lower than the precision with which he can forecast the actions of other individuals of type t in a homogeneous community. If $\theta > 1 + \sigma - \sqrt{\sigma + \sigma^2}$ then there exists a value $\widehat{\lambda} \in (0, 1)$ such that this is the case whenever $\lambda^{t'} > \widehat{\lambda}$.*

Proof: See Appendix.

Lemma 3 shows that heterogeneity may impose some direct costs, since it introduces individuals with preferences that are less predictable than those of people similar to us. However, heterogeneity also has benefits. These can be divided in two categories. First, there are the benefits that derive from the fact that community composition affects individual behavioral choices. As seen in the previous section, the move from a homogeneous to a heterogeneous community has a coordination-enhancing effect, in that it decreases the weight put by incumbent agents on their preferences when selecting their

[18] Since, as will become clear below, the actions of individuals of type t become *easier* to predict the smaller λ^t (equivalently, the larger $\lambda^{t'}$) this actually holds when considering the actions of type t in *any* community. However, here we are discussing whether heterogeneity is desirable or not, so the appropriate benchmark is the predictability of actions of type t in a homogeneous community.

actions. In turn, this implies that their actions are less affected by the realizations of type-specific and idiosyncratic shocks to preferences, and are therefore easier to predict.

Note that being part of a mixed community affects the behavior of the newcomers in the same way, namely, by making their actions less dependent on their personal preferences. The extent to which this happens is inversely proportional to their share in the mixed community. If the share of newcomers in the community is sufficiently small (and if θ is sufficiently large) their actions may actually end up being easier to predict for the incumbents than the actions of others of their same type in a homogenous community.[19] In this case, heterogeneity benefits the incumbents in two ways: it makes the actions of other incumbents easier to predict, and it introduces people who are themselves very predictable. In both instances, the gain stems from the effect of heterogeneity on individual behavior.

To sum up, one of the benefits of heterogeneity is that affects individual behavior, causing people to put less weight on their personal preferences (which decreases strategic uncertainty). This is summarized in lemma 4.

Lemma 4 (First Benefit of Heterogeneity) *The precision with which an individual of type $t = A, B$ can forecast the actions of others of his same type in his interacting group is increasing in $\lambda^{t'}$, the share of individuals of type $t' \neq t$ in the community. Moreover: if $\theta > 1 + \sigma - \sqrt{\sigma + \sigma^2}$, then there exists a value $\widehat{\lambda} \in (0, 1)$ such that, when $\lambda^{t'} < \widehat{\lambda}$, the precision with which an individual of type t can forecast the actions of type $t' \neq t$ is greater than the precision with which he can forecast the actions of other individuals of type t in a homogeneous community.*

Proof: See Appendix.

There is also another benefit of heterogeneity, which does not depend on the effect that community composition has on people's choice of actions, but is purely a diversification effect. Since type-specific shocks are uncorrelated, mistakes incurred when predicting the average action of type A and those incurred when predicting that of type B partially offset each other. In other words, since the range of behaviors that coexist in mixed groups is wider, the risk that your behavior is entirely out of line with the average is lower than in homogeneous groups. Hence, the probability of becoming an "outlier" is smaller. In this sense, therefore, mixed communities can be seen as more tolerant. This cultural diversification effect is summarized in lemma 5.

Lemma 5 (Second Benefit of Heterogeneity) *Suppose the weight assigned by individuals on their personal preferences is independent of community composition. Then introducing a sufficiently small fraction of people of a different type in an otherwise homogeneous community reduces aggregate strategic uncertainty.*

Proof: See Appendix.

[19] However, as highlighted in lemma 3, this may never occur once $\lambda^{t'}$ surpasses a certain threshold.

We are now ready to state our result on preferences over community composition. Consider an individual i of type $t = A, B$. It is straightforward to verify that, for this individual, $E\left(Var\left(\overline{a}_j \mid \alpha_i\right)\right)$ is strictly convex in $\lambda^{t'}$ and reaches a minimum at $\lambda^{t'} = \lambda^* \in (0, 0.5)$. Hence, some heterogeneity unambiguously decreases the amount of strategic uncertainty faced by individual i. As $\lambda^{t'}$ increases, however, the direct cost described in lemma 3 becomes gradually more important, and eventually takes over. The strategic uncertainty faced by i is minimized when the share of individuals belonging to a different type than i is somewhere between zero and one-half.

Proposition 2 (Preferences over Community Composition) *Agents always prefer to interact in mixed as opposed to homogeneous communities. However, individuals exhibit a self-bias, in that their expected utility is maximized when the share of people of their type in the community is greater than one-half.*

Proof: See Appendix.

Proposition 2 establishes that individuals of each type would ideally like to include a positive but minoritarian fraction of the other type in their communities. The rationale for the first result – namely, that individuals would like to include a positive fraction of the other type in their community – arises because the possible cost of introducing a very small fraction of strangers in the community is only second-order, since it is proportional to their share in the community. In contrast, the benefits are first-order. Hence, introducing an arbitrarily small share of the other type in a homogeneous community is unambiguously welfare-improving: it generates first-order benefits, and only a second-order cost (if at all).

Given perfect symmetry between the two types, the variance of actions encountered in a community where, say, $\lambda^A = 0.3$, $\lambda^B = 0.7$ is identical to that of a community where $\lambda^A = 0.7$, $\lambda^B = 0.3$. The amount of strategic uncertainty faced by individuals, however, is different. A type A individual will face less strategic uncertainty in the latter case than the former (and vice-versa for a type B). This is essentially the rationale for the second result – namely, that individual preferences exhibit a self-bias, in that their ideal community is one where their type is majoritarian.

The self-bias identified in proposition 2 describes an effect that shares some similarity with the literature on discrimination (such as Becker 1957), in that it shows that people would rather not introduce too many strangers in their community. However, rather than assuming from the outset that people suffer from a "natural aversion to heterogeneity" (Alesina and La Ferrara 2002: 225), here the self-bias effect is derived endogenously. This allows us to gain better understanding of why this effect may emerge. Moreover, in contrast with previous literature, we find that aversion to strangers is not constant, but varies with community composition. Indeed, the self-bias effect only kicks in when the share of people of different type present in the community is sufficiently high. As shown in proposition 2, communities unambiguously benefit from having a few strangers around. Hence, our model predicts that homogeneous communities should generally welcome the arrival of individuals of a different type. As the share of those strangers in the community grows, however, attitudes should become less welcoming. This latter prediction is a consequence of the self-bias effect.

We now explore whether it is possible to construct communities that are simultaneously preferred by both types over homogeneous communities, composed exclusively of individuals of their own type. This is important, as it gives a measure of the sustainability of mixed communities in practice.

Proposition 3 (Simultaneous Preference for Mixed vs. Homogeneous Communities) *The necessary and sufficient condition for the existence of mixed communities that are simultaneously preferred by both types over homogeneous communities (composed only of their own type) is that* $\theta > 1-\sqrt{\sigma+2\sigma^2}$.

Proof: See Appendix.

The condition lain out in proposition 3 is easier to satisfy (i) the larger θ and (ii) the larger σ. The reason for (i) is essentially the same as in proposition 1. Now consider (ii). As seen in corollary 1, a greater σ reduces the cohesion-enhancing effect of strangers, and this should make the result harder to obtain. However, a large σ also has other effects. First, it makes individuals of the other type more predictable, since they pay less attention to their idiosyncratic preferences.[20] Moreover, in some cases, a larger σ may make individuals of one own's type less predictable (since one's own personal preferences are a noisier signal of those of others of the same type).[21] The overall outcome is that a large σ makes mixed communities relatively more desirable.

Note that $1-\sqrt{\sigma+2\sigma^2} < 1+\sigma-\sqrt{\sigma+\sigma^2}$. Hence, proposition 3 holds also in the worst possible case, namely when – as seen in lemma 3 – θ is sufficiently small to ensure that the actions of strangers are always less predictable than those of individuals of our own type (and this is true even if we consider the actions of those of our own type in a homogeneous community).

To sum up, therefore, proposition 3 establishes that mixing in the same community may simultaneously benefit both types. Although individuals would ideally like to live in communities where their type is majoritarian, the benefits generated by mixing may be sufficiently large to ensure that people prefer mixed over homogeneous communities even when their type is *not* majoritarian. This is more likely to hold in environments where people care sufficiently about coordinating with others, and/or the preferences of people of the same type are sufficiently dispersed.

4 Implications and Extensions

Implications Our analysis has highlighted how heterogeneity produces benefits, as well as costs. A question that may naturally arise is whether it is possible to eliminate the costs associated with mixed

[20] The logic for this is highlighted in the discussion following corollary 1. Formally, the effect can be seen by noticing that, as shown in the Appendix, the strategic uncertainty faced by type t when forecasting the actions of type t' is $(k^{t'})^2$ decreasing in σ.

[21] This can be seen by noticing that, as shown in the Appendix, the strategic uncertainty faced by type t when forecasting the actions of others of type t is equal to $\left(k^t\right)^2 \frac{\sigma}{\sigma+1}$. Although k^t is decreasing in σ, $\frac{\sigma}{\sigma+1}$ is an increasing function of σ. The net result of these opposing effects depends on parameter values.

communities without affecting their benefits. Our analysis suggests that this may unfortunately not be possible. The benefits and costs of mixed communities are two faces of the same coin: strangers generate benefits *precisely* because their preferences are uncorrelated with those of incumbents – a feature that, in turn, makes their preferences harder to predict. Policies aimed at decreasing the direct costs of heterogeneity – such as policies aimed at assimilation, namely the homogenization of preferences across types[22] – will therefore also eliminate its less visible benefits.

Another interesting conclusion that emerges from our model is that the costs of heterogeneity have their roots in strategic uncertainty – namely, individual ability to make predictions of others' preferences – rather than differences between types *per se.* Assimilation is therefore only one route through which these costs can be softened. Campaigns aimed at familiarizing people with those belonging to different types would lower the uncertainty faced when dealing with them, and would therefore achieve the same objective. However, as noted above, policies that reduce the direct costs of heterogeneity will also have the effect of decreasing its less direct benefits (although to a lesser extent than policies of assimilation, since preferences would nonetheless remain uncorrelated).

Dynamic Extensions One way in which our model can be extended to a dynamic framework is to have people utilize past experience to makes predictions about the behavior of others. Although a full analysis is beyond the scope of this paper, it seems reasonable to conjecture that, at least at first, people should have more familiarity with individuals of their same type. This provides an additional rationale for why they may be able to better assess their preferences and therefore actions. As time goes by, however, people accumulate interactions with the "newcomers". This suggests that, over time, the precision of people's predictions on the preferences of those of a different type should improve. Hence, the direct cost associated with heterogeneity should decrease. Importantly, this happens even in the absence of homogenization of preferences across types.

Although more precise, the predictions that people make of the preferences of those of another type should over time also become more dispersed, since they will be based on an individual's private experience. As different types interact more and more with one another, therefore, both the direct cost associated with heterogeneity – arising from the unpredictability of strangers – and the indirect benefit identified in lemma 4 – arising from their coordination-enhancing effect – will lose strength.

More Than Two Types A natural question that may arise concerns the effect of having more than two types interact with one another. Although a full analysis of this case is beyond the scope of this work (and is therefore left to future research), here we can sketch a few intuitions. Intuitively, if strangers belong to several different types, incumbents should modify their behavior using a weighted

[22] For instance, in the UK, policy makers have recently abandoned the dominant approach of multiculturalism in favor of what some call "a return to assimilation" – Letki (2007), citing Cheong and al. (2005). Another example – borrowed from Kuran and Sandholm (2007) citing Schlessinger (1991) – is the policy pursued in the US in the early 20th century. In one of his speeches, president Woodrow Wilson argued that "A man who thinks of himself as belonging to a particular national group in America has not yet become an American". Similarly, during the fascist regime in Italy, people of foreign origin were strongly encouraged to assimilate with the natives, even to the extent of "italianizing" their family names.

average of their prior information on these types (where weights reflect the relative share of each type). All types will therefore move their actions in a direction that is somewhere in the middle of the other types' predicted preferences, but that does not necessarily reflect any particular type's.

Another implication of having more than two types is that the self-bias effect should now take a milder form: individuals would like to live in communities where their type has a larger share than any of the other types – though not necessarily exceeding one-half.

Asymmetries To keep the analysis as simple as possible, we have restricted attention to symmetric environments. More generally, however, the values of θ – capturing concern for coordination – and σ – the (relative) variance of idiosyncratic shocks to preferences – may be type-dependent. This type of analysis may be especially interesting for providing information as to which types are more willing to introduce strangers in their communities. Are these characterized by high or low values of θ and/or σ? Consider the limit as $\theta^t \to 1$ and/or $\sigma^t \to 0$. Intuitively, when $\theta^t \to 1$, people of type t only care about coordinating with others. In a homogeneous community, composed only of type t, we then have a continuum of equilibria, all characterized by people perfectly coordinating with one another. Similarly, when $\sigma^t \to 0$, the personal preferences of all individuals of type t are identical. Again, in a homogeneous community, perfect coordination can then be achieved. In both instances, type–t homogenous communities are characterized by no strategic uncertainty at all. For type t, the ideal share λ^* of people of type t' is therefore equal to zero. By continuity, for types characterized by very high θ^t and/or very small σ^t, the value of λ^* is correspondingly very small – in fact smaller than for all other types.[23] Homogeneous communities where people care very much about coordination and/or where people have very similar preferences are thus the least welcoming towards strangers. By a similar logic , this should also generally apply to communities where there is very little uncertainty over individual preferences. Consider a dynamic setting, where idiosyncratic shocks to preferences are correlated over time. Under these circumstances, uncertainty about individual preferences should be lower in small environments, where the same people keep on interacting with each other. This would match casual evidence that very small communities are generally less open to newcomers than larger communities.

Dispersion of Prior Beliefs In our analysis, we have assumed that both μ^A and μ^B are common knowledge to all players. An implication of this is that the forecast that type individuals t make of the preferences of type $t' \neq t$ is the same for all (since it corresponds to the prior $\mu^{t'}$). This strengthens the coordination-enhancing effect of strangers, and could at first glance appear to be a major driving force for our results. This may be troublesome, since, in real life, it seems reasonable that beliefs over the preferences of strangers may differ across individuals. It is therefore important to explain what would happen if beliefs on $\mu^{t'}$ differed across type t individuals. Consider the following modification of our model. All type $t = A, B$ individuals know the value of μ^t. However, their beliefs over $\mu^{t'}$ are dispersed, since each individual i only receives a signal $\mu_i^{t'} = \mu^{t'} + x_i$ where $x_i \sim N(0, X)$ for a

[23] Note however that this does not imply that λ^* is always decreasing in θ^t and/or increasing in σ^t. In some ranges, we may get non-monotonicities. Details are omitted but are avaible from the authors upon request.

positive constant X, with $x_i \perp x_j$. Suppose that the common prior over $\mu^{t'}$ for all type t individuals is uniform over $\mathbb{R}$. This implies that, for each individual i of type t, his best estimate of $\mu^{t'}$ is given by his private signal $\mu_i^{t'}$. It is straightforward to verify that, with this modification, our results would remain qualitatively unchanged. In particular, proposition 2 would still hold, implying that individuals would always benefit from introducing a positive share of strangers in their community. Although the presence of type t' introduces noise in type t individual's actions (since $\mu_i^{t'}$ differs across individuals), this is only a second-order effect. To see this, note that in this case expression (4) becomes

$$a_i^t = k^t \left(e^t + \varepsilon_i\right) + \lambda^{t'}\theta\mu_i^{t'} + (1 - \lambda^{t'}\theta)\mu^t \tag{6}$$

The weight put by individual i on $\mu_i^{t'}$ is proportional to $\lambda^{t'}$, the share of type t' in the community. Hence, for small $\lambda^{t'}$, the additional variation in actions introduced by heterogeneous $\mu_i^{t'}$ is correspondingly small. In contrast, as seen above, the benefits of mixed communities are first-order.

Note however that variations in $\mu_i^{t'}$ would introduce an extra cost of mixed communities, in addition to that discussed in the main text.

5 Concluding Remarks

What happens when different types of people are mixed in the same community? Is there any reason why different types of people should wish to be mixed in the same community? Is it true that individuals exhibit a "taste for discrimination", in that they are always better off when they interact only with other individuals of their same type? We believe that our analysis may have provided a useful theoretical contribution to these debates. Starting from preferences that are not directly defined over community composition, we have shown that individuals behave differently in mixed as opposed to homogeneous environments. From this, we have identified the benefits and costs of heterogeneity from an individual welfare perspective. We have shown that preferences over community composition depend on strategic uncertainty – the ability to predict the actions of others. Strangers have preferences that are uncorrelated to our own, and may therefore be costly to interact with. Importantly, however, this direct cost of mixed communities arises not because strangers have different preferences *per se*, but because, as a result of them having different preferences, their actions may be more difficult to predict. Mixed communities also have benefits, which are however more subtle than their costs. First, strangers act as a coordination device, and make coordination among incumbents easier. Second, they also generate what we call a cultural diversification effect, by making aggregate actions less dependent on type-specific shocks. The benefits of heterogeneity are sufficiently strong to ensure that communities always gain from introducing some "different" individuals in their midst, and that individuals of both types may simultaneously be better off in mixed as opposed to homogeneous communities.

References

[1] **Akerlof, G. A.** (1997) Social Distance and Social Decisions, *Econometrica* 65 (5).

[2] **Akerlof, G. A. and R. Kranton** (2000) Economics and Identity, *Quarterly Journal of Economics* 115 (3).

[3] **Alesina A. and E. La Ferrara** (2000) Participation in Heterogeneous Communities, *Quarterly Journal of Economics* 115 (3).

[4] **Alesina A. and E. La Ferrara** (2002) Who Trusts Others? *Journal of Public Economics*, 85 (2).

[5] **Alesina A. and E. La Ferrara** (2005) Ethnic Diversity and Economic Performance, *Journal of Economic Literature* 43 (3).

[6] **Angeletos, G. M. and A. Pavan** (2004) Transparency of Information and Coordination in Economies with Investment Complementarities, *American Economic Review* 94 (2).

[7] **Angeletos, G. M. and A. Pavan** (2007(a)) Efficient Use of Information and Social Value of Information, *Econometrica*, forthcoming.

[8] **Angeletos, G. M. and A. Pavan** (2007(b)) Socially Optimal Coordination: Characterization and Policy Implications, *Journal of the European Economic Association*, forthcoming.

[9] **Arrow, K. J.** (1973) *The Theory of Discrimination*, in Orley Ashenfelter and Albert Rees, eds., Discrimination in Labor Markets, Princeton NJ: Princeton University Press.

[10] **Becker, G. S.** (1957) *The Economics of Discrimination*, Chicago: University of Chicago Press.

[11] **Bénabou, R.** (1993) Workings of a City: Location, Education, and Production *Quarterly Journal of Economics*,108 (3).

[12] **Bénabou, R.** (1996) Heterogeneity, Stratification, and Growth: Macroeconomic Effects of Community Structure, *American Economic Review*, 86 (3).

[13] **Bisin, A., U. Horst and O. Özgür** (2006) Rational Expectations Equilibria of Economies with Local Interactions, *Journal of Economic Theory*, 127 (1).

[14] **Cheong, P. H., R. Edwards, H. Goulbourne and J. Solomos** (2005) Immigration, Social Cohesion and Social Capital: A Critical Review, mimeo.

[15] **Cooper, R. and A. John** (1988) Coordinating Coordination Failures in Keynesian Models, *Quarterly Journal of Economics*, 103 (3).

[16] **Costa, D. L. and M. E. Kahn** (2003) Civic Engagement and Community Heterogeneity: An Economist's Perspective, *Perspectives on Politics*, 1 (1).

[17] **Durlauf, S.** (1996a) A Theory of Persistent Income Inequality, J*ournal of Economic Growth*, 1 (1).

[18] **Durlauf, S.** (1996b) *Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality*, in Dynamic Disequilibrium Modeling: Proceedings of the Ninth International Symposium on Economic Theory and Econometrics, ed. by W. Barnett, G. Gandolfo, and C. Hillinger. Cambridge, UK: Cambridge University Press.

[19] **Glaeser, E., H. D. Kallal, J. A. Scheinkman and A. Shleife**r (1992) Growth in Cities, *Journal of Political Economy* 100 (6).

[20] **Glaeser, E. and J. A. Scheinkman** (2001) *Measuring Social Interactions,* in Social Dynamics, ed. by S. Durlauf and P. Young. Cambridge: MIT Press.

[21] **Glaeser, E. and J. A. Scheinkman** (2002) *Non-Market Interactions*, in Advances in Economics and Econometrics: Theory and Applications, Eight World Congress of the Econometric Society, edited by M. Dewatripont, L. P. Hansen and S. Turnovsky, Cambridge UK: Cambridge University Press.

[22] **Glaeser, E., Sacerdote, B. and J. A. Scheinkman** (1996) Crime and Social Interactions, *Quarterly Journal of Economics*, 111 (2).

[23] **Glaeser, E., Sacerdote, B. and J. A. Scheinkman** (2003) The Social Multiplier, *Journal of the European Economic Association*, 1 (2-3).

[24] **Hong L. and Page S.E.** (2001) Problem Solving by Heterogeneous Agents, *Journal of Economic Theory*, 97 (1).

[25] **Kuran, T. and W. H. Sandholm** (2007) Cultural Integration and Its Discontents, *Review of Economic Studies,* forthcoming.

[26] **Lazear, E. P.** (1999) Globalization and the Market for Team-Mates, *Economic Journal* 109 (454).

[27] **Letki, N.** (2007) Does Diversity Erode Social Cohesion? Social Capital and Race in British Neighbourhoods, *Political Studies*, forthcoming.

[28] **Manski, C. F.** (2000) Economic Analysis of Social Interactions, *Journal of Economic Perspectives* 14 (3).

[29] **Morris, S. and H. S. Shin** (2002) Social Value of Public Information, *American Economic Review* 92 (5).

[30] **Morris, S. and H. S. Shin** (2005) *Heterogeneity and Uniqueness in Interaction Games*, in The Economy as an Evolving Complex System III, edited by L. Blume and S. Durlauf, Santa Fe Institute Studies in the Sciences of Complexity, New York: Oxford University Press.

[31] **Schlessinger, A. M.** (1991) *The Disuniting of America: Reflections on a Multicultural Society.* Knoxville, Tenn.: Whittle Books.

Appendix

Proof of Lemma 1: We first characterize the linear equilibrium of the game. We then argue that this is also the unique equilibrium.

Linear Equilibrium: An individual i of type t selects his action to minimize

$$E\left[\left(a_i - \theta\overline{a}_j - (1-\theta)\alpha_i^t\right)^2 \mid \alpha_i^t\right] \tag{7}$$

The first order condition yields:

$$a_i = \theta E\left(\overline{a}_j \mid \alpha_i^t\right) + (1-\theta)\,\alpha_i^t \tag{8}$$

In a linear equilibrium, the actions of a generic individual of type t can be written as[24]

$$a_j^t = k^t \alpha_j^t + \delta^t \tag{9}$$

The share of individuals of type A in the community is λ^A, and that of type B is λ^B (where $\lambda^A + \lambda^B = 1$). From (8) we can therefore write

$$a_i^A = \theta E\left(\lambda^B a_j^B + \lambda^A a_j^A \mid \alpha_i^A\right) + (1-\theta)\alpha_i^A \tag{10}$$

which gives after replacing the actions by their expressions as a function of k^A, k^B, δ^A, δ^B, α_i^A and α_i^B

$$a_i^A = \theta\left[\lambda^B\left(k^B E\left(\alpha_j^B \mid \alpha_i^A\right) + \delta^B\right) + \lambda^A\left(k^A E\left(\alpha_j^A \mid \alpha_i^A\right) + \delta^A\right)\right] + (1-\theta)\,\alpha_i^A \tag{11}$$

Now, because the preferences of individuals of different types are uncorrelated, we have

$$E\left(\alpha_j^B \mid \alpha_i^A\right) = E\left(\mu^B + e^B + \varepsilon_j \mid \alpha_i^A\right) = \mu^B \tag{12}$$

The preferences of a generic type A individual do not inform him on the preferences of a generic type B individual, and consequently type A can only predict that the preferences of the type B agents he will be matched with are the average preference, μ^B. In contrast, the preferences of individuals of the same type are correlated. Hence:

$$\begin{aligned} E\left(\alpha_j^A \mid \alpha_i^A\right) &= \mu^A + E\left(e^A \mid \alpha_i^A\right) + E\left(\varepsilon_j \mid \alpha_i^A\right) \\ &= \mu^A + E\left(e^A \mid \alpha_i^A\right) \text{ as } E\left(\varepsilon_j \mid \alpha_i^A\right) = 0 \end{aligned} \tag{13}$$

where the prediction of the type-specific shock to preferences of type A individuals, e^A, is given by the linear regression of e^A against the information available to agent i, namely α_i and μ^A (see Morris and

[24] It is straightforward to verify that, since there is continuum of individuals, all have the same first order condition. Hence, all equilibria must be symmetric.

Shin 2002 or Angeletos and Pavan 2007 (a)),

$$E\left(e^A \mid \alpha_i^A\right) = \frac{1}{\sigma+1}\left(\alpha_i^A - \mu^A\right) \tag{14}$$

Substituting in (11) we obtain

$$a_i^A = \alpha_i^A\left[1-\theta+\frac{\lambda^A k^A}{\sigma+1}\right] + \theta\lambda^B(k^B\mu^B+\delta^B) + \theta\lambda^A k^A\frac{\sigma}{\sigma+1}\mu^A + \theta\delta^A \tag{15}$$

We therefore have.

$$k^A = \left(\frac{\lambda^A}{\sigma+1}\theta k^A + 1-\theta\right) \text{ i.e. } k^A = (1-\theta)\frac{\sigma+1}{\sigma-\theta\lambda^A+1} \tag{16}$$

and

$$\delta^A = \theta\left(\lambda^B(k^B\mu^B+\delta^B) + \lambda^A\left(k^A\frac{\sigma}{\sigma+1}\mu^A+\delta^A\right)\right) \tag{17}$$

Now consider an individual of type B. By analogy, we have

$$E\left(\alpha_j^A \mid \alpha_i^B\right) = \mu^A \tag{18}$$

and

$$E\left(\alpha_j^B \mid \alpha_i^B\right) = \frac{\sigma\mu^B+\alpha_i^B}{\sigma+1} \tag{19}$$

Hence:

$$\begin{aligned} a_i^B &= \theta\left[\lambda^B\left(k^B\left(\frac{\sigma}{\sigma+1}\mu^B+\frac{1}{\sigma+1}\alpha_i^B\right)+\delta^B\right)+\lambda^A\left(k^A\mu^A+\delta^A\right)\right]+(1-\theta)\,\alpha_i^B & (20)\\ &= \alpha_i^B\left[1-\theta+\frac{\theta\lambda^B k^B}{\sigma+1}\right]+\theta\lambda^B k^B\frac{\sigma}{\sigma+1}\mu^B+\theta\delta^B+\theta\lambda^A\left(k^A\mu^A+\delta^A\right) & (21)\end{aligned}$$

It follows that

$$k^B = \left(\frac{\lambda^B}{\sigma+1}\theta k^B + 1-\theta\right) \text{ i.e. } k^B = (1-\theta)\frac{\sigma+1}{\sigma-\theta\lambda^B+1} \tag{22}$$

and

$$\delta^B = \theta\left(\lambda^B\left(k^B\frac{\sigma}{\sigma+1}\mu^B+\delta^B\right)+\lambda^A\left(k^A\mu^A+\delta^A\right)\right) \tag{23}$$

We now have a system, composed of (11) and (23), where k^A and k^B are given by (16) and (22), that allows us to solve for δ^A and δ^B. It is straightforward to verify that, substituting for $\lambda^{t'} = 1-\lambda^t$, the solution of the system gives

$$\delta^t = \mu^{t'}\lambda^{t'}\theta + \mu^t\left(1-k^t-\lambda^{t'}\theta\right)$$

for $t = A, B$ (and $t' = B, A$).

Uniqueness: To prove the uniqueness of the linear equilibrium, we follow the route explored by Morris and Shin (2002) and generalized by Angeletos and Pavan (2007 (a)). In order to make sure that the expected action chosen in equilibrium across agents is finite, we add an extra assumption consisting in assuming that all action belong to the bounded and closed interval $[-M, M]$,

$$a_i \in [-M, M] \text{ for any } i, \tag{24}$$

where M is a positive real number. Moreover we assume that the number of agents interacting is finite, equal to J, and we let J^A and J^B be the number of agents of each type. The proportion λ^t is consequently equal to J^t/J. The proof consists then of two parts: first, under the assumption that $\theta \in (0, 1)$, we show that under some conditions on the distribution of unobserved individual preferences $\alpha^t_{i\,i=1,\ldots,J}$, the game has a unique symmetric equilibrium in which all agents play according to the strategy derived above. Then we show that when M and J go to infinity, the probability of the distribution of preferences under which this equilibrium exists goes to 1, ruling out asymmetric equilibria.

Under the assumption that actions must belong to the closed interval $[-M, M]$, and since $\theta > 0$, the reaction function of a generic individual i is given by

$$a_i = \begin{cases} -M & \text{if } E\left(\overline{a}_j \mid \alpha^t_i\right) \leq -\frac{M}{\theta} - \frac{1-\theta}{\theta}\alpha^t_i \\ \theta E\left(\overline{a}_j \mid \alpha^t_i\right) + (1-\theta)\,\alpha^t_i & \text{if } E\left(\overline{a}_j \mid \alpha^t_i\right) \in \left(-\frac{M}{\theta} - \frac{1-\theta}{\theta}\alpha^t_i, \frac{M}{\theta} - \frac{1-\theta}{\theta}\alpha^t_i\right) \\ M & \text{if } E\left(\overline{a}_j \mid \alpha^t_i\right) \geq \frac{M}{\theta} - \frac{1-\theta}{\theta}\alpha^t_i \end{cases} \tag{25}$$

where $E\left(\overline{a}_j \mid \alpha^t_i\right)$ denotes the expected action of the average individual with whom individual i is matched, given its preferences α^t_i. Given our assumptions on the number of agents, $\overline{a}^t_j$, which denotes the average action of individuals of type t excluding agent i and $\overline{a}^{t'}_j$, which denotes the average action of individuals of type t', are equal to

$$\overline{a}^t_j = \sum_{j=1, j\neq i}^{J^t} a^t_j/(J^t - 1) \ \text{ and } \ \overline{a}^{t'}_j = \sum_{j=1}^{J^{t'}} a^{t'}_j/J^{t'}. \tag{26}$$

Therefore the average action of agents $j \neq i$, which is equal to

$$E\left(\overline{a}_j \mid \alpha^t_i\right) = (\lambda^t - 1/J)E\left(\overline{a}^t_j \mid \alpha^t_i\right) + \lambda^{t'} E\left(\overline{a}^{t'}_j \mid \alpha^t_i\right) \tag{27}$$

simplifies into

$$E\left(\overline{a}_j \mid \alpha^t_i\right) = \frac{1}{J}\sum_{j=1, j\neq i}^{J^t} E\left(a^t_j \mid \alpha^t_i\right) + \frac{1}{J}\sum_{j=1}^{J^{t'}} E\left(a^{t'}_j \mid \alpha^t_i\right). \tag{28}$$

Substituting for the average action of individuals j, and using agents best responses given by the second line of the expression above, action a_i rewrites

$$\begin{aligned}
a_i = \quad & \tfrac{\theta}{J}\textstyle\sum_{j=1,j\neq i}^{J^t} E\left(a_j^t \mid \alpha_i^t\right) + \tfrac{\theta}{J}\sum_{j=1}^{J^{t'}} E\left(a_j^{t'} \mid \alpha_i^t\right) + (1-\theta)\alpha_i^t \\
= \quad & \tfrac{\theta}{J}\textstyle\sum_{j=1,j\neq i}^{J^t} E\left(\theta E(\overline{a}_{-j} \mid \alpha_j^t) + (1-\theta)\alpha_j^t \mid \alpha_i^t\right) \\
& +\tfrac{\theta}{J}\textstyle\sum_{j=1}^{J^{t'}} E\left(\theta E(\overline{a}_{-j} \mid \alpha_j^{t'}) + (1-\theta)\alpha_j^{t'} \mid \alpha_i^t\right) + (1-\theta)\alpha_i^t \\
= \quad & \tfrac{\theta^2}{J}\textstyle\sum_{j=1,j\neq i}^{J^t} E\left(E\left(\overline{a}_{-j} \mid \alpha_j^t\right) \mid \alpha_i^t\right) + \tfrac{\theta(1-\theta)}{J}\sum_{j=1,j\neq i}^{J^t} E\left(\alpha_j^t \mid \alpha_i^t\right) \\
& +\tfrac{\theta^2}{J}\textstyle\sum_{j=1}^{J^{t'}} E\left(E\left(\overline{a}_{-j} \mid \alpha_j^{t'}\right) \mid \alpha_i^t\right) + \tfrac{\theta(1-\theta)}{J}\sum_{j=1}^{J^{t'}} E\left(\alpha_j^{t'} \mid \alpha_i^t\right) + (1-\theta)\alpha_i^t
\end{aligned} \tag{29}$$

As we noted before, $E\left(\alpha_j^{t'} \mid \alpha_i^t\right) = \mu^{t'}$ and $E\left(\alpha_j^t \mid \alpha_i^t\right) = \mu^t + (\alpha_i^t - \mu^t)/(\sigma+1)$. The equation above simplifies therefore into

$$\begin{aligned}
a_i = \; & \tfrac{\theta^2}{J}\textstyle\sum_{j=1,j\neq i}^{J^t} E\left(E\left(\overline{a}_{-j} \mid \alpha_j^t\right) \mid \alpha_i^t\right) + \tfrac{\theta^2}{J}\sum_{j=1}^{J^{t'}} E\left(E\left(\overline{a}_{-j} \mid \alpha_j^{t'}\right) \mid \alpha_i^t\right) \\
& +\tfrac{\theta(1-\theta)}{J}(J^t-1)\left(\tfrac{\alpha_i^t}{\sigma+1} + \tfrac{\sigma\mu^t}{\sigma+1}\right) + \tfrac{\theta(1-\theta)}{J}J^{t'}\mu^{t'} + (1-\theta)\alpha_i^t
\end{aligned} \tag{30}$$

We can reiterate the process of substituting the average action of individuals $\overline{a}_{-j}$ with its expression as a function of private tastes and conditional expectation of the actions individual $-j$ is interacting with ... as it is done in Morris and Shin (2002). In the limit of this substitution process, the action of agent i is equal to the sum of a coefficient that goes to 0 times the conditional expectation of the conditional expectation... up to infinity ... of the average action of individual other than i, plus a combination of μ^t, $\mu^{t'}$ and α_i^t. As all actions must belong to $(-M, M)$ for this computation to hold, we know that the infinite conditional expectation is bounded (it belongs to $(-M, M)$), and therefore the product term that contains it disappears, as it is equal to an element that goes to 0 times an element which is finite. Action a_i is therefore equal to a combination of μ^t, $\mu^{t'}$ and α_i^t, equal to what we found before:

$$a_i^t = k^t\alpha_i^t + \lambda^{t'}\theta\mu^{t'} + (1-\lambda^{t'}\theta - k^t)\mu^t \tag{31}$$

For this symmetric equilibrium to occur, it must be the case that the distribution of unobserved tastes verify

$$\sup_{i,t} a_i^t \leq M \text{ and } \inf_{i,t} a_i^t \geq -M \tag{32}$$

giving

$$\sup_{i,t} \alpha_i^t \leq \min\{(M - \lambda^B\theta\mu^B + (1-\lambda^B\theta - k^A)\mu^A)/k^A, (M - \lambda^A\theta\mu^A + (1-\lambda^A\theta - k^B)\mu^B)/k^B\} \tag{33}$$

and

$$\inf_{i,t} a_i^t \geq \max\{(-M - \lambda^B\theta\mu^B + (1-\lambda^B\theta - k^A)\mu^A)/k^A, (-M - \lambda^A\theta\mu^A + (1-\lambda^A\theta - k^B)\mu^B)/k^B\} \quad (34)$$

If the order statistics do not verify these conditions, which is possible for M and J finite, there may be asymmetric equilibria in which some individuals select the upper or the lower bound of the action space, while some others select an action strictly in between the two bounds. It may also be the case that there exist no equilibrium in pure strategy. We proved none of these two results. However, as the tastes are normality distributed, when M and J go to infinity the order statistics $\sup_{i,t} \alpha_i^t$ and $\inf_{i,t} a_i^t$ verify the two inequalities above with probability 1 (with $M > J$ to avoid the convergence of the order statistic quicker to infinity than the bound of the action space). Therefore the symmetric equilibrium we derived is unique when M and J go to infinity. ■

Proof of Lemma 2: The mean type t–action in a community is

$$\begin{aligned} E\left(a_i^t\right) &= k^t E\left(\alpha_i^t\right) + \theta\lambda^{t'}\mu^{t'} + (1-\theta\lambda^{t'} - k^t)\mu^t \\ &= \left(1-\theta\lambda^{t'}\right)\mu^t + \theta\lambda^{t'}\mu^{t'} \text{ since } E\left(\alpha_i^t\right) = \mu^t \end{aligned} \quad (35)$$

The variance of the action of an individual of type t around mean type t–action in a community is

$$E\left[\left(a_i^t - E\left(a_i^t\right)\right)^2\right] \quad (36)$$

where

$$a_i^t - E\left(a_i^t\right) = k^t\left(\alpha_i^t - \mu^t\right) = k^t\left(e^t + \varepsilon_i\right) \quad (37)$$

Hence, we have

$$E\left[\left(a_i^t - E\left(a_i^t\right)\right)^2\right] = (k^t)^2(1+\sigma) = \left((1-\theta)\frac{\sigma+1}{\sigma-\theta\left(1-\lambda^{t'}\right)+1}\right)^2(1+\sigma) \quad (38)$$

given that $k^t = (1-\theta)\frac{\sigma+1}{\sigma-\theta\left(1-\lambda^{t'}\right)+1}$ from lemma 1. The derivative of the righthandside of (38) with respect to $\lambda^{t'}$ is

$$-2\theta\left(1-\theta\right)^2\frac{(\sigma+1)^3}{\left(\sigma-\theta(1-\lambda^{t'})+1\right)^3} < 0 \quad (39)$$

since, under the assumption that θ and $\lambda^{t'}$ are both in between 0 and 1, the denominator is strictly negative.■

Proof of Corollary 1: The value of $\mid k^t(\lambda^{t'})/d\lambda^{t'} \mid$ is $\frac{(\sigma+1)(1-\theta)\theta}{(\sigma-\theta+\theta\lambda+1)^2}$. Straightforward calculations show that this is strictly decreasing in σ, and it is concave in θ, reaching a maximum at $\theta = \frac{\sigma+1}{2\sigma+\lambda^{t'}+1}$. ■

Proof of Proposition 1: The average action in the community composed of a share λ^A of type A and a share λ^B of type B is

$$\overline{a} = \lambda^B \left(k^B \mu^B + \delta^B\right) + \lambda^A (k^A \mu^A + \delta^A) \tag{40}$$

The variance of actions in the community is

$$E\left[(a_j - \overline{a})^2\right] = \lambda^B E\left[\left(k^B\left(\mu^B + e^B + \varepsilon_j\right) + \delta^B - \overline{a}\right)^2\right] + \lambda^A E\left[\left(k^A\left(\mu^A + e^A + \varepsilon_j\right) + \delta^A - \overline{a}\right)^2\right] \tag{41}$$

Substituting for $\overline{a}$, letting $\lambda^A = 1 - \lambda^B$ and rearranging we obtain

$$\begin{aligned} E\left[(a_j - \overline{a})^2\right] &= \lambda^B E\left[\left(\left(k^B \mu^B + \delta^B - k^A \mu^A - \delta^A\right)(1 - \lambda^B) + k^B\left(e^B + \varepsilon_j\right)\right)^2\right] + \\ &\quad \left(1 - \lambda^B\right) E\left[\left((k^A \mu^A + \delta^A - k^B \mu^B - \delta^B)\lambda^B + k^A\left(e^A + \varepsilon_j\right)\right)^2\right] \end{aligned} \tag{42}$$

Substituting for δ^A and δ^B and rearranging

$$\begin{aligned} E\left[(a_j - \overline{a})^2\right] &= \lambda^B E\left[\left((\mu^B - \mu^A)(1-\theta)(1-\lambda^B) + k^B\left(e^B + \varepsilon_j\right)\right)^2\right] + \\ &\quad \left(1 - \lambda^B\right) E\left[\left(((\mu^A - \mu^B)(1-\theta)\lambda^B + k^A\left(e^A + \varepsilon_j\right)\right)^2\right] \\ &= \lambda^B (1 - \lambda^B)(1-\theta)^2 (\mu^A - \mu^B)^2 + \left[E\left(\varepsilon_j^2\right) + E\left(\left(e^B\right)^2\right)\right] \lambda^B \left(k^B\right)^2 + \\ &\quad \left[E\left(\varepsilon_j^2\right) + E\left(\left(e^A\right)^2\right)\right](1 - \lambda^B)\left(k^A\right)^2 \end{aligned} \tag{43}$$

which finally simplifies to

$$E\left[(a_j - \overline{a})^2\right] = \lambda^B (1 - \lambda^B)(1-\theta)^2 (\mu^A - \mu^B)^2 + (1+\sigma)\left[\lambda^B \left(k^B\right)^2 + (1 - \lambda^B)\left(k^A\right)^2\right] \tag{44}$$

Evaluated at $\lambda^B \to 0$, the derivative of the righthandside of (44) with respect to λ^B is negative when

$$\left(\mu^A - \mu^B\right)^2 < \theta \frac{\sigma + 1}{(\sigma - \theta + 1)^3}\left[(\sigma + 1)(4(\sigma + 1) - 3\theta) + \theta^2\right] \tag{45}$$

Condition (45) ensures that by introducing a small share of individuals of type $t' = B, A$ in a community of individuals of type $t = A, B$, community variance decreases.■

Proof of Lemma 3: Let $\overline{a}_j^t$ denote the average action of individuals of type $t = A, B$ in an interacting group. We have

$$\overline{a}_j^t = k^t \left(\mu^t + e^t\right) + \delta^t \tag{46}$$

for $t = A, B$. This takes care of average actions. What about expectations of average actions? Consider without loss of generality an individual i if type A. We have

$$E\left(\overline{a}_j^B \mid \alpha_i^A\right) = k^B \mu^B + \delta^B \tag{47}$$

Hence, the variance of $\overline{a}_j^B$ around $E\left(\overline{a}_j^B \mid \alpha_i^A\right)$ is equal to

$$Var\left(\overline{a}_j^B \mid \alpha_i^A\right) = \left(k^B\right)^2 \tag{48}$$

Now take $E\left(\overline{a}_j^A \mid \alpha_i^A\right)$. This is given by

$$\begin{aligned} E\left(\overline{a}_j^A \mid \alpha_i^A\right) &= k^A\left(\mu^A + E\left(e^A \mid \alpha_i^A\right)\right) + \delta^A \\ &= k^A\left(\mu^A + \frac{e^A + \varepsilon_i}{\sigma + 1}\right) + \delta^A \end{aligned} \tag{49}$$

Hence, the variance of $\overline{a}_j^A$ around $E\left(\overline{a}_j^A \mid \alpha_i^A\right)$ is equal to

$$Var\left(\overline{a}_j^A \mid \alpha_i^A\right) = \left(k^A\right)^2 \frac{\sigma}{\sigma + 1} \tag{50}$$

Consider a homogeneous community, composed only of type A. Substituting for k^A, we have

$$Var\left(\overline{a}_j^A \mid \alpha_i^A\right)_{\lambda^B = 0} = \frac{\sigma (1-\theta)^2 (\sigma + 1)}{(\theta - \sigma - 1)^2} \tag{51}$$

Now consider an heterogeneous community. Suppose that $\lambda^B = \lambda > 0$. From (48), and substituting for $\left(k^B\right)^2$, we see that the ease with which i can forecast the actions of people of type B is given by

$$Var\left(\overline{a}_j^B \mid \alpha_i^A\right)_{\lambda^B = \lambda} = \frac{(1-\theta)^2 (\sigma + 1)^2}{(\theta\lambda - \sigma - 1)^2} \tag{52}$$

Now,

$$\frac{d\left(Var\left(\overline{a}_j^A \mid \alpha_i^A\right)_{\lambda^B=0} - Var\left(\overline{a}_j^B \mid \alpha_i^A\right)_{\lambda^B=\lambda}\right)}{d\lambda} = -\frac{2\theta (1-\theta)^2 (\sigma+1)^2}{(1 + \sigma - \theta\lambda)^3} < 0 \tag{53}$$

implying that, as λ increases, $Var\left(\overline{a}_j^A \mid \alpha_i^A\right)_{\lambda^B=0} - Var\left(\overline{a}_j^B \mid \alpha_i^A\right)_{\lambda^B=\lambda}$ may only change sign from above. Note that

$$\lim_{\lambda \to 1}\left(Var\left(\overline{a}_j^A \mid \alpha_i^A\right)_{\lambda^B=0} - Var\left(\overline{a}_j^B \mid \alpha_i^A\right)_{\lambda^B=\lambda}\right) = -(\theta - 1)^2 \frac{\sigma + 1}{(\sigma - \theta + 1)^2} < 0 \tag{54}$$

and

$$\lim_{\lambda \to 0}\left(Var\left(\overline{a}_j^A \mid \alpha_i^A\right)_{\lambda^B=0} - Var\left(\overline{a}_j^B \mid \alpha_i^A\right)_{\lambda^B=\lambda}\right) = \frac{(\theta - 1)^2\left[(2\theta - 1)(\sigma + 1) - \theta^2\right]}{(\sigma - \theta + 1)^2} \tag{55}$$

If $\theta < \sigma + 1 - \sqrt{\sigma + \sigma^2}$ the righthandside of (55) is negative, implying that $Var\left(\overline{a}_j^A \mid \alpha_i^A\right)_{\lambda^B=0} < Var\left(\overline{a}_j^B \mid \alpha_i^A\right)_{\lambda^B=\lambda}$ for all values of λ. If $\theta > \sigma + 1 - \sqrt{\sigma + \sigma^2}$ the righthandside of (55) is positive, implying that there exists a value $\widehat{\lambda}$ such that $Var\left(\overline{a}_j^A \mid \alpha_i^A\right)_{\lambda^B=0} > (<)\ Var\left(\overline{a}_j^B \mid \alpha_i^A\right)_{\lambda^B=\lambda}$ for $\lambda < (>)$ $\widehat{\lambda}$. This proves the result.■

Proof of Lemma 4: Again, without loss of generality consider an individual i if type A. Following the same procedure as above, we may derive $Var\left(\overline{a}_j^A \mid \alpha_i^A\right) = \left(k^A\right)^2 \frac{\sigma}{\sigma+1}$. From lemma 1, we know that $k^A = (1-\theta)\frac{\sigma+1}{\sigma-\theta\lambda^A+1} = (1-\theta)\frac{\sigma+1}{\sigma-\theta(1-\lambda^B)+1}$, given $\lambda^A = 1-\lambda^B$. Since k^A is decreasing in λ^B, the first result follows. The second result is proved above (proof of lemma 3).■

Proof of Lemma 5: The amount of strategic uncertainty faced by individual i, $Var\left(\overline{a}_j \mid \alpha_i\right)$, is given by

$$E\left[E\left(\left(\overline{a}_j - E\left(\overline{a}_j \mid \alpha_i\right)\right)^2 \mid \alpha_i\right)\right] \tag{56}$$

Now, since $\lambda^A = 1-\lambda^B$, $\overline{a}_j - E\left(\overline{a}_j \mid \alpha_i\right)$ can be written as

$$\left(1-\lambda^B\right)\left(\overline{a}_j^A - E\left(\overline{a}_j^A \mid \alpha_i\right)\right) + \lambda^B\left(\overline{a}_j^B - E\left(\overline{a}_j^B \mid \alpha_i\right)\right) \tag{57}$$

Hence, we have

$$Var\left(\overline{a}_j \mid \alpha_i\right) = \left(1-\lambda^B\right)^2 Var\left(\overline{a}_j^A \mid \alpha_i\right) + \left(\lambda^B\right)^2 Var\left(\overline{a}_j^B \mid \alpha_i\right) \tag{58}$$

Suppose that the behavior of individuals is fixed with respect to community composition, so that $Var\left(\overline{a}_j^t \mid \alpha_i\right)$, $t = A, B$ is independent of λ^B. Then

$$\lim_{\lambda^B \to 0} \frac{dVar\left(\overline{a}_j \mid \alpha_i\right)}{d\lambda^B} = -2Var\left(\overline{a}_j^A \mid \alpha_i\right) < 0 \tag{59}$$

Keeping the behavior of individuals constant, introducing an infinitesimally small fraction of type B in a community composed only of type A decreases aggregate strategic uncertainty. The same argument applies for an infinitesimally small fraction of type A in a community composed only of type B.■

Proof of Proposition 2: We wish to prove that, for an individual of type t, $E\left[E\left(\left(\overline{a}_j - E\left(\overline{a}_j \mid \alpha_i\right)\right)^2 \mid \alpha_i\right)\right]$ is strictly convex in $\lambda^{t'}$ and reaches a minimum at $\lambda^{t'} = \lambda^* \in (0, 0.5)$. Suppose without loss of generality that individual i is of type A, and denote as λ^B the share of individuals of type B in the community. We then have

$$\overline{a}_j = \lambda^B\left(k^B\left(\mu^B + e^B\right) + \delta^B\right) + \left(1-\lambda^B\right)\left(k^A\left(\mu^A + e^A\right) + \delta^A\right)$$

and

$$\begin{aligned}E\left(\overline{a}_j \mid \alpha_i^A\right) &= \lambda^B\left(k^B\mu^B + \delta^B\right) + (1-\lambda^B)\left(\tfrac{k^A}{\sigma+1}\left(\sigma\mu^A + \alpha_i^A\right) + \delta^A\right) \\ &= \lambda^B\left(k^B\mu^B + \delta^B\right) + (1-\lambda^B)\left(k^A\left(\mu^A + \tfrac{e^A+\varepsilon_i}{\sigma+1}\right) + \delta^A\right) \\ &\quad \text{after substituting for } \alpha_i^A = \mu^A + e^A + \varepsilon_i\end{aligned} \tag{60}$$

Hence, we can write

$$\overline{a}_j - E\left(\overline{a}_j \mid \alpha_i^A\right) = \lambda^B k^B e^B + \frac{1-\lambda^B}{\sigma+1}k^A(\sigma e^A - \varepsilon_i) \tag{61}$$

so that

$$E\left[E\left(\left(\overline{a}_j - E\left(\overline{a}_j \mid \alpha_i^A\right)\right)^2 \mid \alpha_i^A\right)\right] = \left(\lambda^B\right)^2\left(k^B\right)^2 + \frac{\left(1-\lambda^B\right)^2\left(k^A\right)^2\sigma}{\sigma+1} \tag{62}$$

Now,

$$\frac{d^2\left(\left(\lambda^B\right)^2\left(k^B\right)^2\right)}{d\left(\lambda^B\right)^2}=2\left(1-\theta\right)^2\frac{\left(\sigma+1\right)^3}{\left(\sigma-\theta\lambda^B+1\right)^4}\left(\sigma+2\theta\lambda^B+1\right)>0 \tag{63}$$

and

$$\frac{d^2\left(\left(1-\lambda^B\right)^2\left(k^A\right)^2\sigma\right)}{d\left(\lambda^B\right)^2}=2\sigma\left(1-\theta\right)^2\frac{\left(\sigma+1\right)^3}{\left(\sigma-\theta(1-\lambda^B)+1\right)^4}\left(1+\sigma+2\theta\left(1-\lambda^B\right)\right)>0 \tag{64}$$

Hence, $E\left[E\left(\left(\overline{a}_j-E\left(\overline{a}_j\mid\alpha_i\right)\right)^2\mid\alpha_i^A\right)\right]$ is undoubtedly convex in λ^B. Moreover:

$$\lim_{\lambda^B\to 0}\frac{d\left(\left(\lambda^B\right)^2\left(k^B\right)^2+\frac{\left(1-\lambda^B\right)^2\left(k^A\right)^2\sigma}{\sigma+1}\right)}{d\lambda^B}=-2\sigma\left(1-\theta\right)^2\frac{\left(\sigma+1\right)^2}{\left(\sigma-\theta+1\right)^3}<0 \tag{65}$$

and

$$\lim_{\lambda^B\to 0.5}\frac{d\left(\left(\lambda^B\right)^2\left(k^B\right)^2+\frac{(1-\lambda)^2\left(k^A\right)^2\sigma}{\sigma+1}\right)}{d\lambda^B}=8\left(\sigma+1\right)^2\frac{\left(1-\theta\right)^2}{\left(2\sigma-\theta+2\right)^3}>0 \tag{66}$$

This proves that, for an individual of type A, $E\left[E\left(\left(\overline{a}_j-E\left(\overline{a}_j\mid\alpha_i\right)\right)^2\mid\alpha_i^A\right)\right]$ reaches a minimum at $\lambda^B=\lambda^*\in(0,0.5)$. By symmetry, the argument can be applied that, for an individual of type B, $E\left[E\left(\left(\overline{a}_j-E\left(\overline{a}_j\mid\alpha_i\right)\right)^2\mid\alpha_i^B\right)\right]$ reaches a minimum at $\lambda^A=\lambda^*\in(0,0.5)$.■

Proof of Proposition 3: A sufficient condition for the proposition to hold is that both types are strictly better off in a mixed community with 50/50 type composition than in a homogeneous community. Consider an individual i of type A. From (62), this is the case if

$$\lim_{\lambda^B\to 0}\left(\left(\lambda^B\right)^2\left(k^B\right)^2+\frac{\left(1-\lambda^B\right)^2\left(k^A\right)^2\sigma}{\sigma+1}\right)-\lim_{\lambda^B\to 0.5}\left(\left(\lambda^B\right)^2\left(k^B\right)^2+\frac{\left(1-\lambda^B\right)^2\left(k^A\right)^2\sigma}{\sigma+1}\right)>0 \tag{67}$$

Similarly, by symmetry, this holds for an individual i of type B if

$$\lim_{\lambda^A\to 0}\left(\left(\lambda^A\right)^2\left(k^A\right)^2+\frac{\left(1-\lambda^A\right)^2\left(k^B\right)^2\sigma}{\sigma+1}\right)-\lim_{\lambda^A\to 0.5}\left(\left(\lambda^A\right)^2\left(k^A\right)^2+\frac{\left(1-\lambda^A\right)^2\left(k^B\right)^2\sigma}{\sigma+1}\right)>0 \tag{68}$$

Given perfect symmetry between the two types, (67) and (68) are also necessary for the proposition to hold. Solving out, we obtain

$$\begin{aligned}
&\lim\nolimits_{\lambda^B\to 0}\left(\left(\lambda^B\right)^2\left(k^B\right)^2+\tfrac{\left(1-\lambda^B\right)^2\left(k^A\right)^2\sigma}{\sigma+1}\right)-\lim\nolimits_{\lambda^B\to 0.5}\left(\left(\lambda^B\right)^2\left(k^B\right)^2+\tfrac{\left(1-\lambda^B\right)^2\left(k^A\right)^2\sigma}{\sigma+1}\right)\\
&=\lim\nolimits_{\lambda^A\to 0}\left(\left(\lambda^A\right)^2\left(k^A\right)^2+\tfrac{\left(1-\lambda^A\right)^2\left(k^B\right)^2\sigma}{\sigma+1}\right)-\lim\nolimits_{\lambda^A\to 0.5}\left(\left(\lambda^A\right)^2\left(k^A\right)^2+\tfrac{\left(1-\lambda^A\right)^2\left(k^B\right)^2\sigma}{\sigma+1}\right)\\
&=\left(1-\theta\right)^2\tfrac{(\sigma+1)^2\left(2\theta+\sigma-\theta^2+2\sigma^2-1\right)}{\left(\sigma-\theta+1\right)^2\left(2\sigma-\theta+2\right)^2}>0\text{ if }\left(2\theta+\sigma-\theta^2+2\sigma^2-1\right)>0
\end{aligned} \tag{69}$$

It is straightforward to verify that $\left(2\theta+\sigma-\theta^2+2\sigma^2-1\right)>0$ if $\theta>1-\sqrt{\sigma+2\sigma^2}$.■