Discussion Papers in Economics

On the Strategic Benefits of Diversity

By

Martin Kaae Jensen
(University of Surrey).

DP 14/19

School of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web https://www.surrey.ac.uk/school-economics
ISSN: 1749-5075

On the Strategic Benefits of Diversity*

Martin Kaae Jensen†

July 18, 2019

Abstract

This paper studies the relationship between functional diversity and team performance. The main question is whether diversity may entail strategic benefits that enable diverse teams to outperform homogenous teams even if the homogenous teams are more skilled on average, or diversity entails a direct efficiency loss a la Benabou (1996). Both ability diversity and cognitive diversity (Johnson-Laird (1983), Page (2008)) are studied, and the paper also considers the role of Becker and Murphy (1992)-type coordination costs. In all cases, the main message is that effort adjustments set off by greater diversity may significantly change the outcome in comparison with an assessment based on the more familiar direct effects. For example, a diverse team may outperform a homogenous team even if the elasticity of substitution is positive and less capable individuals therefore "drag down" the more capable individuals productivities; and under the same condition, a "superstar" may outperform a cognitively diverse team even though a positive elasticity of substitution implies decreasing returns to talent in the sense of Rosen (1981). The paper discusses the implications of these findings for the general diversity debate, for optimal team selection, and for market salaries. The main insights, as well as the tools developed to reach those insights, are very general and extend to other contexts where diversity plays a role.

Keywords: Functional diversity, ability diversity, cognitive diversity, extrinsic value theory of diversity, performance indices, teamwork, does diversity trump ability, aggregative games.

JEL Classification: C72, D40, D80, M10, Z13.

---

*I would like to thank Andre Cunha, Kevin Reffett, Esra Kaytaz, and various conference participants for very useful input and discussion. I am especially thankful to Daron Acemoglu whose comments and suggestions very significantly impacted the paper. As usual, such acknowledgments imply no responsibility for biases, errors or omissions.

†School of Economics, University of Surrey.

1 Introduction

This paper studies the relationship between functional diversity and team performance. Simply put, the question is whether diversity may bestow strategic benefits which enable diverse teams to outperform homogenous ones even if the homogenous teams are more skilled on average, or diversity entails a direct efficiency loss due, for example, to a Benabou (1996)-type complementarity. This issue ties in with a large strategic management literature on the functional benefits of diversity (*e.g.* Roland and Galunic (2004), Horwitz and Horwitz (2007), Reynolds and Lewis (2017)), and also with the general public debate and growing cross-disciplinary literature on whether diversity "trumps" ability or the other way around (*e.g.*, Hong and Page (2004), Page (2008), Zollman (2010), Landemore (2012), Thompson (2014), Thoma (2015), Brennan (2017), Singer (2019)).

To appreciate the basic mechanisms, imagine two individuals who contribute to a joint project. For example, think of a micro- and a macroeconomist who coauthor a paper, or a clerical worker and an engineer who produce a physical good. When the individuals' efforts are equated with their labor inputs and determined exogenously to the teamwork process itself, the functional role of diversity depends exclusively on the production technology. In the simplest case, the technology is linear in ability adjusted efforts (effective labor inputs), hence output only depends on the average ability and diversity plays no functional role. If the elasticity of substitution is positive, however, then diversity matters: A less able microeconomist will "drag down" the productivity of a more able macroeconomist (Benabou (1996), p.587). In this case, the more diverse the two individuals' abilities are, the lower is the output. If the team is "cognitively homogenous", consisting of two macroeconomist and no microeconomist, say, it will likewise suffer an efficiency loss since two macroeconomists equal one macroeconomist with the combined ability, and no microeconomist equals a microeconomist with zero ability.

With a positive elasticity of substitution, then, less diverse abilities and greater cognitive diversity entail efficiency gains; and so the ideal team is ability-wise balanced and cognitively diverse (a micro- and a macroeconomist who are equally skilled).[1]

If the elasticity of substitution is instead negative a la Rosen (1981) (see also Jones (2011)), then more diverse abilities and less cognitive diversity entail efficiency gains. Then a single-authoring "superstar" will outperform a team consisting of a less capable macroeconomist and a very capable microeconomist who will, in turn, outperform a macro- and a microeconomist of equal ability.

This paper's main message is that the previous description misses an important aspect of diversity's role in teams by equating efforts with exogenously determined ("fixed") labor supplies.[2] Fixed efforts is an adequate description of very large teams or macroeconomic relationships (*e.g.*, Benabou (1996)), but it is hardly a good description of academic coauthors, executive teams, or

[1] Of course in reality, this may not be feasible — for example, a clerical worker may never be as productive as an engineer. But the conclusion remains the same: the less (functionally) diverse the team, the greater the output.

[2] In particular, the models used in the cross-disciplinary literature cited a moment ago does this by treating individuals' "mechanistically" (*e.g.*, Hong and Page (2004) and Singer (2019)).

smaller work units. Academics, executives, and other "problem-solvers" allocate their time across multiple (joint) projects as well as other necessary tasks; workers more generally may spend 8 hours on the job and work effectively for only 4; soldiers may push to the front line or they may run away. In each case, the effort decision is endogenous, and depends on own ability, (opportunity) costs, as well as the expected gain. The gain depends, in turn, on everyone else's ability adjusted efforts and therefore teamwork is fundamentally a strategic endeavor guided by incentives.

Taking effort decisions into account, one would still expect less capable individuals to contribute less to teams' success in ability-adjusted terms. The question is whether, in equilibrium, the more capable individuals will over- or under-compensate for this. What this paper does is identify simple and economically interpretable conditions for both of these possibilities and investigate when such over- or under-compensation dominate the more familiar direct efficiency effects described a moment ago. The paper also defines and studies cognitive diversity and shows how the different types of diversity relate and jointly determine the outcome.

The endogenous efforts perspective has important implications for how we assess what Nehring and Puppe (2002) aptly call diversity's "use value" (p.1156); and it also has implications for our understanding of functional diversity more generally. Here are the key findings, focusing first on diverse abilities, and then on cognitive diversity. More specialized findings are discussed along the way in the paper and summarized in the concluding section.

In strategic versions of Benabou (1996) as well as the natural generalizations in the spirit of *e.g.* Holmstrom (1982) or Becker and Murphy (1992), diverse abilities *always* bestows strategic benefits on teams under standard convex cost assumptions. Thus more capable individuals systematically *over*-compensate when teamed up with less capable colleagues, and therefore more diverse teams will outperform less diverse teams unless diversity entails a sufficiently "crippling" direct efficiency loss. In particular, diversity may improve teams' performance even if the elasticity of substitution is positive and less capable individuals thus "drag down" the productivity of more capable ones. Absent sufficiently strong direct efficiency losses, it also follows that homogenous teams must be more skilled on average to perform as well as diverse teams (hence ability-diversity may "trump" average ability).

If (ability-adjusted) effort costs are instead concave, there will be teamwork equilibria where diversity entails a strategic performance loss, and the previous conclusions are then reversed. For example, an ability-wise balanced team may outperform a "superstar" even if the elasticity of substitution is negative. But even in this case, there will always be at least one equilibrium where diversity provides a performance boost.

The magnitudes are as interesting as the directions. In the coauthor model sketched a moment ago, I give an example where under fairly innocuous parameter values, a diverse team must be 3% more able on average to match a homogenous team due to a Benabou-complementarity (specifically, the less capable microeconomist impedes the macroeconomist's productivity). Taking incentives and strategic interaction into account, this outcome is reversed and the homogenous team must now be $21\ \%$ more able to match the diverse team. So the strategic performance boost completely overpowers the direct efficiency loss and makes diversity strongly beneficial even though the more familiar logic based on direct efficiency effects implies that diversity is (functionally) disadvantageous. The strategic performance boost also has the capacity to overwhelm the direct efficiency effect when the two pull in the same direction. An illustration of this, inspired by the Duke of Wellington's decision to combine weak and strong forces at the Battle of Waterloo, is presented in Section 5.1.

Concerning cognitive diversity, I show that when the elasticity of substitution is positive and cognitive diversity therefore entails direct efficiency gains, a maximally cognitively diverse team will outperform an entirely homogenous high-ability team as long as the "ability gap" is not too great and Becker and Murphy (1992)-type coordination costs not too prohibitive. This, roughly, amounts to a reconfirmation of Hong and Page (2004)'s "diversity trumps ability theorem" in a strategic setting with coordination costs. The intuition is that "deleting" a cognitive group — which is done at least once when a maximally diverse team is compared with an entirely homogenous team — imposes a direct efficiency loss on the cognitively homogenous, high-ability team without offering it a way to compensate strategically along the lines described before.

However, with endogenous efforts, quantity (a cognitive group's total effort) is *not* a perfect substitute for ability — and even when cognitive diversity entails direct efficiency gains, it is therefore possible that a cognitively homogenous team with diverse abilities outperforms a cognitively diverse team with homogenous abilities. In particular, a single "superstar" may outperform an ability-wise balanced, cognitively diverse team. Again this contrasts with the conclusions we get with fixed efforts: Then a "superstar" can only outperform a cognitively diverse team if the elasticity of substitution is negative as in Rosen (1981) and cognitive diversity therefore entails a direct efficiency loss.

Again the magnitudes are interesting. I give an example with no coordination costs where the direct efficiency gain to cognitive diversity is reduced by half because the individuals are not equally able. In the "diversity trumps ability" result just described, it is also interesting that the maximum ability gap can be very large indeed if the cognitive division is fine enough, despite sizable coordination costs. For example, I calculate, again under fairly unimposing parametric restrictions, that with 100 cognitive groups — think "crowdsourcing" or "peer-production" — a cognitively homogenous team must be roughly 200.000 times smarter on average to beat a maximally cognitively diverse team, although the teams are of the same size and each cognitive group increases coordination costs by 10 percentage points. With more draconian coordination costs on the other hand, quadratic in the number of groups for example, enough disincentives are caused for the resulting equilibrium adjustments to ruin the direct efficiency gains, no matter how large they are. This is very intuitive: Even if the direct efficiency gains to, say, an anthropologist and an economist coauthoring a paper are very substantial, the two will not produce a successful piece of work if they face insurmountable effort costs because they find it exceedingly difficult to communicate their ideas to each other.

These findings' significance for the functional diversity debate is obvious: If we treat individuals "mechanistically" and abstract from teams' social interactions, as do the "diversity literature" referenced above and again in a moment, we risk getting both the directions and the magnitudes wrong when assessing the costs and benefits of diversity. As mentioned already, the errors' magnitudes may be of the first order.

In Section 2 and again in more detail in Section 5.1, I also discuss the significance for optimal team selection in market settings, including the relationship with market salaries, Rosen (1981) and the associated "superstar literature" (which is in turn closely related to the literature on stratification and especially to Benabou (1996) already mentioned several times previously). Aside from labor, organizational, and institutional economics, the analysis also overlaps with a much wider economics literature because it directly involves a number of fundamental economic concepts such as "returns to talent", "workforce complementarity", "returns to specialization", and "returns to variety" (*e.g.*, Ethier (1982), Romer (1990), Becker and Murphy (1992), Jones (2011)).

Why then do diverse teams receive a strategic performance boost? The next section's main purpose is to answer this as simply and intuitively as possible. Nonetheless, to give the main intuition already here, imagine we begin in teamwork equilibrium with an entirely homogenous team where everyone exerts the same effort. Fix this level of effort and subject the (identical) abilities to a mean-preserving spread, thus "replacing" the homogenous team with a more ability-diverse comparison team. The direct efficiency effects now kick in and reduce or increase output/performance as appropriate. If efforts are taken to be exogenous, the analysis stops here. With endogenous effort formation, however, identical efforts cannot be a teamwork (Nash) equilibrium of the diverse team because with convex costs, in ability adjusted terms the more capable individuals' marginal contributions will be greater than the less capable individuals'.

In response, the more capable individuals will increase their efforts and the less capable ones will decrease theirs until the marginal contributions are equal (with convex costs, this is one way to define a teamwork equilibrium). Crucially, because these adjustments' impact *on the team's* performance are weighted by the abilities, the more capable individuals' adjustments overpower the less capable individuals' in terms of the team's output. Thus the more capable individuals end up "overcompensating" conditional on the direct efficiency effects. Consider a football team: A very good striker will pursue any ball lost if he conditions on the defense being weak. Meanwhile, a weak defender might "slack off" if he conditions on the attack being strong, but the impact of this is less significant for the team because he is already a poor footballer. Similarly, a brilliant executive will double check every angle if the board contains executives who are more prone to misjudgments — and because she is brilliant, this is very valuable to the executive team and goes unmatched by less capable executives' input reductions. This insight is quite general. A crack soldier will stand firm if he cannot depend on the next person holding the line, an excellent macroeconomist will carefully check a less skilled microeconomist's reasoning, etc., etc.

Note that because the strategic performance boost is conditional on the direct efficiency effects, it may still be overpowered by a direct efficiency loss. But equally, as mentioned, it may not.

The broader significance of these findings is discussed throughout the paper, but briefly, it derives from the policy and related implications where, to be sure, a lot is at stake. The mentioned "diversity trumps ability theorem" of Hong and Page (2004) and Page (2008) has been cited in a brief to the Supreme Court of the United States supporting the promotion of diversity in the armed forces, in support of a diversity requirement by the UCLA as well as *e.g.* by NASA and the US Geological Survey.[3] The management branch of the functional diversity literature has influenced best practice so immensely that the reader will almost certainly be familiar with the consequences (see *e.g.* Roland and Galunic (2004), Reynolds and Lewis (2017), and especially the empirical meta-study and survey by Horwitz and Horwitz (2007)). In political science, the issues have even set off a debate about the vices and virtues of democracy (Landemore (2012), Brennan (2017)).

To this debate, which is about direct efficiency effects, the paper adds the incentivized and strategic perspectives inherent in a more economic approach. But equally, the paper introduces this debate into economics.

The organization is as follows: Section 2 considers an illustration while paying particular attention to the key intuitions and fundamental insights. Section 3 contains the game-theoretic basics and key abstract definitions while Section 4 contains the general results on the relationship between diversity and outcomes. The paper has been structured so that, in a first reading, it is possible to skip directly from Section 2 to Section 5, which contains the main discussion of diversity and team performance. Section 5 also contains the main conclusions, while the final section discusses more broadly and points out some directions for future work.

2 An Illustration

A macroeconomist and a microeconomist work on a joint project. Their efforts are $s_1, s_2 \geq 0$ and their abilities $c_1, c_2 > 0$. The team production function is $[(c_1 s_1)^{1+\alpha} + (c_2 s_2)^{1+\alpha}]^{1/(1+\alpha)}$ and the project succeeds with probability $P([(c_1 s_1)^{1+\alpha} + (c_2 s_2)^{1+\alpha}]^{1/(1+\alpha)}) \in [0,1]$ where P is a smooth, strictly increasing and strictly concave success probability function.

Following first Benabou (1996), let efforts be exogenously determined ("fixed") and equal to the labor supply $L > 0$. Then compare a homogenous team $c_1 = c_2 = \bar{c}$ with a diverse team $\tilde{c}_1 \neq \tilde{c}_2$ that is equally capable on average, $\bar{c} = (\tilde{c}_1 + \tilde{c}_2)/2$. If the homogenous team puts in efforts $s_1 = s_2 = L$, its probability of succeeding is $P([c_1^{1+\alpha} + c_2^{1+\alpha}]^{1/(1+\alpha)} L)$. To match this, the diverse

[3] For the first two citations, see Singer (2019), p.179. The last two are mentioned by Thompson (2014), p.1029. Note also that in terms of the narrowly related public debate, a search for *does diversity trump ability* returns 18.8 million pages at the moment of writing.

team must exert effort

$$\tilde{L} = \frac{[2(\bar{c})^{1+\alpha}]^{1/(1+\alpha)}}{[\tilde{c}_1^{1+\alpha} + \tilde{c}_2^{1+\alpha}]^{1/(1+\alpha)}} \cdot L \,. \tag{1}$$

If $\alpha < 0$, then by Jensen's inequality $([2(\bar{c})^{1+\alpha}]^{1/(1+\alpha)})/([\tilde{c}_1^{1+\alpha}+\tilde{c}_2^{1+\alpha}]^{1/(1+\alpha)}) > 1$. So the diverse team must put in strictly greater effort to match the homogenous team. This *direct efficiency loss* to diversity arises because the elasticity of substitution $\sigma = -1/\alpha$ is positive, and therefore the low productivity of the microeconomist "drags down" the productivity of the macroeconomist (Benabou (1996), p.587). For example, if $\alpha = -0.5$ and the macroeconomist is three times as able as the microeconomics, the diverse team must work $2^2/[1.5^{0.5} + 0.5^{0.5}]^2 - 1 = 7\ \%$ more than the homogenous team to offset this loss. It is also easy to see that the closer the elasticity substitution is to 0 (the smaller α is), the larger the efficiency loss; intuitively, because the weaker author's productivity drag gets bigger if the stronger author has difficulties stepping in and substituting.

If $\alpha > 0$ then the elasticity of substitution is negative, and so the more able macroeconomist "pulls up" the productivity of the weaker microeconomist (Rosen (1981), Jones (2011)). So $([2(\bar{c})^{1+\alpha}]^{1/(1+\alpha)})/([\tilde{c}_1^{1+\alpha} + \tilde{c}_2^{1+\alpha}]^{1/(1+\alpha)}) < 1$, hence $\tilde{L} < L$, and diversity therefore entails a *direct efficiency gain*.

The final case is when the team production function is linear ($\alpha = 0$) and diversity does not impact performance.

Comparing teams that exert the *same* effort L, we conclude:

Observation 1 *Given fixed efforts, the diverse team will outperform the homogenous team if $\alpha > 0$, the homogenous team will outperform the diverse team if $\alpha < 0$, and if $\alpha = 0$ the teams are equally likely to succeed.*

Note that the focus here is entirely on abilities. Cognitive diversity is returned to in Section 5.2.

As will be shown later, Observation 1's statements remain valid if the word "homogenous" is replaced with "less diverse". For example, a clerical worker may never be as able as an engineer (Benabou (1996), p.587) and the relevant comparison is then between a less diverse team (a team with a relatively more able clerical worker) and a more diverse team (a team with a relatively more able engineer). Note that in addition to such upper bounds on specific individuals' abilities, we would typically also want to impose lower bounds to capture that teams might not function if an individual is "too inept". This limits how diverse teams can be just as upper bounds limit how homogenous they can be.

Let us now consider what happens when efforts are endogenous. To simplify, these are treated as fully endogenous but it is best to think of them as endogenous conditional on the exogenously determined labor supplies ($s_i = \omega_i(L)$ where ω_i maps "hours at work" to "hours actually worked"). As discussed in the Introduction, the distinction between s_i and L could simply reflect that only a fraction of time at work is actually spent productively rather than by the water cooler, or it could reflect individuals' discretion over how much they focus or concentrate.

Both individuals value success equally highly $v > 0$, incur the same constant marginal costs $\gamma > 0$, take the coauthor's effort as given, and maximize

$$vP([(c_1s_1)^{1+\alpha} + (c_2s_2)^{1+\alpha}]^{1/(1+\alpha)}) - \gamma s_i,\ i = 1, 2\ .$$

Note that it suffices to take either the colleague's ability-adjusted effort c_js_j, or her effective input into the project $(c_js_j)^{1+\alpha}$ as given which is fortunate because c_j and s_j are likely not individually observable (it is difficult to tell if someone has proved a theorem because she thought very hard or because her skills enabled her to prove it with little effort).

In teamwork equilibrium the two individuals thus provide the *same marginal contribution* to the project, γ. If we express decisions in terms of the effective inputs $\hat{s}_i = (c_is_i)^{1+\alpha}$, teamwork equilibrium instead involves equalization of the marginal effective input costs (this perspective turns out to be useful in some cases). As we shall see, there are also other interpretations, including "cooperative" ones.

From the first-order conditions, one immediately obtains the ratio of the equilibrium efforts $s_1^*/s_2^* = (c_2/c_1)^{(1+\alpha)/\alpha}$. Inserting back into an individual first-order condition then yields $v \cdot [c_1^{-(1+\alpha)/\alpha} + c_2^{-(1+\alpha)/\alpha}]^{-\alpha/(1+\alpha)}P'([(c_1s_1^*)^{1+\alpha} + (c_2s_2^*)^{1+\alpha}]^{1/(1+\alpha)}) = \gamma$. Hence the larger $[c_1^{-(1+\alpha)/\alpha} + c_2^{-(1+\alpha)/\alpha}]^{-\alpha/(1+\alpha)}$ is, the more likely a team is to succeed. Using again Jensen's inequality we conclude:

Observation 2 *In teamwork equilibrium, the diverse team will outperform the homogenous team if $\alpha > -1/2$, the homogenous team will outperform the diverse team if $\alpha < -1/2$, and if $\alpha = -1/2$ the teams are equally likely to succeed.*

In terms of the elasticity of substitution, the diverse team thus outperforms the homogenous team if $\sigma > 2$. As shown a moment ago, diversity entails a direct efficiency loss if $\sigma > 0$. Thus diversity trumps homogeneity provided the individual tasks are not too specialized and the direct efficiency loss therefore not "too severe". As I now explain, the reason is that diversity provides a *strategic performance boost* which counteracts the direct efficiency loss.

The argument simplifies considerably by noting that teamwork equilibrium can also be viewed as the outcome of the individuals jointly maximizing

$$vP([(c_1s_1)^{1+\alpha} + (c_2s_2)^{1+\alpha}]^{1/(1+\alpha)}) - \gamma \cdot (s_1 + s_2)\ . \tag{2}$$

As discussed in the footnote, this observation also has implications for the foundation of the equilibrium concept.[4] If (s_1^*, s_2^*) is a teamwork equilibrium, it therefore minimizes (total) costs

[4] As seen, if the individuals maximize utilitarian social welfare cooperatively, *i.e.*, if they maximize $2vP([(c_1s_1)^{1+\alpha} + (c_2s_2)^{1+\alpha}]^{1/(1+\alpha)}) - \gamma \cdot (s_1 + s_2)$, then the only difference is that the valuations are doubled. So nothing changes provided the efforts remain interior to the feasible sets (which they might not, of course, but that is insubstantial because this could equally well happen if the valuation is v).

$\gamma \cdot (s_1 + s_2)$ conditional on the equilibrium output Y^*. The "cost function" is

$$\mathcal{C}(Y; c_1, c_2) = \min_{\{(s_1,s_2)\geq 0: Y \leq [(c_1 s_1)^{1+\alpha} + (c_2 s_2)^{1+\alpha}]^{1/(1+\alpha)}\}} \gamma \cdot (s_1 + s_2) \ . \tag{3}$$

This (dual) perspective is both very useful and very intuitive. It is intuitive because it equates teamwork with utilitarian cost sharing which may be interpreted as the "essence" of team behavior in this model. The perspective is useful since by a standard comparative statics argument, costs are one-to-one with performance: If a team A with abilities (c_1, c_2) produce team output Y^* in equilibrium, then a team B with abilities $(\tilde{c}_1, \tilde{c}_2)$ will produce $\tilde{Y}^* > Y^*$ in equilibrium if and only if $\mathcal{C}(Y^*; \tilde{c}_1, \tilde{c}_2) < \mathcal{C}(Y^*; c_1, c_2)$. In words, team B will outperform team A if and only if it can produce team A's equilibrium output at lower costs.

Consider now the homogenous team, noting that the solution to (3) entails equal efforts, $s_1^* = s_2^* = s_H^*$. Given these, the homogenous team's cost is $\mathcal{C}(Y^*; \bar{c}, \bar{c}) = 2s_H^*/\bar{c}$ where Y^* is its equilibrium output. As shown previously, one way the diverse team can produce Y^* is by exerting efforts $s_D^* = ([2(\bar{c})^{1+\alpha}]^{1/(1+\alpha)})/([\tilde{c}_1^{1+\alpha} + \tilde{c}_2^{1+\alpha}]^{1/(1+\alpha)}) \cdot s_H^* > s_H^*$. Thus (s_D^*, s_D^*) satisfies feasibility in (3). If we "force" (s_D^*, s_D^*) on the diverse team, we force upon it the costs

$$\frac{[2(\bar{c})^{1+\alpha}]^{1/(1+\alpha)}}{[\tilde{c}_1^{1+\alpha} + \tilde{c}_2^{1+\alpha}]^{1/(1+\alpha)}} \cdot \mathcal{C}(Y^*; \bar{c}, \bar{c}) > \mathcal{C}(Y^*; \bar{c}, \bar{c}) \ .$$

Note the exact parallel with (1). What we are seeing here is thus again the direct efficiency loss, except now expressed as a cost disadvantage from a utilitarian cost sharing team's point of view.

The crucial point is that if we condition on the output Y^* and then as a *first step* impose (s_D^*, s_D^*) on the diverse team, it will not find itself in a teamwork equilibrium. The reason is simply that the two individuals' marginal contributions are unequal since their abilities differ. The team will therefore, in a *second step* adjust its efforts in order to equalize the marginal contributions. Since the strategic adjustment amounts to solving (3) and (s_D^*, s_D^*) is feasible, this adjustment *necessarily* reduces the costs, *i.e.*,

$$\mathcal{C}(Y^*; \tilde{c}_1, \tilde{c}_2) < \frac{[2(\bar{c})^{1+\alpha}]^{1/(1+\alpha)}}{[\tilde{c}_1^{1+\alpha} + \tilde{c}_2^{1+\alpha}]^{1/(1+\alpha)}} \cdot \mathcal{C}(Y^*; \bar{c}, \bar{c}) \ .$$

So while the direct efficiency effect increases costs, the strategic adjustment reduces it. Since costs are one-to-one with team performance as explained a moment ago, this second adjustment will consequently provide a *strategic performance boost* after the first step performance loss induced by the Benabou-type complementarity.[5] This is the *strategic benefit of diversity*.

This paper's most important message is that the existing literature on diversity has, by keeping efforts fixed, ignored this benefit, and in so doing missed a number of essential aspects of diversity's functional role. The one described so far amounts to saying that if we keep efforts fixed,

[5]In light of Observation 2, it is not surprising that this strategic cost reduction overpowers the direct efficiency loss, *i.e.*, that $\mathcal{C}(Y^*; \tilde{c}_1, \tilde{c}_2) < \mathcal{C}(Y^*; \bar{c}, \bar{c})$, if and only if $\alpha > -1/2$. This will be established much more generally in the paper. Similarly, it can be shown that in the limit case $\alpha = -1/2$, the strategic performance boost precisely cancels the direct efficiency loss out, $\mathcal{C}(Y^*; \tilde{c}_1, \tilde{c}_2) = \mathcal{C}(Y^*; \bar{c}, \bar{c})$; and finally, if $\alpha < -1/2$ then the direct efficiency loss is so great that the strategic boost cannot cope, $\mathcal{C}(Y^*; \tilde{c}_1, \tilde{c}_2) > \mathcal{C}(Y^*; \bar{c}, \bar{c})$.

we run the risk of "underestimating" diversity's potential benefits. As we shall see both in a moment when looking at the quantitative side of things, and later when the paper turns to cognitive diversity, this is just the "tip of the iceberg".

Upon reflection, the strategic benefit of diversity expresses something very familiar. Just as the "invisible hand" leads decentralized perfectly competitive economies to optimality; teamwork — provided it has a decentralized element to it (decisions guided by incentives and self-interest) — is a very advantageous institution for adapting optimally to individual differences and idiosyncracies, including not just ability diversity but any type of functionally relevant diversity. In fact, it is clearly immaterial for the previous argument that the individual characteristics are abilities — they might as well have been, say, personal cost parameters which could in turn be a function of more deeply seated characteristics (identity diversity). All that really matters is whether we keep the efforts fixed or not. If we keep them fixed, we ignore the adjustments diversity stimulates endogenously. Mathematically, this was expressed above by the fact that the "second step" cost adjustment necessarily reduces the team's costs. In Section 4, this will be established much more generally in terms of abstract (possibly multi-dimensional) characteristics.

Note also that whether we consider teamwork or market economies, the "invisible hand" statements require the absence of "market failures", and also, it requires convexity as above since marginal contribution (or marginal cost) equalization can lead to very bad outcomes in the presence of non-convexities (Guesnerie (1975)).[6] Another aspect of the above that intuitively ought to matter is that the individuals' objectives are fully aligned (they have a common team objective). Finally, one wonders what the implications are for markets: What if a manager hires the workers from a pool of candidates and must pay salaries that are increasing in the abilities? Specifically, if the market is pricing diversity correctly as in Benabou (1996) (where a manager is indifferent between more or less diverse teams), does it also price diversity correctly when efforts are endogenous. This question is interesting because if the answer is no, then managers can "reap benefits" by selecting on diversity.

Finally, it is worth pointing out the previous analysis' quantitative implications which are as important as the qualitative implications focused on so far. To prove Observation 2 above, I derived the *performance index*,

$$\mathcal{P}(c_1, c_2) = [c_1^{-(1+\alpha)/\alpha} + c_2^{-(1+\alpha)/\alpha}]^{-\alpha/(1+\alpha)} ,$$

and noted that a team will outperform another team if and only if its index is greater. By comparison, with fixed labor inputs the performance index is just "Benabou's H" (Benabou (1996), p.587),

$$H(c_1, c_2) = [c_1^{1+\alpha} + c_2^{1+\alpha}]^{1/(1+\alpha)} .$$

For explicit parameter values, we can now proceed to calculate, for example, how much more able a homogenous team must be on average to match a diverse team. The paper derives a number of

[6]The same can happen in teamwork (Section 5.4).

similar performance indices that cover also *e.g.* "cognitively differences". In the current case, if $\alpha = -0.25$ and as in the first numerical example above, the macroeconomist is three times as able as the microeconomist, the answer is that the homogenous team must be

$$\frac{[c_1^3 + (3c_1)^3]^{1/3}}{[(2c_1)^3 + (2c_1)^3]^{1/3}} - 1 = \frac{[1^3 + 3^3]^{1/3}}{[2 \cdot 2^3]^{1/3}} - 1 = 20.5\ \%$$

more skilled to overcome the diverse team's strategic performance boost. In comparison, when we fix the labor supplies and evaluate according to the performance index H, we find that the diverse team must be

$$\frac{[(2c_1)^{0.75} + (2c_1)^{0.75}]^{1/0.75}}{[c_1^{0.75} + (3c_1)^{0.75}]^{1/0.75}} - 1 = \frac{[2 \cdot 2^{0.75}]^{1/0.75}}{[1^{0.75} + 3^{0.75}]^{1/0.75}} - 1 = 3.4\ \%\ ,$$

more skilled to match the homogenous team.

So the strategic performance boost turns as $3.4\ \%$ "diversity disadvantage" into a $20.5\ \%$ advantage. The reason is, intuitively, that even though the homogenous team does not suffer from one person dragging the other's productivity down as reflected by the $3.4\ \%$ direct efficiency loss, balance reduces the need for "vigilance" and concentration. Simply put, homogeneity fosters placidity. In contrast, the macroeconomist on the diverse team is induced to strongly focus the mind when reading through the paper because the weaker microeconomist is prone to errors and misjudgments. When marginal contributions are equalized, this causes the macroeconomist to "overcompensate" conditional on the direct efficiency loss which in turn stimulates the weaker microeconomist to work more too (in equilibrium). What the previous numbers show is that the consequences of these strategic adjustments may be even more important quantitatively than the more familiar direct efficiency effects.

3 Preliminaries

This section and the next contain this paper's general results on how diversity affects economic outcomes. While this paper's focus is on teamwork, a key point is that the fundamental insights and intuitions span across disparate situations, and so are not exclusive to the paper's specific models, or even to teamwork. Indeed, the following "tool kit" may be used to analyze the functional role of diversity in a variety of settings. For example, a reader will have no problems using it to address whether, say, a diverse population is more innovative than a homogenous population, or a homogenous industry has more market power than a heterogenous industry.

Although the mathematical tools are naturally needed to appreciate the proofs and detailed calculations in the main discussion (Section 5), the broader intuitions from Section 2 will be found to map over. It is therefore, in a first reading, possible to rely on these and at this point jump straight to Section 5 (in particular, that section has been structured to make this feasible).

The key concept is "quasi-convex differences" (Definition 3) which ensures that the combined effect of the strategic adjustment and direct effect is positive. This leads to this paper's main abstract result, Theorem 2 which says that under quasi-convex differences, diversity increases outcomes. The second key result is Theorem 3 which derives general performance indices which enable us to make quantitative comparisons.

The game-theoretic setting is Selten (1970)'s class of aggregative games. This class easily covers ability diversity in teamwork models while models featuring cognitive diversity are "isomorphic" to aggregative games where the players are cognitive groups (see *e.g.* the proof of Proposition 4). Crucially, in aggregative games, performance (or "use value" more generally) is well-defined and comparable across different groups independently of the set of individuals and, in particular, their number and their characteristics. This feature is of course critical for any study of functional diversity.

There is a finite set of $I \in \{2, 3, 4, \ldots\}$ individuals (or agents) who interact strategically when choosing their actions $s_i \in S_i \subseteq \mathbb{R}$, $i = 1, \ldots, I$. Their objectives are to maximize the payoff functions

$$\pi(s, c_i) = \Pi(s_i, \sum_j s_j, c_i) \ , \ i = 1, \ldots, I \ .$$

Here Π is the "reduced payoff function" (see *e.g.* Jensen (2018b)). The individuals' characteristics are $c = (c_1, \ldots, c_I)$ where $c_i \in C \subseteq \mathbb{R}^N$ and C is the set of characteristics. For example, in the "coauthor game" of the previous section but with the general cost functions $\gamma(1+\beta)^{-1} \cdot s_i^{1+\beta}$ and salaries w_i, $i = 1, \ldots, I$, we can write the payoffs in terms of the effective inputs $\hat{s}_i = (c_i s_i)^{1+\alpha}$,

$$\pi(\hat{s}, c_i) = vP([\sum \hat{s}_j]^{1/(1+\alpha)}) - \mathcal{C}(\hat{s}_i, c_i) + w_i \ , i = 1, \ldots, I \ . \tag{4}$$

where

$$\mathcal{C}(\hat{s}_i, c_i) = \gamma \cdot (1+\beta)^{-1} c_i^{-(1+\beta)} \hat{s}_i^{\frac{1+\beta}{1+\alpha}} \ . \tag{5}$$

So in that game, the reduced payoff function is $\Pi(s_i, Q, c_i) = vP(Q^{1/(1+\alpha)}) - \gamma \cdot \frac{1}{1+\beta} c_i^{-(1+\beta)} s_i^{(1+\beta)/(1+\alpha)} + w_i$, and the team output $Y = Q^{1/(1+\alpha)}$.

A (pure strategy Nash) *equilibrium* $s^* \in \prod_i S_i$ is defined as usual and the associated *equilibrium aggregate* is denoted by $Q^* = \sum_i s_i^*$. Since Q^* and the team's equilibrium output Y^* are one-to-one, "use value" is one-to-one with the equilibrium aggregate (the same is true if "use value" is measured by the success probability $P(Y^*)$ as long as P is increasing).

It is sufficient to focus on games that are "nice" in the sense of Acemoglu and Jensen (2013).[7] Note that not all games considered in this paper are nice, but in all cases the arguments are easy to adapt.

Definition 1 ***(Nice Aggregative Games)*** *The aggregative games* $(\Pi, S_i, c_i)_{i=1}^I$ *are nice aggregative games if the action sets* $S_i \subseteq \mathbb{R}_+$ *are convex (intervals), the payoff functions* $\pi(s, c_i) =$

[7] Note that the exclusion of boundary points is strictly unnecessary for the results and made purely to simplify the exposition (by allowing us to use first-order characterizations of optimal actions). See Acemoglu and Jensen (2013) for parallel arguments that do not rely on boundary conditions.

$\Pi(s_i, \sum_j s_j, c_i)$ *are twice continuously differentiable in the joint actions* $s \in S$*, pseudo-concave in own actions* $s_i \in S_i$*, and the boundary points of* S_i *(if any) are never optimal.*[8]

In nice aggregative games, define

$$\Psi(s_i, Q, c_i) = D_{s_i}\Pi(s_i, Q, c_i) + D_Q\Pi(s_i, Q, c_i)\ . \tag{6}$$

With (4) for example, we get:

$$\Psi(s_i, Q, c_i) = vQ^{1/(1+\alpha)-1}P'(Q^{1/(1+\alpha)}) - D_{s_i}\mathcal{C}(s_i, c_i)\ . \tag{7}$$

The importance of the function defined in (6) has been known at least since Corchón (1994) who explicitly defines it and assumes $D_{s_i}\Psi_i(s_i, Q, c_i) < 0$ for all s_i, Q, and c_i (Corchón (1994), p.155). It turns out that Corchon's condition is too exacting in the current context (specifically, it rules out that diversity can be a strategic liability). In its place we take the following variant of one of the main definitions in Acemoglu and Jensen (2013).

Definition 2 ***(Local solvability)*** *The nice aggregative games* $(\Pi, S_i, c_i)_{i=1}^I$*,* $(c_1, \ldots, c_I) \in C^I$ *satisfy the local local solvability condition if either (i)* $\Psi(s_i, Q, c_i) = 0 \Rightarrow D_{s_i}\Psi(s_i, Q, c_i) < 0$ *for all* $(Q, s_i, c_i) \in \mathbb{R}_+ \times S_i \times C$*, or (ii)* $\Psi(s_i, Q, c_i) = 0 \Rightarrow D_{s_i}\Psi(s_i, Q, c_i) > 0$ *for all* $(Q, s_i, c_i) \in \mathbb{R}_+ \times S_i \times C$*.*

With (7), a sufficient condition for *(i)* is that $\mathcal{C}$ is strictly convex in effort s_i while *(ii)* holds if $\mathcal{C}$ is strictly concave in s_i.

Note that Definition 2, just as the definition of nice games, places a restriction on the set of games (plural). The next definition — which is new — similarly places restrictions across games with varying characteristics, and again the conditions are expressed in terms of the function Ψ at points where $\Psi = 0$. It should be mentioned that, appearances aside, Definition 3 has nothing in common with the similarly named condition in Jensen (2018a) (which concerns individual optimization problems). In particular, neither is necessary or sufficient for the other.

Definition 3 ***(Quasi-concave and Quasi-convex Differences)*** *The nice aggregative games* $(\Pi, S_i, c_i)_{i=1}^I$*,* $(c_1, \ldots, c_I) \in C^I$ *exhibit quasi-concave differences if for every* $Q > 0$*,* $\Psi(s_i, Q, c_i)$ *is quasi-concave in* $(s_i, c_i) \in \{\mathcal{N}(s_i') \times \mathcal{N}(c_i') \subseteq S_i \times C : \Psi(s_i', Q, c_i') = 0\}$*.*[9] *If instead,* $\Psi(s_i, Q, c_i)$ *is quasi-convex in* $(s_i, c_i) \in \{S_i \times C : \Psi(s_i, Q, c_i) = 0\}$ *for all* $Q > 0$*, the games exhibit quasi-convex differences.*

[8] A differentiable payoff function π is *pseudo-concave* in s_i if $(s_i' - s_i)^T D_{s_i}\pi(s_i, s_{-i}, c_i) \leq 0 \Rightarrow \pi(s_i', s_{-i}, c_i) \leq \pi(s_i, s_{-i}, c_i)$ for all $s_i, s_i' \in S_i$ (Mangasarian (1965)). Clearly, any concave function is pseudo-concave; and any strictly increasing concave transformation of a pseudo-concave function is again pseudo-concave (so for example, any log-concave function is pseudo-concave).

[9] $\mathcal{N}(\cdot)$ denotes a small open neighborhood of the point evaluated.

Definition 3 is this paper's main abstract definition. It is normally straight-forward to verify because we may use the bordered Hessian criterion. If $C \subseteq \mathbb{R}$, this amounts to checking a single inequality because $\Psi(s_i, Q, c_i)$ is then quasi-convex in (s_i, c_i) if and only if

$$(8) \qquad 2D_{s_i}\Psi \cdot D_{c_i}\Psi \cdot D^2_{s_i c_i}\Psi \geq [D_{s_i}\Psi]^2 D^2_{c_i^2}\Psi + [D_{c_i}\Psi]^2 D^2_{s_i^2}\Psi \,.$$[10]

In the coauthor example of the previous subsection, $\Psi(s_i, Q, c_i) = vQ^{1/(1+\alpha)-1}P'(Q^{1/(1+\alpha)}) - \gamma(1+\alpha)^{-1}c_i^{-1}\hat{s}_i^{-\frac{\alpha}{1+\alpha}}$ and as is easily verified, (8) $\Leftrightarrow \alpha \geq -1/2$. Comparing with Observation 2 this is precisely the condition for the direct efficiency loss to not be so severe that it dominates the strategic performance boost (the degree of specialization not "too extensive" as measured by the elasticity of substitution $\sigma = -1/\alpha$). As we shall see in Section 5, this is what quasi-convex differences ensures *in general* in teamwork games. If the games instead exhibit quasi-concave differences, in particular, if the inequality in (8) is reversed, then the strategic performance boost is dominated by the direct efficiency loss. Note that this could happen either because the direct efficiency loss is severe, or because the strategic adjustment actually works against a team's performance (Subsection 5.4).

As shown next, the previous observations are true in any nice aggregative game under the local solvability condition. In particular, quasi-convex differences can always be interpreted as ensuring that the strategic adjustment dominates the direct effect though, of course, the precise economic interpretation will differ from game to game. The next section proves that this leads to the expected outcomes in terms of the equilibrium aggregate.

Consider the best-response function of an individual with characteristic c_i, $r(x_{-i}, c_i) = \arg\max_{s_i \in S_i} \Pi(s_i, x_{-i} + s_i, c_i)$. Assume that this function is at least thrice differentiable and consider the direct best-response effects of a small increase in everyone else's effort $D_{x_{-i}}r(x_{-i}, c_i) = \partial r(x_{-i}, c_i)/\partial x_{-i}$, and a small increase in the characteristic $D_{c_i}r(x_{-i}, c_i) = \partial r(x_{-i}, c_i)/\partial c_i$. Now consider the backward response function b (Selten (1970)): $s_i = b(Q, c_i) \Leftrightarrow s_i = r_i(Q - s_i, c_i)$. The *direct backward response* effect of a small increase in the characteristic can then be expressed as

$$(9) \qquad D_{c_i}b(Q, c_i) = \underbrace{\frac{1}{1 + D_{x_{-i}}r(Q - b(Q, c_i), c_i)}}_{\text{Strategic adjustment}} \cdot \underbrace{D_{c_i}r(Q - b(Q, c_i), c_i)}_{\text{Direct effect}} \,.$$

As the following result demonstrates, the elasticities of the two terms on the right-hand side precisely capture the general trade-off between direct efficiency effects and strategic adjustments.

Lemma 1 *Consider nice aggregative games that satisfy the local solvability condition and assume that $C \subseteq \mathbb{R}$. Let $\epsilon_{\text{D.E.}} = \frac{\partial D_{c_i}r(Q-b(Q,c_i),c_i)}{\partial c_i} / \left(D_{c_i}r(Q - b(Q, c_i), c_i)/c_i\right)$ denote the elasticity of the direct effect and $\epsilon_{\text{S.A.}} = \frac{\partial \frac{1}{1+D_{x_{-i}}r(Q-b(Q,c_i),c_i)}}{\partial c_i} / \left(\frac{1}{1+D_{x_{-i}}r(Q-b(Q,c_i),c_i)}/c_i\right)$ the elasticity of the strategic*

[10] Note that since quasi-convexity of Ψ is only required to hold locally when $\Psi(s_i, Q, c_i) = 0$, a sufficient condition for quasi-convex differences is that (8) holds with strict inequality whenever $\Psi(s_i, Q, c_i) = 0$. This allow us to plug the equation $\Psi(s_i, Q, c_i) = 0$ into (8) before evaluating.

adjustment, and assume $D_{s_i}\Psi < 0$. Then the games exhibit quasi-convex differences if and only if $\epsilon_{\text{S.A.}}(Q, c_i) + \epsilon_{\text{D.E.}}(Q, c_i) \geq 0$ for all $Q > 0$ and $c_i \in C$.

Proof. $\frac{\partial D_{c_i} b(Q,c_i)}{\partial c_i} \Big/ \frac{D_{c_i} b(Q,c_i)}{c_i} = \epsilon_{\text{S.A.}} + \epsilon_{\text{D.E.}} \Rightarrow D^2_{c_i} b(Q, c_i) = (\epsilon_{\text{S.A.}} + \epsilon_{\text{D.E.}}) \cdot D_{c_i} b(Q, c_i)$. We consider only the case of quasi-convex differences. There are two cases. If $D_{s_i}\Psi < 0$ then $S.A. > 0$ and so $D^2_{c_i} b(Q, c_i) \geq 0 \Leftrightarrow (\epsilon_{\text{S.A.}} + \epsilon_{\text{D.E.}})\text{sign}[\text{D.E.}(Q, c_i)] \geq 0$. If $D_{s_i}\Psi > 0$ then $S.A. < 0$ and so $D^2_{c_i} b(Q, c_i) \leq 0 \Leftrightarrow (\epsilon_{\text{S.A.}} + \epsilon_{\text{D.E.}})\text{sign}[\text{D.E.}(Q, c_i)] \geq 0$. Since $\text{D.E.}(Q, c_i)$, $\text{S.A.}(Q, c_i)$ and $\text{D.E.}(Q, c_i)$ are all evaluated at $s_i = b(Q, c_i) \Leftrightarrow \Psi(s_i, Q, c_i) = 0$, it follows by Lemma 2 that $(\epsilon_{\text{S.A.}} + \epsilon_{\text{D.E.}})\text{sign}[\text{D.E.}(Q, c_i)] \geq 0$ is both necessary and sufficient for quasi-convex differences. ■

As the proof also demonstrates, the critical condition becomes instead $\epsilon_{\text{S.A.}}(Q, c_i) + \epsilon_{\text{D.E.}}(Q, c_i) \leq 0$ if the local solvability condition holds in the form $D_{s_i}\Psi > 0$.

4 General Results

In the coauthor example in Section 2, it was possible to establish a general relationship between diversity and performance (Observation 2), and also to find a "performance index" (p. 9) through straight-forward direct arguments based on first-order conditions. This section establishes similar results that apply much more generally, and in particular to more realistic descriptions of teamwork where variability stems both from abilities and cognitive traits, and where teams are of any size.

For ease of exposition, focus is on situations with a unique equilibrium. Everything (including the proofs) remains valid when there are multiple equilibria when the results apply to the smallest and largest equilibrium aggregates (*e.g.*, teams' best and worst possible performance).[11]

It is useful to first delineate situations where diversity does not matter, functionally speaking, *i.e.*, where outcomes *only* depend on teams' average characteristics $\bar{c} = (\bar{c}^1, \ldots, \bar{c}^N) \in C$, $\bar{c}^n = (\sum_i c_i^n)/I$, $n = 1, \ldots, N$.

Theorem 1 ***(Diversity and Outcomes: The Representative Agent)*** *Consider nice aggregative games $(\Pi, S_i, c_i)_{i=1}^I$, $c \in C^I$ which satisfy the local solvability condition and where $S_i = S$ for all i. The equilibrium aggregate depends only on the mean characteristics $\bar{c} = (\bar{c}^1, \ldots, \bar{c}^N)$ if and only if the games exhibit quasi-monotone differences (i.e., both quasi-concave and quasi-convex differences).*

Proof. In the Appendix. ■

[11] With multiplicity, the statement that $\succeq$-diversity increases the aggregate should thus be interpreted as saying that $c \succeq \tilde{c} \Rightarrow [Q_S(c) \geq Q_S(\tilde{c})$ and $Q_R(c) \geq Q_R(\tilde{c})]$ where $Q_R(c)$ and $Q_S(c)$ denote, respectively, the largest and smallest equilibrium aggregates given the characteristics c.

Theorem 1 parallels Gorman (1968)'s Theorem on "representative consumers" in demand theory, and is closely related to Bergstrom and Varian (1985), who also ask when Nash equilibria are independent of individuals' characteristics.[12] When there is a single functionally relevant characteristic, $C \subseteq \mathbb{R}$, the theorem applies mechanically whenever (8) holds with equality. For example, the co-author games from Section 2 exhibit quasi-monotone differences if and only if $\alpha = -1/2$ (compare with Observation 2). The intuition is that in this limit case the strategic performance boost precisely balances out the direct efficiency loss.

For two populations with characteristics $c \in C^I$ and $\tilde{c} \in C^I$, say that the former is more diverse than the latter in the *convex diversity order*, written $c \succeq_{cx} \tilde{c}$, if $\frac{1}{I}\sum_i f(c_i) \geq \frac{1}{I}\sum_i f(\tilde{c}_i)$ for any convex function $f : \mathbb{R}^N \to \mathbb{R}$. With a single (functionally relevant) characteristic $N = 1$, this simply means that c is a mean-preserving spread of $\tilde{c}$ (Rothschild and Stiglitz (1970)).

There is a lot to be said about how the words "more diverse" can and should be defined, and the interested reader is invited to consult the online appendix (Appendix B) which lays out a fully developed descriptive theory of diversity, and also discusses the relationship with the existence-value theory of Nehring and Puppe (2002). The bottom line is that the convex diversity order is the most restrictive, meaningful way of attaching content to the statement that a population is "more diverse" than another population. Because it is the most restrictive way, it is also the *least controversial*: If c is more diverse than $\tilde{c}$ in the convex order, then every sensible person ought to agree that c is "more diverse" than $\tilde{c}$. If we are presented with a *different* way of defining the words more diverse — which could be another order $\succeq \subseteq C^I \times C^I$ or even an index $D : C \to \mathbb{R}$ given which c is more diverse than $\tilde{c}$ if and only if $D(c) \geq D(\tilde{c})$ — we can then ask if this passes the minimum sensibility requirement of not contradicting the convex diversity order ($c \succeq_{cx} \tilde{c} \Rightarrow c \succeq \tilde{c}$, or $c \succeq_{cx} \tilde{c} \Rightarrow D(c) \geq D(\tilde{c})$ in the case of an index). A diversity order/index that passes this test is said to be "*sensible*". As I show in Appendix B, all of the familiar orders/indices from the literatures on inequality/uncertainty/information are sensible: the Gini index, the Lorenz order, all Atkinson and Shorrocks indices, the mean-variance order, the Rawlsian index, the Strongest Link index, etc., etc.

Theorem 2 ***(Diversity and Outcomes: The Convex Diversity Order)*** *Consider nice aggregative games $(\Pi, S_i, c_i)_{i=1}^I$, $c \in C^I$ that satisfy the local solvability condition.*

1. *If the games exhibit quasi-convex differences then convex order diversity increases the equilibrium aggregate, and unless the equilibrium aggregate depends only on the mean characteristics, there does not exist a sensible diversity order $\succeq$ such that $\succeq$-diversity reduces the equilibrium aggregate.*

[12]Their answer is "if and only if backward response functions [Definition 4] are affine"; in particular they explicitly calculate the backward response function in a public good provision game and verify this. Emphasis and presentation aside, the mathematical contribution is thus to provide necessary and sufficient conditions for affine backward response functions.

2. *If the games exhibit quasi-concave differences then convex order diversity reduces the equilibrium aggregate and unless the equilibrium aggregate depends only on the mean characteristics, there does not exist a sensible diversity order $\succeq$ such that $\succeq$-diversity increases the equilibrium aggregate.*

In light of Lemma 1, Theorem 2 shows that when the combined effect of the strategic adjustment and the direct effect is positive (quasi-convex differences), diversity increases the equilibrium aggregate. While the interpretation of this statement will in general depend on the specific game or context; in teamwork models, it involves the tradeoff between the strategic performance boost and direct efficiency effect already discussed extensively at this point. Using the bordered Hessian criterion in (8) to check for quasi-convex differences immediately yields conditions for diversity to improve or reduce performance (see Proposition 1 whose proof equals these calculations).

The second important message of Theorem 2 is that quasi-convex/quasi-convex differences are what Milgrom and Roberts (1994) call *critical* conditions. If games exhibit quasi-convex differences and one accepts that diversity orders must be sensible, then there is no scope for finding a sensible diversity order given which diversity *reduces* the equilibrium aggregate. Similarly, if the games exhibit quasi-concave differences then no sensible description of relative diversity is compatible with diversity improving the equilibrium aggregate. The economic consequences of these observations are returned to in the next section. In practical "tool kit" terms, the main implication is that quasi-convex/quasi-concave differences must always be considered as a "first pass" when asking a diversity comparative statics question. If the games pass, then, *but only then* does it make sense to query stronger conclusions. The next result does precisely that by turning to the very exacting concept of a *performance index*, *i.e.*, a sensible diversity index that is one-to-one with the equilibrium aggregate.

Note that in the teamwork specification (4), $\Psi(s_i, Q, c_i) = vP'(Q) - D_{s_i}\mathcal{C}(s_i, c_i)$ (compare with (6)), so it is clear that the following result has bite in this case.

Theorem 3 ***(Performance Indices)*** *Consider nice aggregative games $(\Pi, S_i, c_i)_{i=1}^I$, $c \in C^I$ that satisfy the local solvability condition.*

- *The aggregate is increasing in diversity in the sense of a sensible diversity index $D : C^I \to \mathbb{R}$ if and only if $\Psi(s_i, Q, c_i)$ (modulo algebraic manipulations) takes one of the following functional forms:*

  1. $\Psi(s_i, Q, c_i) = h_1(Q)g_1(c_i) + g_2(c_i)(s_i + h_3(Q))^\epsilon$ *where* $\psi(c) = -\mathrm{sign}[\epsilon g_2(c_i)](\frac{g_1(c)}{g_2(c)})^{\epsilon^{-1}}$ *is well-defined and convex, and* $h_1(Q) < 0$.
  2. $\Psi(s_i, Q, c_i) = h_1(Q)(g_1(c_i))^{h_2(Q)} - (g_2(c_i))^{h_2(Q)}\exp(s_i - h_3(Q))$ *where* $h_1(Q) > 0$ *and either* $(g_2(c_i))^{h_2(Q)}h_2(Q) > 0$ *and* $\psi_1(c_i) = \log(\frac{g_1(c)}{g_2(c)})$ *is convex, or else* $(g_2(c_i))^{h_2(Q)}h_2(Q) < 0$ *and* $\psi_2(c_i) = \log(\frac{g_2(c)}{g_1(c)})$ *is convex.*

3. $\Psi(s_i, Q, c_i) = h_1(Q) + g_1(c_i) + \log(s_i - h_3(Q))$ *where* $\psi(c_i) = -\exp(g_1(c_i))$ *is convex.*

- *The aggregate is decreasing in diversity in the sense of a sensible diversity index* $D : C^I \to \mathbb{R}$ *if and only if* Ψ *takes one of the functional forms 1.-3. with the function* ψ *(or in the case of 2. either* ψ_1 *or* ψ_2*) being concave.*

Proof. In the appendix. ■

The corollary explicitly characterizes the diversity indices implied by Theorem 3.

Corollary 1 *If diversity increases the aggregate in Theorem 3, we may use the sensible diversity index* $D(c) = \sum_i \psi(c_i)$*, or any monotone transformation hereof (where in case 2.* $\psi = \psi_1$ *if* $h_2(Q) < 0$ *and* $\psi = \psi_2$ *if* $h_2(Q) > 0$*); and if diversity reduces the aggregate, we may use the sensible diversity index* $D(c) = -\sum_i \psi(c_i)$ *or any monotone transformation hereof (where in case 2. again* $\psi = \psi_1$ *if* $h_2(Q) < 0$ *and* $\psi = \psi_2$ *if* $h_2(Q) > 0$*).*

Theorem 2 and its corollary play major roles in the next section because they enable us to compute performance indices in teamwork models.

Now, it will come as no surprise that not all games (and also not all teamwork games) imply diversity indices that are one-to-one with the equilibrium aggregate. This raises the question of how robust any investigation based on a performance index is, since it necessarily entails a special case in terms of functional forms. The question of how far conclusions are removed from assumptions is, arguably, important in the context of the diversity debate due to the rather high stakes and correspondingly high potential for controversy. See Thompson (2014), who makes precisely this point in her rejection of Hong and Page (2004) and Page (2008).

The answer, surprisingly, is that performance indices are highly robust because they can *always* be found if: (i) we have a "reference" population whose outcome we wish to compare with the expected outcome of a different population, (ii) the equilibrium is unique, (iii) the games either exhibit quasi-convex or quasi-concave differences. (ii) is not a problem in any of this paper's models. Since (iii) are critical conditions, it is at this point obvious why this is required. That we require a "reference" population is more subtle, and to fully appreciate why, the reader is referred to the proof of the theorem. It should be noted that this proof has much in common with the proof of one of the main theorems in Acemoglu and Jensen (2018) although, obviously, there is no economic/intuitive relationship what-so-ever between the two set-ups.

Theorem 4 ***(Performance Indices with a Reference Population)*** *Consider nice aggregative games that satisfy the local solvability condition and assume that the equilibrium is unique (for all* $c \in C^I$*). Fix a reference population with characteristics* $c \in C^I$*.*

- *If the games exhibit quasi-convex differences, there exists a sensible diversity index* $D : C^I \to \mathbb{R}$ *such that any comparison population with characteristics* $\tilde{c} \in C^I$ *will have a higher* *equilibrium aggregate if and only if $D(\tilde{c}) \geq D(c)$. Specifically, if the equilibrium aggregate of the reference population is denoted by Q^*, such an index is given by $D(\tilde{c}) = \sum_i \psi(\tilde{c}_i)$ where $\psi(c) = -\text{sign}[D_{s_i}\Psi] \cdot \{s_i : \Psi(s_i, Q^*, c) = 0\}$.*

- *If the games exhibit quasi-concave differences, there exists a sensible diversity index $D : C^I \to \mathbb{R}$ such that any comparison population with characteristics $\tilde{c} \in C^I$ will have a lower equilibrium aggregate if and only if $D(\tilde{c}) \geq D(c)$. Specifically, if the equilibrium aggregate of the reference population is denoted by Q^*, such an index is given by $D(\tilde{c}) = \sum_i \psi(\tilde{c}_i)$ where $\psi(c) = \text{sign}[D_{s_i}\Psi] \cdot \{s_i : \Psi(s_i, Q^*, c) = 0\}$.*

Theorem 4 may be interpreted as an "implicit function theorem" for performance indices: Given a local solution (a reference population's equilibrium aggregate Q^*), one (implicitly) solves $\{s_i : \Psi(s_i, Q^*, c) = 0\}$ from which a performance index can be found. This index is "local" in the sense that it only applies for the given reference population, but it is global in the sense that it allows the reference population's performance to be compared to any other population's performance. In this paper's explicit models, this procedure will naturally return the performance indices of Theorem 3. It is also a standard exercise to show robustness from these observations: Given "sufficient smoothness" of the functional forms involves, the performance index determined in this way varies smoothly with the functional form.

5 Teamwork, Diversity and Performance

The previous section contains the results needed to analyze the relationship between diversity and performance both generally and robustly. This section begins where Section 2 left off by looking at "ability diversity" in the general Becker-Murphy-Holmstrom-Benabou-type setting where I derive general performance indices, discuss the quantitative aspects, and also study optimal team selection and market wages.

The model is then extended to allow for "cognitive diversity" and I ask when a cognitively diverse team will outperform a cognitively homogenous team with the same abilities, and study the relationship between cognitive and ability diversity in detail. The discussion then follows Hong and Page (2004) and Page (2008) in asking whether cognitive diversity "trumps" ability, *i.e.*, whether a cognitively diverse team of moderate ability can outperform a cognitively homogenous team of high ability.

Finally, I briefly consider situations where diversity may be a "strategic liability" rather than providing a performance boost which turns all of the results on their heads, and arguably approximates dysfunctional teams fairly well.

5.1 Do Teams of Lightweights and Superstars Outperform Teams of "Average Joes"?

Consider a principal, such as a head of department, production manager, research council, military general, or a football coach. The principal must select a team for a project or a task that may or may not succeed. Each team has $I > 1$ members and completes the project successfully with probability $P([\sum_{i=1}^{I} \frac{1}{1+\alpha}(c_i s_i)^{1+\alpha}]^{1/(1+\alpha)}) \in [0, 1]$, where $P : \mathbb{R}_+ \to [0, 1]$ is a smooth and strictly increasing success probability function, $s = (s_1, s_2, \ldots, s_I) \in \mathbb{R}_+^I$ are the team's efforts, and $c = (c_1, \ldots, c_I) \in \prod_i C_i \subseteq \mathbb{R}_{++}^I$ the team's task relevant abilities.

The function $[\sum_{i=1}^{I} (c_i s_i)^{1+\alpha}]^{1/(1+\alpha)}$ is the team production function, which, given the abilities and any suppressed inputs such as capital equipment, maps efforts into the relevant team output (see *e.g.* Alchian and Demsetz (1972), Holmstrom (1982), Becker and Murphy (1992)).

It should be remarked that the specification in terms of a success probability function is for clarity and concreteness only. All results remains valid if $P : \mathbb{R}_+ \to \mathbb{R}_+$ is a deterministic production function as in *e.g.* Holmstrom (1982) or Becker and Murphy (1992).

Focusing for concreteness on a manager and teams of workers, the former is authorized to pay the latter a bonus $v > 0$ if the project succeeds, and also pays basic salaries $w_i > 0$, $i = 1, \ldots, I$. The bonus payment and resulting individual valuations are kept fixed across individuals in order to focus on the influence of the abilities.[13] Effort incurs personal (opportunity) costs $\mathcal{C}(s_i) = \gamma(1 + \beta)^{-1} s_i^{1+\beta}$, $i = 1, \ldots, I$, where $\beta > -1$ and $\gamma > 0$. If worker i's team is chosen, her expected payoff given the team's efforts $s \in \mathbb{R}_+^I$ is consequently

$$\pi(s_i, s_{-i}, c) = vP([\sum_{j=1}^{I} (c_j s_j)^{1+\alpha}]^{1/(1+\alpha)}) - \gamma \frac{1}{1+\beta} s_i^{1+\beta} + w_i \,. \tag{10}$$

In teamwork (Nash) equilibrium, *i.e.*, when workers maximize (10) taking other workers' efforts as given, this determines equilibrium efforts $s^* \in S$, team output $Y^* = [\sum_{i=1}^{I} \frac{1}{1+\alpha}(c_i s_i^*)^{1+\alpha}]^{1/(1+\alpha)}$, and consequently, the equilibrium probability that the team completes the project successfully, $P^* = P(Y^*)$.

If the teamwork equilibrium s^* is unique, then $s^* = \arg\max_{s \in S} vP([\sum_{j=1}^{I} (c_j s_j)^{1+\alpha}]^{1/(1+\alpha)}) - \sum_i \gamma \frac{1}{1+\beta} s_i^{1+\beta} + \sum_i w_i$. So we may also think of teamwork equilibrium as a cooperative endeavor. In fact, it is seen that if we replace v with $I \cdot v$, we get the (utilitarian) social optimum. Ipso facto, everything that follows also applies to socially optimal teamwork.[14]

The question is now which team the manager should choose in order to maximize P^* if the only difference is that one team has more diverse abilities than the other: *Does a combination of "lightweights" and "superstars" outperform a team of "average Joes"*?

[13] If valuations and/or the cost parameter defined next differ, one gets a composite index in Proposition 2 below. This raises no new abstract/mathematical difficulties but it obviously makes it more challenging to interpret the results because ability diversity is combined with other sources of heterogeneity.

[14] To simplify, I nearly always ignore upper boundary constraints because in all cases, the corrections implied are trivial. Concerning the Pareto optimum, it is clear that when the valuation is increased from v to $v \cdot I$, the equilibrium is affected and, in particular, it is more likely that efforts will "hit" maximum feasible efforts. Since this could already happen with v instead of $v \cdot I$, *i.e.*, in teamwork equilibrium, there is no economic substance to this difference.

Recall the convex diversity order discussed on page 15, which in the current situation simply says that a team A is more diverse than another team B, if team A's abilities $c^A = (c_1^A, \ldots, c_I^A)$ are a mean-preserving spread of team B's abilities $c^B = (c_1^B, \ldots, c_I^B)$. A function $S : \prod_i C_i \to \mathbb{R}$ is a *performance index* if team A outperforms team B when $S(c^A) \geq S(c^B)$ and strictly outperforms it when $S(c^A) > S(c^B)$.

For this section's results, P is not required to be (strictly) concave but if it is, and if $\beta \geq 0 > \alpha$, the conditions of the following result are satisfied.

Proposition 1 ***(Diverse Abilities and Performance)*** *Assume that $\beta \geq \alpha$, that $(YP''(Y))/P'(Y) < \beta$, and let diversity be described by the convex diversity order. Then diversity improves performance if $\alpha \geq -\frac{1}{2+\beta}$ and reduces it if $\alpha < -\frac{1}{2+\beta}$. In both cases diversity strictly affects performance unless $\alpha = -\frac{1}{2+\beta}$, when performance only depends on teams' average ability, or $\alpha = \beta$, when performance is one-to-one with the most able worker's ability and the "Strongest Link Index" $S(c) = \max\{c_1, \ldots, c_I\}$ therefore is a performance index.*

Proof. When $\alpha \neq \beta$, this is a direct application of Theorems 1-2. One calculates directly from (4)-(5) that the function Ψ defined in (6) is given by $\Psi(s_i, Q, c_i) = v\hat{P}'(Q) - \gamma \cdot (1+\alpha)^{-1} c_i^{-(1+\beta)} s_i^{(\beta-\alpha)/(1+\alpha)}$ where $\hat{P}(Q) = P(Q^{1/(1+\alpha)})$ has been substituted to simplify the calculations. By (8), it is then straight-forward to verify that Ψ is quasi-convex in s_i and c_i if and only if $\beta \geq \alpha \geq -1/(2+\beta)$. Since these conditions are if and only if, they are also necessary and sufficient for the games to exhibit quasi-convex differences. If $-1/(2+\beta) = \alpha < \beta$, Ψ is quasi-monotone (both quasi-concave and quasi-convex) and the local solvability condition holds. Hence by Theorem 1, performance depends only on the mean abilities. If $\beta > \alpha$, the games are nice, the local solvability conditions holds, and by Theorem 2 convex order diversity therefore improves performance if $\alpha > -1/(2+\beta)$ and reduces it if $\alpha < -1/(2+\beta)$. If $\alpha = \beta$, the games are not nice. In this case, $\hat{P}$ must be strictly concave (see the proof of Lemma 3 in the appendix), and in equilibrium, $\Psi(s_i, Q, c_i) = v\hat{P}'(Q) - \gamma \cdot (1+\alpha)^{-1}(\max\{c_1, \ldots, c_I\})^{-(1+\beta)} = 0$. It follows that performance is increasing in $\max\{c_1, \ldots, c_I\}$, and since the maximum is a convex function this implies that convex order diversity improves performance also when $\alpha = \beta$. ■

By this proposition, a manager who faces teams that are more and less diverse in the sense of the convex diversity order should choose the more diverse team if $\alpha > -\frac{1}{2+\beta}$ and the less diverse team if $\alpha < -\frac{1}{2+\beta}$. Only in the limit case $\alpha = -\frac{1}{2+\beta}$ should the manager ignore diversity and focus on mean abilities. The interpretation and intuition is completely analogues to the special case studied in Section 2: Diversity provides a strategic performance boost because more capable individuals "overcompensate" for the weaker individuals, conditional on the direct efficiency effects. This performance boost overpowers the direct efficiency loss and ensures that diversity beats homogeneity as long as the elasticity of substitution is not too small/the individual tasks are not "too specialized". As seen, the strength of the performance boost will generally depend on the cost function's curvature β. If $\beta = 0$ we get Observation 2 from Section 2, while if $\beta \to +\infty$, the boost so weakens that diversity is not beneficial unless it actually entails a direct efficiency gain, $\alpha \geq 0$. On the other hand, if $\beta < 0$ ("learning by doing") and close enough to -1, then diversity beats homogeneity as long as no individual's input is indispensable, *i.e.*, as long as $\alpha > -1$ or equivalently, the elasticity of substitution greater than one. As returned to in subsection 5.4, the condition that $\beta \geq \alpha$ is critical: If $\beta < \alpha$, diversity may impose a strategic performance loss rather than a boost, and this performance loss can even reverse direct efficiency gains. Another new thing we see is that in the limit case $\alpha = \beta$, team performance is driven entirely by the "strongest link". This is perhaps the simplest example one can think of where diversity impacts performance (if abilities are subjected to a mean-preserving spread then, obviously, the "strongest link" becomes stronger). However, it is a very special example — in all other cases, a team's performance will depend on everyone's abilities.

By Theorem 2's critical statements, we can be fairly confident about these findings' robustness. To elaborate, imagine that the direct efficiency loss is not "too severe", $\alpha > -\frac{1}{2+\beta}$, and that someone (*e.g.*, a manager) claims that diversity reduces performance in the sense of some diversity measure which she, for whatever reasons, thinks appropriately captures the meaning of the words "more diverse". We can then conclude that she is not basing her assessment on a sensible diversity order as defined on page 15. In particular, if she advances the claim with reference to a diversity order that *is* sensible — which basically means with reference to any known measure such as the standard deviation, the Gini coefficient, a Shorrocks or Atkinson index, etc., etc.— we can "refute her claim" by presenting her with teams A and B where in her view A is more diverse than B but where nonetheless, A is going to outperform B (in expectation).

Moving on, it is clear that the convex diversity order is a sufficiently restrictive way of describing relative diversity that a manager is in practice likely to face teams which are neither more or less diverse than each other in this sense. This happens if the manager has a fixed salary budget $B > 0$ and must hire I workers from a pool of candidates, and pay the salary $w_i = w(c_i) = \theta + \rho \cdot c_i$ to a worker with ability c_i (where $I \cdot \theta < B$ and $\rho > 0$).[15] The same is true with other salary schemes like the one in Benabou (1996) returned to in a moment.

For the "critical conditions" reason just discussed, Proposition 1 remains important but evidently it can no longer be used to generate firm decision rules. One of the more notable findings of this paper is just how much we can improve on this state of affairs without giving up on robustness.

Note that by Proposition 1, diversity does not affect performance if $\alpha = -\frac{1}{2+\beta}$ and if $\alpha = \beta$, the manager already has a performance index at her disposal, namely the "Strongest Link index". These two cases may therefore be ignored from now on.

[15]Note that such a manager will necessarily assemble a team with average ability $(\sum_i c_i)/I = B/(I\rho) - \theta/\rho$, so any difference in performance traces to abilities' variability rather than any average advantage.

Proposition 2 ***(Performance Indices)*** *Assume $\beta > \alpha \neq -\frac{1}{2+\beta}$ and consider the (absolute)* Atkinson *index,*

$$D_\rho^{\mathrm{At}}(c) = (\sum_i c_i^\rho)^{\frac{1}{\rho}}, \text{ where } \rho = \frac{(1+\alpha)(1+\beta)}{\beta-\alpha} \,. \tag{11}$$

Then D_ρ^{At}-index diversity improves performance if $\alpha > -\frac{1}{2+\beta}$ and reduces performance if $\alpha < -\frac{1}{2+\beta}$.

Proof. From the previous proof, $\Psi(s_i, Q, c_i) = v\hat{P}'(Q) - \gamma \cdot (1+\alpha)^{-1} c_i^{-(1+\beta)} s_i^{(\beta-\alpha)/(1+\alpha)}$. This fits 1. of Theorem 3 with $h_1(Q) = -vP'(Q) < 0$, $g_1(c_i) = -1$, $g_2(c_i) = -\gamma \cdot (1+\alpha)^{-1} c_i^{-(1+\beta)}$, $h_3(Q) = 0$, and $\epsilon = \frac{\beta-\alpha}{1+\alpha}$. Since $\psi(c) = -\mathrm{sign}[\epsilon g_2(c)](\frac{g_1(c)}{g_2(c)})^{\epsilon^{-1}} = \theta \cdot c^{\frac{(1+\alpha)(1+\beta)}{\beta-\alpha}}$, where $\theta > 0$ is a constant, ψ is convex if $\alpha \geq -\frac{1}{2+\beta}$ and concave if $\alpha \leq -\frac{1}{2+\beta}$. We can thus use the index $D'(c) = \sum_i c_i^{\frac{(1+\alpha)(1+\beta)}{\beta-\alpha}}$ when $\alpha \geq -\frac{1}{2+\beta}$ and $-D'$ when $\alpha \leq -\frac{1}{2+\beta}$ which combine to the absolute Shorrocks index $D_\rho^{\mathrm{Sh}}(c) = \frac{1}{I\cdot\rho\cdot(\rho-1)}[\sum_i c_i^\rho - 1]$, where $\rho = \frac{(1+\alpha)(1+\beta)}{\beta-\alpha}$, or since performance indices are only determined up to monotone transformations, to the Atkinson index (11). ■

According to Proposition 2, the Atkinson index D_ρ^{At} is a performance index when the direct efficiency loss is not too severe, $\alpha > -\frac{1}{2+\beta}$, while if $\alpha < -\frac{1}{2+\beta}$, it is instead an "underperformance index" which is to say that $-D_\rho^{\mathrm{At}}$ is a performance index. Since the Atkinson index measures diversity, these directions are what we would expect from Proposition 1, but the current statements are much sharper.

Having a performance index is very useful to a manager (or a modeler) because it allows her to compare *any* two teams. In particular, she can select/assemble a team from a pool of candidates given a fixed budget and any given salary scheme $w_i = w(c, c_i)$ by solving the optimization problem

$$\max_{\{c:\sum_i w_i \leq B\}} \sum_i (c_i^{\frac{(1+\alpha)(1+\beta)}{\beta-\alpha}})^{\frac{\beta-\alpha}{(1+\alpha)(1+\beta)}} \,. \tag{12}$$

If effort levels are fixed, this "*team selection problem*" should instead be based on the performance index described on page 9 ("Benabou's H"),

$$\max_{\{c:\sum_i w_i \leq B\}} (\sum_i c_i^{1+\alpha})^{1+\alpha} \,. \tag{13}$$

If the salary scheme $w_i = w(c, c_i)$ is such that the manager when solving a team selection problem is indifferent between any two teams that exhaust her budget, say that there is "*no diversity arbitrage*".

If there is no diversity arbitrage, no performance gains can be made by selecting on diversity. In the aggregate market setting of Benabou (1996) one would expect this to obtain, and indeed the salary scheme in equation (6), p.587 of Benabou (1996),

$$w_i = w(c_i, c) = \delta(H(c))^{-\alpha} c_i^{1+\alpha}, \delta > 0\ ,\ H(c) = (\sum_i c_i^{1+\alpha})^{1+\alpha}\ , \tag{14}$$

precisely ensures this outcome as long as the efforts are exogenously determined since $\sum_i w_i = \delta(H(c))^{-\alpha} \sum_i c_i^{1+\alpha} = \delta(H(c))^{-\alpha}(H(c))^{1+\alpha} = \delta H(c)$. By exactly the same argument, we get the following when efforts are endogenous and the relevant team selection problem therefore is (12):

Proposition 3 (No Arbitrage) *A salary scheme implies no diversity arbitrage if and only if it is of the form*

$$w_i = w(c, c_i) = \delta \cdot [D_\rho^{\mathrm{At}}(c)]^{1-\rho} c_i^\rho \tag{15}$$

where $\rho = \frac{(1+\alpha)(1+\beta)}{\beta-\alpha}$ *and* $\delta > 0$.

The significance of this result is that if the salaries in an economy do not conform with the no-arbitrage condition, then a manager can improve upon team performance by "selecting on diversity" — or to put it in a more balanced manner; a manager should be advised to take diversity into consideration along with all the other factors that influence team performance.

Crucially, if there is no diversity arbitrage "on aggregate", *i.e.*, if salaries are given by (14), then $\{c : \sum_i w(c_i) \leq B\}$ is a convex set. If the performance index $D_\rho^{\mathrm{At}}(c)$ is a strictly concave function — which happens if and only if $\alpha < -\frac{1}{2+\beta}$ — then (12) therefore has an interior solution. Hence, by symmetry, the manager will optimally choose a homogenous team if $C_1 = C_2 = \ldots = C_I$. If there are individual bounds on abilities as in the stratification setting of Benabou (1996), the manager will instead choose c that attains those upper bounds and so assemble a team that is as close to being completely homogenous as possible (in Benabou's example, the most able clerical worker affordable).

If, on the other hand, there is no diversity arbitrage "on aggregate" and $D_\rho^{\mathrm{At}}(c)$ is strictly convex, *i.e.*, if $\alpha > -\frac{1}{2+\beta}$, then the team selection problem entails a boundary solution (again, this would normally be marked by minimum and maximum bounds on individuals' abilities). So a good manager should now select *in favor* of diversity. A simple example is helpful to illustrate what this might entail in practice.

Example 1 ***(Linear-Quadratic Teamwork Games and Mean-Variance Diversity)*** *Assume that the teamwork game is linear-quadratic ($\alpha = 0$, $\beta = 1$), and consider two teams where neither has an absolute advantage (equal average abilities). Then the team whose abilities have the highest variance,*

$$V(c) = \frac{1}{I} \sum_i (c_i - \bar{c})^2\ ,$$

will strictly outperform the other team and should therefore be chosen by the manager even if there is no diversity arbitrage "on aggregate".

Proof. Follows directly from the previous discussion, Proposition 3 and Proposition 2 since $\rho = \frac{(1+\alpha)(1+\beta)}{\beta-\alpha} = 2$ and $0 = \alpha > -1/(2+\beta) = -1/3$. ■

Note that in this example $\alpha = 0$, so there is no direct efficiency effect and the diverse teams' outperformance is driven entirely by the performance boost.

Next let us turn to the quantitative side of the discussion. In Section 2, I gave an example where the strategic benefits of diversity turned a direct performance loss into a substantial performance gain (page 10). Similar examples are easily constructed with larger teams and/or arbitrary cost curvature. Such examples are of course of major significance because they demonstrate that the strategic benefits of diversity can overturn the direct efficiency effects also in a quantitatively significant way, and so they demonstrate that the strategic performance boost is not of the "second order". Nonetheless, to save space, they are left to the reader. Instead, the following example demonstrates that the strategic performance boost is (generally) also not of the second order when the direct and strategic effects go in the same direction.

Example 2 ***(Strategic Magnification of Direct Efficiency Gains)*** *Consider two armies (or army squads). A diverse army, half of whose soldiers are $\rho > 1$ times more able than the other half, and a homogenous army all of whose soldiers are $\eta > 1$ times more able than the diverse army's average soldier. By Proposition 4's performance index, the diverse army will beat the homogenous one if, and only if,*

$$\eta < \eta^{\max} \equiv \left[\frac{1 + \rho^{\frac{(1+\alpha)(1+\beta)}{\beta-\alpha}}}{2(0.5 + 0.5\rho)^{\frac{(1+\alpha)(1+\beta)}{\beta-\alpha}}} \right]^{\frac{\beta-\alpha}{(1+\alpha)(1+\beta)}} . \tag{16}$$

For example, assume the diverse army's most capable soldiers are three times as able as the less capable ones ($\rho = 3$), and take $\alpha = 0.25$. This implies a direct efficiency gain of 3% ($[(1+3^{1.25})/(2 \cdot 2^{1.25})]^{1/1.25} = 1.03$). Compare with the incentivized situation taking $\beta = 0.5$: $\eta^{\max} = [(1+3^{7.5})/(2 \cdot 2^{7.5})]^{1/7.5} = 1.37$. Hence if the assessment is based on exogenous efforts, the homogenous army must be $3\ \%$ more able to match the diverse army. If the strategic boost is taken into account, the homogenous army must be $37\ \%$ more able to match the diverse army.

This example is inspired by a historical episode. At the battle of Waterloo, the Duke of Wellington famously mixed his "*polyglot and varied force [...] just as he had done so successfully during the Peninsular war*" (Glover (2014), p.33). This pitted ability-wise highly diverse units against Napoleon's uniform conscript army. It seems reasonable to assume that combining weaker and stronger units could entail direct efficiency gains (the veterans "pulling up" the fresh recruits' abilities). But since the opposite case could be argued as well, these gains are likely to be small and uncertain. The point the example makes is that, nonetheless, the Duke of Wellington's decision to combine weak and strong forces could have been quite significant because of the strategic performance boost.

5.2 Does Cognitive Diversity Trump Homogeneity, and "Superstars"?

Next let us consider what happens when more than one individual can be assigned to each subtask or specialty on a team. This simple extension has far-reaching consequences because it allows us to formalize and study the influence of so-called "*cognitive diversity*" (Johnson-Laird (1983), Page (2008)).[16]

While this topic has received little attention in economics journals (to the best of my knowledge, none), it has as described in the Introduction inspired a large and hugely influential cross-disciplinary literature spanning management science, mathematics, political science, psychology, and philosophy. Crucially, the relationship between the current model and that literature precisely parallels the relationship with the fixed effort description in the previous subsection: All of the (non-empirical) contributions to the "cognitive diversity literature" treat individuals "mechanistically" if in great detail (*e.g.*, Hong and Page (2004), Singer (2019), and the lengthy footnote below), *i.e.*, they ignore endogenous effort formation. The question becomes again, then, how incentives and strategic interaction modulate the direct efficiency effects which are now caused by cognitive diversity.

Each individual now has two characteristics; ability $c_i \in C_i$ as previously, and then a *cognitive index*, $t_i \in \{1, \ldots, N\}$ which assigns the individual to one of $N > 1$ "*cognitive groups*". Cognitive grouping might reflect "areas of specialty" (macro and micro) or "academic disciplines" (economics and mathematics), as well as cognitive traits or "thinking styles" such as "perspective", "knowledge processing approach", "representation", or "heuristic" (Hong and Page (2004), Reynolds and Lewis (2017)). As Singer (2019) points out (pp. 179-180), cognitive traits may also reflect more deeply seated characteristics such as age, gender, or ethnicity ("identity diversity"). Note, however, that while *e.g.* Roland and Galunic (2004), Horwitz and Horwitz (2007), and Reynolds and Lewis (2017) all find strong evidence that cognitive diversity improves management teams' performance, none of these studies find any evidence that identity diversity affects performance. The study of Norway's mandated female board representation by Ahren and Ditmar (2012) similarly finds no positive relationship (in fact, a negative relationship although mediated by experience). This suggests that there is no systematic correlation between identity diversity and functionally relevant types of cognitive diversity for management tasks (which of course does not preclude such a relationship for other tasks).

What sets cognitive groups apart functionally is that they provide different "*input varieties*" in the spirit of *e.g.* Ethier (1982). Hence the team production function is now

$$[\sum_{n=1}^{N} e_n^{1+\alpha}]^{1/(1+\alpha)} ,$$

where $e_n = \sum_{i:t_i=n} c_i s_i$ is the ability-adjusted effort of cognitive group n.

[16]Note that the current focus on cognitive diversity obviously does not preclude other interpretations. Everything applies equally to, for example, "occupational roles" (stratification), "playing styles" on a football team, "fighting styles" in armies, and so on.

If efforts or labor supplies are decided exogenously, this description is (mathematically speaking) a special case of the simpler model in the previous subsection. The reason is that with fixed efforts, a single macroeconomist with ability $c_1 + c_2$ is functionally equivalent to two macroeconomists with abilities c_1 and c_2, as well as to a single macroeconomist with the ability $(c_1+c_2)/2$ working twice as much. This formal equivalence is very useful to us because we can transfer the fixed efforts performance index from (12) over, subject to the obvious modifications,

$$H^{\text{cog}}(c,t) = [\sum_n \hat{c}_n^{1+\alpha}]^{1/(1+\alpha)} , \tag{17}$$

where $\hat{c}_n = \sum_{i:t_i=n} c_i$ is cognitive group n's combined ability.

From the performance index H^{cog} is seen, in particular, that cognitive diversity entails a direct efficiency gain if $\alpha < 0$, a direct efficiency loss if $\alpha > 0$, and it does not impact performance if $\alpha = 0$. The origin and interpretation of these direct efficiency effects are discussed in detail in Page (2008) as well as *e.g.* Thoma (2015) and Singer (2019). For example, if $\alpha = -0.25$, then two macroeconomists with average ability $\bar{c}$ will, due to their more limited understanding of economic behavior, have to work $[2\bar{c}^{0.75}]^{1/0.75}/[(2\bar{c})^{0.75}]^{1/0.75} - 1 = 26\ \%$ more to match a micro- and a macroeconomist both with ability $\bar{c}$.

The case where $-1 < \alpha < 0$ is the most interesting one because under this parameter restriction, cognitive groups' inputs are complementary (the elasticity of substitution $-1/\alpha$ is positive), and teams are able to function without any one of them as befits cognitive diversity as opposed to Becker and Murphy (1992)-type task specialization (the elasticity of substitution is greater than 1). Furthermore, inter-group complementarities dwarf intra-group complementarities (in our model, intra-group efforts are perfect substitutes in ability-adjusted terms). In other words, by restricting α to lie between -1 and 0 we precisely capture the main tenets of the existing cognitive diversity literature (*e.g.* Hong and Page (2004), Page (2008), Zollman (2010), or Singer (2019)).[17]

[17] Due to space-constraints, this claim must unfortunately be taken on face value by the reader. But to give an indication, consider Hong and Page (2004) who model team problem solving as the maximization of a function with multiple local optima. In the simplest version, each individual searches in a specific direction (her "perspective") and finds the nearest local maximum in the direction of search. The team searches sequentially, updating at each step if there is an improvement, and when no one's "perspective" can increase the value, the process stops. This description is easily extended so that levels of effort are not fixed (*e.g.*, individuals could try more "perspectives" by exerting more effort). Our "black box" team production function here simply maps teams' group frequencies and efforts into the maximum values they attain (or into the average of these across a set of initial values). The team output could also be random which would correspond to a random initial condition. As long as P maps the resulting random variable into a deterministic quantity by use of a law of large numbers, this does not greatly complicate the game theoretic analysis as long as we remain with in the aggregative games framework (see Acemoglu and Jensen (2010), Acemoglu and Jensen (2015)). To get to a team production function with a specific functional form, we would "reverse engineer" the mechanical description (values, search directions, and so on). Crucially, the key features of this model ("complementarity between perspectives" and "decreasing returns to perspectives") will necessarily imply $\alpha < 0$ if we rationalize a CES functional form. Note that it is less substantial that with our within-group specification (perfect substitutes), an individual's productivity never falls to 0, it merely goes towards zero as the individual's group grows in size. Hong and Page (2004) take the maximum within groups: the second member of a cognitive group is unproductive because she searches in the same direction as the first member (to quote from Page (2008): "Two heads are not only not better than one in this case — they *are* one", p.153).

In terms of this paper's general model, the expected payoff of an individual team member i is now

$$\pi(s,c) = vP([\sum_{n=1}^{N}(\sum_{j:t_j=n} c_j s_j)^{1+\alpha}]^{1/(1+\alpha)}) - \gamma\frac{1}{1+\beta}s_i^{1+\beta} + w_i \ . \tag{18}$$

In teamwork equilibrium, each individual maximizes this payoff function taking other individuals' (ability-adjusted) efforts as given, including the efforts of his own cognitive group. The question is then whether "diversity trumps homogeneity": *Will a cognitively diverse team outperform a cognitively homogenous team if the teams have the same abilities?*

This question addresses, for example, the situation facing a macroeconomist with ability c_1 who must decide whether to work with a micro- or a macro-economist both with the same ability c_2 (or an economics department that must hire an additional member of staff facing the same choice). Importantly, by assuming that "the abilities follow the individuals, not the groups", we are approaching the issue under the implicit assumption that no cognitive group innately is more productive than any other group (the next subsection departs from precisely this feature).

Proposition 4 ("Does Cognitive Diversity Trump Homogeneity?") *Let the assumptions of Proposition 1 hold. Consider a cognitively homogenous team with abilities $(c_1, \ldots, c_I)$, and a cognitively diverse team with the same abilities $(c_1, \ldots, c_I)$ and one individual in each of N $(= I)$ cognitive groups. Then the cognitively diverse team outperforms the homogenous team if $\alpha < 0$, if $\alpha = 0$, performance is independent of cognitive diversity, and if $\alpha > 0$, the cognitively homogenous team outperforms the diverse team.*

Proof. Replace P with $P(Q^{1/(1+\alpha)})$ to simplify notation. Express the cognitively diverse team's payoff functions as $vP(\sum_i \hat{s}_i^D) - \gamma \cdot (1+\beta)^{-1}c_i^{-(1+\beta)}(\hat{s}_i^D)^{(1+\beta)/(1+\alpha)}$ where $\hat{s}_i^D = (c_i s_i)^{1+\alpha}$, and the homogenous team's payoff functions as $vP((\sum_i \hat{s}_i^H)^{1+\alpha}) - \gamma \cdot \frac{1}{1+\beta}c_i^{-(1+\beta)}(\hat{s}_i^H)^{1+\beta}$ where $\hat{s}_i^H = c_i s_i$. The diverse team's aggregate backwards response function is $b^D(Q,c) = [(1+\alpha)v\gamma^{-1}P'(Q)]^{(1+\alpha)/(\beta-\alpha)}\sum_i c_i^{((1+\alpha)(1+\beta))/(\beta-\alpha)}$. An individual on the homogenous team has first-order condition $v(1+\alpha)(\sum_i(\hat{s}_i^H))^{\alpha}P'((\sum_i \hat{s}_i^H)^{1+\alpha}) = \gamma \cdot c_i^{-(1+\beta)}(\hat{s}_i^H)^{\beta}$. Insert $Q = (\sum_i \hat{s}_i)^{1+\alpha}$ and rearrange to get the individual backward response function $b^H(Q,c_i) = [v\gamma^{-1}(1+\alpha)Q^{\alpha/(1+\alpha)}P'(Y)]^{1/\beta}c_i^{(1+\beta)/\beta}$. So Q is an equilibrium aggregate for the homogenous team if and only if $Q = [v\gamma^{-1}(1+\alpha)Q^{\alpha/(1+\alpha)}P'(Q)]^{(1+\alpha)/\beta}[\sum_i c_i^{(1+\beta)/\beta}]^{1+\alpha} \Leftrightarrow Q = b^H(Q,c)$ where $b^H(Q,c) = [v\gamma^{-1}(1+\alpha)P'(Q)]^{(1+\alpha)/(\beta-\alpha)}[\sum_i c_i^{(1+\beta)/\beta}]^{[(1+\alpha)\beta]/(\beta-\alpha)}$. Let Q_D and Q_H denote the teams' unique equilibrium outputs. Then $Q_D \geq Q_H$ if and only if $Q_D = [(1+\alpha)v\gamma^{-1}P'(Q_D)]^{(1+\alpha)/(\beta-\alpha)}\sum_i c_i^{[(1+\alpha)(1+\beta)]/(\beta-\alpha)} \geq v\gamma^{-1}(1+\alpha)P'(Q)]^{(1+\alpha)/(\beta-\alpha)}[\sum_i c_i^{(1+\beta)/\beta}]^{[(1+\alpha)\beta]/(\beta-\alpha)}$. Since $\beta > \alpha$, this inequality holds if and only if $D^{\mathrm{At}}_{[(1+\alpha)(1+\beta)]/(\beta-\alpha)}(c) = [\sum_i c_i^{[(1+\alpha)(1+\beta)]/(\beta-\alpha)}]^{(\beta-\alpha)/[(1+\alpha)(1+\beta)]} \geq [\sum_i c_i^{(1+\beta)/\beta}]^{\beta/(1+\beta)} = D^{\mathrm{At}}_{(1+\beta)/\beta}(c)$. It is clear that if $\alpha = 0$, this equality holds. Let us now show that if $0 < p < p'$ then $D^{\mathrm{At}}_p(c) < D^{\mathrm{At}}_{p'}(c)$. Since we may divide through with $\min c_i$, there is no loss of generality in assuming that $c_i \geq 1$ for all i and $c_i = 1$ for some i. Differentiating $[\sum_i c_i^p]^{1/p}$ with respect to p, it then follows that $[\sum_i c_i^p]^{1/p} \log[\sum_i c_i^p][\sum_i c_i^p \log(c_i)] > 0$. Since $(\beta - \alpha)/[(1+\alpha)(1+\beta)] > \beta/(1+\beta)$, the proposition's statement follows. ■

Corollary 2 *(Performance Indices) The cognitively homogenous team has performance index $D^{\mathrm{At}}_{(1+\beta)/\beta}(c)$ and this is directly comparable with the performance index $D^{\mathrm{At}}_{((1+\alpha)(1+\beta))/(\beta-\alpha)}(c)$ of Proposition 2. In particular, performance can be compared whether or not the teams have the same abilities.*

Proposition 4 tells us that as long as $\beta \geq \alpha$ and the abilities are conditioned upon, incentivized behavior and strategic interaction does not qualitatively overturn the direct efficiency effects associated with cognitive diversity. This is useful to have on record vis-a-vis the existing cognitive diversity literature, however, the more interesting, and certainly the more surprising observations come out of the Corollary's quantitative implications and the relationship with Proposition 1.

Specifically, the Corollary and the direct efficiency index in (17) enable us to assess to what extent incentives and strategic interaction modulate the direct efficiency effects of cognitive diversity. With endogenous effort formation, a cognitively homogenous team will have to work $[\sum_i c_i^{[(1+\alpha)(1+\beta)]/(\beta-\alpha)}]^{(\beta-\alpha)/[(1+\alpha)(1+\beta)]}/[\sum_i c_i^{(1+\beta)/\beta}]^{\beta/(1+\beta)}$ times as much to match a cognitively diverse team. If efforts are fixed, the number is instead $[\sum_i c_i^{1+\alpha}]^{1/(1+\alpha)}/[(\sum_i c_i)^{1+\alpha}]^{1/(1+\alpha)}$. The next example explores the consequences.

Example 3 *("Strategic Dampening") Consider two teams as follows: A microeconomist together with a three times more able macroeconomist, and a macroeconomist together with a three times more able macroeconomist. Take $\alpha = -0.25$, $\beta = 0.5$. The direct efficiency effect is then $[1 + 3^{0.75}]^{1/0.75}/4 = 1.22$, so with fixed efforts, the cognitively homogenous team (the two macroeconomists) will have to work 22 % more to match the cognitively diverse team. With endogenous efforts, we get instead $[1 + 3^{1.5}]^{1/1.5}/[1 + 3^3]^{1/3} - 1 = 11$ %. So incentivized behavior and strategic interaction cut the direct efficiency gain to cognitive diversity in half.*

The example is an instance of the following more general observations: If everyone is equally able, $c_i = \bar{c}$ all i, then $[\sum_i c_i^{[(1+\alpha)(1+\beta)]/(\beta-\alpha)}]^{(\beta-\alpha)/[(1+\alpha)(1+\beta)]}/[\sum_i c_i^{(1+\beta)/\beta}]^{\beta/(1+\beta)} = [\sum_i c_i^{1+\alpha}]^{1/(1+\alpha)}/[(\sum_i c_i)^{1+\alpha}]^{1/(1+\alpha)}$, and so the fixed and endogenous effort models have the same quantitative implications. But if abilities are diverse as in the example, then incentives and strategic interaction will *dampen* the benefits of cognitive diversity. The explanation traces to the previous subsection: Because cognitive diversity reduces ability diversity and ability diversity provides a strategic performance boost, cognitive diversity — even when on balance good for performance — entails a strategic performance loss. In this sense, an assessment based purely on exogenous effort formation systematically overestimates the benefits of cognitive diversity (unless everyone is exactly equally able).

This raises the question, then, of whether diverse abilities may in fact in some situations be "more important" for team performance than cognitive diversity. Again, it is easiest to approach this by way of an example.

Example 4 ***("Ability Diversity vs. Cognitive Diversity")*** *Consider three teams as follows: One team (Team A) consists of two macroeconomists with average ability $\bar{c}$, one team (Team B) of a micro- and a macro-economist both with ability $\bar{c}$, and the third "team" (Team C) of a single "superstar" macroeconomist with ability $2\bar{c}$. Note that these three teams are on average equally capable. Take $\alpha = -0.25$ and $\beta = 0.5$. Whether efforts are fixed or endogenously determined, the cognitively homogenous Team A will have to work $2^{1/0.75-1} - 1 = 26\%$ more to match the cognitively diverse Team B. If efforts are fixed, there is no functional difference between Team A and Team C (see page 26). Hence the "superstar" Team C will also have to work $26\ \%$ more to match the cognitively diverse Team A if we assess the situation without taking endogenous effort formation into account. But under endogenous effort formation, the relative performance index is $\frac{[2\cdot\bar{c}^{1.5}]^{1/1.5}}{[(2\bar{c})^3]^{1/3}} = 2^{1/1.5-1} = 0.79$. So unless the "superstar" in Team C works $21\ \%$ less, he will outperform the cognitively diverse Team B (and therefore also Team A).*

What the previous example shows is that although cognitive diversity improves performance (when $\alpha < 0$), the performance boost teams receive from ability diversity can be even more substantial. If we consider the example's "superstar vs. cognitively diverse team" assessment more generally, the ratio of the performance indices is:

$$\frac{\left[2 \cdot \bar{c}^{[(1+\alpha)(1+\beta)]/(\beta-\alpha)}\right]^{(\beta-\alpha)/[(1+\alpha)(1+\beta)]}}{\left[(2\bar{c})^{(1+\beta)/\beta}\right]^{\beta/(1+\beta)}} = 2^{(\beta-\alpha)/[(1+\alpha)(1+\beta)]-1} \leq 1 \Leftrightarrow \alpha \geq -1/(2+\beta)\ .$$

Comparing with Proposition 1, all falls into place: If $\alpha < 0$, cognitive diversity improves performance and if $\alpha < -1/(2+\beta)$, ability diversity reduces performance too. The ideal team is then maximally cognitively diverse and "completely balanced" in terms of the groups' effective inputs. Such an ideal team can be assembled in a variety of functionally equivalent ways, *e.g.*, by placing an equal number of equally capable individuals in each cognitive group; or by "substituting quantity for quality", *e.g.*, if the engineer in Benabou (1996) is five times as able as the clerical worker, the ideal team has five times as many clerical workers as engineers. But if $-1/(2+\beta) < \alpha < 0$, the strategic performance boost teams receive from having diverse abilities overpowers the direct efficiency loss *whether this efficiency loss stems from ability diversity or cognitive homogeneity*.[18] In this case, "ability diversity trumps cognitive diversity".

[18]Here is a detailed explanation: When comparing teams that are not equally cognitively diverse, the direct efficiency effects remain straight-forward because ability-diversity and cognitive diversity are mathematically speaking the same thing. For example, instead of comparing a team with two macroeconomists with ability c (Team A) to a team with a macroeconomist and a microeconomist with ability c (Team B), we might compare a team with a macroeconomist with ability $2c$ and a microeconomist with ability 0 (Team C) to a team with a macroeconomist and a microeconomist with ability c (Team B, again). If the elasticity of substitution is positive and we keep efforts fixed, whether we say that the cognitively diverse Team B outperforms the cognitively homogenous Team A; or we say that the ability-homogenous team B will outperform the ability-diverse Team C amounts to the same.

The most important take-away is that when individuals are driven by incentives and influenced by strategic considerations, *quantity is not a perfect substitute for quality* unlike when efforts are fixed and a highly capable individual is functionally equivalent to two half-as-able individuals. Hence a (cognitively homogenous) "superstar" may perform even better than a cognitively diverse team although, on average, the two are equally able and there are direct efficiency gains to cognitive diversity. Now, of course, if a team cannot function without all perspectives, the ideal team must be cognitively diverse to function. It then does not consist just of a single "superstar" with ability c^H, but of a very capable macroeconomist with ability $c^H - \epsilon$ and a microeconomist with ability $\epsilon > 0$ where ϵ is the minimum ability compatible with the microeconomist not consistently turning correct proofs into wrong proofs. Nonetheless, the general take-away remains the same.

The mechanism of action is as described at the end of Section 2 except the highly capable individuals might now be situated in a cognitively diverse team (to reap the benefits from cognitive diversity): More able individuals are induced to concentrate very hard on integrating and checking the inputs from a team of mixed abilities. With a balanced team — whether cognitively diverse or not — this role is absent, which fosters placidity and diminishes performance conditional on the direct efficiency effects.

5.3 Does Cognitive Diversity Trump Ability?

Having investigated the relationship between cognitive and ability diversity, let us now turn to a question that has spurred a rather heated debate in the existing literature on functional diversity, namely whether "diversity trumps ability" (Hong and Page (2004), Thompson (2014)).

To briefly motivate this question, consider the United Kingdom where a sizeable fraction of the important decision makers have taken a "thinking style degree" called a *PPE*. Anyone who holds a *PPE* is also certain to be of exceptionally high ability.[19] It may be asked then, whether the United Kingdom's tendency to select on this "thinking style" (or indeed, on ability) is good or bad for team performance.

The simplest way to capture this aspect of the debate is to assume "perfect correlation" be-

When efforts are endogenously determined, the distinction between Team A and Team C is not just down to "language" but we can still trace the strategic effect to ability-diversity because the strategic effect conditions on the direct efficiency effects which (as we have just seen) can be expressed equivalently in terms of abilities only. For example, if the elasticity of substitution is positive and we compare Team A to Team B above, then we can immediately conclude that Team B will outperform team A because there is a direct efficiency gain to cognitive diversity and no change in ability-diversity (so there is no strategic performance effect). In Example 3 we compared a microeconomist and a three times more able macroeconomist (Team D) with a macroeconomist together with a three times more able macroeconomist (Team E). To compare entirely in the "ability-dimension", we write Team D as a microeconomist with ability c, a macroeconomist with ability $3c$; and Team E as a microeconomist with ability 0, a macroeconomist with ability $4c$. As mentioned, this has no effect on the direct efficiency effect. Clearly, team E is ability-wise more diverse than team D, so in comparison with team D, it receives a performance boost.

[19] PPE stands for "Philosophy, Politics and Economics" and it is an undergraduate degree taught at the University of Oxford. The PPE has incredibly low admission rates.

tween cognitive group and ability ($t_i = t_j \Rightarrow c_i = c_j$). This implies the following payoff functions:

$$\pi(s,c) = vP([\sum_{n=1}^{N}(\sum_{j:t_j=n} c_n s_j)^{1+\alpha}]^{1/(1+\alpha)}) - \gamma \cdot C(N) \cdot \frac{1}{1+\beta} s_i^{1+\beta} + w(c_i) \ , \ i = 1, \ldots, I \ . \tag{19}$$

The term $C(N)$ reflects Becker and Murphy (1992)-type *coordination costs* (C is an increasing function of the number of non-empty cognitive groups N). The motivation for introducing such coordination costs is precisely the same as in Becker and Murphy's study: Realistically, cognitively diverse teams must face additional obstacles due to differences in use of formalism, language, expression, and so on. The empirical literature on cognitive diversity strongly supports the importance of such coordination costs (Simons, Pelled and Smith (1999), p.663, Horwitz and Horwitz (2007), pp.996-998).

The question is now whether a "moderate ability", cognitively diverse team (the "non-PPEs") will outperform a more able cognitively homogenous team (the "PPEs"). When answering, we may normalize $C(1) = 1$ without loss of generality so that cognitively homogenous teams face no coordination costs.

Proposition 5 ***(Does Diversity Trump Ability?)*** *Let the assumptions of Proposition 1 hold. Consider two teams as follows: A cognitively homogenous team featuring I workers with ability $c_H > 0$ and no coordination costs $C(N) = 1$, and a cognitively diverse team with $I/N \in \mathbb{N}$ workers in each of N cognitive groups all of which have ability $c_L < c_H$ and face coordination costs $C(N)$. Denote by $\omega = c_H/c_L > 1$ the "ability gap". Then the cognitively diverse team will outperform the high-ability team if, and only if,*

$$\omega < N^{-\alpha/(1+\alpha)} \cdot C(N)^{-1/(1+\beta)} \ . \tag{20}$$

In particular, regardless of how large the ability gap is and no matter the coordination costs, the cognitively diverse team will outperform the high-ability team if complementarities are strong enough (α close enough to -1). Further, if there are no coordination costs ($C(N) = 1$ for all N) then given any ability gap, there exists a fine enough epistemic division of labor (i.e., a large enough N) such that the cognitively diverse team will outperform the high-ability team.

Proof. Omitting the details to save space, one can as in the proof of Proposition 4 derive the following performance index $D(c,g) = \sum_n (C(N) c_n^{-(1+\beta)} g_n^{-\beta})^{\frac{1+\alpha}{\alpha-\beta}}$ where $g = (g_1, \ldots, g_N)$ are the group frequencies. Since $c_L = \omega c_H$, the cognitively diverse team will consequently outperform the homogenous team if and only if $\sum_n [C(N)^{\frac{1}{\alpha-\beta}} c_L^{\frac{1+\beta}{\beta-\alpha}} (I/N)^{\frac{\beta}{\beta-\alpha}}]^{1+\alpha} > [(\omega c_L)^{\frac{1+\beta}{\beta-\alpha}} I^{\frac{\beta}{\beta-\alpha}}]^{1+\alpha} \Leftrightarrow$ (20). ■

To elaborate on what this result says, take $\alpha = -0.5$ and consider ten relatively low-ability cognitive groups in addition to a high-ability group, and assume the team size is a multiple of 10 so it becomes possible to place the same number of individuals in each group on the diverse team. Without coordination costs, the cognitively diverse team will then outperform the high-ability team if and only if the ability gap is less than $10^{0.5/0.5} = 10$, *i.e.*, provided the high-ability team is less than ten times as skilled as the cognitively diverse team.

If an additional cognitive group increases coordination costs by 10 percentage points, $C(N) = 0.1N + 0.9$ (note that $C(1) = 1$ must hold by the normalization) and $\beta = 0.5$, the high-ability team will outperform if (and only if) it is more than $10^1 \cdot 1.9^{-1/1.5} = 6.5$ times as able as the cognitively diverse team. If we reduce α to -0.75, these numbers increase to $10^{0.75/0.25} = 10.000$ and $10^{0.75/0.25} \cdot 1.9^{-1/1.5} = 6.519$, respectively, and if there are 100 cognitive groups — think "*crowdsourcing*" and "*peer-production*" — then we get, respectively, $100^{0.75/0.25} = 1.000.000$ and $100^{0.75/0.25} \cdot 10.9^{-1/1.5} = 203.414$. If, on the other hand, we reduce the direct efficiency loss to ability diversity and cognitive homogeneity by taking $\alpha = -0.05$, the maximum ability gap without coordination costs is $10^{0.05/0.95} = 1.13$, or 13 %, and with just two cognitive groups as in Thoma (2015)'s study of explorer and extractor scientists, it decreases to 4 %.

With more exacting coordination costs, the balance shifts decisively in favor of ability. If $C(N) = N$, *i.e.*, if doubling the epistemic division of labor doubles the costs, we get in the first case above a maximum ability gap $10^1 \cdot 10^{-1/1.5} = 2.15$, and with one hundred crowd-sourced groups, the number becomes roughly four. With more draconian coordination costs, diversity seizes to be an advantage all-together. For example, if $C(N) = N^2$, the maximum allowable gap $N^1 \cdot N^{-2/1.5}$ is below 1 for all $N > 1$. So in this case, coordination costs ruin any benefits of cognitive diversity. Note that what accounts for this is that coordination costs directly disincentivize individuals.

What is clear from these calculations is that in broad terms, the (non-strategic) conclusions of Hong and Page (2004), Page (2008), Zollman (2010), Thoma (2015), Singer (2019), etc. carry over to situations where workers are driven by incentives and respond to each other's actions (conditional on coordination costs). Importantly, our model is not subject to any of the critiques that have been leveled against Hong and Page (2004) (*e.g.*, Thompson (2014)).

5.4 When Diversity is a "Strategic Liability"

Let us finish the discussion of diversity and teamwork by returning to the parallel between the strategic benefits of diversity and the "invisible hand" discussed on page 9, and specifically, the potentially very harmful effect of non-convexities (Guesnerie (1975)).

There is little gained from generality, so we return to the "coauthor game" in the joint representation (2) but allow now costs to have any curvature $\beta > -1$:

$$vP([(c_1 s_1)^{1+\alpha} + (c_2 s_2)^{1+\alpha}]^{1/(1+\alpha)}) - \gamma \cdot \frac{1}{1+\beta}(s_1^{1+\beta} + s_2^{1+\beta}) \,. \tag{21}$$

To bring the point out most clearly, it is helpful to set $\hat{P}(Q) = P(Q^{1/(1+\alpha)})$ and rewrite the previous objective in terms of "effective inputs", $\hat{s}_i = (c_i s_i)^{1+\alpha}$, $i = 1, 2$:

$$v\hat{P}(\hat{s}_1 + \hat{s}_2) - \gamma \cdot \frac{1}{1+\beta}(c_1^{-(1+\beta)} \hat{s}_1^{\frac{1+\beta}{1+\alpha}} + c_2^{-(1+\beta)} \hat{s}_2^{\frac{1+\beta}{1+\alpha}}) \,. \tag{22}$$

Conditioning on the team output $Q = \hat{s}_1 + \hat{s}_2$, a necessary condition is then that the team minimizes the costs:

$$\min_{\{(\hat{s}_1,\hat{s}_2):Q=\hat{s}_1+\hat{s}_2\}} \gamma \cdot \frac{1}{1+\beta}(c_1^{-(1+\beta)}\hat{s}_1^{\frac{1+\beta}{1+\alpha}} + c_2^{-(1+\beta)}\hat{s}_2^{\frac{1+\beta}{1+\alpha}})$$

If $\beta \geq \alpha$, this is a *convex minimization problem* and if $\beta > \alpha$ the unique interior solution involves the equalization of marginal costs. In particular, this description applies to the unique teamwork equilibrium studied in the paper so far and all of the previous results go through.

But imagine now that $\beta < \alpha$. The previous minimization problem's objective is now *concave*, so an optimizer will lie at the boundary meaning that cost minimizing sharing of efforts entails one of the two team members doing all of the work. If the team adopts this equilibrium, the paper's conclusions continue to hold. There is then no difference from the case where $\beta \geq \alpha$.

The "problem" is that equalization of the marginal costs,

$$\gamma \cdot \frac{1}{1+\alpha} c_1^{-(1+\beta)} \hat{s}_1^{\frac{\beta-\alpha}{1+\alpha}} = \gamma \cdot \frac{1}{1+\alpha} c_2^{-(1+\beta)} \hat{s}_2^{\frac{\beta-\alpha}{1+\alpha}},$$

leads to a teamwork equilibrium as well. Since marginal costs are equalized — rather than one person doing all of the work — this equilibrium is arguably a realistic prediction by a (casual) fairness argument.

Since the objective is concave, this equilibrium will *maximize* costs — in particular, the strategic adjustment must now *increase* costs conditional on direct efficiency effects. Hence by the same argument as in Section 2, diversity entails a *strategic performance loss* — diversity is a "*strategic liability*".[20] It follows that if diversity imposes a direct efficiency loss ($\alpha < 0$), then a balanced team will necessarily outperform a diverse team. Detailed calculations are left to the reader, but in light of the magnitudes encountered in the previous examples it will come as no surprise that a manager is in this situation well advised to pick a homogenous team. In fact, even if $\alpha > 0$ so that diversity entails an efficiency gain, the strategic performance loss may (significantly) reverse it.

Evidently, cost maximization is a very inefficient outcome for a team. When marginal costs are equalized, a more able individual (higher c_i) will exert less effort in "effective" units than a less able individual (lower c_i). Although inefficient, it is hardly an unrealistic description — what happens is simply that the more capable individuals reduce their efforts to "match" the marginal contributions of less capable individuals.

Note finally that this discussion has limitations: When $\alpha > \beta$, it is likely that individuals will reach the upper bound of their feasible efforts unless P is "sufficiently concave" (P' decreasing sufficiently rapidly as a function of the team's output). Since efforts are (effectively) exogenous if individual always choose to exert maximum effort, we are then back in the exogenous efforts framework.

[20]Technically speaking, this is because when $\alpha > \beta$, the local solvability condition now holds in the form covered by Lemma 2.c. in the Appendix, and so a shift up in the aggregate backward response function will *reduce* the equilibrium aggregate. Note, conveniently, that since the "equal marginal cost" equilibrium is interior, all of this paper's general/abstract results remain valid although the conclusions are "reversed".

6 Concluding Remarks

When assessing diversity's impact on performance, it matters whether less capable individuals impede more capable colleagues or are "pulled up" by them. Differences in approach and perspective matter too since how much "a person improves a solution depends on how her tools combine with and differ from the tools of the other problem solvers" (Page (2008), p.132). At the same time, diversity may impose costs since less similar people tend to expend more time and energy communicating with, and understanding one another.

This paper's first key observation is that such aspects of diversity reflect *direct efficiency effects* which can be formalized simply and directly by relying on the frameworks and insights of Rosen (1981), Becker and Murphy (1992), and especially, Benabou (1996). Cognitive diversity can then be integrated by equating cognitive groups' contributions with different input varieties a la Ethier (1982) or Romer (1990), and exploiting the mathematical equivalence between ability- and cognitive diversity in models where effort levels are fixed.[21] Obviously, the reduction to simple production relationships, expressible in terms of the elasticity of substitution, entails a loss of depth and generality. But as a first step, this loss is outweighed by the advantages of being able to frame the diversity discussion in simple and familiar economic models. In fact, a statement along the lines of "the analysis is too simplistic because the model does not take 'xyz' into account", is not fully justified in light of Sections 3-4. The general economic intuitions and insights — for example, the role of direct efficiency effects just described; as well as the "toolkit" developed to reach those insights and most importantly the main abstract Theorems 2-3, extend to arbitrary (multidimensional) characteristics, considerably more elaborate models of teamwork, and even to otherwise unrelated situations such as *e.g.* oligopolistic competition or models of innovation. So whatever 'xyz' equals (say, mechanism design aspects), this paper should be helpful in addressing the resulting issues even if additional work may be needed too.

Direct efficiency effects form the natural starting point of any functional analysis of diversity. Still, there are situations where they by themselves, provide too limited a perspective. When a small number of individuals are engaged in a "problem-solving" task as the one Page talks about in the previous quote, they will — conditional on showing up at work — be driven by incentives and their strategic interaction will form an integral part of the teamwork task. A manager, even a very powerful one, is unlikely to be able to fully control these efforts, in part because she will, like the teacher who must mark a group assignment, be unable to pinpoint precisely who did what when the output is delivered.

When individual team members have discretion over their efforts, direct efficiency effects will be modulated by "knock-on" effort adjustments, *i.e.*, the direct efficiency effects will be followed by strategic adjustments as a reflection of "team behavior" (Section 2). Those adjustments might, conceivably, have belonged in the category of complicating but less important factors one ought

[21] On this, see page 26 and footnote 18.

to abstract from in a first analysis. This paper's most important take-away is that they do not. The strategic adjustments may reverse, and quantitatively overwhelm, the direct efficiency effects. Making sense of this in order to fully and comprehensively grasp when and why it occurs, is perhaps this paper's main contribution to the diversity literature.

That the strategic adjustments should not be thought of as "second order", is significant for modelers, managers and for the diversity debate at large. If a manager has to assemble a team given a limited wage budget and can choose between candidates with different abilities; an assessment based purely on direct efficiency effects might make her hire a homogenous team when in fact, a more diverse team would perform better. This example ties seamlessly in with the general "diversity debate" which precisely revolves around whether or not it is advisable to select in favor of diversity. All of this — including the relationship with the voluminous "diversity literature" (*e.g.*, Hong and Page (2004), Page (2008), Zollman (2010), Landemore (2012), Thompson (2014), Thoma (2015), Brennan (2017), Singer (2019)) — has been discussed extensively in the paper, for example in Sections 5.2 and 5.3. The findings also have much more specific implications for, for example, labor economics and empirical work. Some of these are worth stressing again:

- Salary structures that are convex, perhaps "strongly so" in abilities can arise even if the elasticity of substitution is positive and ability-diversity, including "superstar" teams, therefore impose a performance loss by a direct efficiency argument. (Proposition 3; discussion p. 23)
- Combining weak and strong individuals in teams can be hugely beneficial even if the direct efficiency effects are "small and uncertain". (Example 2; Proposition 2)
- Cognitive diversity may trump ability even when there are significant coordination costs (Proposition 5; discussion p. 32); but even with no coordination costs, cognitive diversity may in turn be trumped by "superstars". (Example 4 and associated discussion)
- A structural empirical assessment of cognitive diversity could, when teams have varying abilities, significantly overestimate the benefits of cognitive diversity. (Example 3; Corollary 2)
- Because cognitive and ability diversity cannot be fully separated at the direct efficiency level, isolating the effect of one or the other empirically requires a structural framework. (p. 26 and footnote 18)

With each of these statements, it is true to say that this paper is closer to opening the issue than to closing it. Each thus represents a direction for future research. On that note, as mentioned *e.g.* on page 25, the empirical evidence that "diversity matters" is substantial, yet in light of this paper's results, a more structural approach is both desirable and feasible. Another important issue is how all of this relates to the normative aspects of diversity (*e.g.*, Nehring and Puppe (2002)). This is briefly discussed in the Supplement but otherwise, this paper's approach is strictly consequentialist. However, diversity clearly matters for other reasons than its economic "use value".

References

Acemoglu, D. and M. K. Jensen (2010): "Robust Comparative Statics in Large Static Games", *IEEE Proceedings on Decision and Control* CDC 2010, 3133-3139.

Acemoglu, D. and M.K. Jensen (2013): "Aggregate Comparative Statics", *Games and Economic Behavior* 81, 27-49.

Acemoglu, D. and M.K. Jensen (2015): "Robust Comparative Statics in Large Dynamic Economies", *Journal of Political Economy* 123, 587-640.

Acemoglu, D. and M.K. Jensen (2018): "Equilibrium Analysis in the Behavioral Neoclassical Growth Model", *NBER Working Paper* #25363.

Alchian, A.A. and H. Demsetz (1972): "Production, Information Costs, and Economic Organization", *American Economic Review* 62, 777-795.

Ahren, K.R. and A.K. Ditmar (2012): "The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation", *Quarterly Journal of Economics* 127, 137-197.

Atkinson, A.B. (1970): "On the Measurement of Inequality", *Journal of Economic Theory* 2, 244-263.

Becker, G.S. and K.M. Murphy (1992): "The Division of Labor, Coordination Costs, and Knowledge", *Quarterly Journal of Economics* 107, 1137-1160.

Benabou, R. (1996): "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance", *American Economic Review* 86, 584-609.

Bergstrom, T. and H.L. Varian (1985): "When Are Nash Equilibria Independent of the Distribution of individuals' Characteristics?", *Review of Economic Studies* 52, 715-718.

Brennan, J., *Against Democracy*, Princeton University Press, 2017.

Corchón, L. (1994): "Comparative Statics for Aggregative Games. The Strong Concavity Case", *Mathematical Social Sciences* 28, 151-165.

Dubey, P., O. Haimanko, and A. Zapechelnyuk (2006): "Strategic Complements and Substitutes, and Potential games", *Games and Economic Behavior* 54, 77-94.

Ethier, W.J. (1982): "National and International Returns to Scale in the Modern Theory of International Trade", *American Economic Review* 72, 389-405.

Fields, G. S. and J.C.H. Fei (1978): "On Inequality Comparisons", *Econometrica* 46, 303-316.

Glover, G., *Waterloo: Myth and Reality*, Pen and Sword Military, South Yorkshire, 2014.

Gorman, W.M. (1968): "The Structure of Utility Functions", *Review of Economic Studies* 35, 367-390.

Guesnerie, R. (1975): "Pareto Optimality in Non-Convex Economies", *Econometrica* 43, 1-30.

Hinnosaar, T. (2018): "Optimal Sequential Contests", *Working Paper*, Collegio Carlo Alberto.

Holmstrom, B. (1982): "Moral Hazard in Teams", *Bell Journal of Economics* 13, 324-340.

Hong, L. and S.E. Page (2004): "Groups of Diverse Problem Solvers Can Outperform Groups of High-ability Problem Solvers", *Proceedings of the National Academy of Sciences* 16, 16385-16389.

Horwitz, S.K. and I.D. Horwitz (2007): "The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography", *Journal of Management* 33, 987-1015.

Jensen, M.K. (2018a): "Distributional Comparative Statics", *Review of Economic Studies* 85, 581-610.

Jensen, M.K. (2018b): "Aggregative Games", in *Handbook of Game Theory and Industrial Organization* (eds. L.C. Corchón and M.A. Marini), Edward Alger.

Johnson-Laird, P.N., *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*, Harvard University Press, 1983.

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

Keynes, J.M., *The General Theory of Employment, Interest and Money*, Macmillan, London, 1936.

Kimball, M. (1990): "Precautionary Saving in the Small and in the Large", *Econometrica* 58, 53-73.

Kimball, M. (1993): "Standard Risk Aversion", *Econometrica* 61, 589-611.

Lambert, P.J., *The Distribution and Redistribution of Income*, 3rd Edition, Manchester University Press, 2001.

Lancaster, K.J. (1966): "A New Approach to Consumer Theory", *Journal of Political Economy* 74, 132-157.

Landemore, H., *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*, Princeton University Press, 2012.

Leland, H.E. (1968): "Saving and Uncertainty: The Precautionary Demand for Saving", *Quarterly Journal of Economics* 82, 465-473.

Mangasarian, O.L. (1965): "Pseudo-Convex Functions", *SIAM Journal of Control* 3, 281-290.

Milgrom, P. and J. Roberts (1994): "Comparing Equilibria", *American Economic Review* 84, 441-459.

Nehring, K. and C. Puppe (2002): "A Theory of Diversity", *Econometrica* 70, 1155-1198.

Page, S.E., *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton University Press, 2008.

Rawls, J., *A Theory of Justice*, Harvard University Press, 1971.

Reynolds, A. and D. Lewis (2017): "Teams Solve Problems Faster When They're More Cognitively Diverse", *Harvard Business Review* March 30.

Roland, S. and C. Galunic (2004): "More than Network Structure: How Knowledge Heterogeneity Influences Managerial Performance and Innovativeness", *Strategic Management Journal* 25, 541-562.

Romer, P. (1990): "Endogenous Technological Change", *Journal of Political Economy* 98, 71-102.

Rosen, S. (1981): "The Economics of Superstars", *American Economic Review* 71, 845-58.

Rothschild, M. and J.E. Stiglitz (1970): "Increasing Risk I: A Definition", *Journal of Economic Theory* 2, 225-243.

Rothschild, M. and J.E. Stiglitz (1971): "Increasing Risk II: Its Economic Consequences", *Journal of Economic Theory* 3, 66-84.

Scarsini, M. (1998): "Multivariate Convex Orderings, Dependence, and Stochastic Equality", *Journal of Applied Probability* 35, 93-103.

Selten, R., *Preispolitik der Mehrproduktenunternehmung in der Statischen Theorie*, Springer Verlag, Berlin, 1970.

Shorrocks, A.F. (1980): "The Class of Additively Decomposable Inequality Measures", *Econometrica* 48, 613-625.

Simons, T., L.H. Pelled and K.A. Smith (1999): "Making Use of Difference: Diversity, Debate, and Decision Comprehensiveness in Top Management Teams", *The Academy of Management Journal* 42, 662-673.

Singer, D.J. (2019): "Diversity, Not Randomness, Trumps Ability", *Philosophy of Science* 86, 178-191.

Thoma, J. (2015): "The Epistemic Division of Labor Revisited", *Philosophy of Science* 82, 454-72.

Thompson, A. (2014): "Does Diversity Trump Ability?", *Notices of the American Mathematical Society* 61, 1–24.

Tobin, J. (1965): "The Theory of Portfolio Selection", in *The Theory of Interest Rates* (Hahn, E. and F. Brechling, Eds.), Macmillian, London.

Zollman, K.J.S. (2010): "The Epistemic Benefit of Transient Diversity", *Erkenntnis* 72, 17-35.

A Appendix: Proofs Omitted From the Main Text

A.1 Some Preliminary Developments

Definition 4 ***(Backwards Response Functions)*** *An action $s_i \in S_i$ is called a backward response to the aggregate $Q \in \mathbb{R}$ for an individual with characteristics $c_i \in C$, and we write $s_i \in B(Q, c_i)$, if $s_i \in R(Q - s_i, c_i)$ where $R(x_{-i}, c_i) = \arg\max_{s_i \in S_i} \Pi(s_i, x_{-i} + s_i, c_i)$. If $B(Q, c_i) = \{b(Q, c_i)\}$ for all Q and c_i, this defines a function b called the backward response function, and a function $\sum_i b(Q, c_i)$ called the aggregate backward response function.*

In an aggregative game, Q^* is an equilibrium aggregate if and only if it is a fixed point of the aggregate backward response function:

$$Q^* = \sum_i b(Q^*, c_i) \; . \tag{23}$$

It follows that if we can characterize backward response functions, then we can map diversity into outcomes.

Lemma 2 ***(Key Properties of Nice Aggregative Games)*** *Consider the nice aggregative games $(\Pi, S_i, c_i)_{i=1}^I$, where $(c_1, \ldots, c_I) \in C^I$, and assume that these satisfy the local solvability condition. Then:*

1. *The backward response function $b(Q, c_i)$ is a well-defined continuous function for all i, and $s_i = b(Q, c_i) \Leftrightarrow \Psi_i(s_i, Q, c_i) = 0$.*

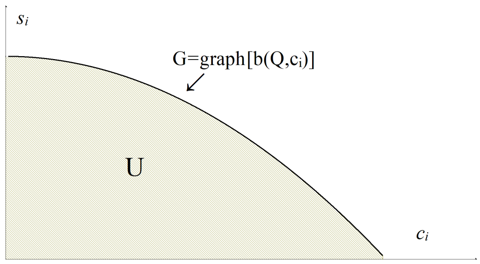
2. *If $D_{s_i}\Psi < 0$ then (a) the backward response function $b(Q, c_i)$ is concave in the individuals' characteristics $c_i \in C$ if and only if the games exhibit quasi-concave differences, (b) the backward response function $b(Q, c_i)$ is convex in the individuals' characteristics $c_i \in C$ if and only if the games exhibit quasi-convex differences, and (c) the aggregate backward response function $\sum_i b(Q, c_i)$ begins above and ends below the $45°$-line.*

3. *If $D_{s_i}\Psi > 0$ then (a) the backward response function $b(Q, c_i)$ is concave in the individuals' characteristics $c_i \in C$ if and only if the games exhibit quasi-convex differences, (b) the backward response function $b(Q, c_i)$ is convex in the individuals' characteristics $c_i \in C$ if and only if the games exhibit quasi-concave differences, and (c) the aggregate backwards response correspondence begins below and ends above the $45°$-line and intersects it precisely once*

Proof.

1. That the backward response function is well-defined and continuous in Q is a standard result under the stated conditions (see *e.g.* Jensen (2018b), Propositions 2).

2. Note first that since $b(Q,c_i) \in \mathrm{Int}(S_i)$, $b(Q,c_i) = s_i \Leftrightarrow \Psi(s_i,Q,c_i) = 0$. Hence the graph of the backward response function is $G = \mathrm{graph}[b(Q,c_i)] = \{(s_i,c_i) \in S_i \times C : \Psi(s_i,Q,c_i) = 0\}$, its upper contour set is $U = \{(s_i,c_i) \in S_i \times C : \Psi(s_i,Q,c_i) \geq 0\}$, and its lower contour set is $L = \{(s_i,c_i) \in S_i \times C : \Psi(s_i,Q,c_i) \leq 0\}$.



Fix Q. If $D_{s_i}\Psi \leq 0$, then, as in the figure, the upper contour set U equals the lower epigraph of G: $(s_i,c_i) \in U \Rightarrow (\tilde{s}_i,c_i) \in U$ for all $\tilde{s}_i \leq s_i$. Hence $b(Q,c_i)$ is concave in $c_i \in \mathcal{N}(c_i')$ if and only if $U \cap (\mathcal{N}(c_i') \times \mathcal{N}(b(Q,c_i')))$ is convex, which in turn is equivalent to $\Psi(s_i,Q,c_i)$ being quasi-concave in $(s_i,c_i) \in \mathcal{N}(c_i') \times \mathcal{N}(b(Q,c_i'))$. Hence $b(Q,\cdot)$ is concave, and by quasi-concave differences this holds for all Q (and this is clearly also necessary). Note that this argument applies whether or not s_i is monotone in c_i or c_i is one-dimensional. The case of quasi-convex differences is proved by repeating the previous argument for the lower contour set L (which is the upper epigraph of $b(Q,\cdot)$ when $D_{s_i}\Psi \leq 0$). That the aggregate backward response function begins above and ends below the 45 degree line when $D_{s_i}\Psi < 0$ is proved in Acemoglu and Jensen (2013).

3. If $D_{s_i}\Psi \geq 0$, the lower contour set L equals the lower epigraph of G (this situation corresponds to replacing U with L in the figure); and the upper contour set U equals the upper epigraph of G. Using these observations, the concavity and convexity of $b(Q,\cdot)$ is proved in the same manner as under 2. Since $s_i = b(Q,c_i) \Leftrightarrow \Psi(Q,s_i,c_i) = 0 \Leftrightarrow s_i = r(Q - s_i,c_i)$ it follows from the implicit function theorem that $D_Q b(Q,c_i) = r'/(1+r')$ and that $r' < -1 \Leftrightarrow D_{s_i}\Psi(Q,s_i,c_i) > 0$. Hence in a nice aggregative game, $D_{s_i}\Psi(Q,s_i,c_i) > 0 \Leftrightarrow D_Q b(Q,c_i) = r'/(1+r') > 1$. Therefore the aggregate backward response function must have slope strictly greater than 1 at all points, in particular it is increasing and can intersect with the 45-degree line at most once. ■

To clarify the contribution, 1. and 2(c) are well known, and 3(c) follows straight-forwardly from the definitions and standard arguments of the type found *e.g.* in Corchón (1994). The rest of the observations are new. Note that the paper's definition of quasi-convex/quasi-concave differences must not be confused with the similarly named definitions in Jensen (2018a). Jensen (2018a)'s definitions ensure that *best-response* functions are convex or concave which is a non-strategic property of behavior that is neither necessary or sufficient for backward response functions to be convex or concave.

A.2 Proof of Theorem 1

Say that nice aggregative games $(\Pi, S, c_i)_{i=1}^I$, $c \in C^I$ have a *symmetric representation* if there exists a set of symmetric aggregative games $(\tilde{\Pi}, s, \bar{c})_{i=1}^I$, where $\bar{c} \in \{I^{-1} \cdot (\sum_i c_i) : c \in C^I\} \subseteq \mathbb{R}^N$ is any feasible vector of mean characteristics in the original game, such that $\sum_i b(Q, c_i) = I\tilde{b}(Q, \bar{c})$ for all $c \in C^I$ and $Q \in X$ when b denotes the backward response function of the original games and $\tilde{b}$ the backward response function of $(\tilde{\Pi}, s, \bar{c})_{i=1}^I$.[22] Clearly, if aggregative games have a symmetric representation then for all $c \in C^I$, the equilibrium aggregate coincide with the equilibrium aggregate of the symmetric representation $(\tilde{\Pi}, S, \bar{c})_{i=1}^I$ where $\bar{c} = I^{-1} \cdot (\sum_i c_i) \in \mathbb{R}^N$. By definition, there exists a symmetric representation if for some $\tilde{b}$ and all $c \in C^I$, $\sum_i b(Q, c_i) = I\tilde{b}(Q, \bar{c})$. It immediately follows that the games have a symmetric representation if and only if $b(Q, c_i) = A(Q) + B(Q)c_i$ (see also Bergstrom and Varian (1985)). Since $D_{s_i}\Psi(s_i, Q, c_i) \neq 0$, the upper and lower contour sets of Ψ must locally coincide with (the relevant local restrictions of) $\{(s_i, c_i) : s_i \leq A(Q) + B(Q)c_i\}$ and $\{(s_i, c_i) : s_i \geq A(Q) + B(Q)c_i\}$ where either one could be the lower or upper contour. Hence if the games admit a symmetric representation, Ψ is quasi-monotone in the neighborhood of any point with $\Psi = 0$. That quasi-monotonicity of Ψ implies affine backward responses follows by the same argument considering again the upper and lower contour sets.

A.3 Proof of Theorem 2

To prove 1. we must consider two separate cases according to whether the local solvability condition holds with $D_{s_i}\Psi < 0$ or $D_{s_i}\Psi > 0$. "$D_{s_i}\Psi < 0$": Since the games exhibit quasi-convex differences, the backward response function $b(Q, c_i)$ is convex in c_i for every Q by 2.b. of Lemma 2. Therefore $\sum_i b(Q, c_i) \geq \sum_i b(Q, \tilde{c}_i)$ whenever $c \succeq_{cx} \tilde{c}$ by the definition of the convex diversity order. Since by Lemma 2, the aggregate backward response function is continuous and begins above and ends below the 45 degree line, it follows that the smallest and largest equilibrium aggregates must increase. "$D_{s_i}\Psi > 0$": Since the games exhibit quasi-convex differences, the backward response function $b(Q, c_i)$ is concave in c_i by 3.b of Lemma 2. Hence $c \succeq_{cx} \tilde{c}$ implies $-\sum_i b(Q, c_i) \geq -\sum_i b(Q, \tilde{c}_i) \Leftrightarrow \sum_i b(Q, c_i) \leq \sum_i b(Q, \tilde{c}_i)$. Since the aggregate backward response function begins below and ends above the 45 degree line by Lemma 2, this implies that the (unique) equilibrium aggregate increases. For the second statement, let $\succeq$ be a sensible diversity order and assume that $\succeq$-diversity reduces the smallest equilibrium aggregate. Since $\succeq_{cx}$-diversity increases the aggregate and the game by Theorem 1 does not exhibit quasi-monotone differences, there exist by 1. $c \succeq_{cx} \tilde{c}$ such that $Q_S(c) > Q_S(\tilde{c})$. But since $\succeq$ is sensible, $c \succeq_{cx} \tilde{c} \Rightarrow c \succeq \tilde{c}$ and therefore by assumption $Q_S(c) \leq Q_S(\tilde{c})$. A contradiction. Precisely the same argument applies to the largest equilibrium aggregate. The proof of 2. is similar and is therefore omitted.

[22]Here $X \subseteq \mathbb{R}$ denotes the set of feasible equilibrium aggregates.

A.4 Proof of Theorem 3

We begin with sufficiency. In each case, we establish both the case where index diversity increases and reduces the aggregate.

1. Since the backward response correspondence is invariant to algebraic manipulations of Ψ, the condition that $h_1(Q) < 0$ entails no loss of generality. By direct calculations, $b(Q, c_i) = [\frac{-h_1(Q)g_1(c_i)}{g_2(c_i)}]^{\epsilon^{-1}} - h_3(Q)$ and $\text{sign}[D_{s_i}\Psi(s_i, Q, c_i)] = \text{sign}[g_2(c_i)\epsilon] \neq 0$ by the local solvability condition. It follows that $\sum_i b(Q, c_i) \geq \sum_i b(Q, \tilde{c}_i) \Leftrightarrow (\star)\ \sum_i (\frac{g_1(c_i)}{g_2(c_i)})^{\epsilon^{-1}} \geq \sum (\frac{g_1(\tilde{c}_i)}{g_2(\tilde{c}_i)})^{\epsilon^{-1}}$. By the argument used in the proof of Theorem 2, we therefore conclude that the (smallest and largest) equilibrium aggregate increases if either (i) $\text{sign}[g_2(c_i)\epsilon] < 0$ and ($\star$) holds, or (ii) $\text{sign}[g_2(c_i)\epsilon] > 0$ and ($\star$) holds with the inequality sign reversed. Hence the aggregate increases in the diversity index

$$D(c) = -\text{sign}[\epsilon g_2(c_i)] \sum_i (\frac{g_1(c_i)}{g_2(c_i)})^{\epsilon^{-1}}$$

$D(c)$ is a sensible diversity index if and only if $-\text{sign}[\epsilon g_2(c_i)](\frac{g_1(c_i)}{g_2(c_i)})^{\epsilon^{-1}}$ is convex. If instead $-\text{sign}[\epsilon g_2(c_i)](\frac{g_1(c_i)}{g_2(c_i)})^{\epsilon^{-1}}$ is concave, $-D(c)$ is a sensible diversity index, and clearly increased diversity in the sense of the index $-D(c)$ reduces the aggregate.

2. By direct calculations, $b(Q, c_i) = \log(h_1(Q)) + h_2(Q)[\log(g_1(c_i)/g_2(c_i))] - h_3(Q)$, $D_{s_i}\Psi(s_i, Q, c_i) = -(g_2(c_i))^{h_2(Q)} \exp(s_i - h_3(Q)) \neq 0$ (by local solvability), and $\sum_i b(Q, c_i) \leq \sum_i b(Q, \tilde{c}_i) \Leftrightarrow (\star)\ h_2(Q) \sum_i [\log(g_1(c_i)/g_2(c_i))] \leq h_2(Q) \sum_i [\log(g_1(\tilde{c}_i)/g_2(\tilde{c}_i))]$. If $D_{s_i}\Psi > 0$, the equilibrium aggregates increase if ($\star$) holds, and if $D_{s_i}\Psi < 0$, the (smallest and largest) equilibrium aggregate increases if ($\star$) holds with the inequality sign reversed. Hence if $\log(g_1(c_i)/g_2(c_i))$ is convex and $(g_2(c_i))^{h_2(Q)}h_2(Q) > 0$ $[(g_2(c_i))^{h_2(Q)}h_2(Q) < 0]$, it thus follows that the aggregate is increasing [decreasing] in the sensible diversity index

$$\sum_i [\log(g_1(c_i)) - \log(g_2(c_i))]$$

If $-\log(g_1(c_i)/g_2(c_i))$ is convex and $(g_2(c_i))^{h_2(Q)}h_2(Q) > 0$ $[(g_2(c_i))^{h_2(Q)}h_2(Q) < 0]$, the aggregate is decreasing [increasing] in the sensible diversity index

$$\sum_i [\log(g_2(c_i)) - \log(g_1(c_i))] \,.$$

3. In this case $D_{s_i}\Psi(s_i, Q, c_i) = (s_i - h_3(Q))^{-1} > 0$ and $b(Q, c_i) = \exp(h_1(Q))\exp(g_1(c_i)) + h_3(Q)$. So when $-\exp(g_1(c_i))$ is convex, the aggregate increases in the sensible diversity index

$$D(c) = -\sum \exp(g_1(c_i))$$

and when $-\exp(g_1(c_i))$ is concave, the aggregate decreases in the sensible diversity index $-D(c)$.

As for necessity, for $\alpha(Q) + \beta(Q)\gamma(c_i)$ we must have $\frac{\partial b(Q,c_i)/\partial c_i}{\partial b(Q,\tilde{c}_i)/\partial \tilde{c}_i} = M(c_i)/M(\tilde{c}_i)$ (or something like that). Key will be to write in terms of Ψ and argue that s_i cannot enter (if it did, it would enter as $b(Q, c_i)$ and so we'd have to cancel Q out which we can only do if $b(Q, c_i) = \beta(Q)\gamma(c_i)$, etc.). What do we get again if we divide a demand function with respect to the price? (direct price effect? So relative price effect independent of income...) As a trick we might use that $b(Q, c_i) = \arg\max_{s_i \in [0,Q]} \int_{-\infty}^{s_i} \Psi(Q, \tau, c_i)\, d\tau$.

A.5 Proof of Theorem 4

Consider $D(\tilde{c}) = \sum_i \psi(\tilde{c}_i)$ where $\psi(c) = -\text{sign}[D_{s_i}\Psi]b(Q^*, c)$. If the games exhibits quasi-convex differences, then $b(Q^*, \cdot)$ is convex if $D_{s_i}\Psi < 0$ and concave if $D_{s_i}\Psi > 0$ (Theorem 2). Hence ψ is convex, and the index D sensible. If the game exhibits quasi-concave differences, we instead take $\psi(c) = \text{sign}[D_{s_i}\Psi]b(Q^*, c)$ and note by Theorem 2 that $b(Q^*, \cdot)$ is concave if $D_{s_i}\Psi < 0$ and convex if $D_{s_i}\Psi > 0$. So again ψ is convex and D therefore sensible. We shall prove only the case where the games exhibit quasi-convex differences (the quasi-concave differences case involves only obvious modifications). We must show that for any comparison population $\tilde{c}$ with equilibrium aggregate $\tilde{Q}^*$, we have $\tilde{Q}^* \geq Q^*$ if and only if $D(\tilde{c}) \geq D(c)$. There are two cases: $D_{s_i}\Psi < 0$ and $D_{s_i}\Psi > 0$. In the first case, the aggregate backwards response function begins above the 45-degree line (Theorem 2) and since by assumption it intersects the 45-degree line only once, it obviously ends below it. Furthermore, $D(\tilde{c}) \geq D(c) \Leftrightarrow \sum_i b(Q^*, \tilde{c}) \geq \sum_i b(Q^*, c)$. By the argument used in the proof of Theorem 1 in Acemoglu and Jensen (2018), this holds if and only if $\tilde{Q}^* \geq Q^*$. When $D_{s_i}\Psi > 0$, the aggregate backwards response function begins below the 45-degree line, ends above it and intersects it precisely once (Theorem 2). Furthermore, $D(\tilde{c}) \geq D(c) \Leftrightarrow \sum_i b(Q^*, \tilde{c}) \leq \sum_i b(Q^*, c)$. Leaving the details to the reader (this case is not relevant in the setting of Acemoglu and Jensen (2018)), the proof in that paper is straight-forwardly adapted to this dual case where the aggregate backwards response function instead shifts down but the intersection is from below, hence $\sum_i b(Q^*, \tilde{c}) \leq \sum_i b(Q^*, c) \Leftrightarrow \tilde{Q}^* \geq Q^*$.

A.6 Proofs omitted from Section 5

Pseudo-concavity and Uniqueness.

Lemma 3 *Assume $\alpha \leq \beta$ and $(OP''(O))/P'(O) < \beta$ for all $O > 0$. Then payoff functions are strictly pseudo-concave and there exists a unique pure strategy Nash equilibrium in the models of Section 5.*

Proof. To simplify the calculations, replace output with total effective input, $\hat{P}(Q) = P(Q^{1/(1+\alpha)})$. Note that $(OP''(O))/P'(O) < \beta$ for all $O > 0$ if and only if $(Q\hat{P}''(Q))/\hat{P}'(Q) < (\beta-\alpha)/(1+\alpha)$ for all $Q > 0$. Since payoff functions are smooth, it is necessary and sufficient for strict pseudo-concavity that $v\hat{P}'(\sum_{j=1}^{I} s_i) - \tilde{\gamma} s_i^{(\beta-\alpha)/(1+\alpha)} = 0 \Rightarrow v\hat{P}''(\sum_{j=1}^{I} s_i) - \tilde{\gamma}(\beta-\alpha)/(1+\alpha) s_i^{(\beta-\alpha)/(1+\alpha)-1} < 0$ where $\tilde{\gamma}$ is a constant. Since $\hat{P}' > 0$, this can also be written as $v\hat{P}'(\sum s_i) - \tilde{\gamma} s_i^{(\beta-\alpha)/(1+\alpha)} = 0 \Rightarrow (\star)$ $(s_i\hat{P}''(Q))/\hat{P}'(Q) < (\beta-\alpha)/(1+\alpha)$. If $\hat{P}'' \leq 0$, this inequality holds automatically and if $\hat{P}'' > 0$, it holds because $s_i \leq Q$. If $\hat{P}'' < 0$, the aggregate backward response function is decreasing and the equilibrium therefore unique. When $\hat{P}'' \geq 0$, we calculate $(D_Q b(Q, c_i)Q)/b(Q, c_i) = (Q\hat{P}''(Q)(1 +$

$\alpha)[\hat{P}'(Q)]^{(1+\alpha)/(\beta-\alpha)-1})/((\beta-\alpha)[\hat{P}'(Q)]^{(1+\alpha)/(\beta-\alpha)}) = (1+\alpha)\cdot(Q\hat{P}''(Q))/((\beta-\alpha)\hat{P}'(Q)) < 1$. The last inequality is implied by ($\star$) since $s_i \leq Q$. Hence the share correspondence $b(Q, c_i)/Q$ is strictly decreasing which by Jensen (2018b), Theorem 6 implies the existence of a unique pure strategy Nash equilibrium. ■

Example 5 *Consider $P(O) = AO^\rho$ which is a CES production game, $P(O) = A(\sum_i s_i^{1+\alpha})^{\rho/(1+\alpha)}$. Since $(OP''(O))/P'(O) = \rho - 1$, the condition is that $\rho < 1 + \beta$. If $\rho = 1$ (constant returns), this implies $\beta > 0$ but if $\rho < 1$ (decreasing returns), β can be negative.*

Note that if $\alpha > \beta$, then the condition that $\frac{OP''(O)}{P'(O)} < \beta$ is necessary for pseudo-concavity (this is seen from the previous proof noting that ($\star$) must hold). In particular, P must be strictly concave. The condition $\frac{OP''(O)}{P'(O)} < \beta$ still, by the same proof, implies existence of a unique *interior* Nash equilibrium (when $\alpha > \beta$ it is easy to see that there will be multiple boundary equilibria as well). ■

B A Descriptive Theory of Diversity (Supplement)

The starting point, just as in the paper's models, is a "Lancasterian view" of individuals as collections of characteristics $c_i \in C$, where $C \subseteq \mathbb{R}^N$ is a given *set of characteristics*.[23] That characteristics are vectors of real numbers entails no loss of generality because we may label any non-quantitative characteristic such as, say, "marital status" by taking $c_i^n = 1$ if individual i is married, and $c_i^n = 0$ if he is not. That characteristics are drawn from the *same* set C is on the other hand restrictive. It signifies that we (only) aim to compare populations whose members are sufficiently similar in the first place to be "cooked from the same basket of ingredients" (a biologist would call this a *species*). The restriction is fully adequate for our purposes and completely aligned with modern use of the word *diverse* whereby, for example, groups of women and men can be more or less diverse (gender diversity) but at the same time be seen as comparable in all functionally relevant dimensions — as opposed to belonging to different "species" with fundamentally incomparable characteristics. Crucially, we cannot, and also do not aim to, compare "an angel and a stone" with "two angels", or "6.000.130.000 humans" with "6.000.000.000 humans and 130.000 chimpanzees" (Nehring and Puppe (2002), p.1155). This marks a first difference from intrinsic or "existence" value theories of diversity. The second and more substantial difference is that we remain purely descriptive: To us "more diverse" never implies any value or expresses any preference in and of itself.[24]

What follows takes its cue from consumption theory, although as already mentioned, our orderings do not express preference (which is not only different from intrinsic value theories, but also from consumption theory). Note that for clarity, we focus first on (direct) diversity of innate characteristics, and then map the relevant concepts and results over to situations where characteristics are functionally relevant because they separate individuals into groups (cognitive diversity).

Definition 5 (Diversity Orders) *Let $C \subseteq \mathbb{R}^N$ denote the set of characteristics and consider a population of size $I \geq 2$. A diversity order $\succeq$ is a reflexive, transitive, and permutation invariant binary relation on C^I, and when $c \succeq \tilde{c}$, we say that c is more $\succeq$-diverse than $\tilde{c}$.*

The requirements imposed here are minimal and they are clearly reasonable in light of the descriptive interpretation. In particular, permutation invariance ensures that a comparison of populations does not depend on the order in which we list the individuals.[25] Now imagine that

[23] Of course Lancaster (1966) studies consumption baskets and not people. Note that, just as Lancaster, we take $N < +\infty$ but this is just a simplification (everything that follows is easily extended to an infinite number of characteristics). Finally, note as a subtlety that the actual index i may carry content (for example, i may specify a specific position at an assembly line). This information is, however, *not* intrinsic to individuals but part of the strategic framework of Section 3.

[24] To elaborate, we focus squarely on what Nehring and Puppe (2002) call the "use value" of diversity (p.1156). This "use value" depends on the context ("the rules of the game") and the principal's objective. In particular, different principals may have different and possibly conflicting objectives, so to us the "value" of diversity is always in the eyes of the beholder. This is in sharp contrast to intrinsic value theories such as Nehring and Puppe (2002).

[25] Formally, if $c = (c_1, \ldots, c_I)$ is more diverse than $\tilde{c} = (\tilde{c}_1, \ldots, \tilde{c}_I)$ then $c_p = (c_{p(1)}, \ldots, c_{p(I)})$ is more diverse than $\tilde{c}_q = (\tilde{c}_{q(1)}, \ldots, \tilde{c}_{q(I)})$ for arbitrary bijective mappings, $p, q : \{1, \ldots, I\} \to \{1, \ldots, I\}$.

there is just a single functionally relevant characteristic $C \subseteq \mathbb{R}$ such as ability or motivation level, and then consider the following claim:

(DC) "*If population A's characteristics are a mean-preserving spread of population B's characteristics, then population A is more diverse than population B*".

The validity of this statement is a key premise of this paper. Recall that Rothschild and Stiglitz (1970) formally define $c \in \mathbb{R}^I$ to be a *mean-preserving spread* of $\tilde{c} \in \mathbb{R}^I$ if "c's distribution equals the distribution of $\tilde{c}$ plus uncorrelated variation". This very statement clearly lends support to (DC) but as we shall see in the first paragraph after Lemma 7, we have a much stronger argument in its favor. Rothschild and Stiglitz also prove that c is a mean-preserving spread of $\tilde{c}$ if and only if ($\star$) $\frac{1}{I}\sum_i f(c_i) \geq \frac{1}{I}\sum_i f(\tilde{c}_i)$ for any convex function $f : \mathbb{R}^N \to \mathbb{R}$, where in their case $N = 1$. In this paper we shall (slightly) abuse terminology and also when $N > 1$ call c a mean-preserving spread of $\tilde{c}$ if ($\star$) holds. As we explain in the footnote, this actually makes (DC) even more plausible when $N > 1$ than when $N = 1$.[26]

Accepting (DC) as an "axiom" significantly restricts the set of meaningful diversity comparisons because we can discard any candidate diversity order that is not sensible in the following sense:

Definition 6 (Sensible Diversity Orders) *A diversity order $\succeq$ is sensible if $c \succeq \tilde{c}$ whenever c is a mean-preserving spread of $\tilde{c}$.*

The following are all sensible diversity orders: The *convex diversity order*, $c \succeq_{cx} \tilde{c} \Leftrightarrow \frac{1}{I}\sum_i f(c_i) \geq \frac{1}{I}\sum_i f(\tilde{c}_i)$ for any convex function $f : \mathbb{R}^N \to \mathbb{R}$; the *Lorenz order*, $c \succeq_{Lorenz} \tilde{c} \Leftrightarrow \frac{1}{I}\sum_i f(c_i/\bar{c}) \geq \frac{1}{I}\sum_i f(\tilde{c}_i/\bar{\tilde{c}})$ for any convex function $f : \mathbb{R} \to \mathbb{R}$ where the bar takes the mean; the *moments order*, $c \succeq_M \tilde{c} \Leftrightarrow [\sum c_i^k \geq \sum \tilde{c}_i^k,\ k = 1, 2, 3, \ldots]$, and the *k'th moment order* ($k = 1, 2, \ldots$), $c \succeq_{M,k} \tilde{c} \Leftrightarrow \sum c_i^k \geq \sum \tilde{c}_i^k$. The convex diversity order says of course precisely that c is a mean-preserving spread of $\tilde{c}$ and as such it is the uniquely most restrictive diversity order that is sensible. Given (DC), the convex diversity order therefore expresses the *uniquely least controversial interpretation of the words "more diverse"*, *i.e.*, the only way to define "more diverse" which every sensible person must agree with. The Lorenz, moments and k'th moment diversity orders (all of which require $N = 1$) are all less restrictive than the convex diversity order and they are therefore sensible. But equally, they are more controversial: Not every sensible person would necessarily agree, for example, that a population is more diverse than another population simply because its characteristics have a greater second moment ($\succeq_{M,2}$), even if the means are equal, *i.e.*, even if the characteristics are more diverse in the sense of the *mean-variance order* from the theory of portfolio

[26]With multiple characteristics, ($\star$) implies that each characteristic is subjected to a mean preserving spread as seen by taking $f(c) = \tilde{f}(c^n)$ where $\tilde{f} : \mathbb{R} \to \mathbb{R}$ is convex. In particular, the mean of each characteristic must be the same. But ($\star$) also imposes any number of "cross-restrictions" (Scarsini (1998)), hence amounts to an extremely strong comparison criterion in the multi-dimensional case. Note also that the main reason for our (abuse of) terminology is that it simplifies language considerably throughout the paper by allowing us to separate (DC) from the convex order defined in a moment.

selection (*e.g.*, Tobin (1965)). Nonetheless, this does not preclude the possibility that the (sensible) mean-variance diversity order *precisely* captures the functional "use" value of diversity in specific situations (Example 1).

If we define "more diverse" through the convex, Lorenz, moments, or mean-variance diversity orders, there will be populations which are neither more or less diverse than other populations. With *complete* diversity orders such as the k'th moment order, a population is in contrast either more diverse or less diverse than any other population. As complete preferences in consumption theory admit utility representations, complete diversity orders admit numerical diversity representations, $R : C^I \to \mathbb{R}$: $c \succeq \tilde{c} \Leftrightarrow R(c) \geq R(\tilde{c})$. For example, the Gini coefficient $R_G(c) = (\sum_i \sum_j |c_i - c_j|)/(2I \sum_i c_i)$ is a diversity representation, and the diversity order it induces is sensible because the Lorenz order is sensible, $c \succeq_{cx} \tilde{c} \Rightarrow c \succeq_{Lorenz} \tilde{c}$, and by Atkinson (1970), $c \succeq_{Lorenz} \tilde{c} \Rightarrow R_G(c) \geq R_G(\tilde{c})$. Although diversity representations is an interesting topic, it turns out that for the class of games considered in this paper (aggregative games) we only need a very special type of additive representations which are closely related to the decomposable inequality measures of Shorrocks (1980). To define these, let $\bar{c} = (\bar{c}^1, \ldots, \bar{c}^N) \in \mathbb{R}^N$ denote the means of a population's vector of characteristics $c \in C^I$.

Definition 7 (Diversity Indices) *A diversity index is a function $D : C^I \to \mathbb{R}$ of the form $D(c) = \frac{1}{\theta(\bar{c})} \sum_i [\psi(c_i) - \varsigma(\bar{c})]$ where $\theta(\bar{c}) > 0$ and ψ is a continuous function. A diversity index induces a (complete) diversity order under the convention that*

$$c \succeq \tilde{c} \Leftrightarrow D(c) \geq D(\tilde{c}) \; . \tag{B1}$$

If the induced diversity order is sensible, the index D is called a sensible diversity index.

For example, the k'th moment order is induced by a diversity index with $\theta(\bar{c}) = 1$, $\varsigma(\bar{c}) = 0$, and $\psi(c_i) = c_i^k$, and since the k'th moment order is sensible, this index is sensible.

When $N = 1$ and $\psi = \varsigma$ in Definition 7, D is a diversity index if and only if it is an additively decomposable inequality measure in the sense of Shorrocks (1980), p.617. Although the restriction $\psi = \varsigma$ has no proper foundation in the diversity setting, the relationship with Shorrocks' work is nonetheless extremely helpful to us because so much has been written about inequality measures (see *e.g.* Lambert (2001)) and much of the intuition is not profoundly different from situations where heterogeneity stems from innate characteristics rather than income or wealth differences.

Lemma 4 *Let $\succeq$ be induced by the diversity index $D(c) = \frac{1}{\theta(\bar{c})} \sum_i [\psi(c_i) - \varsigma(\bar{c})]$. The following statements are equivalent:*

1. *$\succeq$ is sensible*
2. *ψ is convex*
3. *$D(c) \geq D(\frac{1}{I} \sum_i c_i, \frac{1}{I} \sum_i c_i, \ldots, \frac{1}{I} \sum_i c_i)$ for all $c \in C^I$.*

Proof. "2. $\Rightarrow$ 1.": If ψ is convex and $c \succeq_{cx} \tilde{c}$ then $\bar{c} = \bar{\tilde{c}}$ and $\sum_i \psi(c_i) \geq \sum_i \psi(\tilde{c}_i)$. Therefore $c \succeq \tilde{c}$. "1. $\Rightarrow$ 2.": assume that $c \succeq_{cx} \tilde{c} \Rightarrow D(c) \geq D(\tilde{c})$. Pick any c and let $\tilde{c} = (\frac{1}{I}\sum_i c_i, \ldots, \frac{1}{I}\sum_i c_i)$. Clearly $c \succeq_{cx} \tilde{c}$ and $\bar{c} = \bar{\tilde{c}}$, hence $D(c) \geq D(\tilde{c}) \Leftrightarrow \sum_i \psi(c_i) \geq \sum_i \psi(\tilde{c}_i) \Leftrightarrow \frac{1}{I}\sum_i \psi(c_i) \geq \psi(\frac{1}{I}\sum_i c_i)$. Since ψ is continuous, it is therefore convex. "2 $\Leftrightarrow$ 3": Since c and $(\frac{1}{I}\sum_i c_i, \frac{1}{I}\sum_i c_i, \ldots, \frac{1}{I}\sum_i c_i)$ have the same mean and $\theta(\frac{1}{I}\sum_i c_i) > 0$, $D(c) \geq D(\frac{1}{I}\sum_i c_i, \frac{1}{I}\sum_i c_i, \ldots, \frac{1}{I}\sum_i c_i) \Leftrightarrow \frac{1}{I}\sum_i \psi(c_i) \geq \psi(\frac{1}{I}c_i)$. Since ψ is continuous, this last inequality is both necessary and sufficient for ψ to be convex. ∎

3. says that a population is never less diverse than the homogenous averaged version of itself.[27] This statement is impossible to disagree with, hence when focus is on diversity indices there can be no doubt that the claim (DC) must be accepted and attention restricted to sensible diversity orders. Since — as we shall see in Section 4 — we are precisely able to do so in aggregative games, any objections to (DC) are of little relevance for this paper's results.

The literatures on variability/uncertainty/inequality/information are full of indices all of which induce sensible diversity orders because they are monotone transformations of sensible diversity indices. One such class turns out to play a particularly prominent role in teamwork models.

Definition 8 ***(Shorrocks Diversity Indices)*** *Assume that $C \subseteq \mathbb{R}$. A Shorrocks diversity index is a diversity index of the form,*

$$D^{\mathrm{Sh}}_{\rho}(c) = \begin{cases} \frac{1}{I\rho(\rho-1)}\sum_i[(\frac{c_i}{\theta(\bar{c})})^{\rho} - 1] & \rho \neq 0, 1 \\ \frac{1}{I}\sum_i(\frac{c_i}{\theta(\bar{c})})\log(\frac{c_i}{\theta(\bar{c})}) & \rho = 1 \\ -\frac{1}{I}\sum_i \log(\frac{c_i}{\theta(\bar{c})}) & \rho = 0 \end{cases}, \tag{B2}$$

where $\theta : \mathbb{R} \to \mathbb{R}_{++}$. If $\theta(\bar{c}) = \bar{c}$, such an index is said to be mean independent*, and if $\theta(\bar{c}) = 1$, it is said to be* absolute*.*

In each case, it is easy to write D^{Sh}_{ρ} in the form $D(c) = \frac{1}{\theta(\bar{c})}\sum_i[\psi(c_i) - \varsigma(\bar{c})]$ where ψ is convex. Hence by Lemma 4, Shorrocks diversity indices are sensible. Mean independent Shorrocks indices were (also) first studied by Shorrocks (1980), who derived them axiomatically as cardinal income inequality measures. In the literature, the mean independent case is often also referred to as *generalized entropy indices* (and when $\rho \in \{0, 1\}$ as *Theil entropy indices*, see Shorrocks (1980), p.613). Note, however, that this terminology is slightly misleading because, as Shorrocks also points out, these indices in fact measure redundancy and not entropy. The Shannon entropy, for example, is *not* a diversity index — it is a group diversity index as explained below.[28] Shorrocks also proves that mean-independent measures forms the unique class of monotone transformations of one of the single most important classes of (normatively derived) inequality measures, namely additively

[27] In the decomposable case where $\psi = \varsigma$, it is straight-forward to adapt the proof of Theorem 2 in Shorrocks (1980) to show that ψ is non-negative and strictly convex if and only if: $D(c) \geq 0$ for all c with equality if and only if $c_i = c_j$ for all i, j (this is Assumption 2, p.614 in Shorrocks (1980)).

[28] Similarly, what ecologists call diversity (or bio-diversity) indices are normally group diversity indices in the sense of Definition 9 below.

decomposable *Atkinson indices* (Atkinson (1970), Lambert (2001)):

(B3)
$$D^A_\epsilon(c) = \begin{cases} 1 - \left[\frac{1}{I}\sum_i (\frac{c_i}{\bar{c}})^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}} & \text{for } \epsilon \in \mathbb{R}_{++}\backslash\{1\} \\ 1 - \frac{(\prod c_i)^{1/I}}{\bar{c}} & \text{for } \epsilon = 1 \end{cases}$$

In particular, any Atkinson index is a sensible diversity index as is the scale invariant (absolute) version where the mean $\bar{c}$ is replaced with 1. Note that if in this we take the limit, we get $\lim_{\epsilon\to+\infty}[1 - ((\sum_i c_i^{1-\epsilon})^{1/(1-\epsilon)})/I] = 1 + R(c)$ where

(B4)
$$R(c) = -\min\{c_1, \ldots, c_I\}$$

Since $R(c) \geq R(\tilde{c}) \Leftrightarrow \min_i c_i \leq \min_i \tilde{c}_i$, this index might be called the *Rawlsian diversity index* (*cf.* Rawls (1971)). The Rawlsian diversity index is a sensible diversity index because the pointwise limit of a sequence of sensible diversity indices is a sensible diversity index.[29] The reflexion of this index, $S(c) = R(-c) = \max\{c_1, \ldots, c_I\}$ could be called the *Strongest Link diversity index* since it measures diversity by the strongest link of the chain. Since a reflexion of a sensible diversity index is a sensible diversity index, $S(c)$ is sensible. It is intuitive that these diversity indices have roles to play in teamwork models because projects may be such that performance depends entirely on the best (or worst) workers. At the same time, the Rawlsian and Strongest Link indices illustrate particularly well why less restrictive diversity orders than the convex diversity order potentially garner controversy; certainly, not everyone would agree that a team is more diverse than another team simply because the best worker is more able or more highly motivated.

Finally, let us transfer the previous concepts and results over to situations where the functional relevance of innate characteristics derives from their induced division of individuals into (cognitive) groups. With I individuals and N groups, a *group frequency* is a vector $g = (g_1, g_2, \ldots, g_N) \in \{\{0, 1, \ldots, I\}^N : \sum_n g_n = I\}$ that specifies how large the N groups are. To make diversity comparisons with reference to such group frequencies we need to take into account that more diverse innate characteristics implies more homogenous group frequencies. To this end, let g' denote the permutation of g which lists the groups in order of increasing size, $g'_1 \leq g'_2 \leq \ldots \leq g'_N$ (for example, if $g = (1, 3, 2, 0)$ then $g' = (0, 1, 2, 3)$), and then consider:

Definition 9 ***(Diversity on Group Frequencies)*** *A population with group frequencies $\tilde{g}$ is convex order more diverse than a population with group frequencies g if $g' \succeq_{cx} \tilde{g}'$. A diversity order $\succeq$ defined on group frequencies is sensible if $g' \succeq_{cx} \tilde{g}' \Rightarrow \tilde{g} \succeq g$.*

Since these conventions simply reverse the order $\succeq_{cx}$ in comparison with our previous developments, everything immediately maps over. In particular, it follows from Lemma 4 that a *group diversity index*, $I(g) = \sum_{n:g_n>0} \psi(g_n)$, is sensible if and only if ψ is concave.[30] Hence any of the

[29] More generally, as may be shown, the space of sensible diversity indices is a pointwise closed topological vector space.

[30] Note that we must always exclude empty groups because these do not represent any individuals' innate characteristics.

diversity indices we have seen previously become sensible group diversity indices by multiplication with -1. For example, we get from D_1^{Sh} the *Shannon entropy* $-D_1^{\mathrm{Sh}}(g/N)$ which is therefore a sensible group diversity index on the relative frequencies in the mean independent case, and on the frequencies in the absolute case. A second example stems from the Rawlsian diversity index which by multiplication with -1 becomes $L(g) = \min\{g_n : g_n > 0\}$. Thus the *Smallest (non-empty) Group Size* is a sensible group diversity index. Perhaps the simplest example of them all is when $\psi(g_n) = 1$ because then $I(g) = \sum_{n:g_n>0} 1$ simply counts the number of non-empty groups. To finish, note that convex order group diversity is in itself an extremely intuitive concept. The reason is that g' is a mean-preserving spread of $\tilde{g}'$ if and only if we can get from g' to $\tilde{g}'$ by *sequentially moving individuals from larger to smaller groups without at any point changing the relative size ranking*.[31]

[31] Mathematically, this is because each step in such a sequence is a *progressive Pigou-Dalton transfer*, and by Theorem 2.1 in Fields and Fei (1978), g is a mean-preserving spread of $\tilde{g}$ if and only if it can be reached by a sequence of progressive Pigou-Dalton transfers.