Concentration in knowledge output to social network: Evidence from India

Bino Paul G D[*] Krishna M[+]

July 2008

(*Paper Prepared for the conference on* ***Globalisation of Knowledge Development and Delivery*** *to be held at the Institute for the Study of Industrial Development (ISID) Delhi, October 17-18, 2008*)

*Associate Professor, School of Management and Labour Studies, Tata Institute of Social Sciences, Mumbai, Maharashtra, India email: *bino@tiss.edu*

[+] M Phil Scholar, School of Social Sciences, Tata Institute of Social Sciences, Mumbai, Maharashtra, India email: *krishna.m@tiss.edu*

Abstract

This paper examines two important aspects of knowledge output: degree of concentration in knowledge output and nature of social network. Taking the *Economic and Political Weekly* (EPW), a known scholarly periodical in the social sciences, as a case, we analyse the author data, looking at frequency distribution of authors against the number of papers published and the social network of authors during the period 1996–2005. While the degree of concentration is assessed using a power law namely the Lotka's law, co-author data is transformed to a symmetric sociogram. The study covers 1803 authors who published research articles in the *EPW*, including contributions by social scientists, policy professionals and scientists. It appears progress in Indian social Science has apparent link with the history of *EPW*, accounting for a significant part of intellectual output and scholarly discourses in Indian Social Science. Taking cues from scholarly lineages, which investigate knowledge as a social phenomenon including sociology of science, economics of knowledge and social network theory, we examine an important question: What is the degree of concentration prevailing in the *EPW*, during 1996-2005? Does the social network of authors, captured from the data, provide explanations for the degree of concentration? Our analysis shows high degree of author concentration and a relatively organized social network of authors co-exist, indicating concentration in knowledge output, perhaps, is inherently linked to the structure.

JEL Classification: Z13, D85,

Key Words: Degree of author concentration, Social Network

1 Introduction

A basic characteristic of the knowledge output is the frequency distribution of this variable is likely to have an asymmetric distribution. As shown by Lotka (1926), frequency distribution of number of authors plotted against number of published journal papers, for a period, is likely to be an asymmetric curve, exhibiting a power law. Essence of a power law is lion's share of authors contributing just one paper while a few publish more than one article. What lies beneath the power law? Taking cues from Klamer and van Dalen (2002), collective nature of knowledge activity co-exists with concentration in knowledge output. The collective nature of knowledge is manifest in social network of authors, which is a collection of ties. In fact, the tie between authors plays pivotal role in the knowledge activity, varying from visible ties like co-authorship to relatively informal ties like knowledge sharing. Interestingly, members in a social network vary in number of ties and degree of organization. An asymmetric distribution of frequency of authors against the number of ties, as shown by Barabasi and Bonabeau, (2003), represents a small world where a large majority of authors has a few ties, indicating the network consists of the core and the periphery. This dichotomy, core and periphery, indicates the network is organized and, perhaps, sensitive to counterfactual possibilities like removal of one author making the structure relatively sparse. Moreover, core component in a network consists of cliques of authors, which have double-edged role in knowledge activity, either enabling or constraining (Cowan and Jonard, 1999). Quite obviously, an investigation about this phenomenon would look at the basic unit of analysis i.e. the author. Like many human activities, the knowledge activity is collective in nature involving a network of authors and related institutions. This paper examines the link between the degree of author concentration and underlying forms of social network. For investigating this link, we look at to what extent the social network of authors is organized., at both macro and micro levels.

The rest of the paper is organized as follows: Section 2 provides a brief review of literature on concentration of knowledge output and role of social network in knowledge activity. Section 3 deals with the analysis of author concentration. Section 4 discusses findings from the social network analysis. Section 5 concludes the paper.

2 Knowledge Output: Concentration and Networks

Knowledge has an important role in economic growth of a country. Considerable attention has been devoted in contemporary economic growth theories (Romer 1990) on the link between knowledge and economic growth. Broadly, this scholarly lineage discusses themes such as knowledge embedding, creation of new combination using existing ideas, production of knowledge, knowledge as a public good, knowledge as a factor of production, taxonomy of knowledge and so on. Moreover, knowledge is also conceived as capability, it goes beyond confining knowledge as a factor of production (Bino & Sankaran, 2007). Although defining knowledge is a complex task, it is important to represent knowledge by appropriate measures. Viewing knowledge as output is quite popular among researchers. For instance, journal publication is often considered as knowledge output (Lovell, 1973).

2.1 Concentration in Knowledge Output

It is important to distinguish exchange of knowledge output from conventional exchange of goods and services. First, along with writers and readers, institutions such as university and journals also have an active role in the exchange. Second, networks of authors and readers are pivotal in shaping the knowledge output and explaining the visibility of it. As argued by Hodgson and Rothman (1999), institutional oligopoly is a salient feature of exchange of knowledge output, resulting in transaction costs. In fact, it is debatable if the role of institutions, more importantly resultant hierarchies, positively impact aggregate knowledge output[1]. Moreover, relatively higher degree of concentration in the distribution of knowledge output –journal article-, for units of analysis such as region, area of specialization, university, and author, is a stylized fact. It is important to note that, as shown by Bino et al (2005), frequency distributions for above-mentioned units of analysis show higher degrees of concentration. Interestingly, the literature on distribution of knowledge output including scholarly works in scientometry, economics, and sociology of knowledge provides ample evidence for distributions forming patterns resembling power laws, clearly indicating a

[1]Recent literature on publication process does not provide convincing evidence for constructive role institutions play in journal publication. For instance, Ellison (2002) shows publication lag, which is often assumed as a strategy to ensure quality of publication, may rather affect quality of idea reflected in a journal article.

picture of inequality. This implies, as reflected by Cox and Chung (1991), an author with multiple papers is likely to publish again than an author with just one publication. Similar pattern follows for citations as well. The literature, based on testing if publication distribution, in particular author distribution, follows power laws such as Lotka's Law, clearly shows observed distribution, though not conforming to predicted pattern, approximates asymmetric curve. Further, Bino et al (2005) based on 704 economic journals, shows high degree of concentration prevailing for three units of analysis – region, University, and sub discipline.

2.2 Role of Networks

It is important to note that the scholarly tradition of testing the similarity between observed frequency distribution and predictions by power laws, in particular research on economic journals, appears to have inadequate explanations for what accounts for concentration of knowledge output. Interestingly, exploring perspectives from Merton (1968), Klamer and vanDalen (2002) identifies the role networks in knowledge activity. Klamer & van Dalen (2002) shows the role of network in knowledge output. According them, there are four types of networks: (a) lone wolves, (b) the science ideal, (c) technology leader sets the standard, (d) learning from neighbors and (e) minimal network structure with a core. In the case of lone wolves, individuals with little interaction with the scientific community produce knowledge and may even emerge as trendsetters. However, it is a costly mode of production. The model b represents complete interaction within the community. Although this is an ideal situation, it is far off the reality. In model c, a leader sets the standard to the followers. There are instances of dominant institutions setting norms in the knowledge production. This is a quite realistic situation in economics. Although this form of knowledge production can reduce the cost of production incurred by the followers, it has many hidden costs such as sacrifice of innovation. Further, the leadership, by curtailing innovation from followers, may reduce the production possibility of followers in the long run. Model d portrays a picture of cooperative learning. Unlike model b, this model does not require one to one contact between all the scholars to circulate knowledge. The model d is circular in nature so that just a link between two scholars is enough to circulate knowledge to entire community. Producing knowledge through this model incurs lesser cost in comparison with models a and b. However, this model is not common in economics. Finally, the model e refers to an arrangement in which

there is core set of journals producing fundamental knowledge and supply this to specialized journals. This model is quite prevalent in economics. An important characteristic of this model is that there will be a hierarchy of prestige in which fundamental ones occupying the top and specialized in the lower tier or vice versa. Again, this is also quite cost effective in the sense that by importing theories from core journals to specialised journals saves resources, and in return, core journals get top tier in the prestige hierarchy

Cowan and Jonard (1999) pinpoint the role of network in the creation and diffusion of knowledge, showing high-density network structure aids faster growth of knowledge. In other words, interactions play a pivotal role in translating human actions in to knowledge output. Moreover, underlying structure of knowledge networks, as reflected by Barabasi and Bonabeau (2003), resembles small world where a few actors account for majority of interaction. Taking cues from this work, power law appears to be an inherent characteristic of knowledge output, emanating from structural properties of knowledge network. Relating the network and the innovation, Cowan and Jonard (1999) show cliquishness, a salient feature of small world, can positively impact knowledge production. However, as reflected by them, this positive relation is sensitive to the inflow of knowledge into the clique from the outside. Hodgson and Rothman (1999) argues that the high degree of institutional concentration in knowledge output is an indication of lock-in; implying institutions resist flow of new ideas. To quote them " 'lock-in' may occur, where specific institutions defend specific, and possibly outdated, ideas and approaches. In these circumstances, it will be quite difficult for alternative or innovative approaches to establish themselves." (p 182) Interestingly, there are instances like article having innovative ideas are turned down by the journals by citing inappropriate language as a reason (Gans & Shepherd, 1994)[2]. As shown by Fray (2003), scarcity of publication opportunity[3] puts editors and referees preferences ahead of authors'; authors learn likes and dislikes of referees and editors and accordingly adapt their knowledge output.

[2] George Akerloaf, who won Nobel Prize, is a classic case for such a phenomenon (Gans & Shepherd, 1994).
[3] evidently impacted by its structural aspects of knowledge output including institutions and network of authors

Goyal et al, (2006) identifies the existence of small world properties in economics discipline, based on journals listed in the Ecolit database, for the period 1970-2000. The study examines the phenomenon of small world, using four measures viz. growth in the number of authors during the period of analysis, temporal change in the size of the giant component, the average distance between the nodes in the giant component, and clustering coefficient[4]. Based on the above four measures, the study displays small world effect in economics i.e. the number of authors has grown substantially over the period 1970 to 2000, significant increase in the size of giant component during 1970 to 2000, the giant component has become smaller in terms of distances, increase in clustering coefficient. These findings, according to authors, indicate economics is an emerging small world. An important sign of small world effect is cliquishness in the network, implying a few nodes account for a majority of interactions. Taking cues from Cowan and Jonard (1999), while cliquishness can be catalyst to innovation, it might resist the inflow of new ideas into the network.

2.3 Rationale

Broadly, literature on knowledge output, reviewed in this paper, is of three types: measurement of author concentration, testing of small world, and explaining knowledge activity using network model, which generate the context for this paper (figure 1). In figure 1, the box representing concentration in knowledge output contains two articles, which measures author concentration in economics. While Cox and Chung (1991) measures degree of author concentration for a sample of twenty journals in economics, Bino et al (2005) assesses degree of author concentration for seven economics journals published from India. These two articles based on data from different geographic background view that degree of author concentration is likely to go up with degree of specialization. Although works of this nature provide cues about pattern of knowledge output, a major gap is that no apparent attempt has been made to link degree of author concentration and social network of authors. However, noted contributions in fields such as sociology of science (Merton, 1968), economic sociology (Granovettor, 1985), economics of knowledge (Cowan and Jonard, 1999; Klamer and van Dalen, 2002) view social networks play pivotal role in knowledge

[4] A giant component within a network is the maximal connected sub-graph or largest network within a network (Scott, 2000). Clustering coefficient refers to what percentage of a person's co-authors who are co-authors of each other. Distance is defined as the path length between authors.

activity, treating knowledge output as an outcome of interactions happening in a network. Quite interestingly, role of networks in knowledge activities, in particular in scientific disciplines, are reflected in recent empirical research (Albert and Barabasi, 2002; Goyal et al (2006)). This lineage observes that knowledge activity, covering select sciences and social sciences, shrinks in to a small world of connections where a few of nodes accounts for a majority of interactions. In fact, investigating small world effect in knowledge and fitting distribution of knowledge output are complementary to each other, aiding endeavors to link knowledge activities and network (figure 1).

In this paper, we examine the link between knowledge output and social network, taking the *Economic and Political Weekly* (EPW) as a case study. The *EPW* is a known journal in social sciences in India. We look at frequency distribution of authors and network of co-authors during the period 1996–2005. While the degree of author concentration is assessed using power law namely Lotka's law, co-author data is transformed into a sociogram. The study covers 1803 authors who published research articles in the *EPW*, including contributions by social scientists, policy professionals and scientists. It appears progress in Indian social Science has apparent link with the history of *EPW*, accounting for a significant part of intellectual output and scholarly discourses in Indian Social Sciences.

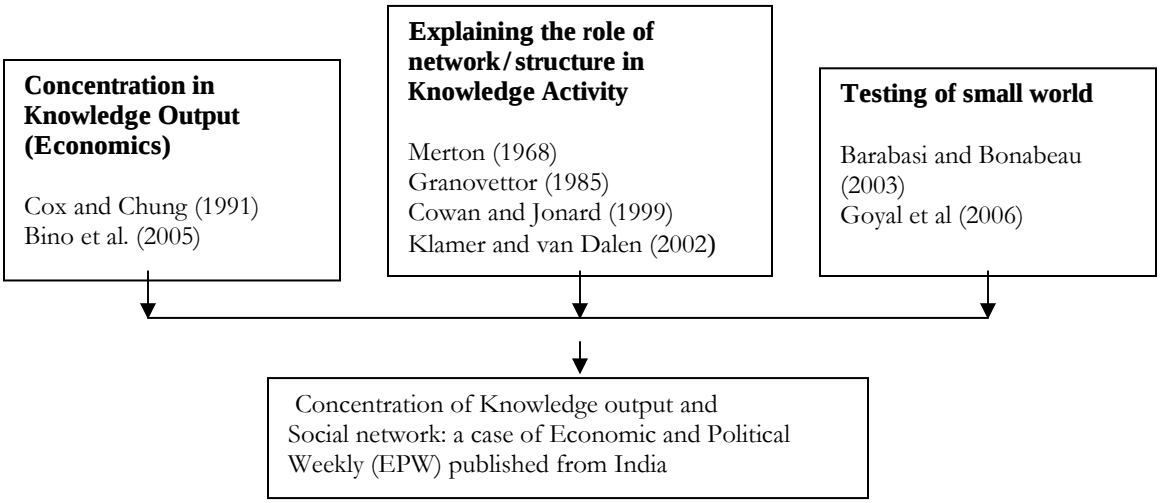


Figure 1. Literature and Context of this study

2.4 Indian Scenario

It is important to note that India's knowledge output is much lower compared to developed and emerging economies (Dahlman & Utz, 2005). However, in India's case the lower knowledge output seems to be explained by relatively lower density of researchers. Basu and Nagpal (1998) shows that majority of scientific publication from India are published in journals having low impact factor, which is a bibliometric measure of scientific visibility. For an Indian author, chance of publishing in journal with high impact factor such as Science is much lesser than that of an author from United States of America (USA). Interestingly, the networks defined in terms links such as co-authorship with scientists from the other region, both the inter state and international, have significant role in knowledge production in Indian Science. As noted by Basu and Nagpal (1998), a major trend in Indian Science is that there seem to be a direct relation between the size of network and knowledge output.

In Indian science, each discipline has its institutional order. The order consists of a few islands of excellence and many less visible institutions. Being the part of same educational system, Social Science in India has certain similarities with the science. Particularly, this is quite evident in Economics profession in India. As observed by Bino et al. (2005), a few institutions dominate the distribution of knowledge output in prominent journals such as the *Indian Economic Review*[5]. It is important to note that there is dearth of empirical research, which integrates concentration of knowledge output and social network of authors, using publication data sourced from Social Science Journals published from India. Perhaps, this gap in the literature is due to lack of scientometric measures derived from social science journals published from India. In fact, for some social science journals published from India; basic data such as author's institutional affiliation and subject code are not reported. A prominent case representing Indian social science research, playing pivotal role in academic research as well as policy debates related to contemporary socio economic issues, is the *EPW*. This journal covers scholarly works from a variety of disciplines such as economics, sociology, anthropology, political economy, science, and society and so on. More succinctly, this journal is not a specialised one, rather covers a wide spectrum of themes. In this paper,

[5] Bino et al (2005) covers seven economic journals i.e. *Indian Economic Review, Indian Economic journal, Journal of Quantitative Economics, Indian Economic and Social history Review, Indian Journal of Economics, Indian Journal of Agricultural Economics, Artha Vijnana.*

we examine author concentration and social network of authors in social science, taking the *EPW* as a case.

2.5 The Economic and Political Weekly (EPW): A brief profile

The *EPW*, a weekly publication, is published from Indian metropolis Mumbai. Tracing EPW's past, Sachin Chaudhuri started erstwhile version of this journal, known as the *Economic weekly*, in 1949. In August 1966, the journal changed its name to the *Economic and Political Weekly*. Following founder's death in 1966, R K Hazari, a well-known economist, took over the editorship of the *EPW*. His tenure ended in 1969, succeeded by Krishna Raj, who remained as editor until 2004 (Table 1). As viewed by Guha (2004), during the early stage of the *EPW*, in particular 1960s, articles in the *EPW* were of two types having approximately equal representation i.e. articles with liberal orientation and articles with inclination towards the leftist ideology. However, this ideological mix in the *EPW*, in 1970s and 1980s, underwent a change, indicating a drift towards left based writing. Interestingly, this change evoked mixed reactions, including voices of dissent from noted intellectuals. Guha (2004), citing a letter from the historian Dharma Kumar, which was printed in the *EPW*, notes changes that germinated in the *EPW* content in response to the criticism against Marxist hegemony in the journal. As recourse to the hegemonic pattern in the *EPW*, post 1990 saw more liberal-oriented economists and other social scientists contributing to the *EPW*. In fact, unleash of economic reforms in India accentuated participation of scholars with neo-classical orientation in the *EPW*. Going by Guha's observation, post 1990 appears as if the pattern of knowledge output was subject to a structural break, departing from the ideological hegemony.

Table 1. Editors of EPW and tenure

Name of Editor	**Tenure**
Economic Weekly (1949-1966)	
Sachin Chaudhuri	1949-1966
EPW since 1966	
R. K Hazari	1967-1969
Krishna Raj	1969-2004
Padma Prakash	2004-05
Rammanohar Reddy	Since 2005

Taking cues from the changes happened during post 1990; we look at the pattern of knowledge output and social network of the *EPW* during 1996–2005, covering both distributional and structural aspects. Presumably, the *EPW* is the epitome of contemporary social science research in India. A few questions emerge; is there any evidence for the hegemony -distributional, structural, and ideological- in the *EPW* during 1995-2005? Is there any link between concentration in knowledge output and social network, representing distributional pattern and structure, respectively? Are top authors inclined towards a particular ideology? Can knowledge network in the *EPW* be reduced to a small world of connections? Exploring these questions, we examine the degree of author concentration and social network using database of authors compiled from the *EPW*, for the period 1996 - 2005.

It appears the *EPW*, combining rigor of a social science journal and flamboyance of discourses, fairly represents the progress in Indian social science, enveloping advancement in methodology, new data, and contemporary perspectives reflected by scholars. Interestingly, the *EPW* seems to have no close substitutes, which emulate its diversity and scholarly lineages. Moreover, the *EPW* has a key role in integrating the social science knowledge being produced in India. Ascribing the role of core institutions in social science knowledge output in India to the *EPW*, largely, is well grounded in reality. It is also important to note that, perhaps, no other social science journal is of either comparable size or having wider spectrum of social science themes. On other hand, the *EPW* is not as specialized as the *Indian Journal of Agricultural Economics* or the *Sociological Bulletin* in the coverage of respective themes of results. Indeed, if we reflect on the history of the *EPW*, its aim is not only just deepening social science but also widening its scope.

3 Analysis of Concentration Indicators

We assess the degree of author concentration, for the *EPW* during 1996-2005, using a model derived from the Lotka's law. Here, the unit of analysis is the author. For examining the degree of author concentration, names of authors who have published research articles[6] in the *EPW* have been compiled, forming the database for the analysis. For our research, we have considered two types of authors: authors who have coauthored at least one research

[6] We exclude comments and rejoinders published in the *EPW*.

paper and authors who have authored at least an article without any coauthors. Following Cox and Chung (1991), we give equal credit to all coauthors. Aggregating both the categories of authors, there are 1803 authors, who published 2073 research articles during 1996-2005. An important question is if the distribution of authors, for number of papers authored, follows a pattern. If the distribution is asymmetric, implying the proportion is inversely related to the number of papers, presumably, the author distribution provides useful cues for assessing the degree of author concentration. A frequently cited asymmetric distribution is the Lotka's law.

Lotka's Law

Lotka (1926) found the frequency distribution of authors against the number of papers published forms a pattern, which is represented by a law, known as Lotka's Law. The Lotka's law states the ratio of the number of authors published *n* papers to the number of authors published only one paper is inverse of n squared. The law is expressed:

$$a_n = \frac{a_1}{n^2}; \quad n=1, 2, 3\ldots. \qquad (1)$$

Where

a_n = the number of authors publishing *n* papers
a_1 = number of authors publishing one paper
n = number of papers

Following Niles and Haborak (1971), Cox and Chung (1991) generated a theoretical probability distribution for the Lotka's law[7], for testing if the observed author distribution follows the theoretical one. Here, the null hypothesis is there is no difference between the observed distribution and the theoretical one, implying the regularity in the distribution of authors. On the other hand, alternative hypothesis states observed and theoretical distribution are different (Table A1 & Figure 2).

[7] The theoretical distribution is computed using the formula i.e. $\frac{6}{\pi^2}\left[\frac{1}{n^2}\right]$.

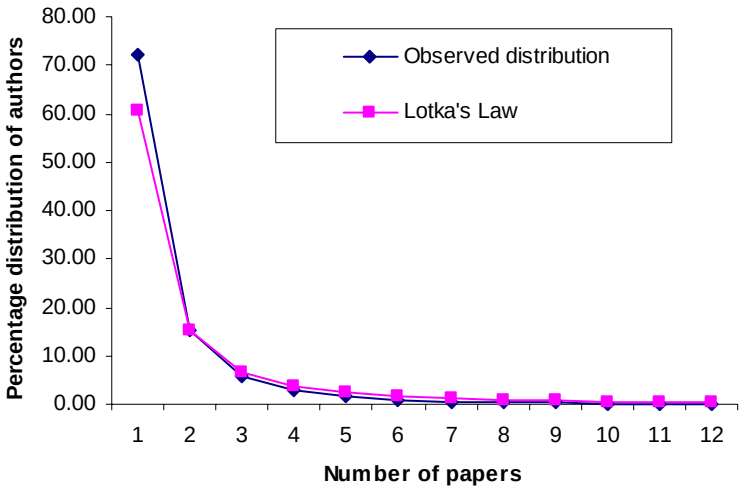


Figure 2. Observed distribution versus Lotka's law

From figure 2, both the distributions are asymmetric, declining throughout. While the observed distribution, for the authors with not more than two papers, lies above the theoretical distribution, the second distribution exceeds the first for the authors who have more than two papers. Although both the distributions appear to have marginal difference, these curves intersect. Moreover, for the author with just one paper, the observed distribution is 72 percent, which is 11 percent greater than the theoretical observation. On the other hand, beyond the intersection, the theoretical distribution is above the observed distribution along the horizontal axis[8].

Taking cues from the intersection of two distributions given in figure 2, we test the null hypothesis that there is correspondence between both the distributions. The chi-square statistic, based on the difference between two distributions, is too large, and we reject the null hypothesis at one percent level of significance[9]. Quite interestingly, our result is consistent with Cox and Chung (1991) and Bino et al (2005), both the studies rejecting the null hypothesis. Obviously, both the distributions are asymmetric, while they differ in

[8] Beyond the intersection, though the difference between upper lying distribution and lower lying one appears too small, difference measured in rate of change i.e. difference between the upper lying points and lower lying ones divided by lower lying ones is quite pronounced, varying from 24 % to 661%. However, both the curves slope downward, tending towards the horizontal axis.

[9] The chi-square value is 96.6 (degree of freedom =11).

respective values along the vertical axis. Perhaps, by converting n's exponent i.e. two to a parameter 'β' we can estimate a plot, which corresponds to the observed distribution. Equation 1 becomes:

$$a_n = \frac{a_1}{n^{\beta}} \qquad (2)$$

To estimate equation 2

$$\log \frac{a_n}{a_1} = - \beta \log n + u \qquad (3)$$

Where

β is parameter, u is error

The model (equation 3) has an interesting implication, as shown in Cox and Chung (1991); higher the β[10] lower will be the degree of author concentration and vice versa. Further, the degree of concentration is sensitive to the extent of specialization of the journal, lower β values for specialised journals and higher values for general ones. For the *EPW*, during 1996 – 2005, the β value is -2.78[11]. Perhaps, the degree of concentration reported in this study, as shown by the β value that is in proximity of two, is too high for a journal such as *EPW*, which is not a specialised journal[12]. It is important to note that only 13 percent of authors have published more than two articles, clearly indicating high degree of asymmetry in the distribution. Presumably, a distribution with pronounced asymmetry indicates the network of authors is a scale-free one, rather than a random network. As shown by Barabasi and Bonabeau (2003), author networks are scale-free implying a few authors account for a majority of co-authorship links. On the other hand, in a random network the distribution is relatively symmetric. Quite interestingly, the plot of observed distribution (figure 2) is quite asymmetric, clearly conveying a power law. Barabasi and Bonabeau (2003) views "power laws are quite different from the bell shaped distribution that characterizes random networks. Specifically, a power law does not have a peak, as a bell curve does, but is instead

[10] Here, we refer to absolute value of beta.
[11] R^2 = .98, F = 617, significant at the 1% level.
[12] Although the EPW often publishes special issues covering select themes, the extent of specialization is much lower than that of the specialized journals such as the Journal of Financial Economics.

describing by continuously decreasing function" (p. 52). Moreover, scale-free networks exhibit structural properties such as 'small world' where connections play a pivotal role in human action. Thus, it is important to link the author concentration in knowledge output, as shown by our analysis, and the social network of authors.

4 Social Network of authors

Is the social network involved in the knowledge output? The phenomenon of 'co-authorship' provides vital cues why social networks play critical role in knowledge output. A coauthored article means two or more authors jointly publish an article. Visibly, a coauthored paper indicates the tie between authors, while ties other than the co-authorship that are not so visible may also have role in the publication. Succinctly, co-authorship is a visible tie. Such ties form a social network of coauthors. Moreover, the coauthor-ship as a tie is relatively embedded in nature, compared to arms length ties like hiring a cab[13]. However, the co-authorship is not as embedded as the friendship between two persons is. Perhaps, the co-authorship is an outcome of a sequence of processes wherein factors such as institutional affiliation, ideology, and research interest have core roles. It is important to note, as shown by Hudson (1996)[14], the percentage of co-authored papers is on increase with respect to time, over decades. Interestingly, for the *EPW*, during 1996-2005, the percentage of co-authored papers has increased from 16 % to 26 %, showing 62% rate of increase during the decade (Table A2, Figure 3). Presumably, increase in the percentage of co-authored articles indicates increasing relevance of social network of authors in the knowledge output.

[13] Granovetter (1985) classifies ties into two: embedded and arms length. While the first type is more trust based, involving more time, institutions and so on, the second type is relatively impersonal, something like transaction between seller and buyer in a perfectly competitive market.

[14] Hudson (1996) shows, for eight top rated economics journals, phenomenal increase in the percentage of co-authored articles during 1950-1993.

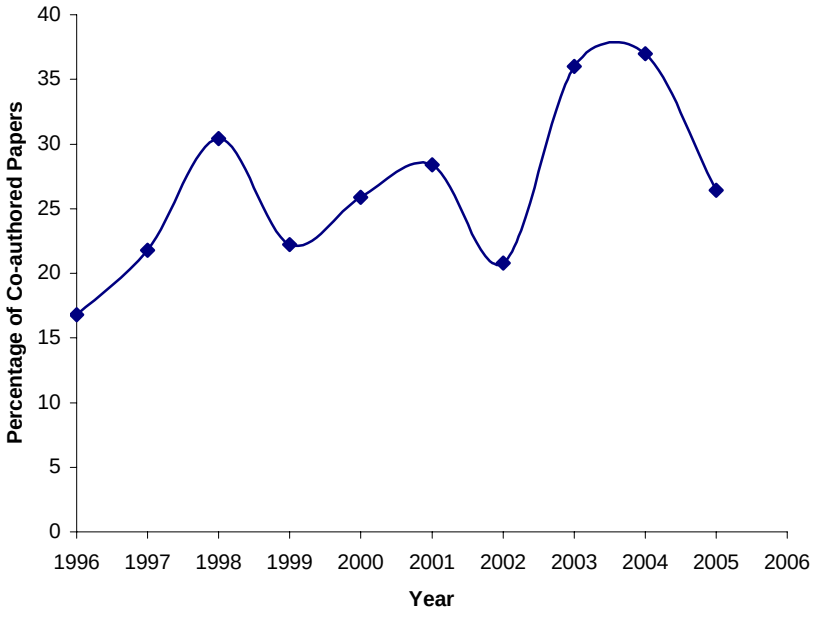


Figure 3. Year wise Percentage of Co-author Papers

We examine the social network of 1803 authors who published research articles[15], covering both the authors with no coauthors and the authors with coauthors. To begin, we visualize a network of n authors[16]. This network is a set of n authors, with maximum ${}^{n}C_{2}$ ties between authors[17]. A tie is between two authors. For the co-authorship, a tie is undirected, implying a tie has neither source nor end. On the other hand, if the tie is a directed one, tie originates from a source and joins the end. Most of relational data such as co-authorship are undirected while most of economic transactions fall in the category of directed[18]. To cite an example, a social network of three authors, viz. A, B and C, ties are undirected, indicating three are coauthors of a paper (figure a) while figure b portrays s social network of two authors (see figure 4). From figures given below, the tie is represented by a two edged arrow, not conveying where is end and where is source. Further, the distance is not physical in nature[19],

[15] According to EPW parlance, research articles are called special articles. This set does not include comments and perspectives.

[16] For this, we use Ucinet (Borgatti et al. 2002).

[17] This implies $\frac{n(n-1)}{2}$ ties.

[18] For directed ties, maximum number of ties is ${}^{n}P_{2}$ i.e. $n(n-1)$.

[19] The arrow is not a Euclidean Distance.

but a graphic one. For a three member social network, as given in figure a, there is an equivalent matrix structure (figure A1). From the matrix, maximum number of ties are computed by summing the number cells after deducting self reflexive ties, represented by number of diagonal cells, and one of the off diagonal triangles, either low or upper, since the tie between A and B is equivalent to the tie between B and A. Figures a[20] and b are called symmetric sociogram.

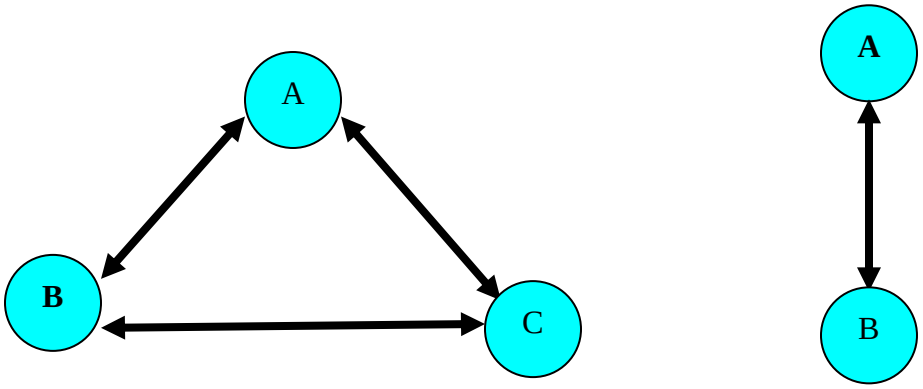


Figure 4. (a) Social network of three coauthors publishing at least one article or transitivity or each collaboration produces at least one paper (b): Social network of two coauthors publishing one article

In this paper, we have created a matrix of 1803 authors[21] who published at least one article in the *EPW*. This matrix contains both the authors who have co-authored with at least one among the other 1802 authors and those who have not coauthored at all during 1996-2005. Each cell in the matrix, representing the social network of authors, represents a tie between two authors, conveying that they have coauthored at least one paper. The cell can have two types of data: either dichotomous data or a ratio scale. If the data is dichotomous, say one or zero, one means two authors have jointly authored at least one paper in the *EPW* and zero indicates no co-authorship. Otherwise, a ratio scaled data, also called valued graph, says

[20] For figure 4a, there are two options: (a) three authors jointly publishing one article (b) transitive ordering i.e. A coauthors with B, and B with C, and A with C, each collaboration produces at least one paper.
[21]The equivalent word for the author is actor in sociology, node in graph theory, and vertex in physics. The tie between two authors is called line in graph theory and edge in physics.

frequency of coauthored publication by these two authors. We study both the whole network as well as the subset of the network, called component. While, for assessing the whole network, we use dichotomous data, components are examined using valued graphs.

Table 3 gives descriptive statistics of social network of authors in the *EPW* during 1996-2003, based on the dichotomous data. From the data, we compute two important aggregate measures representing the whole network viz. inclusiveness and density. Inclusiveness is the number of authors with at least one coauthors i.e. connected author as a proportion of total number of authors in a network. Density refers to the ratio of ties in a network to maximum possible ties in a network. While inclusiveness measures the degree of connectedness, density indicates the degree of completeness of a network. Both these measures vary from zero to one, indicating lowest to highest. As appears from the table, slightly more than a half of authors have at least one coauthor, reflected in the measure inclusiveness. However, the network is relatively incomplete in nature, implying the inter tie connection is weak. For example, figure a (figure 4) shows a network having both these indicators equal to one, conveying complete connectedness as well as completeness of the network. Supposing we remove the tie between A and C, inclusiveness will be one while density drops to two third.

Table 3. Descriptive Statistics of Social Network of authors (1996-2005)

Number of Authors	1803
Number connected Authors	977
Isolated authors as percentage	46 %
Inclusiveness	.54
Maximum possible ties[22]	1624503
Number of ties in the network	940
Density[23]	0.0006

Source: Computed by authors from compiled data from the EPW

[22]If there are n authors in the network, maximum number of ties is ${}^{n}C_{2}$ ties i.e.

$\frac{n(n-1)}{2} = \frac{1803 \times 1802}{2} = 1624503.$

[23] Density = $\frac{l}{{}^{n}C_{2}} = \frac{2l}{n(n-1)}$; l = number of ties, n = Number of authors

The degree of an author is another important indicator, which measures relative importance of an author in the network. Scott (2000) defines the degree of a node as measure of the size of its neighborhood. In other words, referring to our study, the degree of author *x* is the number of coauthors with whom *x* has published. The degree of author *x* varies from zero to number of authors in the network except author *x*. Authors with zero degree are called isolates. There is a relationship between density and degree. Since the sum of degree is two fold of the number ties in a network, density is the mean degree divided by n-1 authors[24]. Obviously, more mean degree implies more density. The mean degree, for the *EPW* during 1996-2005, is 1.04[25], explaining why the density is low (Table A3). This result is similar to Goyal et al (2006), who report average degree of 1.67[26] for the journals listed in *Econolit* database during 1990-1999. Figure 5 shows the percentage distribution of degree (Also see Table A3). Interestingly, degrees not exceeding two constitute slightly above 85% of the total, while degrees above eight forms not even 0.05%. . The distribution represented by steeply falling curve is asymmetric. In fact, we see figures 2 and 5 are similar, indicating the prevalence of the power law.

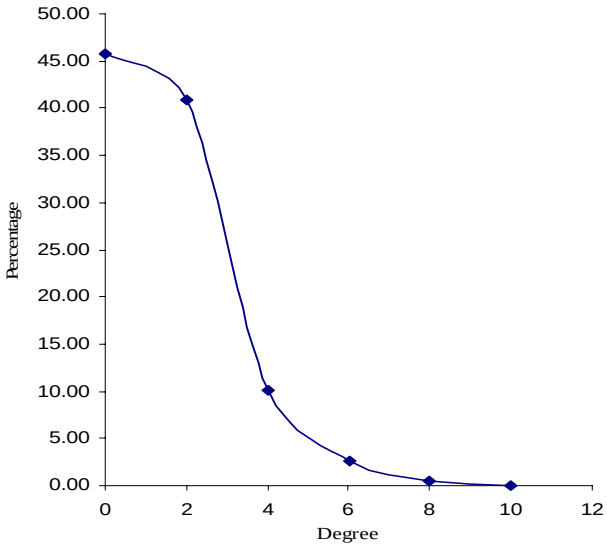


Figure 5. Percentage distribution of degree in the social network of the EPW (1996-2005)[27]

[24] $\frac{l}{{}^{n}C_{2}} = \frac{2l}{n(n-1)}$; Since $2l = \Sigma$ Degree, Density $= \frac{2l}{n(n-1)} = \frac{\Sigma\, Degree}{n(n-1)}$ = Mean degree divided by n-1.

[25] Standard deviation is 1.35.

[26] Goyal et al (2006) reports a standard deviation of 1.69.

[27] Horizontal axis is measured in interval scale. Zero includes zero. Two includes one and two. Four includes three and four and so on.

Quite obviously, as shown by Barabasi and Bonabeau (2003), figure 5 indicates the author network of the *EPW* is scale-free, not a random one. Taking cues from their research on large network such as *World Wide Web* scale-free networks are sensitive to organized attack while accidental error, though affect the periphery, do not influence the core. Applying to this inference to the *EPW*, if relatively more organized ideological orchestration, left or right would influence the core, the journal would then be infested by monologues, which eventually crowd out dialogues. It is important to note if there are centrally positioned authors, they are vulnerable to the organized attack from sources such as ideology, school of thought and so on. This leads to a question: Are there centrally positioned authors in the *EPW*? Assessing the data, we found a range of arrangements varying from relatively centrally positioned authors to dispersed ones. Applying our heuristics, we classify relatively centrally positioned authors into three (Figure 6): (a) Star network[28] (Wasserman and Faust, 1994, p 171) (b) relatively author centric network with at least one triangle[29] (c) relatively author centric network with at least one interlocking[30] (Scott, 2000, p 45). While the *EPW* network has only one star network, b and c are samples.

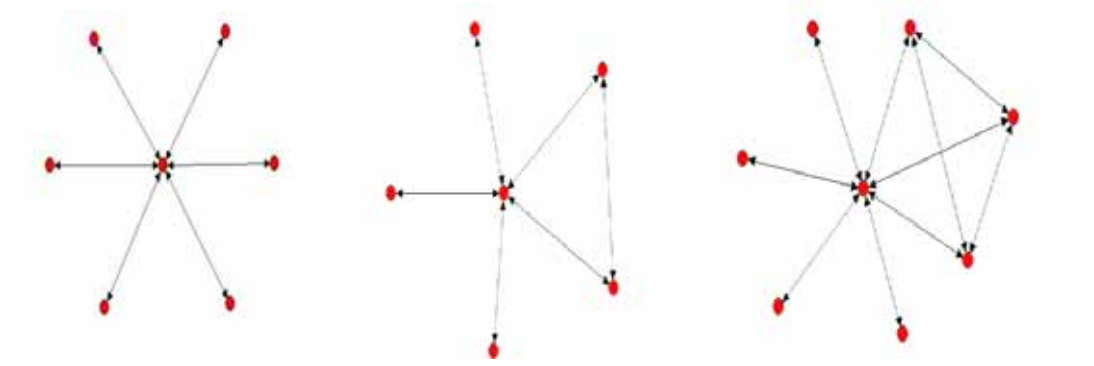


Figure 6. Different types of author centric networks

[28] In a star network, the author who is centrally positioned, the star, has six degrees (a in figure 6)
[29]see figure 4a (also see b in figure 6)
[30] An interlocking is of triangular shape with intersecting diagonals (see c in figure 6).

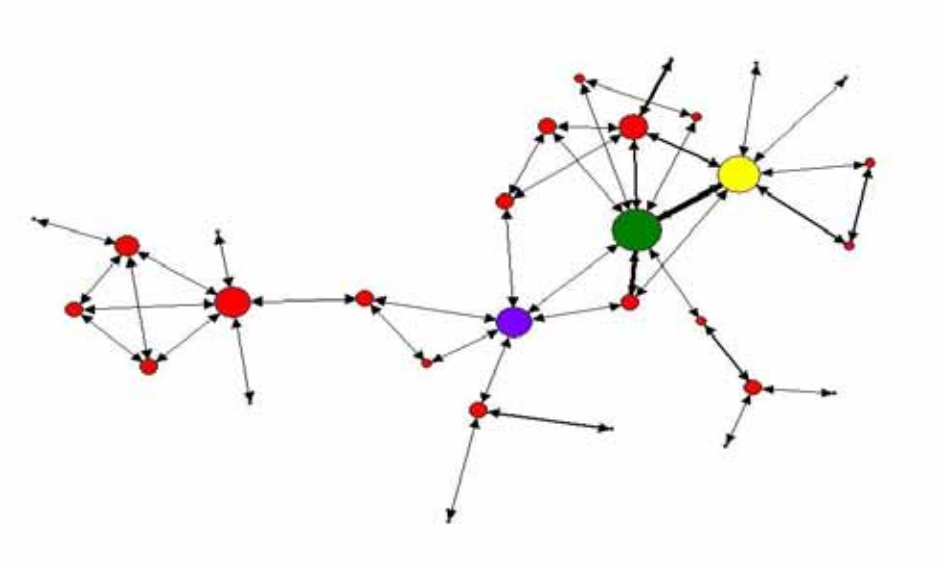
Taking cues from the star network, the star author who is centrally positioned may have decisive role as a firm in organizing coauthors, which are not linked with each other, for publishing articles. This is quite similar to the economic metaphor of entrepreneur, who coordinates other factors of production for producing the output. As reflected by Klamer and van Dalen (2002), the flow of communication implicit in a star network is one of the efficient arrangements, considering high degree of organization in the network. Moreover, they view that the networks, consisting of journals as nodes, provide evidence for the star network, implying peripheral journals import ideas from core journals. In the case of the *EPW*, we found only one star network. Interestingly, this is a network of economists with an eminent development economist at its centre. All members in this network share interest in development related issues. In contrast to the star network, removal of the author, with the highest degree, does not render the network a set of isolated authors. While the removal results in density of star network falling to zero, for b and c at least one tie remains. For b the outcome of removal is a network with one tie and three isolated authors. In fact, the removal has lesser impact on c, compared to a and b; after the removal, c will be reduced to a triangle and four isolated authors. Presumably, impact[31] of the removal of the author with highest degree is sensitive to the centrality of a network; higher will be the impact for the higher centrality and vice versa. Quite interestingly, what integrates above three networks is economists, who share interest in areas like development economics and macroeconomics with special focus on Indian economy, form these networks.

We take one-step further from the author centric organization to a macro level, which is interwoven by a common path connecting all the authors either directly or through adjacent paths. Such a macro unit of network is called the main component. Scott (2000) defines the component as a maximal connected sub-graph[32]. In other words, the component consists of a set of authors who are connected by a main path and another set of authors who are connected to paths, which are adjacent to main path. The main component is the largest component in the network. Sequel to our previous analysis of author as a unit of organization, we explore to what extent main component is organized. For examining the

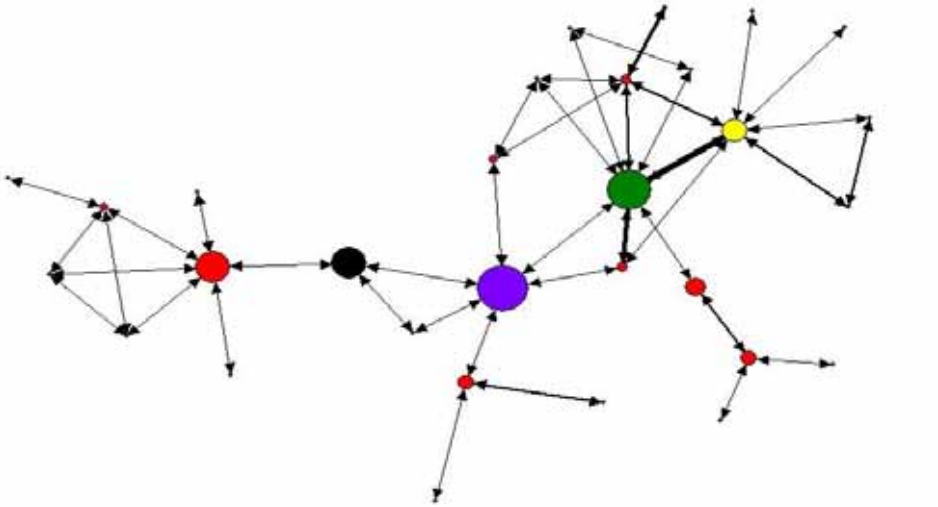
[31] Impact implies future course of the network after the removal of the author with highest degree.

[32] Scott (2000, p 101) elaborates "a sub-graph, like a graph, is 'connected' when all of its points are linked to one another through paths: all points in a connected sub-graph can 'reach' one another through one or more paths, but they have no connections outside the sub-graph".

main component, we weight tie between authors by frequency of publications. Secondly, authors are weighted by two measures of centrality -degree and Betweenness-[33] (Figure 7).



(a)



(b)

(a) Authors weighted by degree[34] and (b) Authors weighted by betweenness

Figure 7. Main Component[35]

[33] Betweenness refers to the number of times an author occurs on a shortest path between authors in the network. Here an author has some control over the interactions (Wasserman and Faust , 1994).
[34] In figure 7, some ties are thicker than others are. This indicates the author has higher degrees.
[35]Inclusiveness and density of binary network of figure 7 are 1 and 0.09, respectively.

As shown in figure 7, the main component consists of 30 authors. In terms of the number of degrees, three authors, marked green yellow, and blue, are ranked first, second and third respectively (Figure a), while the same authors weighted by betweenness (Figure b), blue, green, and yellow taking first, second and third positions respectively. It is important to note blue coloured author, in both the figures, through an author connects a relatively bigger network with a smaller one. This is quite akin to connecting a central business district in a metropolitan city with a small island. In both the figures, using counterfactual reasoning, if blue marked author is removed from the main component, the network becomes sparse, reduced to three islands varying in size -large, medium and small-. Quite interestingly, this author has lesser number of co-authors than green and yellow have (figure a). Further, though one of the low degree authors, black coloured one has relatively higher degree of Betweenness, this author's removal splits the large component into two i.e. a large island and a small island. Compared to blue one, rest of the network, including those with higher degrees, green and yellow, do not make the structure as lesser connected as the blue does. What makes the main component organized by a common thread? In fact, clique[36] is an integral part of the main component. Here, we choose cliques in which all pair of authors is directly connected at distance one, called 1-clique[37]. Figure 7 has nine 1-cliques[38]. This shows the core of the network is cliquish, indicating a structure that is conducive to the knowledge output. However, Cowan and Jonard (1999), while showing how cliquishness of a knowledge network aids innovation, also cautions about cliques' potential to emerge as a barrier to the inflow of new ideas.

4 Concluding Remarks

Two phenomena co-exist: high degree of author concentration in knowledge output and a social network of authors with a core component and relatively sparse subsets. Our examination of the author data of the *EPW* for 1996-2005 unravels the power law and the underlying structure. Perhaps, the power law, which represents the distribution of authors against the number of papers published, is inevitable, being strongly influenced organized

[36] Following Scott (2000), a clique is a sub-set of authors in which a tie directly connects every possible pair of author and the clique is not contained in any other clique.
[37] A triangular network is an example for 1-clique.
[38] The *EPW* network has 133 1-cliques.

components within the social network of authors. Presumably, concentration is a salient feature of knowledge output, reinforced by a scale-free network of authors. Moreover, micro-macro structures within a in a network -author centric, clique and main component- are relatively organized, indicating that there is a small world in the author network of the *EPW*, where a few authors form the core of the network. It is important to ask if authors can be organized in a relatively random manner, contrary to the scale-free network. Sequel to this, would such an alternative brings better results -new idea, new data, and scientific progress and so on-? Answering these questions needs more insights about how the form of network influences the content of the journal. Future research, with in-depth inquiry into the content and institutional aspects of the author, may explore underlying linkages between the structure and the knowledge output.

Appendix 1

Table A1. Frequency distribution of authors by number of papers in Economic Political Weekly for the period 1996-2005

Number of Papers	Frequency of Authors	Observed Distribution (%)	Predicted Distribution by Lotka's Law (%)
1	1300	72.10	60.79
2	276	15.31	15.20
3	106	5.88	6.75
4	55	3.05	3.80
5	27	1.50	2.43
6	13	0.72	1.69
7	9	0.50	1.24
8	6	0.33	0.95
9	4	0.22	0.75
10	3	0.17	0.61
11	3	0.17	0.50
12	1	0.06	0.42
Above 12	0	0.00	4.86
Total Number Of Authors	1803	100	100
Total Number of Papers	2073	____	____

Source: Computed from list of authors compiled from Economic and Political weekly

Table A2. Year wise Percentage of Co-author Papers

Year	**Total Number of Papers**	Co-authored Papers (%)
1996	161	16.8
1997	170	21.8
1998	191	30.4
1999	203	22.2
2000	212	25.9
2001	208	28.4
2002	231	20.8
2003	239	36.0
2004	276	37.0
2005	182	26.4
1996-2005	2073	27.3

Source: Computed from list of authors compiled from Economic and Political weekly

	A	B	C
A	$\{A,A\}$	A,B	A,C
B	$\{B,A\}$	$\{B,B\}$	B,C
C	$\{C,A\}$	$\{C,B\}$	$\{C,C\}$

Figure A1. Matrix form of figure a (Figure 4)[39]

Table A3. Frequency Distribution of Degree

Degree	Frequency
0	826
1	506
2	232
3	143
4	39
5	34
6	13
7	5
8	4
9	0
10	1
Mean Degree	1.04
Total	1803

Source: Computed from list of authors compiled from Economic and Political weekly

[39] Number of Maximum number ties is $\frac{n(n-1)}{2}$ i.e. 3. The $\frac{n(n-1)}{2}$ is equivalent to number of cells in the matrix minus diagonal elements and one off-diagonal triangles, here, the lower triangle, or, number cells minus bracketed cells.

References

Albert, R and A-L Barabasi (2002) 'Statistical mechanics of complex networks', *Review of Modern Physics*, 74, 47-97

A-L, Barabasi and Bonabeau, Eric (2003) 'Scale-free Networks', *Scientific American* 288, 50-59

Basu, Aparna & P S Nag Paul (1998) 'National Mapping of Science: A Bibliometric Assessment of India's Scientific Publications based on Science Citation Index', *National Institute of Science, Technology, and Development Studies*

Bino Paul G.D and Sankaran, K. (2007) 'Indigenous Knowledge and Fast Moving Consumer Goods- Are we seeing some tradeoffs?', *International Journal of Agricultural Resources, Governance and Ecology*, 2007, 6(6), pp. 612-625

Bino, Paul G. D.; Subash S. and Ramanathan A. (2005) 'Concentration in Knowledge Output: A Case of Economics Journals', *The European Journal of Comparative Economics*, 2 (2), pp. 261-278

Borgatti, S.P., Everett, M.G., and L.C (2002) 'Ucinet for Windows: Software for Social Network Analysis', Harvard, M A, Analytic Technologies

Cowan, Robin & Nicolas Jonard (1999), Network Structure and the Diffusion of Knowledge, Published in *Journal of Economic Dynamics and Control* 28 (8), 1557-1575, 2004, MERIT Research Memoranda 1999-028

Cox Raymond A. K and Chung Kee H. (1991) 'Patterns of research Output and Author Concentration in the Economics Literature', *The Review of Economics and Statistics*, pp. 740-746

Dahlman Carl, Anuja, Utz (2005) 'India and the Knowledge Economy', *World Bank,* Washington, D.C

Ellison, Glenn (2002) 'The Slowdown of the Economics Publishing Process', *Journal of Political Economy*, vol.105, pp. 947-993

Frey, Bruno, S. (2003) 'Publishing as Prostitution? – Choosing between one's own ideas and academic success', *Public Choice*, vol. 116, pp 205-223

Gans, Joshua S and G. B. Shepherd (1994) 'How are the Mighty Fallen: Rejected Classic Articles by Leading Economists', *Journal of Economic Perspectives*, vol.8, pp.165-179

Goyal, Sanjeev; Van der Leij, Macro Juri and Moraga-Gonzalez, Jose Luis (2006) 'Economics: an emerging small world', *Journal of Political economy*, vol. 114, pp 403-412

Granovetter, M. (1985) Economic action and social structure: the problem of embeddedness, *The American Journal of Sociology* 91(3) pp 481-510

Guha, Ramachandra (2004), Passing EPW and the thinking Indian, February, accessed from *www.himalmag.com/2004/february/passing.htm*

Hodgson, G. M. and Rothman, Harry (1999) 'The Editors and Authors of Economics Journals: A Case of Institutional Oligopoly', *Economic Journal*, pp F165-F186

Hudson, John (1996) 'Trends in Multi-Authored Papers in Economics', *Journal of Economic Perspective*, vol.10: 3, pp. 153-158

Klamer, A. and Van Dalen, H.P. (2002) 'Attention and the art of scientific publishing', *Journal of Economic Methodology, 9:3, 289-315*

Lotka, A.J. (1926) 'The Frequency distribution of scientific productivity', *Journal of the Washington Academy of science* 16: 317-23

Lovell, Michael, C. (1973) 'The Production of Economic Literature: An Interpretation', *Journal of Economic Literature*, vol. 11, pp. 27-55

Merton, R.K. (1968) 'The Matthew effect in Science', *Science* 159: 6-63

Niles, N. O. and Haborak, G. E. (1971) 'Calculus with Analytic Geometry', Prentice Hall, New Jersey'

Romer, Paul M. (1990) 'Endogenous Technological Change', *Journal of Political Economy,* vol. 98: 5, part 2, pp. s71-s102

Scott, John (2000) 'Social network analysis', a handbook, second edition, *sage publications, London*

Wasserman, Stanley and Faust, Katherine (1994) Social Network Analysis: Methods and Applications, *Cambridge University press*