DEPFID
DEPARTMENT OF ECONOMIC POLICY, FINANCE AND DEVELOPMENT
UNIVERSITY OF SIENA

DEPFID WORKING PAPERS

Mauro Caminati

A knowledge based approach to collaboration in basic research

8 / 2009

DIPARTIMENTO DI POLITICA ECONOMICA, FINANZA E SVILUPPO
UNIVERSITÀ DI SIENA

DEPFID Working Papers – 8 / October 2009

A knowledge based approach to collaboration in basic research

Mauro Caminati

Abstract

This paper suggests a knowledge based approach to the formation of collaboration networks in basic research. Though mainly focused on foundations, it provides the example of a set of knowledge distributions supporting effort allocations that are pairwise equilibria of the collaboration game. These equilibrium outcomes produce a collaboration network consisting of connected quasi stars.

Keywords: ideas, knowledge endowment, modularity, collaboration game, pairwise equilibrium, star network, small world.

JEL Classification: D85, O30

Address for correspondence: caminati@unisi.it

Dipartimento di Politica Economica, Finanza e Sviluppo – Università di Siena

1 Aims and scope of the paper

This paper recomposes under a unifying framework two styles of analysis concerned with the production of ideas, but so far developed in separate and seemingly independent strands of the literature. One is the evolutionary approach to the knowledge based analysis of the division of labour[1]. This takes place not only in the production of goods, but also and most crucially in the production of ideas (a review essay is Marengo, Pasquali and Valente [19]). The second field of enquiry is concerned with the incentive based explanation (as opposed to the statistical explanation) of collaboration networks in scientific research. Recent contributions in the second field are Carayol and Roux [5], Goyal, van der Leij and Moraga [10], [9].

The social network approach to the production of ideas has paid only lip service to the fact that agents are widely heterogeneous with respect to their knowledge endowment, and that the distribution of this endowment is a main determinant of the network architectures which can be sustained as equilibrium outcomes. This paper argues that knowledge heterogeneity is a crucial motivation to scientific collaboration. This view bears a close relation with the recombinant approach to the growth of knowledge (Reiter, [28], [27], Weitzman, [34]) which is here qualified to suggest that a new lineage of useful ideas, say, the family of the different versions of the steam-engine, grows out of recombinations of heterogeneous pre-existing ideas. There are certainly motivations behind collaboration in scientific research, which are not knowledge based, in that they are not concerned with the costs and benefits of exploiting knowledge complementarities. The varying individual preference for team work is a case in point. We shall completely abstract from this type of considerations, in what follows. Our primary aim is to develop a multi-agent foundation of the recombinant approach towards radical discovery. If regarded in this perspective, the rising trend towards scientific collaboration is a direct outcome of the increasing specialization of personal knowledge, which is itself produced by the growth of ideas.

In this paper, agents' quality is exclusively determined by what they know. The specification of the aggregate knowledge stock leads to an endogenous determination of the knowledge fields in the economy, with the expectation that individual knowledge is mostly specialized in a particular field or subfield. The further specification of a knowledge distribution over the set of agents induces a measure of knowledge heterogeneity on this set (as opposed to the geographical distance introduced in Carayol and Roux [5]), which gives rise to a trade off in the allocation of private effort to research projects, carried out in isolation, or in collaboration with other agents. For a given size of i's and j's knowledge endowments, the larger the knowledge heterogeneity between them, the higher the joint competence of their collaboration project. Still, a higher heterogeneity between agents' specializations makes collaboration more demanding in terms of the private effort which is necessary to make that competence effective.

[1] This style of analysis has often availed itself of Stuart Kauffman's $N-K$ fitness landscapes (Kauffman, [15]) and of other tools and concepts borrowed from the natural sciences.

The paper is organized as follows. Section 2 presents a short list of the structural properties shared by the co-authorship networks in a number of scientific disciplines. Section 3 presents the main line of argument. Section 4 builds the basic framework concerning the definition and properties of a knowledge endowment and produces an endogenous partition of the aggregate knowledge stock at a given date into disciplinary fields, based on the notion of modularity. Section 5 introduces a knowledge distribution over the set of agents and provides an accurate distinction between basic research, aimed at discovering new types (lineages) of ideas, and development, the activity of finding improved ideas within the existing lineages. The paper is mainly concerned with the former. Section 6 develops the knowledge based approach to the formation of collaboration networks in basic research. Section 7 concludes.

2 Scientific collaboration networks: empirical properties

The networks of scientific collaboration in a number of disciplines, such as mathematics, physics, biology, medical science, and economics[2] share a number of structural characteristics. A review is available in Newman [23] and in Goyal, van der Leij and Moraga [9]. A scientific collaboration network is built as follows. Every node identifies a researcher, and two nodes are linked if they co-authored a paper in a given time interval. The weight of the link in question is an increasing function of the number of co-authored papers. Scientific collaboration, defined as above, has been growing through time in every discipline, though possibly at a different pace (Laband and Tollison, [17]). We list below the characteristics of the empirical collaboration networks, which are of special interest in the present context.

1. The distribution of collaborators has a fat tail: a small fraction of scientists has a very large number of collaborators[3].

2. The typical collaboration network has 1 giant component comprising the majority of nodes; the other components, if present, are relatively small.

3. The average relational distance[4] between the nodes is small, and the maximum relational distance between a couple of nodes (the diameter) is also small.

4. The average clustering coefficient (the fraction of a node's neighbours that are each-other neighbours) is high, at least compared to random networks.

[2] According to the findings reported in Moody ([20]) sociology may be a partial exception, but the hypothesis seems to require corroboration through more directly comparable tests.

[3] Unlike other networks, the distribution for these collaboration networks do not strictly follow a power-law (are not scale free), although it has been suggested that they may follow a power law with an exponential cut-off (Newman, [23]).

[4] The relational distance between two nodes is the minimum number of links separating the nodes.

5. Nodes with many collaborators have a higher than average probability of being directly connected (positive assortativity), and most of their indirect connections go through a small number of collaborators.

6. Network connectivity is susceptible to the removal of the most connected nodes[5].

The above properties 1-6 are broadly consistent with the preferential attachment model of network formation[6]. Taken together, they suggest the stylized approximation that the undirected representation of a typical co-authorship network is a system of quasi-stars, hierarchically connected, mostly through their centres.

Definition 1 *A N star is a connected undirected network with N nodes and $N-1$ links. The centre node is directly connected to each of the other $N-1$ nodes, which form the periphery. A N quasi star is a connected undirected network with N nodes, such that: the number of links is larger than $N-1$; the centre node is directly connected to each of the other $N-1$ nodes; self-loops are allowed; a periphery node may have $P>1$ links, provided that $\frac{P}{N-1}$ is sufficiently small.*

A study by M. E. Newman and M. Girvan ([25]) extends the analysis to the co-authorship between the researchers attached to a trans-disciplinary scientific institution. They suggest that the set of researchers can be endogenously partitioned into scientific communities on the ground that a node belongs to a community if its within community relations are stronger than its between community relations. Newman and Girvan [25] apply their method of community-structure identification to the collaboration network of the scientists at the Santa Fe Institute, and find that the communities thus identified broadly correspond to scientific fields and sub-fields. Each (sub) field is a system of quasi stars, hierarchically connected, and the different fields joined, mostly by a critical link, joining two centre nodes.

There are two models which claim to provide a successful explanation of the above empirical characterization of a co-authorship network. The preferential attachment model (Barabasi and others [3]) introduces the following hypothesis: the probability that a new published paper contributes new connections to an individual researcher is an increasing function of the number of connections that the individual already has. The incentive based, game theoretic model (Goyal, van der Leij and Moraga [10]) is based on the hypothesis that there are two types of researchers, the smart ones, which can produce high quality ideas (and papers), and the others, which produce low quality ideas, but may contribute to the routine work, which is needed in research.

[5] More precisely, of the nodes with the highest betweenness score (Newman[23]). Moreover, the betweenness scores of collaborators are uncorrelated.

[6] In the preferential attachment model (Barabasi et al., [3]) the probability that a new research project contributes new connections to an individual researcher is an increasing function of the number of connections that the individual already has.

The present article is not directly concerned with paper publication and co-authorship. It has to do with the production of radically new ideas by different coalitions of researchers. Under the bold assumption that the quality-weighted output of a research coalition is correlated with its (quality weighted) paper publication, the data on co-authorship bear some relevance also in our case. On this ground, it seems to be desirable that the predictions of the model to be developed in this paper are prima facie consistent with the generic empirical characteristics of co-authorship networks.

3 The line of argument

Our basic assumption is that there is a selection for modularity in the evolution of knowledge (Simon, [31], Marengo et al. [19]). As a result, the knowledge set characterizing a scientific discipline is nearly decomposable into modules corresponding to more or less specialized fields or subfields. Like in Simon and Ando ([32]) original formulation of near-decomposability, the fitness interactions between ideas belonging to different modules are unfrequent, but, contrary to Simon ([31]), if the interaction occurs, it may have a strong influence on fitness (relative performance). If this is the case, the interaction between the specialized modules cannot be neglected (see Watson, [33]). In the next section, these notions will be made more precise through the language introduced by the following definitions.

Definition 2 *Ideas are embodied in the human brains. An idea is codified knowledge defined by a string* $\mathbf{a} \in \{0,1,s\}^N$ *of* N *elements. An element* a_n *is identified by its location* $n \in \mathbf{N} = \{1,...,N\}$ *on the string. A location* n *is silent (uneffective) for idea* $\mathbf{a}$ *if and only if* $a_n = s$*. The set of non silent locations of* a *is* $\mathbf{NS}(\mathbf{a}) = \{n \in \{1,...,N\}$ *s.t.* $a_n \neq s\}$*. A family, or type, of useful ideas, for instance, the family of the different versions of 'the wheel', is the subspace* $\mathbf{F} \subseteq \{0,1,s\}^N$ *of ideas with identical silent and non silent locations: if* $\mathbf{a} \in \mathbf{F}$ *and* $\mathbf{a}' \in \mathbf{F}$*, then* $a_n = s$ *if and only if* $a'_n = s$*. A type* $\mathbf{F}$ *is uniquely identified by the set* $\mathbf{NS}(\mathbf{F})$*.*

Definition 3 *Development is the search for better configurations of a set of idea types, which leaves the set of types unchanged. Basic research is the activity aimed at discovering new types of potentially useful ideas.*

When a new type is first discovered, it may well be the case that its known configurations are less fit than their competitor ideas already in existence. For this reason, the potential usefulness of types matters. Every agent spends her time endowment in development and basic research according to fixed proportions. The time endowment available for basic research is E, which is uniform across agents.

The benefits of modularity are most relevant in development. To make sure that the benefits from knowledge specialization do not entail the loss of the positive complementarities between the different fields, it is necessary that a set of agents, the knowledge integrators, preserve in their endowment the core ideas which provide the interfaces connecting the specialized fields. The above premises justify a distribution of knowledge, defined as a distribution of types, such that the endowments of two specialists in different fields are sufficiently heterogeneous; the endowments of two specialists in the same field are highly homogeneous; for every specialized endowment, there is a 'knowledge integrator's endowment, which is only partly heterogeneous with respect to the former. To make the notion of knowledge heterogeneity operational, we introduce a measure $d(ij)$ of heterogeneity between i's and j's knowledge endowments. If i and j form a collaboration, their joined competence is increasing with respect to $d(ij)$. The trade off is that higher heterogeneity means that higher effort is required to make the collaboration effective. Under the assumption that there are H researchers in the economy, we introduce the following definitions.

Definition 4 *A collaboration strategy of agent i is a choice of effort allocation* $\mathbf{S}_i = \{e_{i1}, ..., e_{iH}\}$*, where* e_{ij} *is i's effort contribution to the collaboration* (ij)*.*

Definition 5 *A strategy profile* $\mathbf{S}^* = \{\mathbf{S}_1^*, ..., \mathbf{S}_H^*\}$ *is a pairwise equilibrium if: (i)* $\mathbf{S}^*$ *is a Nash equilibrium. (ii)* $\mathbf{S}^*$ *is robust to the formation of two-agents coalitions. For every i and j in* $\mathbf{H}$*, such that* $0 = e_{ij} \in \mathbf{S}_i^*$*,* $0 = e_{ji} \in \mathbf{S}_j^*$*, there is no strategy pair* $(\mathbf{S}_i', \mathbf{S}_j')$*, such that:* $0 < e_{ij} \in \mathbf{S}_i'$*,* $0 < e_{ji} \in \mathbf{S}_j'$ *, and both i and j prefer the strategy profile* $\mathbf{S}' = (\mathbf{S}_{-(i+j)}^*, \mathbf{S}_i', \mathbf{S}_j')$ *the the profile* $\mathbf{S}^*$*.*

We claim that a highly stylized characterization of the knowledge distributions reflecting the historical trend towards rising specialization cum knowledge integration (explained by selection for modularity) supports a quasi-star network of collaboration in basic research as a pairwise equilibrium outcome.

4 Near-decomposability in idea space

An idea is a configuration of a family type. A type $\mathbf{F}$ is useful, if at least one useful idea in $\mathbf{F}$ has been discovered. $\mathbf{\Gamma}_t$ is the set of useful types at t, and $\alpha(\mathbf{\Gamma}_t) = \#\mathbf{\Gamma}_t$ is the number of such types. We fix a labelling of types in $\mathbf{\Gamma}$, such that $\mathbf{\Gamma} = \{\mathbf{F}_1, \mathbf{F}_2, ..., \mathbf{F}_{\alpha(\Gamma)}\}$. A configuration of $\mathbf{\Gamma}$, or knowledge configuration, is a list $\boldsymbol{\tau} = \{\mathbf{a}(\mathbf{F}_1), \mathbf{a}(\mathbf{F}_2), ..., \mathbf{a}(\mathbf{F}_{\alpha(\Gamma)})\}$ specifying one idea configuration $\mathbf{a}_f = \mathbf{a}(\mathbf{F}_f)$ for each family $\mathbf{F}_f$ in $\mathbf{\Gamma}$. $\boldsymbol{\tau} = \{\mathbf{a}_f \cup \mathbf{a}_{-f}\}$, where $\mathbf{a}_{-f}$ is the configuration of the families other than $\mathbf{F}_f$. In general, only a vanishing small fraction of the possible configurations of each idea type is 'useful'. The usefulness of an idea configuration $\mathbf{a}_f$ is, like fitness in biology, a relative, not an absolute concept, and can only be evaluated in the context of the given concomitant knowledge configuration $\mathbf{a}_{-f}$. This is because usefulness is affected by the positive or negative complementarities between ideas. The relative fitness of two ideas $\mathbf{a}_f$ and $\mathbf{a}_f'$ belonging to the same type $\mathbf{F}_f$ is evaluated by a fitness ratio $V(\mathbf{a}_f, \mathbf{a}_{-f})/V(\mathbf{a}_f', \mathbf{a}_{-f})$, where $V()$ is a real function $V : \mathbf{\Gamma} \rightarrow R_+$. At any date t, development aims at selecting and evaluating alternative relevant configurations $(\mathbf{a}_f, \mathbf{a}_{-f})$, $(\mathbf{a}_f', \mathbf{a}_{-f})$ within a given space $\mathbf{\Gamma}_t$. Evaluation takes place through the local 'computation' of $V()$. In contrast, a (radical) discovery at time t defines a new set $\mathbf{\Gamma}_{t+1} \supset \mathbf{\Gamma}_t$. It may also (though not invariably) expand the set of locations from $\mathbf{N}_t$ to $\mathbf{N}_{t+1} = \{1, ..., N_t, ..., N_{t+1}\}$.

Ideas can be used as building blocks for the production of other ideas. If $\mathbf{a}$ is a building block of $\mathbf{a}'$, then $a_n = a'_n$, for every $n \in \mathbf{NS}(\mathbf{a})$. The subset $\mathbf{NS}(\mathbf{a}) \cap \mathbf{NS}(\mathbf{a}') \subseteq \mathbf{N}$ is the overlap between the types $\mathbf{F}$ and $\mathbf{F}'$, such that $\mathbf{a} \in \mathbf{F}$ and $\mathbf{a}' \in \mathbf{F}'$, or between $\mathbf{a}$ and $\mathbf{a}'$ for short[7]. The overlap size is the cardinality of the overlap. Types $\mathbf{F}$ and $\mathbf{F}'$ are independent if and only if they have an empty overlap: $\mathbf{NS}(\mathbf{F}) \cap \mathbf{NS}(\mathbf{F}') = \emptyset$. $\mathbf{\Gamma}$ is *separable* if every couple of types in $\mathbf{\Gamma}$ is independent. $\mathbf{\Gamma}$ is *block separable* (alternatively, *nearly decomposable*), if there exists a partition $\{\mathbf{\Gamma}^1, ..., \mathbf{\Gamma}^Z\}$ of $\mathbf{\Gamma}$, such that the size of the overlap between every couple of subsets $\mathbf{\Gamma}^j$, Γ^h in the partition, with $j \neq h$, $j, h = 1, ..., Z$, is zero (alternatively, sufficiently small). We assume that if the types $\mathbf{F}_f$ and $\mathbf{F}_g$ are independent, then there are no positive or negative complementarities between them. This means that the relative fitness of two different ideas $\mathbf{a}_f$ and $\mathbf{a}'_f$ belonging to $\mathbf{F}_f$ is invariant to a change of $\mathbf{a}_{-f}$, if the latter is produced exclusively by a change $\mathbf{a}_g \to \mathbf{a}'_g$ in the state of $\mathbf{F}_g$.

4.1 Modularity in idea space

Both discovery and development face constraints resulting from the complementarities between non independent ideas. The difficulty arising from a wide potential overlap (interaction) between the types in $\mathbf{\Gamma}$ is that the improvements in the design of one type may conflict with other potential improvements in the design of the types with which it interacts (Simon [31], p. xì).

The difficulty can be reduced if $\mathbf{\Gamma}$ can be partitioned into subsets with (sufficiently) small overlaps between the components. Consistently with the literature (Callebaut and Rasskin-Gutman [**?**]) we call modularity the measure of the extent in which the interaction (complementarity) between the elements (ideas) in the same subset is stronger than the interaction between the subsets. Systems exhibiting this structural property are nearly decomposable (Simon [30], [29]), or modular, and are shown to develop a greater ability to evolve through adaptive change (Altenberg [1], Marengo et al. [19])[8].

[7] By extension, if $\mathbf{B}$ and $\mathbf{C}$ are subsets of $\mathbf{A}$, the set $\{\cup_{\mathbf{c} \in \mathbf{C}} \mathbf{NS}(\mathbf{c})\} \cap \{\cup_{\mathbf{b} \in \mathbf{B}} \mathbf{NS}(\mathbf{b}) \subset \mathbf{N}$ is the overlap between $\mathbf{B}$ and $\mathbf{C}$.

[8] The benefits of modularity should not be overemphasized or misrepresented. Some degree of interaction between modules may yield the optimal trade-off between the advantage of reducing the dimension of the search space, and the benefit resulting from the exploitation of positive between module complementarities. In particular, if the fitness functions embed a sufficiently strong non linearity, the additional gain obtained from solving the coordination between the configurations of two weakly (at a low time frequency) interacting modules $\boldsymbol{\alpha}_h$ and $\boldsymbol{\alpha}_f$ is potentially very large. Watson [33] provides examples of such strongly non linear cases, in which finding the appropriate coordination between the modules may confer a decisive advantage in evolution. Examples of compositional evolution abound in nature and in engineered systems, as shown by sex, by symbiosis, by the relations between a cell and its environment, or by the interactions between a subroutine and the rest of a computer program.

4.1.1 Modularity measures in idea space

The relevance of a modularity measure on $\mathbf{\Gamma}$ is here judged from the view point of the relation between 'modularity' and the minimum number $\beta_{Min}(\Gamma)$ of configurations of $\mathbf{\Gamma}$ which need evaluation, to find the optimal configuration. $\beta_{Min}(\Gamma)$ is the minimum dimension of the search space Γ. Intuitively, the modularity of a search space is a measure of the relative extent in which the dimension of the search space can be reduced.

To avoid the difficulties faced by other measures of modularity in idea space (appendix A), we propose that modularity measures suggested in the network literature are borrowed for the purpose of measuring the decomposability of $\mathbf{\Gamma}$. A preliminary step in this direction is the definition of the network structure between the types in $\mathbf{\Gamma}$.

Definition 6 *We fix an ordering of the types in $\mathbf{\Gamma}$, such that $\mathbf{\Gamma} = \{\mathbf{F}_1, \mathbf{F}_2, ..., \mathbf{F}_{\alpha(\Gamma)}\}$. $\mathbf{L}(\mathbf{\Gamma}) = \{1, 2, ..., \alpha(\mathbf{\Gamma})\}$ is the set of labels corresponding to the types in $\mathbf{\Gamma}$. $\mathbf{NS}_j$ is the set of the active locations corresponding to type $\mathbf{F}_j$. The knowledge network induced by $\mathbf{\Gamma}$ is the weighted directed network $(\mathbf{L}(\mathbf{\Gamma}), \mathbf{W}(\mathbf{\Gamma}))$, where $\mathbf{L}$ is the set of nodes, $\mathbf{W}(\mathbf{\Gamma})$ is the set of weighted links between such nodes, and will be referred to as network $\mathbf{W}(\mathbf{\Gamma})$[9]. The strength w_{hg} of the link from node g to node h is derived as follows. Let $n_{hg} = n_{gh} = \#\{\mathbf{NS}_h \cap \mathbf{NS}_g\}$, for $g \neq h$; $n_{hh} = \#\mathbf{NS}_h$. For every g and h in $\mathbf{L}$, $w_{hg} = \frac{n_{hg}}{\#\mathbf{NS}_g}$. By definition, $0 \leq w_{hg} \leq 1$. For ease of notation, $\mathbf{W}$ shall also denote the $\alpha(\mathbf{\Gamma}) \times \alpha(\mathbf{\Gamma})$ matrix $[w_{hg}]$ of connection weights.*

w_{hg} is a measure of the average frequency with which a configuration change in one active component of type g may affect the relative fitness of type h's configuration. Accordingly, the network $(\mathbf{L}, \mathbf{W})$ is separable if $\mathbf{W}$ is diagonal; it is block-separable if there exists a partition $\{\mathbf{L}^1, ..., \mathbf{L}^Z\}$ of $\mathbf{L}$ and a corresponding permutation matrix $\mathbf{P}$, such that $\mathbf{PWP}^T$ is block diagonal[10].

If the conditions for the block separability of $\mathbf{W}$ are quite demanding, the modularity of $\mathbf{W}$ is more a matter of degree. From an intuitive view point, the modularity of the interaction matrix $\mathbf{W}$ has to do with the possibility of finding a partition of $\mathbf{L}$ into groups $\{\mathbf{L}^1, ..., \mathbf{L}^Z\}$, such that the frequency of interaction between any two different groups of the partition is sufficiently low, relative to the average frequency of interaction within the two groups [11]. This intuitive and quite general idea of modularity admits a quantitative expression, based on recent contributions to the mathematical theory of networks and its applications. The construction of a network between idea types, brings to our disposal the measures of network modularity, such as the Newman and Girvan

[9] The script $(\mathbf{\Gamma})$ will be omitted, if unnecessary.

[10] Here, ^{T} is the transpose operator.

[11] "Another way to describe this structure is to state that the frequencies of interaction among elements in any particular subsystem of a system are an order of magnitude or two greater than the frequencies of interaction between the subsystems. We call systems with this property nearly completely decomposable systems, or for short, nearly decomposable (ND) systems (Simon and Ando [32])". (Simon [31], p. x, citation in the original).

([26]) Q measure, extended by Leicht and Newman ([18]) to weighted directed networks. For each possible partition of $\mathbf{L}(\mathbf{\Gamma})$ Leicht and Newman ([18]) define a corresponding Q measure of modularity. Network modularity $Q(\mathbf{W})$ is the measure corresponding to the Q maximizing partition. If there is no meaningful way of defining different modules in $\mathbf{W}$, $Q(\mathbf{W})$ is zero. If $\mathbf{W}$ is diagonal, $Q(\mathbf{W})$ is maximal for the given $\alpha(\Gamma)$, which is written $Q(\mathbf{W}) = Q_{\alpha(\mathbf{\Gamma})}$, and $\lim_{\alpha(\mathbf{\Gamma})\to\infty} Q_{\alpha(\mathbf{\Gamma})} = 1$ (see appendix B).

The notion of a knowledge network $\mathbf{W}$ induced by $\mathbf{\Gamma}$ and the related operational definition of network modularity, together with de facto observation, motivate the definition of a disciplinary knowledge field as a module in the space $\mathbf{\Gamma}$.[12]

Definition 7 *The set of disciplinary fields* $\{1, ..., Z\}$ *is endogenously defined by the* Q *maximizing partition of* $\mathbf{L}$ *into modules* $\{\mathbf{L}^1, ..., \mathbf{L}^Z\}$.

We developed a number of simulations suggesting that the Q measure performs well as an indicator of the extent in which the dimension of a search space $\mathbf{\Gamma}$ can be reduced to the end of finding its optimal configuration. If $\mathbf{\Gamma}$ and $\mathbf{\Gamma}'$ are such that $\alpha(\Gamma) = \alpha(\Gamma')$, and $\beta_{Min}(\Gamma) < \beta_{Min}(\Gamma')$, then $Q(\Gamma) > Q(\Gamma')$. In this respect, Q performs better than the measure of modularity on technological fitness landscapes recently suggested in Frenken [8] (see appendix A).

5 The carriers of ideas

In what follows, we expand on the fact that ideas are embodied in the human brains. $\mathbf{H}$ is the set of agents, and H is the cardinality thereof. The statement that $\mathbf{a}$ is an element in the set $\mathbf{A}_i$ of i's ideas, does not simply mean that agent i has the information $\mathbf{a}$. It means that i has the capability to carry out a set of operations on $\mathbf{\Gamma}_{i\cdot}$, which use the information $\mathbf{a}$ as input. We may refer here to a standard example, which marks the difference between having a mathematics handbook at one's disposal, and having a full grasp of the proofs and potential applications of the mathematical proposition printed in the text. As a rule, agents' knowledge is specialized in a disciplinary field defined as above and a field identifies a scientific community[13].

An agent i at time t is here simply defined by the set $\mathbf{A}_i(t)$, which defines also the set $\mathbf{\Gamma}_i(t)$ of the types to which such ideas belong. We shall abstract in what follows from innate exogenous differences between agents, concerning their

[12] Some quantitative evidence of the modular organization of knowledge into disciplinary fields, which broadly correspond to application domains, is offered by the network of patent-citation flows connecting the technology fields. The patent citation networks recovered from USPTO data for the periods 1975-1986 and 1987-1999 yield Q measures above 0.6 (Caminati and Stabile[6]).

[13] This means that there exists a partition of $\mathbf{H}$ into M communities, $\{\mathbf{H}^1, ..., \mathbf{H}^M\}$, such that the union of the knowledge sets of the agents belonging to the same community, covers some disciplinary field (formally, there exist $\mathbf{\Gamma}^z$ in $\{\mathbf{\Gamma}^1, ..., \mathbf{\Gamma}^Z\}$, and $\mathbf{H}^m$ in $\{\mathbf{H}^1, ..., \mathbf{H}^M\}$, such that $\mathbf{\Gamma}^z \subseteq \cup_i \mathbf{\Gamma}_i, i \in \mathbf{H}^m$, where $\mathbf{\Gamma}_i$ is agent i's knowledge set).

capabilities in the processing of ideas, or their individual preferences. Every agent engages in both basic-research and development activities, according to fixed proportions.

The knowledge output of a development project is measured, not by the number of the new ideas produced, but by the fitness improvement enabled by these new configurations of $\mathbf{\Gamma}$. In basic research, the fitness improvement corresponding to a new type of idea is only potential, because the new type may need to go through a long development phase, before it can successfully compete with pre-existing ideas. Since usefulness is only potential in basic research, novelty and originality is all what matters. We assume that research output is measured by the number of new types produced by a research project. A research contribution is more 'original' if it produces a larger number of new and potentially useful types.

There are complementarities between development and research. On the ground that modularity reduces the dimension of the search space faced by the activity of development, there is a powerful incentive to specialization in human capital formation. This intuition is made more precise as follows.

5.1 Specialization in development

Given the set $\mathbf{L}$ of labels corresponding to the types in $\mathbf{\Gamma}$, consider the Q maximizing partition $\{\mathbf{L}^1, ..., \mathbf{L}^Z\}$ of $\mathbf{L}$ into label groups, which define the set of disciplinary fields or subfields. If the groups $\mathbf{L}^y$ and $\mathbf{L}^x$ have an empty overlap, the activity of developing the types in $\mathbf{L}^y$ can be carried out independently of the corresponding activity on $\mathbf{L}^x$. This provides a straightforward powerful incentive to specialization in development, and through this, also to specialization in human capital formation. If instead the groups $\mathbf{L}^y$ and $\mathbf{L}^x$ have a low, but non zero, frequency of interaction, the possibility of reducing the dimension of the search space is still provided by the organization of the search activity according to a hierarchic modular design. In particular, if the weak interactions between $\mathbf{L}^y$ and $\mathbf{L}^x$ are not sparse, but are carried by specific nodes (types), then the hierarchic form of modularity is supported by a decentralization of development to specialized agents. We may think of the example in which $\mathbf{L}^y = \{y, y+1, ..., y+m\}$, $\mathbf{L}^x = \{x, x+1, ..., x+n\}$ and the interaction between the two groups takes place through the link connecting the nodes (types) y and x. Conditional on the configuration of y and x, the choice of the optimal configuration of the two subfields $\{\mathbf{L}^y - y\}$ and $\{\mathbf{L}^x - x\}$ can be assigned to different specialized agents, which may have in their knowledge endowment only the fields $\mathbf{L}^y$ and $\mathbf{L}^x$, respectively. The two specialists do not need to communicate between them, if they communicate with a third agent who is endowed with the knowledge of both types y and x. These types contain the core ideas, connecting the fields $\mathbf{L}^y$ and $\mathbf{L}^x$. The agent endowed with the core ideas acts like a knowledge integrator. Upon communication of the best selections separately made by the specialists, the integrator(s) can fix the optimal joint configuration of the types y and x.

This argument provides useful guidelines suggesting what may be relevant distributions of the knowledge endowment across the set of agents. To the extent that the knowledge network is highly, but not perfectly modular, we expect a distribution of the knowledge endowments such that a large number of agents is specialized in a specific field or subfield[14]. The number of types shared by agents i and j will be typically high or low, depending on whether they are specialized in the same or in different fields. Simultaneously, knowledge specialization requires that there agents preserving in their knowledge endowment the core ideas providing the knowledge interfaces connecting the specialized fields or subfields.

6 Collaboration networks in basic research

Discoveries originate from the effort produced by the agents engaged in basic research[15]. Research projects may be carried out in collaboration, or in isolation (self-collaboration). For the sake of simplicity, we assume that at most two members of $\mathbf{H}$ collaborate on a single project. A couple jointly working on a project is also called a collaboration; at any date a collaboration is engaged in a single project, although a single agent can participate in many projects simultaneously.

We hold to the view that new types of ideas grow out of recombinations of pre-existing types (Weitzman [34])[16]. On this premise, we move some steps toward the construction of a knowledge based approach to the explanation of collaboration networks in basic research.

We assume that for every $i \in \mathbf{H}$, the set $\mathbf{\Gamma}_i \neq \emptyset$. $B(i,j)$ is the flow of new basic ideas (new types) produced by the collaboration (ij) between agents i and j. We assume that $B(i,j)$ depends on the effective competence $k(ij) = k(e_{ij}, e_{ji})$ and effective effort $e(ij) \geq 0$ of the coalition (ij). Formally:

$$B(ij) = B(k(ij), e(ij)) = B(k(e_{ij}, e_{ji}), e(e_{ij}, e_{ji})) \tag{1}$$

$$B(k, e) = k^{\lambda} e \ \lambda \geq 1 \tag{2}$$

[14] If there are non negligible interactions between fields $\mathbf{L}^y$ and $\mathbf{L}^x$ in idea space, we expect to find meaningful knowledge flows between the communities working in such fields. These knowledge flows take place in a variety of ways (e.g. access to paper and patent publications, as evidenced by citations), including direct collaboration between the scientists. The point made in the text is that the desirable extent of knowledge specialization is increased, if the need of direct and frequent communication between different-field specialists is replaced by the fixation and adaptation of the appropriate interface standards between the fields. The ideas providing such interfaces are in the knowledge endowment of one or more agents acting as knowledge integrators. This argument is partly reminiscent of the literature on system integration in production technology ([7]).

[15] Occasionally, the discovery of a new type is the serendipitous outcome of a development project. For the sake of simplicity, we rule the possibility of serendipitous discoveries and assume that a research output is the outcome of a deliberate allocation of effort to a research project.

[16] There is some evidence that radical discoveries are occasionally produced by recombination of seemingly obsolete input ideas.

For the sake of simplicity we assume that every project has a maximum scale: $0 \leq e(ij) \leq D$.

The competence of the collaboration (ij) reflects the joint size of i's and j's knowledge contributions, provided that their respective effort in the collaboration is strictly positive.

$$k(e_{ij}, e_{ji}) = \alpha(\mathbf{\Gamma}_i \cup \mathbf{\Gamma}_j) \text{ if } e_{ij} \cdot e_{ji} > 0;\ k(e_{ij}, 0) = \alpha(\mathbf{\Gamma}_i);\ k(0, e_{ji}) = \alpha(\mathbf{\Gamma}_j) \quad (3)$$

where e_{ij}, $0 \leq e_{ij} \leq E$, is the individual effort produced by i in the collaboration (ij).

The joint effort $e(ij) = e(e_{ij}, e_{ji})$ is a performance function $e()$ which depends on the parameters w_{ij} and $d = d(ij)$.

$$e(ij) = [f(d)]^{-1} \left[w_{ij} \left(e_{ij}\right)^{\rho} + (1 - w_{ij}) \left(e_{ji}\right)^{\rho} \right]^{1/\rho} \quad (4)$$

where: $\rho = 1 - d$, $f^{'} \geq 0$, and $\lim_{d \to \infty} f(d) = F$.

w_{ij} enters the performance function to characterize how the relative size of i's and j's knowledge contributions:

$$w_{ij} = \frac{\alpha(\mathbf{\Gamma}_i)}{\alpha(\mathbf{\Gamma}_i) + \alpha(\mathbf{\Gamma}_j)};\ w_{ji} = (1 - w_{ij}) = \frac{\alpha(\mathbf{\Gamma}_j)}{\alpha(\mathbf{\Gamma}_i) + \alpha(\mathbf{\Gamma}_j)};\ w_{ii} = 1 \quad (5)$$

$d(ij)$ is the heterogeneity between i's and j's idea types:

$$d(ij) = \frac{\alpha(\mathbf{\Gamma}_i \cup \mathbf{\Gamma}_j) - \alpha(\mathbf{\Gamma}_i \cap \mathbf{\Gamma}_j)}{\alpha(\mathbf{\Gamma}_i \cap \mathbf{\Gamma}_j)} = \frac{\alpha(\mathbf{\Gamma}_i \cup \mathbf{\Gamma}_j)}{\alpha(\mathbf{\Gamma}_i \cap \mathbf{\Gamma}_j)} - 1;\ d(ii) = 0 \quad (6)$$

$d(ij)$ has the properties: $0 \leq d(ij) \leq \infty$; $d(ii) = 0$; $d(ij) = d(ji)$; $d(ij) = \infty$ if and only if $\mathbf{\Gamma}_i \cap \mathbf{\Gamma}_j = \emptyset$. Is is worth stressing that the function $d(ij)$ may not meet the triangle inequality. For instance, if i and j do not have any idea in common, but each of them shares some idea with h, then $\infty = d(ij) > d(ih) + d(hj)$. $d(ij)$ affects the joint effective effort in two ways. In the first place, the productivity of i's and j's individual efforts is lower (ceteris paribus), if the heterogeneity $d(ij)$ grows above $d(ij) = 1$. Beyond this threshold level, higher knowledge heterogeneity makes communication between i and j more time consuming, and the production of coordinated research effort more demanding. We therefore assume that a parametric increase in $d(ij)$ beyond the threshold $d(ij) = 1$ lowers the effective joint effort $e(ij)$:

$$\frac{\partial e}{\partial d} = 0 \text{ if } 0 \leq d \leq 1;\ \frac{\partial e}{\partial d} < 0 \text{ if } d > 1 \quad (7)$$

More precisely, we assume the following specification of $f(d)$ in equation (4): $f(d) = 1 + \phi(d)$, $\phi(d) = 0$ if $0 \leq d \leq 1$; $\phi' > 0$ if $d > 1$.

In the second place, the rules by which the efforts e_{ij} and e_{ji} cooperate to produce new ideas depend on the circumstances affecting the substitutability or complementarity between i's and j's idea sets. We assume that the higher the heterogeneity $d(i,j)$, the lower the substitutability between i's and j's knowledge, hence, the higher the complementarity between e_{ij} and e_{ji}. If i and j have identical types of ideas, $d(ij) = 0$, and their efforts are perfect complements, if they do not have any idea in common, $d(ij) = \infty$, and their efforts are perfect complements; in the intermediate situation where $d(ij) = 1$, $e()$ is Cobb Douglas (see appendix C). In particular, if $\alpha(\mathbf{\Gamma}_i) = \alpha(\mathbf{\Gamma}_j)$, so that $w_{ij} = w_{ji} = 1/2$, we obtain the restrictions:

$$d(ij) = 1 \ \rightarrow e(ij) = e_{ij}^{1/2} e_{ji}^{1/2} \tag{8}$$

$$d(ij) = 0 \rightarrow e(ij) = \frac{1}{2}(e_{ij} + e_{ji}) \tag{9}$$

$$e(ii) = e_{ii}$$

The $B(\cdot)$ function determines the knowledge output of the R&D effort produced by the collaboration (ij). The reputation payoff earned by each member of the coalition (ij) is $rB(ij)$, where $0 < r \leq 1$, if $i \neq j$, and $r = 1$, if $i = j$. The cost to agent i of her participation in the project (ij) is a linear increasing function $c \cdot e_{ij}$ of the effort produced.

A basic research strategy by agent i is a choice $\mathbf{S}_i = (e_{i1}, ..., e_{iH})$. $e_{ij} = 0$ means that i is not prepared to collaborate with $j \in \mathbf{H}$. Agent i chooses $\mathbf{S}_i$ to maximize net pay-off:

$$\Pi_i = B(ii) - ce_{ii} + \left[\sum_{z \in \mathbf{H}-i} rB(iz) - c \sum_{z \in \mathbf{H}-i} e_{iz} \right] \tag{10}$$

subject to the constraint:

$$\sum_{z \in \mathbf{H}} e_{iz} = E$$

Network formation is formalized as a simultaneous 'collaboration game'. Each agent $i \in \mathbf{H}$ simultaneously announces her strategy $\mathbf{S}_i$. The collaboration (ij) takes place if and only if $e_{ij} > 0$, $e_{ji} > 0$; $g_{ij}(\mathbf{S}) = g(e_{ij})$ for $e_{ij} \in \mathbf{S}_i \in \mathbf{S}$. $\Pi_i(\mathbf{S})$ is i's payoff induced by the strategy profile $\mathbf{S}$.

It is worth summarizing some implications which follow from the knowledge production function $B(ij)$. and the payoff function (10).

a. Since it is never optimal to offer collaboration to an agent who is not reciprocating the offer, a very weak necessary requirement for $\mathbf{S}$ being a pairwise equilibrium is $g_{ij}(\mathbf{S}) > 0$ only if $g_{ji}(\mathbf{S}) > 0$.

b. An agent obtains a competence advantage from collaborating with those agents which contribute with ideas that are not in her field of specialization. In particular, for fixed cardinalities $\alpha(\mathbf{\Gamma}_i)$ and $\alpha(\mathbf{\Gamma}_j)$, competence $k(ij)$ increases with heterogeneity $d(ij)$. If i and j have identical ideas the competence benefit from a positive collaboration (ij) is null: $k(ij) = k(ii) = k(jj)$; moreover, i's and j's efforts are perfect substitutes in the production of joint effort.

c. There is a drawback in collaborating with an agent whose field of specialization is too remote from ours. A productive collaboration (ij) will be one in which i and j share some knowledge background. This is formalized by choosing $f(d)$ so that $\lim_{d\to\infty} e(ij) = 0$.

Definition 8 *The strategy profile* $\mathbf{S}$ *generates the square* $H \times H$ *weight matrix* $\mathbf{g}(\mathbf{S})$*, such that for every* i *and* j *in* $\mathbf{H}$*,* $g_{ij}(\mathbf{S}) = e_{ij}$*. This defines the weighted directed network* $\{\mathbf{H}, \mathbf{g}(\mathbf{S})\}$*, where* $\mathbf{H}$ *is the set of nodes, and there exists a directed link from node* i *to node* j*, if and only if the weight* $g_{ij}(\mathbf{S}) > 0$*.* $\mathbf{g}(\mathbf{S})$ *fully defines the weighted directed network supported by* $\mathbf{S}$ *and will be referred to as network* $\mathbf{g}(\mathbf{S})$ *in what follows.*

Definition 9 *The weighted directed network* $\mathbf{g}(\mathbf{S})$ *defines the* $H \times H$ *adiacency matrix* $\mathbf{G}(\mathbf{S})$*, such that for every* i *and* j *in* $\mathbf{H}$*,* $G_{ij}(\mathbf{S}) = G_{ji}(\mathbf{S}) = 1$*, if and only if* $g_{ij}(\mathbf{S}) > 0$ *and* $g_{ji}(\mathbf{S}) > 0$*;* $G_{ij}(\mathbf{S}) = G_{ji}(\mathbf{S}) = 0$ *otherwise.* $\mathbf{G}(\mathbf{S})$ *uniquely defines the unweighted undirected network* $\{\mathbf{H}, \mathbf{\Lambda}(\mathbf{S})\}$ *such that* ij *is a link in* $\mathbf{\Lambda}(\mathbf{S})$ *if and only if* $G_{ij}(\mathbf{S}) = G_{ji}(\mathbf{S}) = 1$*. For ease of reference,* $\mathbf{G}(\mathbf{S})$ *denotes in what follows the undirected network supported by* $\mathbf{S}$*.*

6.1 Equilibrium in collaboration in a system with field specialists and knowledge integrators

Equipped with the remarks above, we consider the collaboration equilibria which are supported by specialization in the knowledge system of the form discussed above. We assume that the there exists a partition $\{\mathbf{H}^1, ..., \mathbf{H}^M, \mathbf{H}^{M+1}\}$ of $\mathbf{H}$, such that $\mathbf{H}^z$, with $z = 1, ..., M$, is the scientific community of agents which develop the ideas belonging to field z. $\mathbf{H}^{M+1}$ is the community of agents endowed with the knowledge interfaces which may directly or indirectly connect the specialized fields. In our stylized representation, we assume: The size of the knowledge endowment is uniform; every community $\mathbf{H}^z$ $z = 1, ..., M$, contains the same number R of researchers; any two agents belonging to the same community $\mathbf{H}^z$ have identical knowledge endowments; if they belong to different specialized fields, their endowment are 'sufficiently' heterogeneous. We study the equilibrium collaborations which are supported by the varying heterogeneity, as opposed to the varying size, of agents' endowments. In particular, we consider the following special case:

Case 10 *A knowledge distribution with field specialists and knowledge integrators*

1. For every $i \in \mathbf{H}$, $\alpha(\mathbf{\Gamma}_i) = \alpha$.
2. $\#\mathbf{H}^z = R$, $z = 1, ..., M$; $H = RM + P$, where $P = \#\ \mathbf{H}^{M+1} \geq M$.
3. Fix $z \in \{1, ..., M\}$. If $i \in \mathbf{H}^z$, $j \in \mathbf{H}^z$, then, $d(ij) = 0$.
4. There is an ordering $\{h^1, h^2, ..., h^P\}$ of the agents in $\mathbf{H}^{M+1}$, such that: $d(h^q h^{q+1}) = 1$, $q = 1, ..., P-1$; $d(h^P h^1) = 1$; $d(h^x h^y) > 1$, for $1 < \|y - x\| < P - 1$.

5. If $i \in \mathbf{H}^z$, $j \in \mathbf{H}^v$, with $v \neq z$, $\{v, z\} \subset \{1, ..., M\}$, then, $d(ij) \geq \Omega > 1$.

6. Fix $z \in \{1, ..., M\}$. There exists one and only one $j \in \mathbf{H}^{M+1}$such that, if $i \in \mathbf{H}^z$ and $h \in \{\mathbf{H}^{M+1} - j\}$, then: $d(ij) = 1$; $1 < \Omega_z \leq d(ih) \leq \Omega$.

Proposition 11 *Assume the case above. The following restriction holds.*

$$k(ij) = 2\alpha \frac{d(ij)+1}{d(ij)+2} = k(d(ij)) \tag{11}$$

Remark 12 *We may notice from 11 that competence $k(ij)$ increases with heterogeneity $d(ij)$, but at a slower rate; moreover, $d(ij) = 0$ implies $k(ij) = \alpha$; $d(ij) = 1$ implies $k(ij) = \alpha \frac{4}{3}$.*

Definition 13 *A collaboration (ij) is symmetric if and only if $e_{ij} = e_{ji}$.*

Proposition 14 *Assume the case above. If $e_{ii} < D$ and $e_{jj} < D$, a symmetric collaboration project (ij) such that $d(ij) = 0$ can not be sustained as a Nash outcome..*

Proof: For any value of $r \in [0, 1]$, i's marginal benefit from increasing effort in project (ii) is higher than i's marginal benefit from increasing effort in (ij).

$$\alpha^\lambda - c > \frac{1}{2} r \alpha^\lambda - c$$

Proposition 15 *Assume the case above, and $e_{jj} = e_{ii} = D < E$. A symmetric collaboration project (ij) such that $d(ij) = 0$ can be sustained as a Nash outcome only if*

$$\frac{1}{2} r \alpha^\lambda (e_{ij} + e_{ji}) - c e_{ij} = e_{ij} \left[r \alpha^\lambda - c \right] > 0$$

Proposition 16 *Assume that agent i has the opportunity to participate in a collaboration (ih) such that $d(ih) = 1$. i's marginal effort re-allocation away from a symmetric collaboration (ij) such that $d(ij) = 0$, towards (ih) is payoff increasing (decreasing) if:*

$$\frac{e_{ih}}{e_{hi}} < (>) \left(\frac{4}{3} \right)^{2\lambda} \tag{12}$$

i's marginal effort re-allocation away from a project (ii) towards (ih) is payoff increasing (decreasing) if:

$$\frac{e_{ih}}{e_{hi}} < (>) \frac{r^2}{4} \left(\frac{4}{3} \right)^{2\lambda} \tag{13}$$

Proposition 17 *Assume that the restrictions of the case above hold, $\left[r \alpha^\lambda - c \right] > 0$, $r/2 > (3/4)^\lambda$, $P = M$, $2 + \left(\frac{4}{3}\right)^\lambda \geq E/D \geq \left[1 + \left(\frac{4}{3}\right)^\lambda \right]$, and Ω_z is sufficiently large. For any $M \geq 2$, there exists a non empty range of the parameters λ, R, E/D supporting a pairwise equilibrium with the following properties. (i) Every* *collaboration (ij) with $d(ij) = 0$ is symmetric. (ii) For every $h \in \mathbf{H}^{M+1}$ there exists a unique $z \in \{1, ..., M\}$ such that, for every $i \in \mathbf{H}^z$, h enters a size D collaboration (ih) with $d(ih) = 1$. Every $h \in \mathbf{H}^{M+1}$ enters 2 collaborations (ph) such that $d(ph) = 1$, and $p \in \mathbf{H}^{M+1}$. (iii) For each $z \in \{1, ..., M\}$, every $i \in \mathbf{H}^z$ invests simultaneously her effort: in a project (ii) of size D; in one and only one project (ih) such that $d(ih) = 1$ and $h \in \mathbf{H}^{M+1}$; in V projects (ij) such that $d(ij) = 0$ and $j \in \mathbf{H}^z$, where $0 \leq V \leq R - 1$.*

The proof of the proposition is in two steps. In the first step, we introduce the working hypothesis that the Nash equilibria to be described in the first step are robust to the formation of any collaboration (iy), where $i \in \mathbf{H}^z$, $z \in \{1, ..., M\}$, $y \in \mathbf{H}^{M+1}$ and $d(iy) \geq \Omega_z$. The hypothesis is proved in the second step.

1. For any given $z \in \{1, ..., M\}$, every $i \in \mathbf{H}^z$ has a unique opportunity to enter a collaboration (ih) such that $d(ih) = 1$. We introduce the following restrictions on the parameters λ, R, and E/D:

$$R = \left(\frac{4}{3}\right)^{2\lambda} \quad \lambda = \frac{\log R}{2[\log 4 - \log 3]} \approx \frac{\log R}{0.5754} \tag{14}$$

$$E/D \geq 2 + R^{1/2} = 2 + \left(\frac{4}{3}\right)^{\lambda} \tag{15}$$

With this restriction, at $e_{ih}/e_{hi} = R$, the condition (12) holds with strict equality. In this state, agent i strictly prefers project (ii) to project (ih); this, in turn, is indifferent to any project (ij) such that $d(ij) = 0$. The working hypothesis implies that i invests $e_{ii} = D$ in the project (ii) and she is prepared to invest up to $e_{ih} = \left(\frac{4}{3}\right)^{\lambda} D = R^{1/2} D$ in the project (ih) of size D. She invests the residual endowment $E - (R^{1/2} + 1)D \geq 0$ in V projects of type $d(ij) = 0$, where $0 \leq V \leq R - 1$. Such projects of type $d(ij) = 0$ are sustainable as Nash outcomes under the specified conditions. For every $h \in \mathbf{H}^{M+1}$ there exists one and only one $z \in \{1, ..., M\}$, such that $i \in \mathbf{H}^z$ implies that (ih) is of type $d(ih) = 1$. (ih) is h's best investment opportunity at $e_{hi}/e_{ih} = \left(\frac{3}{4}\right)^{2\lambda}$. h invests a total effort $R\left(\frac{3}{4}\right)^{\lambda} D = \left(\frac{4}{3}\right)^{\lambda} D = R^{1/2} D$ in R projects of this type. At $\left(\frac{4}{3}\right)^{\lambda} > 2/r$, h strictly prefers a symmetric collaboration (hp) such that $d(hp) = 1$ to a project (hh). With the restrictions imposed by the case above, and by 15, h enters 2 symmetric collaborations (hp) such that $d(hp) = 1$, $p \in \mathbf{H}^{M+1}$, investing $e_{hp} = D$ in each. This concludes the first step.

2. We are left with the proof that under the chosen parametrization, the Nash equilibria described in step 1 are robust to the formation of any collaboration (iy), where $i \in \mathbf{H}^z$, $z \in \{1, ..., M\}$, $y \in \mathbf{H}^{M+1}$ and $d(iy) \geq \Omega_z$. y is prepared to enter any such collaboration (iy) *only if* $e_{iy}/e_{yi} \geq \theta(\Omega_z)$, which is the critical ratio which makes y indifferent between (iy) and (yy). Here $\theta(\Omega_z)$ is a strictly increasing function, and the assumption that Ω_z is

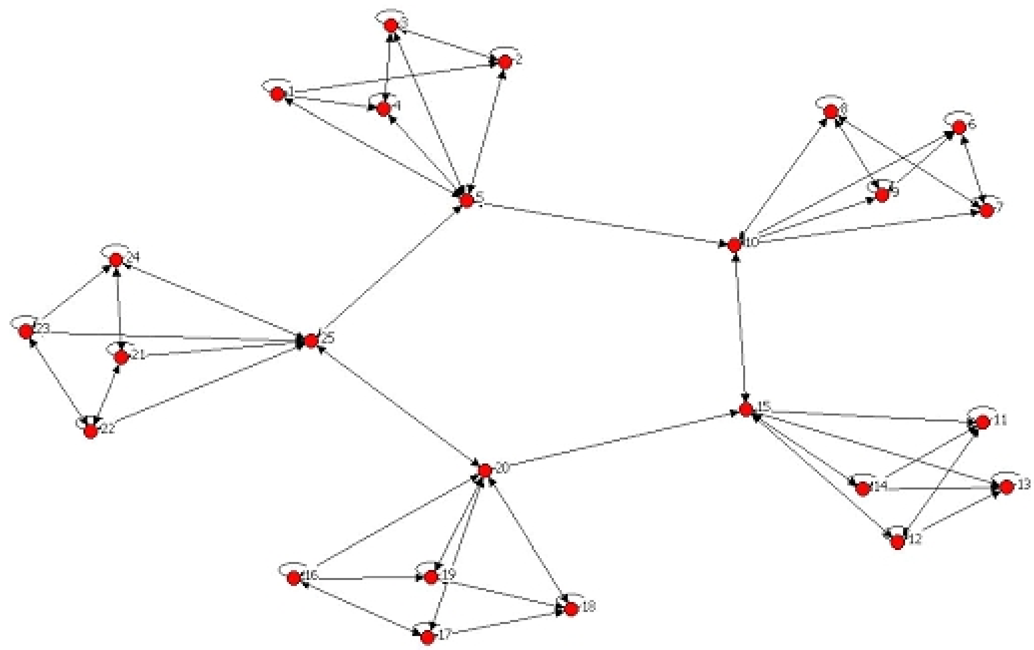

Figure 1: Figure 1. Undirected network representation of a pairwise equilibrium of the collaboration game supported by the case above, for $H = 25$, $R = 4$, $E = 4$, $P = M = 5$, $D = 1$, $\lambda = \frac{\log R}{2(\log 4 - \log 3)} \approx 2.4093$.

sufficiently large implies $\theta(\Omega_z) > R$. Recall that $R = e_{iy}/e_{yi}$ is the critical ratio which makes i indifferent between (ih), such that $d(ih) = 1$ and the symmetric collaboration (ij) such that $d(ij) = 0$. Under the chosen parametrization, $i \in \mathbf{H}^z$ is not constrained by the opportunity to increase her collaboration with some $j \in \mathbf{H}^z$. This shows that collaborations (iy) such that $d(iy) \geq \Omega_z$ can not be sustained as Nash outcomes. This completes the proof.

Figure 1 shows the undirected network representation of a pairwise equilibrium supported by the case above under the parametrization (14), with $H = 25$, $R = E = 4$, $D = 1$, $P = M = 5$.

7 Conclusions and directions of further work

The main goal of this paper is to move some steps towards a foundation of a knowledge based approach to the analysis of the division of labour and of collaboration in research. In this closing section we outline some implications of our approach, and suggest promising directions of further work.

The general intuition behind the propositions of the preceding section is that the heterogeneity of agents' knowledge endowments induces a trade off. For a given size of i's and j's knowledge endowment, the larger the heterogeneity $d(ij)$ between them, the higher the competence $k(ij)$ of their joint project. The other side of the coin is that the private effort which is necessary to make the collaboration effective is weakly increasing with $d(ij)$, and it is strictly increasing, if $d(ij)$ exceeds a minimum threshold. This points to a situation in which two specialists in highly heterogeneous fields will find that direct collaboration between them is too effort-consuming. Simultaneously, collaboration between two specialists in the same narrow field may offer little, if any, competence advantage. Ceteris paribus, the agents which are only weakly specialized, occupy a critical favourable position in the distribution of knowledge[17], which gives them the opportunity to act as knowledge integrators. A knowledge integrator is in the position to collaborate with a specialized agent under favourable terms; as a result, the former will be able to enter a larger number of collaborations than the latter. On this ground, we offer a knowledge based interpretation of the quasi stars which are observed in the empirical scientific-collaboration networks based on co-authorship data.

The rising historical trend in scientific collaboration can be explained, on the same ground, not only (and most obviously) by the fact that the appropriation parameter r may have risen in the meanwhile to a level sufficiently close to 1. More interestingly, the observed growth in the number of collaborators per-capita appears to be a result of the very growth in the number of researchers, in particular, in the average number R of researchers belonging to a specialized field.

To the extent that the model is about knowledge production, it is natural to think of dynamic extensions in which i's knowledge endowment at t is the outcome of her endowment at $t-1$, and of the knowledge output of the projects entered by i at $t-1$. In this dynamic version, the knowledge endowment becomes an endogenous variable.

8 Appendix

8.1 Appendix A. On a pleiotropy based measure of modularity

The structural complementarities imposed by wide overlaps between idea types can be partly described by adapting to our purpose the biologically inspired definition of a system pleiotropy. The number P_j of idea types in $\mathbf{\Gamma}$ sharing the active location j is the pleiotropy of j. Assuming that location j is not always silent in Γ, then $1 \leq P_j \leq \alpha$. Frenken [8] measures system pleiotropy P as:

$$P =^{\alpha} \log \left(\prod_{n=1}^{N} P_n \right) \tag{16}$$

[17] This critical position is partly reminiscent of the notion of a structural hole (Burt, [4]).

Here $^{\alpha}\log$ is base α logarithm. The structural relation between string components and types is described by the $\alpha \times N$ matrix $\mathbf{\Phi} = [\phi_{fn}]$. Each row in this matrix corresponds to a different type in $\mathbf{\Gamma}$, and $\phi_{fn} = 1$, or $\phi_{fn} = 0$ depending on whether n is or is not a active location of the fth type in $\mathbf{\Gamma}$. Biologically inspired problem representations identify a string $\mathbf{a}$ with a phenotype character and label $\mathbf{\Phi}$ the genotype–phenotype map. Following Frenken [8], we may assume a non-separable space $\mathbf{\Gamma}$, and observe that the proposed P measure (16) achieves its minimum $P_{\min}$ when $N-1$ components have $P_j = 1$ and one and only one component has $P_j = \alpha$. This yields $P_{\min} =^{\alpha} \log(\alpha) = 1$. Maximum system pleiotropy $P_{\max}$ obtains when all components j have $P_j = \alpha$, yielding $P_{\max} =^{\alpha} \log(\alpha^N) = N$. Frenken [8] suggests a measure M of system modularity based on the notion that the lower P, the lower the average interdependency between the idea types. More precisely, M depends on the comparison between $P_{\max}$ and the observed value of P.

$$M = 1 - \frac{P}{P_{\max}} = 1 - \frac{P}{N} \tag{17}$$

According to this measure, the modularity of a non-separable system is inversely related with system pleiotropy; its maximum $M_{\max} = \frac{N-1}{N}$ tends to 1 as N tends to ∞; its minimum is $M_{\min} = 0$. Apart from the restriction that M is by definition applicable only to non-separable systems, the main problem with this measure is that, in so far as it depends only on the product of pleiotropy measures P_n, it does not retain more detailed information on the *interdependence structure* between the types in $\mathbf{\Gamma}$. It turns out that two sets $\mathbf{\Gamma}$ and $\mathbf{\Gamma}'$, such that $N(\mathbf{\Gamma}) = N(\mathbf{\Gamma}')$, $\alpha(\mathbf{\Gamma}) = \alpha(\mathbf{\Gamma}')$, and $M(\mathbf{\Gamma}) = M(\mathbf{\Gamma}')$ may nevertheless have $\beta_{Min}(\Gamma) \neq \beta_{Min}(\Gamma')$.

8.2 Appendix B. Modularity measure on weighted directed networks

For the given partition $\{\boldsymbol{\alpha}^1, ..., \boldsymbol{\alpha}^Z\}$ of $\boldsymbol{\alpha}$, the total intensity of an outward link from group h directed to itself or to other groups is $\hat{a}_h = \sum_i \sum_j w_{ij}$, $j \in \boldsymbol{\alpha}^h, i = 1, ..., Z.$. The corresponding total intensity of an inward link to group h from itself or from other groups is $\check{a}_h = \sum_j \sum_i w_{ij}$, $i \in \boldsymbol{\alpha}^h$, $j = 1, ..., Z$. If the total intensity of links in $\mathbf{W}$ is $T = \sum_i \sum_j w_{ij}$, $i, j = 1, ..., Z$, then the average relative frequency with which an outward link in $\mathbf{W}$ originates from, and arrives to, group h is $\hat{e}_h = \frac{\hat{a}_h}{T}$ and $\check{e}_h = \frac{\check{a}_h}{T}$, respectively. The modularity measure Q_h of the links from and to group h in the context of the given network $\mathbf{W}$, is then expressed by the extent to which the frequency of within-group links exceeds the frequency that would be expected from the hypothesis of a random wiring.

$$Q_h = \frac{1}{T}\left[\sum_{i \in \{\mathbf{F}_h\}} \sum_{j \in \{\mathbf{F}_h\}} w_{ij}\right] - \hat{e}_h \check{e}_h \tag{18}$$

The modularity of $\mathbf{W}$ according to the partition $\{\boldsymbol{\alpha}^i\}_{i=1}^Z = \{\boldsymbol{\alpha}^1, ..., \boldsymbol{\alpha}^Z\}$ is then expressed by the sum $Q = \sum_{h=1}^Z Q_h$, which may be negative, if the partition is ill-chosen. Indeed, the relative goodness of two alternative partitions of $\boldsymbol{\alpha}$ is evaluated by choosing the partition yielding a higher value of Q. In this spirit, the modules of the network $\{\boldsymbol{\alpha}, \mathbf{W}\}$ are endogenously determined by selecting the Q-maximizing partition $\{\boldsymbol{\alpha}^{i*}\}_{i=1}^{Z^*}$([26]), and the modularity of $\{\boldsymbol{\alpha}, \mathbf{W}\}$) is the Q-measure induced by such partition. Z^* is the number of modules in the Q maximizing partition. Since the Q modularity of the null partition $\{\boldsymbol{\alpha}\}$ is zero, the Q modularity of $\mathbf{W}$ takes values in the interval $[0, 1]$. If $\mathbf{W}$ is diagonal, then $Q(\mathbf{W}) \to 1$ as $F \to +\infty$. A fast algorithm for the computation of Q in large undirected networks [24] was subsequently extended to weighted undirected and directed networks([18], [6]).

8.3 Appendix C. Complementarity and substitutability in the production of effective collaboration effort

We assume that effective joint effort $e(ij)$ is specified by:

$$e(ij) = = [f(d(ij)]^{-1} \left[w_{ij} \left(e_{ij}\right)^\rho + (1 - w_{ij}) \left(e_{ji}\right)^\rho \right]^{1/\rho} \tag{19}$$

where: $\rho = 1 - d(ij)$; $f(d) = 1 + \phi(d) \cdot d$; $\phi(d) = 0$ if $0 \leq d \leq 1$; $f' > 0$ if $d > 1$; $\lim_{d \to \infty} f(d) = F$. For a given fixed specification of the parameters w_{ij}, and $d(ij)$, the $e(\cdot)$ function 19 belongs to the family of CES functions, linear homogeneous with respect to e_{ij}, e_{ji} (Klump and Preissler [16], p.46), and elasticity of substitution $1/(\rho - 1) = -1/d(ij)$ between e_{ij} and e_{ji}. If agents i and j contribute to the coalition with identical types of ideas, $d(ij) = 0$, and $w_{ik} = w_{ji} = 1/2$. i's and j's efforts are then perfect substitutes, with constant marginal contributions to effective effort $e(ij)$.

$$e(ij) = \frac{w_{ij} e_{ij}}{f(d(ij))} + \frac{w_{ji} e_{ji}}{f(d(ij))} = \frac{e_{ij} + e_{ji}}{2f(d(ij))} = \frac{1}{2}(e_{ij} + e_{ji}) \tag{20}$$

The intermediate case $d(ij) = 1$ implies $\rho = 0$, so that $e(\cdot)$ is Cobb-Douglas:

$$e(ij) = \left(\frac{e_{ij}}{f(d(ij))}\right)^{w_{ij}} \left(\frac{e_{ji}}{f(d(ij))}\right)^{w_{ji}} = e_{ij}^{w_{ij}} e_{ji}^{w_{ji}} \tag{21}$$

Finally, if $d(ij) = +\infty$, then $\rho = -\infty$: when i's and j's knowledge sets are disjoint, their efforts are perfect complements; the performance function is then Leontiev.

$$e(ij) = \min \left[\frac{e_{ij}}{f(d(ij))}, \frac{e_{ji}}{f(d(ij))}\right] = \min \left[\frac{e_{ij}}{F}, \frac{e_{ji}}{F}\right] \tag{22}$$

References

[1] Altenberg L. (2005): "Modularity in Evolution: Some Low Level Questions", in Callebaut W. and Rasskin-Gutman D. (eds.), *Modularity. Un-*

derstanding the Development and Evolution of Natural Complex Systems, Cambridge MA., MIT Press, 99-128.

[2] Baldwin C. Y. and Clark K. B. (2000): *Design Rules. Volume 1: The Power of Modularity*, Cambridge, MA, MIT Press.

[3] Barabasi A. L., Jeong H., Néda Z., Ravasz E., Schubert A., and Vicsek T. (2002): "Evolution of the social network of scientific collaboration", *Physica A*, 311, 590-614.

[4] Burt R. S. (1992): *Structural Holes. The Social Structure of Competition*, Cambridge, MA, Harvard University Press.

[5] Carayol N. and Roux P. (2009): "Knowledge Flows and the geography of networks: A strategic model of small world formation", *Journal of Economic Behavior and Organization*, 71, 414-427.

[6] Caminati M. and Stabile A. (2010): "The Pattern of Knowledge Flows Between Technology Fields", *Metroeconomica*, in press.

[7] Dosi G., Hobday M., Marengo L. and Prencipe A. (2005): "The Economics of System Integration. Towards an Evolutionary Interpretation", in Prencipe A., Davies A. and Hobday M. (eds.),*The Business of System Integration*, New York, Oxford University Press, 95-113.

[8] Frenken K. (2006): "A fitness landscape approach to technological complexity, modularity and vertical disintegration", *Structural Change and Economic Dynamics*, 17, 288-305.

[9] Goyal S., van der Leij M., and Moraga J. L. (2004): "Economics: an emerging small world?", Tinbergen Institute Discussion Paper 04-001/1.

[10] Goyal S., van der Leij M., and Moraga J. L. (2006): Economics: emerging small world, *Journal of Political Economy*, 114, 403-412.

[11] Goyal S. (2007): *Connections. An Introduction to the Economics of Networks*, Princeton, Princeton University Press.

[12] Grossman J. W. and Ion P. D. F. (1995): "On a Portion of a Well-Known Collaboration Graph", *Congressus Numerantium*, 108, 129-131.

[13] Grossman J. W. (2002): "The Evolution of the Mathematical Research Collaboration Graph", *Congressus Numerantium*, 158, 202-212.

[14] Jackson M. O. (2008): *Social and Economic Networks*, Princeton, Princeton University Press.

[15] Kauffman S. (1993): *The Origins of Order*, New York, Oxford University Press.

[16] Klump R. and Preissler H. (2000): "CES Production Functions and Economic Growth", *The Scandinavian Journal of Economics*, 102, 41-56.

[17] Laband D. N. and Tollison R. D. (2000): "Intellectual Collaboration", *Journal of Political Economy*, 108, 632-662.

[18] Leicht E. A. and Newman M. E. J. (2008): "Community Structure in Directed Networks", *Physical Review Letters*, 100, 118703.

[19] Marengo L., Pasquali C., and Valente M. (2005): "Decomposability and Modularity of Economic Interactions", in Callebaut W. and Rasskin-Gutman D. (eds.), *Modularity. Understanding the Development and Evolution of Natural Complex Systems*, Cambridge MA., MIT Press, 383-408.

[20] Moody J. (2004): "The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999", *American Sociological Review*, 69, 213-238.

[21] Newman M. E. J. (2001): "The structure of scientific collaboration networks", *Proceedings of the National Academy of Sciences*, 98, 404-409.

[22] Newman M. E. J. (2004): " Analysis of Weighted Networks", *Physical Review E*, 70, 056131.

[23] Newman M. E. J. (2004): "Coauthorship networks and patterns of scientific collaboration", *Proceedings of the National Academy of Sciences*, 101, 5200-5205.

[24] Newman M. E. J. (2006): "Modularity and Community Structure in Networks", *Proceedings of the National Academy of Sciences*, 103, 8577-8582.

[25] Newman M. E. J. and Girvan M. (2002): "Community structure in social and biological networks", *Proceedings of the National Academy of Sciences*, 99, 7821-7826.

[26] Newman M. E. J. and Girvan M. (2004): "Finding and Evaluating Community Structure in Networks", *Physical Review E*, 69, 026113.

[27] Reiter S. (1992): Knowledge, Discovery and Growth. Discussion Paper #1011, Northwestern University.

[28] Reiter S. (2001): "Knowledge, Discovery and Growth", in Olson G. M., Malone T. W., and Smith J. B. (eds.), *Coordination Theory and Collaboration Technology*, Mahwah, NJ, Lawrence Erlbaum Associates, 193-260.

[29] Simon H. A. (1962): “The Architecture of Complexity”, *Proceedings of the American Philosophical Society*, 106, 467-482; reprint in Litterer J. A. (ed.), *Organizations: Systems, Control and Adaptation*, 2, New York, J. Wiley, (1969), 98-114.

[30] Simon H. A. (1973): “The Organization of Complex Systems”, in Patee H. (ed.), *Hierarchy Theory*, New York, Braziller, 1-27.

[31] Simon H. A. (2005): "The Architecture of Complexity in an Evolving World: The Role of Near Decomposability", in Callebaut W. and Rasskin-Gutman D. (eds.), *Modularity. Understanding the Development and Evolution of Natural Complex Systems*, Cambridge MA., MIT Press, ix-xiii.

[32] Simon H. A. and Ando A. (1961): "Aggregation of Variables in Dynamic Systems", *Econometrica*, 29, 111-138.

[33] Watson R. A. (2006): *Compositional Evolution. The Impact of Sex, Symbiosis, and Modularity on the Gradualist Framework of Evolution*, Cambridge MA, MIT Press.

[34] Weitzman M. L. (1998): “Recombinant Growth”, *Quarterly Journal of Economics*, 113, 331–360.

DEPFID WORKING PAPERS

DIPARTIMENTO DI POLITICA ECONOMICA, FINANZA E SVILUPPO
UNIVERSITÀ DI SIENA
PIAZZA S. FRANCESCO 7 I- 53100 SIENA
http://www.depfid.unisi.it/WorkingPapers/
ISSN 1972 - 361X