DISTRIBUTION AND USE OF KNOWLEDGE UNDER THE "LAWS OF THE WEB"

JOSEF FALKINGER

CESIFO WORKING PAPER NO. 2154
CATEGORY 9: INDUSTRIAL ORGANISATION
DECEMBER 2007

CESifo *Working Paper No. 2154*

DISTRIBUTION AND USE OF KNOWLEDGE UNDER THE "LAWS OF THE WEB"

Abstract

Empirical evidence shows that the perception of information is strongly concentrated in those environments in which a mass of producers and users of knowledge interact through a distribution medium. This paper considers the consequences of this fact for economic equilibrium analysis. In particular, it examines how the ranking schemes applied by the distribution technology affect the use of knowledge, and it then describes the characteristics of an optimal ranking scheme. The analysis is carried out using a model in which agents' productivity is based on the stock of knowledge used. The value of a piece of information is assessed in terms of its contribution to productivity.

JEL Code: D83, L86, B41.

Keywords: global rankings, information and internet services, limited attention, diversity, knowledge society.

Josef Falkinger
University of Zurich
Socioeconomic Institute
Zürichbergstr. 14
8032 Zurich
Switzerland
josef.falkinger@wwi.uzh.ch

21 November 2007
I wish to thank Timo Boppart, Hartmut Egger and Michel Habib for their very valuable comments.

1 Introduction

It has often been observed that one of the handicaps of economics in comparison to science is the lack of empirical laws, that is, robust regularities that restrict the possible outcomes of interactions between individuals. Now it seems that computer science can provide us with such regularities in the form of "the laws of the web"(Huberman [2001]). These newly discovered laws tell us the following: (i) No matter how huge the supply of different pieces of information, user perception is still concentrated on only a relatively small number of documents and sites. (ii) The selection of the set of items on which user attention is focused is significantly affected by the ranking algorithms used by search engines (e.g., PageRank).[1] Obviously, these laws are not confined to World Wide Web communication. They can also be observed in other environments in which a rich supply of information is distributed to a broad range of users. For instance, the distribution of knowledge through journals in international scientific communities displays similar regularities. Citations are heavily concentrated, reader attention focuses on highly ranked journals, and standardized ranking procedures are used to assign priorities.[2] The aim of this paper is to show (i) the consequences of this limitation of perception for economic equilibrium analysis, and (ii) how the design of the distribution system affects application of knowledge and economic performance in an economy in which

---

[1] See Huberman [2001] for a broad overview and potential explanations of power laws in the Internet. Lempel and Moran [2003] summarize evidence about the influence of the order of presentation of information on browsing patterns. Cho and Roy [2004] provide an empirical analysis of the impact of search engines on page popularity.

[2] See Klamer and van Dalen [2002] for evidence.

the productivity of agents is based on the stock of knowledge used.

The model presented in this paper accounts for the observed empirical regularities in the following stylized way. A research sector supplies knowledge in the form of documents published in different information sources. The documents are brought to user attention by means of a distribution system. The probability that a document from a given information source is shown depends on the design of this system. Documents from the same source are treated identically. Users can process a limited number of documents. They allocate this capacity to the different sources in accordance with the visibility of documents implied by the distribution system. The distribution and perception of knowledge feeds back into the supply of knowledge by the research sector: Only those sources of information that attract sufficient user attention survive. The supply and distribution of knowledge is complemented by an application sector in which the productivity of users depends on the set of information sources they use. The model is static. There is no accumulation of knowledge over time.

The fact that user perception of information is influenced by the ranking assigned to different sources in the distribution process naturally leads to the question as to what is an appropriate ranking. For instance, in the context of scientific publications, we are used to asking ourselves what is an adequate weight of core journals relative to field journals in scientific ratings. Likewise, the World Wide Web community is concerned about the influence of the algorithms used by search engines and is looking for ways to improve the ranking systems in operation. For instance, Cho et al. [2005] and Huberman and Wu [2006] propose mechanisms that aim at breaking popularity traps by increasing the chance that new items have of appearing in the top ranks.

Obviously, the fact that information use is affected by the ranking of information in the distribution system also has important economic consequences, in particular for the question as to how to achieve an efficient allocation of economic resources. This paper accounts for the economic consequences by considering information as a productivity-enhancing input. More specifically, I use a nested CES function for modeling user productivity as a function of the diversity and the quality of the information sources used. A distribution technology is optimal if it assigns ranks to different types of information sources in a way that leads to maximal productivity. Information sources can differ in two respects. They may be vertically differentiated in the sense that different sources provide documents of different quality or different degree of specialization. Or they may be horizontally differentiated in the sense that documents from different sources belong to different fields but are equal with respect to quality and specificity. In order to capture these two aspects, I distinguish two types of sources, A and B, where B sources are horizontally differentiated.

The fact that user attention is concentrated on a relatively small number of documents underpins the general point made by Herbert Simon [1971] that "in an information-rich economy there is a poverty of attention" (p. 40). Thus, this paper is related to the literature on limited attention (see, e.g., Kahneman [1973], Sims [2003], Gabaix et al. [2006], Falkinger [2007a, 2007b], and the references cited in these works). However, this literature focuses mainly on user behavior. Exceptions are Falkinger [2007a and 2007b], which look at competition for attention and the equilibrium supply of information. Compared to these works, the innovative contribution of this paper is an economic analysis of the role of priority settings in the distribution of information, in particular, the impact of rankings of information sources on the effective stock of knowledge of an economy.

The paper is structured as follows. The next section presents details of the model. Section 3 describes the optimal distribution system, and Section 4 discusses implications for research policy. Section 5 concludes the paper.

2 The Model

In order to investigate how the distribution of information affects economic performance, I consider an economy consisting of three sectors. One sector produces knowledge and supplies it in the form of documents. The next sector distributes these documents. The third sector consists of users who apply the information they have received and build up their stock of knowledge. This stock of knowledge determines their productivity.

2.1 Supply of knowledge

Researchers describe their findings in documents. The documents appear in sources. The sources are classified. To be concrete, I consider two types of sources. There is a mass 1 of A sources. Each A source provides n_A documents of quality q_A. In addition, there is a set $\mathbf{I} = [0, I]$, $I > 0$, of B sources. For simplicity's sake, it is assumed that B sources are identical with respect to the number and the quality of the documents. Each B source provides n_B documents of quality q_B. However, the documents appearing in the different sources come from different backgrounds ("fields"). Thus, I is a measure of the diversity of the knowledge supplied, whereas the distinction between A and B sources accounts for other aspects of heterogeneity, for instance, quality differentials or general interest topics vs. specialized field knowledge. The analysis is consistent with the interpretation that documents from A sources satisfy higher quality standards than documents from B sources. Alternatively, whatever the quality differential, we may think of A sources as outlets for core knowledge which is of interest to all users, regardless of whatever other more specific information is processed. Quality and number of documents provided in the different sources are the characteristics of the research sector. Diversity I and visibility of A sources relative to B sources depend on the distribution technology.

2.2 Distribution of knowledge

There is a user population of mass 1 looking for information. The users have access to a medium that shows them documents from the $1 + I$ sources of the research sector. All users see the same set of documents. The ranking of the distribution system is modeled as follows. Documents from A sources are shown to the users with probability r_A, documents from B sources with probability r_B. Thus, the number of documents that are visible to a user is given by $r_A n_A + I r_B n_B \equiv \tau$. Each user has the capacity to process $\overline{\tau}$ documents. If $\tau < \overline{\tau}$, all documents will be processed. This would mean that we are in an information-poor economy – for instance, in a small science community with a local distribution technology, where each member may be able to study any received document. Under global distribution technologies like the World Wide Web, there is an abundance of available documents. Therefore, $\tau > \overline{\tau}$ is assumed for the remainder of the analysis. Users allocate their capacity $\overline{\tau}$ according to the relative visibility of information sources. This means that the probability that a user will process a document from a source of type i is given by $m_i = \frac{r_i n_i}{\tau}\,\overline{\tau}$, $i \in \{A, B\}$. This can be rewritten in the form:

$$m_i = \frac{r_i n_i}{r_B n_B} \frac{\overline{\tau}}{z + I}, \qquad z \equiv \frac{r_A n_A}{r_B n_B}, \tag{1}$$

where z reflects the relative visibility of A sources compared to a B source. Since the size of the user population is normalized to one, m_i is also the impact of a source of level i in terms of downloads. I assume that an information source is only viable if it reaches some minimal impact $\overline{m} > 0$. That means that the relative visibility of the A sources and the diversity of the B sources are restricted by the following conditions:[3]

$$z/I \geq \frac{1}{\overline{\tau}/\overline{m} - 1} \quad \text{and} \quad z + I \leq \frac{\overline{\tau}}{\overline{m}}. \tag{2}$$

(Obviously, $\overline{m} < \overline{\tau}$ is required, for otherwise no A source would ever attract sufficient attention to survive.) The variables z and I depend on the design of the distribution system. For given characteristics of the research sector, the relative visibility z of A documents compared to B documents is determined by the probability differential r_A/r_B. This differential is set by the algorithm that controls the presentation of information from different sources to users. The algorithm also controls the diversity of sources visible to the users.

---

[3] According to (1), $m_A = z\overline{\tau}/(z + I)$ and $m_B = \overline{\tau}/(z + I)$. Thus, $m_A \geq \overline{m}$ and $m_B \geq \overline{m}$ are equivalent to the first and second inequalities in (2), respectively.

2.3 Application of knowledge

Users combine the information obtained from the various documents to form their stock of knowledge. The productivity of the user population depends on the quality and the diversity of the documents that have been processed. Let user productivity be denoted by Q, where Q is produced by combining the information inputs from the $1 + I$ sources. In efficiency units, the information input from an A source is $q_A m_A$. The input from a B source is $q_B m_B$. To model the relative importance of the different inputs and their substitutability in producing Q, I use the following nested CES production function:

$$Q = \left[b_A (q_A m_A)^{1-\varphi} + b_B D^{1-\varphi} \right]^{\frac{1}{1-\varphi}} \tag{3}$$

$$D = \left[\int_0^I (q_B m_B)^{1/\rho} \right]^{\rho} = I^{\rho} q_B \, m_B, \tag{4}$$

where $\varphi > 0, \varphi \neq 1$ and $\rho > 1$. Parameter ρ is equal to the elasticity of the knowledge gained from B sources with respect to the diversity of B sources. Thus, ρ reflects the value of diversity. $1/\varphi$ is the elasticity of substitution between information received from A sources and information received from a B source. If φ is very high, then A-level information cannot be substituted by B-level information. If φ approaches zero, then A and B sources are exchangeable. Finally, b_A and b_B allow for differences in the importance of information from A and B sources, respectively.

For $\varphi \to 1$, the technology (3) corresponds to a (generalized) Cobb-Douglas function $Q = (q_A m_A)^{b_A} D^{b_B}$.[4] It is worth noting that this case can also be interpreted as a reduced form of the following two-stage process of knowledge-based production. Q

---

[4] The results derived also apply to this case.

requires the integration of specialized field knowledge. The B sources supply field knowledge in I different fields and together provide the stock of knowledge D as given in (4). Output Q depends both on the size of D and on the users' ability to exploit D. This ability depends in turn on the core knowledge provided by the A sources. This gives

$$Q = q_A m_A D, \tag{5}$$

which is an instance of the generalized Cobb-Douglas function.

3 Design of optimal information distribution

Substituting (1) for m_i into (3), we can express user productivity as

$$Q = \frac{\overline{\tau}}{z+I}\left[b_A(q_A z)^{1-\varphi} + b_B(q_B I^{\rho})^{1-\varphi}\right]^{\frac{1}{1-\varphi}}. \tag{6}$$

The square-bracketed term shows how visibility (z) of A documents and diversity (I) of B documents add to productivity. The factor $\frac{\overline{\tau}}{z+I}$ reflects the conflict between the visibility of A-level items and the promotion of diversity.[5] The variables z and I are determined by the way in which the distribution technology presents information sources to users. Under an optimal distribution technology, z and I maximize user performance. Thus, we have to solve

$$\max_{z,I} Q \quad \text{subject to} \quad (2). \tag{7}$$

---

[5] In an information-poor environment, where users are able to pay attention to any document they receive, this factor boils down to $r_B n_B$. Since both the number of documents and their range of diffusion are small, promotion of documents from different information sources does not generate competition for user attention.

The following proposition characterizes the optimal program:

Proposition 1. *(Optimal information distribution). Assume* $r_B n_B > \overline{m}$*.*

(i) A maximizer of (7) *satisfies*

$$z + I = \overline{\tau}/\overline{m} \tag{8}$$

and $z \geq 1, I \leq \overline{\tau}/\overline{m} - 1 \equiv \overline{I}$*.*

Optimal diversity is given by $I^* = min\left\{\overline{I}, I^0\right\}$*, where* I^0 *is a function of* $s \equiv \overline{\tau}/\overline{m}$ *and* $q \equiv q_A/q_B$*. Optimal visibility is given by* $z^* = max\left\{1, z^0\right\}, z^0 \equiv \overline{\tau}/\overline{m} - I^0$*.* I^0 *and* z^0 *have the following properties:*

(ii) If $\rho(1-\varphi) < 1$*, then*

$$\partial I^0/\partial s > 0, \quad \partial z^0/\partial s > 0. \tag{9}$$

(iii) If $\rho(1-\varphi) \geq 1$*, an optimal program with* $I > 0$ *may not exist. If it exists, then*

$$\partial I^0/\partial s > 0, \quad \partial z^0/\partial s \leq 0. \tag{10}$$

$\partial z^0/\partial s = 0$ *applies if* $\rho(1-\varphi) = 1$*.*

(iv) Finally,

$$\begin{aligned} \partial I^0/\partial q > 0, \quad & \partial z^0/\partial q < 0 \;\; if \;\; \varphi > 1, \\ \partial I^0/\partial q < 0, \quad & \partial z^0/\partial q > 0 \;\; if \;\; \varphi < 1. \end{aligned} \tag{11}$$

For $\varphi = 1, I^0$ *and* z^0 *are independent of* q*.*

Proof: See Appendix.

The assumption of $r_B n_B > \overline{m}$ requires that the distribution technology promotes documents from B sources sufficiently so as to be viable. This guarantees that users are exposed to rich information under an optimal program. Information exposure is given by $\tau = r_A n_A + r_B n_B I = r_B n_B(z + I)$. In view of (8), $r_B n_B > \overline{m}$ implies $\tau > \overline{\tau}$. In other words, there is crowding of information and user attention focuses on a subset of the documents that have been presented. Part (i) of the proposition shows that under an optimal program, there is a conflict between visibility of A documents and diversity. Accounting for $m_A \geq \overline{m}$ in the design of the optimal distribution system means that A documents are considered to be essential. $z \geq 1$ guarantees that A sources are viable.[6] This restricts the space left for the diversity of B sources. The feasible set is defined by the ratio of the user capacity to the minimal impact requirement. Thus, s is a measure for the size of the perception space. The space expands if the focus of user attention ($\overline{\tau}$) is extended or if the survival constraint ($\overline{m}$) for information sources is reduced, that is, fewer downloads are required to stay in business.

Part (ii) deals with the case of diversity of information sources having relatively little importance for user productivity (i.e., ρ is low), or of A sources not being easily substitutable by diversity of B sources (i.e., $1/\varphi$ is low). In particular, Part (ii) applies if $\varphi \geq 1$, which means that the elasticity of substitution between information received from an A source and information received from a B source is equal to or less than one. The conditions in (9) show that in the case of an interior solution, any room for additional information sources should be used to increase both the diversity of B sources and the visibility of A sources.

Remark: Recall that $\varphi \to 1$ corresponds to the case of a generalized Cobb-Douglas production function for Q. In particular, this covers the case of knowledge produc-

---

[6] According to (2), $m_A \geq \overline{m}$ is equivalent to $z(\overline{\tau}/\overline{m} - 1) \geq I$. In view of (8), this inequality reduces to $z \geq 1$.

tion described by (5). For $\varphi \to 1$, the optimal program can be explicitly calculated from condition (A7) in the proof. We have

$$I^0 = \frac{\overline{\tau}\rho}{\overline{m}(b+\rho)}, \quad z^0 = \frac{\overline{\tau}b}{\overline{m}(b+\rho)}. \tag{12}$$

If $\rho > b(\frac{\overline{\tau}}{\overline{m}} - 1)$, then $I^0 > \overline{I}$, $z^0 < 1$, so that the optimal program is $I = \frac{\overline{\tau}}{\overline{m}} - 1, z = 1$ rather than (12).

Part (iii) shows that the space for additional information sources should be used differently if diversity has relatively substantial importance, or if information from A sources is easily substitutable by information from B sources. In this case, according to (10), any additional space should be used to increase the diversity of B sources rather than the visibility of the documents from A sources. However, no optimal program with a positive mass of B sources may exist in the first place; this would be the case, for example, if the relative importance (b) or the quality gap (q) of information from A sources over information from B sources were high.[7] Finally, the impact of a change in the quality differential (q) of information from A sources over information from B sources is shown by part (iv) of the proposition. Whether or not a higher quality gap should lead to greater promotion of documents from A sources and lower diversity of B sources depends on the elasticity of substitution between the two inputs for user productivity.

---

[7] As shown in the proof, sufficient conditions for $dQ/dI < 0$ at any I are: $\rho(1-\varphi) > 1$ and $bq^{1-\varphi}(\overline{\tau}/\overline{m})^{\rho(1-\varphi)-1} \geq \rho(\overline{\tau}/\overline{m})^{\varphi}$. The left-hand side of the inequality increases with b, but also with q, since $\rho(1-\varphi) > 1$ implies that $\varphi < 1$. In this case, it would be optimal to bring only A documents to users' attention.

4 Discussion

In the analysis presented here, user value is generated by the interaction between the research sector and the design of the distribution system that brings knowledge from the research sector to the attention of the users. The research sector is characterized by the number and quality of research documents produced, n_i and q_i, respectively, $i \in \{A, B\}$. The essential parameter controlled by the distribution system is the priority $r \equiv r_A/r_B$ attributed to documents from A sources compared to documents from B sources. A high priority means that documents from A sources have a greater chance than documents from B sources of showing up in the list of items on which user attention focuses. By controlling r, the distribution system controls relative visibility and diversity of information sources. This interaction between production and distribution of knowledge leads to two questions. First, what are the implications for the research sector if the distribution of research to users is based on priority r? Second, how should the design of the distribution system react to given characteristics of the research sector?

4.1 How distribution design affects the impact of research

The crucial variable which is affected by the design of the distribution system is impact. According to (1), the impacts of the various research sources are $m_A = \frac{z\overline{\tau}}{z+I}$ and $m_B = \frac{\overline{\tau}}{z+I}$. For any distribution design satisfying the restrictions $z + I = \overline{\tau}/\overline{m}$ and $z \geq 1$, in particular for the optimal design characterized by Proposition 1, we have

$$m_A = r\frac{n_A}{n_B}\,\overline{m}, \quad m_B = \overline{m}, \tag{13}$$

where $z \geq 1$ requires $rn_A/n_B \geq 1$.

Suppose first that parameter r of the distribution technology is fixed. Then the impact of A sources increases if more documents are published in these sources. As a consequence, their visibility is increased and B sources are crowded out. B sources could counteract by increasing their number of publications. This would increase the diversity of B sources but would not strengthen the impact of any single source. From the point of view of an individual researcher, the equations in (13) confirm that placing documents in A sources leads to a stronger reputation if individual reputation is measured by the impact of the publication outlet. This remains true if individual reputation is measured by the impact of published documents, m_i/n_i, as long as the distribution design gives priority to A sources ($r > 1$).

Things are different if the distribution system reacts to changes in the research sector. If parameter r of the distribution technology is optimally adjusted, then $m_A = \frac{z^*\overline{\tau}}{z^*+I^*}$ and $m_B = \frac{\overline{\tau}}{z^*+I^*}$, which – in view of Part (i) of Proposition 1 – reduces to

$$m_A = z^*\overline{m}, \quad m_B = \overline{m}. \tag{14}$$

Hence, any attempt by the research sector to affect the impact of research by means of the number of published documents is offset by the distribution system. The only variable that is under the control of the research sector and may affect (through z^*) the impact of research on users is the quality of the documents published. For instance, A sources could raise the quality gap over B sources by applying stricter quality control in the selection of published documents. However, Part (iv) of Proposition 1 shows that this does not necessarily improve the visibility of A sources.

Whether or not stricter quality control pays off in terms of impact depends on the substitutability between the different types of information sources for user productivity. In the case of $\varphi \to 1$, resulting for instance under the two-stage process of knowledge application described by (5), the quality gap plays no role in visibility (see Equation (11)). For the individual researcher, the conclusion that it is better to place a document in an A source rather than in a B source remains generally valid under an optimally designed distribution technology – if individual reputation is measured by the impact of the publication source. This is a consequence of the fact that A sources are considered to be essential ($z \geq 1$). More precisely, there is an incentive to publish in an A source as long as $z^* > 1$. If individual reputation is based on the impact of published documents, m_i/n_i, publication in an A source is attractive if $n_A/n_B < z^*$.

A useful example for illustrating the policy implications of these results is the recent debate among economists as to how the European science community can improve its position relative to the United States. One central aspect concerns the "acceptance of publications-*cum*-citations in international journals as a measure of research performance with the implication that incentives be provided on that basis to individuals and departments" (Drèze and Estevan [2007], p. 289). Now, one thing which the above analysis tells us is that individuals and single departments will indeed try harder to publish in A journals if incentives are based on impact. However, a second conclusion is that sooner or later journal rankings may be adjusted when a large population strengthens its efforts to get distributed through an A source. Why and how is discussed in the following subsection.

4.2 Which sources of knowledge should be favored by the distribution technology?

Should distribution technologies give priority to knowledge from A sources when presenting information to their users? Parameter r expresses the weight assigned to A sources relative to B sources. For given characteristics of the research sector, the optimal weight is

$$r = (n_B/n_A)z^*. \tag{15}$$

The optimal adjustment to variations in the number of documents published by a source was already addressed in the previous section. Since z^* does not depend on the number of documents supplied by the different sources, one definite answer follows immediately from Equation (15). An optimal distribution design should adjust priorities in favor of B sources if the number of A documents increases relative to the number of B documents. This would happen, for instance, if the fact that more people target their efforts towards publishing in A sources led to relatively more A publications than B publications in the aggregate. Obviously, this is not the only possibility. Another consequence could be that the quality differential, q_A/q_B, would change. To see the implications of such a change, we have to look at the role of z^*.

According to Proposition 1, z^* is a function of the quality differential q and of the size of the perception space s defined by the users' attention capacity and the minimal impact required for a viable source. Combining the results of Proposition 1 with Equation (15), we see that the impact of both s and q on the optimal ranking of sources is ambiguous. It is not true that, generally, the relative weights assigned to different sources by the distribution system should reflect their quality differential. In particular, in the event that productivity results from integrating knowledge from various fields (see (5)), the presentation of information to the users should not be biased in favor of top quality. The reason is that such a bias has costs in the form of a less diverse set of information sources from which users draw their knowledge. The second determinant for the optimal ranking is s, the size of the perception space. According to Proposition 1, an increase in the users' perception capacity, or an environment that allows sources to survive despite their low impact, should be used to present a larger variety of sources to the users. However, whether or not this should be accompanied by a change in the relative ranking of sources cannot be answered unequivocally. For $\varphi \to 1$, the answer is that the ranking should be invariant with respect to changes in the diversity of B sources resulting from variations in the scale of viable information sources. In this respect, the American Economic Association's plan to launch new field journals will be an interesting experiment. Will this distract attention from other field journals or will it lead to an adjustment of the relative weight of general-interest journals in international rankings?

4.3 What improves user productivity?

Under optimal distribution of information, user productivity is given by Function (6) evaluated at the optimal distribution design z^*, I^*, where $z^* + I^* = \overline{\tau}/\overline{m}$ and $z^* \geq 1$. Let Q^* denote the resulting productivity. Applying the envelope theorem, we obtain for $i \in \{A, B\}$

$$\frac{\partial Q^*}{\partial q_i} > 0, \quad \frac{\partial Q^*}{\partial n_i} = 0, \tag{16}$$

and

$$\frac{\partial Q^*}{\partial s} > 0, \quad s \equiv \overline{\tau}/\overline{m}. \tag{17}$$

If the information technology is adjusted in an optimal way, then raising the quality of research is always good for user productivity. Moreover, and maybe less obviously, increasing the number of publications does not help. This is an implication of the "laws of the web". User attention concentrates on a small subset of the available pieces of information. Therefore, user productivity cannot be increased by supplying more items unless the quality of information is improved. Any increase in the aggregate number of supplied documents is offset by a lower probability that a particular document will be downloaded by a user. In an information-poor environment, this would be different. If $\tau < \overline{\tau}$, any published document is used. In this case, an increase in the number of publications would increase user productivity.

Apart from the quality of research, the size of the perception space also matters for user productivity. For a given user capacity of perception, this size expands if the impact requirement for information sources is relaxed. In practice, this would mean that in the research sector, information sources with low impact are kept alive. We know from the preceding discussion that the consequences for the optimal distribution design are ambiguous. However, according to Proposition 1, diversity definitely increases under the optimal distribution design. And so too does user productivity, according to (17).

5 Conclusion

Kant, when he asked “What can we know?”, came to the conclusion that all our perception is framed by categories. This abstract idea has occasionally been illustrated in more practical terms, for example the observation that what a fisherman catches depends on the size of the mesh in his fishing net. Today the net that allows everyone in the world to visit the philosopher without going to Königsberg has its own laws. Users see what is brought to their attention. The purpose of this paper was to point out some important implications of this fact.

Economic productivity is a function of the stock of knowledge. Knowledge flows from a rich set of sources and is brought to the user by means of a distribution technology based on a ranking scheme. The essential parameters in the design of this technology are the priorities given to different types of sources. They determine the subset of documents on which user attention is concentrated and thus the effective knowledge base for productivity. An optimal design has to set priorities in a way that maximizes the value of information brought to the users' attention, where the value of information is measured in terms of its contribution to productivity. This paper showed how the optimal distribution design has to account for both the quality and the diversity of information. The idea of giving maximal visibility to sources that are accepted as essential – like core journals or top-quality web sites – is generally not the best one. It leads to suboptimal diversity by crowding out horizontally differentiated sources of knowledge.

The analysis presented here has the following implications for the research sector. For a given ranking scheme used by the distribution technology, individual publishers can strengthen their impact on users by increasing the number of published documents, and research communities can intensify their impact by targeting their efforts at more visible outlets. As a consequence, other sources of knowledge will have less impact on users and the mix of applied information will change. In order to reestablish the productivity-maximizing mix, the priorities given to the different sources by the distribution system have to be adjusted. Under an optimal ranking scheme, the research sector cannot influence the impact on users by changing the number of publications. The reason is that the knowledge actually applied by the users is only a subset of the supplied information. Which subset is used depends on the design of the distribution technology. Finally, the mere number of documents provided by a given type of information source is also not essential for the users' productivity. What counts is the quality of the information and the size of the perception space. This size is determined by the users' capacity to process information in interaction with the minimal impact required from an information source to participate in knowledge production.

These conclusions are based on the assumption that information is abundant – an assumption motivated by the “laws of the web”, which indicate that users concentrate on a narrow subset of items. Obviously, things are different if the information supply is scarce. Then users can process each piece of information that has been produced.

A Appendix: Proof of Proposition 1

The Lagrangian for *max* Q *s.t.* (2) is

$$\mathscr{L} = Q + \lambda\left(z - \tfrac{I}{\overline{\tau}/\overline{m}-1}\right) + \mu\left(\tfrac{\overline{\tau}}{\overline{m}} - z - I\right),$$

where Q is given by (6) and $\lambda, \mu \geq 0$. The first-order conditions are:

$$Q_z + \lambda - \mu = 0 \quad \text{and} \quad Q_I - \frac{\lambda}{\overline{\tau}/\overline{m} - 1} - \mu = 0, \tag{A1}$$

where

$$Q_z = Q\left[\frac{-1}{z+I} + \frac{b_A(q_A z)^{1-\varphi}/z}{b_A(q_A z)^{1-\varphi} + b_B(q_B I^\rho)^{1-\varphi}}\right], \tag{A2}$$

$$Q_I = Q\left[\frac{-1}{z+I} + \frac{\rho b_B(q_B I^\rho)^{1-\varphi}/I}{b_A(q_A z)^{1-\varphi} + b_B(q_B I^\rho)^{1-\varphi}}\right]. \tag{A3}$$

(Subscript notation is used for partial derivatives.)

Proposition 1 is proved in three steps. First, I show that no interior maximizer exists, i.e., $z + I = \overline{\tau}/\overline{m}$ or $z(\frac{\overline{\tau}}{\overline{m}} - 1) = I$. Second, I show that (z, I) satisfying $z+I < \overline{\tau}/\overline{m}$ and $z(\overline{\tau}/\overline{m}-1) = I$ cannot be a maximizer. These two steps prove Part (i) of the proposition. Step 3 characterizes the maximizer subject to the restriction $z + I = \overline{\tau}/\overline{m}$.

Step 1. Defining $x \equiv \frac{b_B}{b_A}\left(\frac{q_B I^\rho}{q_A z}\right)^{1-\varphi}$, we obtain from (A2) and (A3) the conditions

$$Q_z \gtreqless 0 \quad \text{if} \quad I \gtreqless zx, \tag{A4}$$

$$Q_I \gtreqless 0 \quad \text{if} \quad zx \gtreqless I\left(\frac{1+x}{\rho} - x\right). \tag{A5}$$

Now, for $\mu = \lambda = 0$, a necessary condition for a maximum is $Q_z = Q_I = 0$. According to (A4) and (A5), this cannot hold since $\rho > 1$. Hence, no interior solution exists.

Step 2. Suppose that $\lambda > 0$, i.e., $z\left(\frac{\overline{\tau}}{\overline{m}} - 1\right) = I$ and $z + I = z\overline{\tau}/\overline{m} = I\frac{\overline{\tau}}{\overline{\tau}-\overline{m}}$. Thus, (6) can be written in the form

$$Q = \tfrac{\overline{\tau}-\overline{m}}{I}\left[b_A(q_A\tfrac{\overline{m}}{\overline{\tau}-\overline{m}}I)^{1-\varphi} + b_B(q_B I^{\rho})^{1-\varphi}\right]^{\frac{1}{1-\varphi}}.$$

Differentiating Q with respect to I, we have

$$\begin{aligned} Q_I &= -\tfrac{\overline{\tau}-\overline{m}}{I^2}\tfrac{QI}{\overline{\tau}-\overline{m}} + \frac{b_A\left(\frac{q_A\overline{m}}{\overline{\tau}-\overline{m}}\right)^{1-\varphi} I^{-\varphi} + b_B q_B^{1-\varphi}\rho I^{\rho(1-\varphi)-1}}{b_A\left(\frac{q_A\overline{m}}{\overline{\tau}-\overline{m}}I\right)^{1-\varphi} + b_B(q_B I^{\rho})^{1-\varphi}} + Q \\ &= \tfrac{Q}{I}\left\{-1 + \tfrac{1+\rho y}{1+y}\right\} \text{with} \quad y \equiv \tfrac{b_B}{b_A}\left[\tfrac{q_B I^{\rho-1}(\overline{\tau}-\overline{m})}{q_A\overline{m}}\right]^{1-\varphi}. \end{aligned}$$

Hence, $Q_I > 0$, so that I is extended as long as $z + I < \overline{\tau}/\overline{m}$.

Step 3. Suppose $\mu > 0$, i.e., $z + I = \overline{\tau}/\overline{m}$. Since restriction $z \geq \frac{I}{\overline{\tau}/\overline{m}-1}$ must hold as well, we have

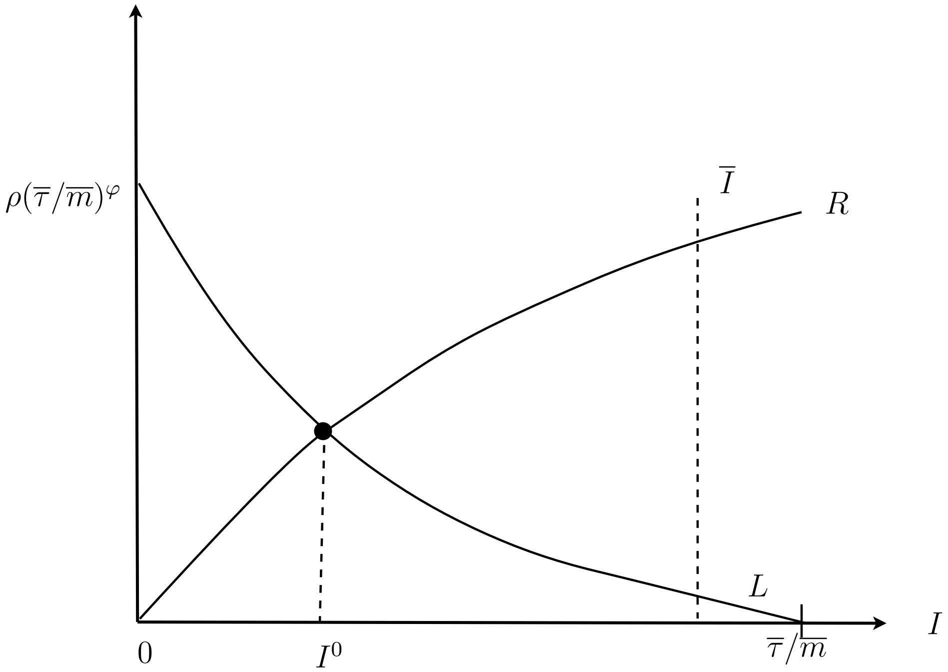
$$I \leq \overline{\tau}/\overline{m} - 1 \equiv \overline{I} \quad \text{and} \quad z \geq 1. \tag{A6}$$

The assumption $r_B n_B > \overline{m}$ implies $\tau = r_B n_B(z+I) > \overline{\tau}$, that is, we are indeed in an information-rich environment. We now look for a maximizer of Q under the restriction $z + I = \overline{\tau}/\overline{m}$. According to (6), $Q = \overline{m} b_B^{\frac{1}{1-\varphi}} q_B \left[b q^{1-\varphi}\left(\frac{\overline{\tau}}{\overline{m}} - I\right)^{1-\varphi} + I^{\rho(1-\varphi)}\right]^{\frac{1}{1-\varphi}}$. Thus, $Q_I \gtreqless 0$ if

$$\rho\left(\frac{\overline{\tau}}{\overline{m}} - I\right)^{\varphi} \gtreqless b q^{1-\varphi} I^{1-\rho(1-\varphi)}, \tag{A7}$$

where $q \equiv q_A/q_B$ and $b \equiv b_A/b_B$. Figures A and B show condition (A7) for the cases $1 > \rho(1-\varphi)$ and $1 \leq \rho(1-\varphi)$, respectively. (They correspond to Parts (ii) and (iii) of the proposition.) Curve L represents the left-hand side of (A7), and Curve R shows the right-hand side. At I^0, condition (A7) holds with equality.

Figure A: $1 > \rho(1-\varphi)$.



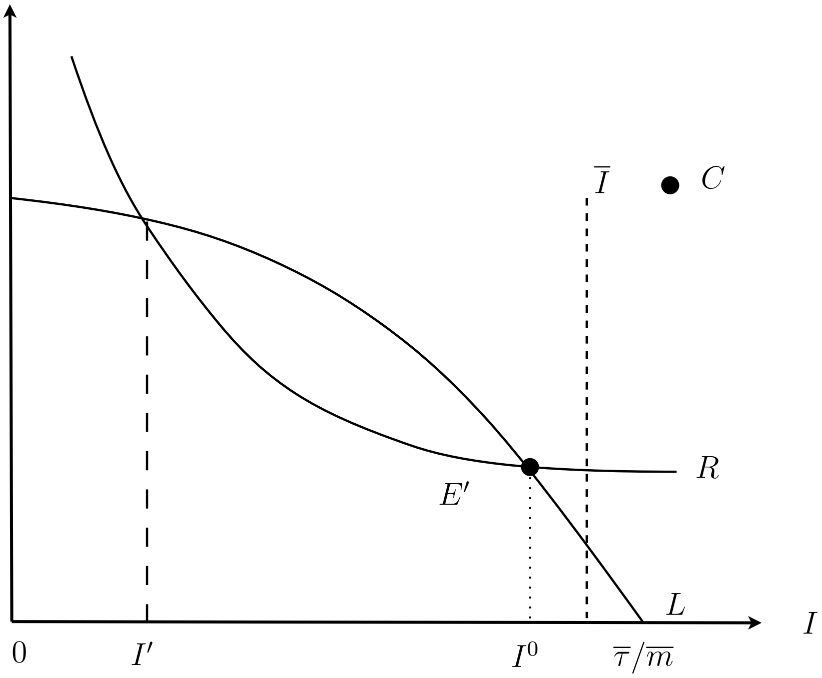
For $1-\rho(1-\varphi) > 0$, the R curve has a positive slope, starting at zero for $I = 0$. If $\varphi \geq 1$, the L curve is convex. If $\varphi < 1$, it is concave. But in any case the L curve is negatively shaped. According to (A7), $Q_I > 0$ if $I < I^0$ and $Q_I < 0$ for $I > I^0$. Thus, $min\left\{\overline{I}, I^0\right\}$ is the maximizer of our problem.

An increase in $s \equiv \overline{\tau}/\overline{m}$ shifts the L curve outwards. Thus, $\partial I^0/\partial s > 0$. An increase in b takes the R curve upwards so that $\partial I^0/\partial b < 0$. Furthermore, an increase in q rotates the R curve upwards if $\varphi < 1$ and downwards if $\varphi > 1$. Thus, $\partial I^0/\partial q^0 < 0$ if $\varphi < 1$ and $\partial I^0/\partial q > 0$ if $\varphi > 1$. For the effects of $\overline{\tau}/\overline{m}, b$, and q on z^0, we recall $z^0 = \overline{\tau}/\overline{m} - I^0$. Thus, $sign\left[\partial z^0/\partial b\right] = -sign\left[\partial I^0/\partial b\right]$ and $sign\left[\partial z^0/\partial q\right] = -sign\left[\partial I^0/\partial q\right]$. Moreover,

$$\rho(z^0)^{\varphi} = bq^{1-\varphi} + (I^0)^{1-\rho(1-\varphi)}. \tag{A8}$$

Thus, $sign\left[\partial z^0/\partial s\right] = sign\left[\partial I^0/\partial s\right], s \equiv \overline{\tau}/\overline{m}$.

Figure B: $1 \leq \rho(1-\varphi)$.



Since $\varphi < 1$, the L curve is concave in this case, whereas the R curve is negatively sloped and convex for $1-\rho(1-\varphi) < 0$ and is a horizontal line if $1 = \rho(1-\varphi)$. According to (A7), $Q_I < 0$ if $I < I'$ or $I > I^0$ and $Q_I > 0$ if $I \in (I', I^0)$. Thus, $I = 0$ or $min\left\{\overline{I}, I^0\right\}$ is the maximizer of our problem. However, the R curve may be lying to the northeast of the L curve. (For instance, if $bq^{1-\varphi}(\overline{\tau}/\overline{m})^{\rho(1-\varphi)-1} \geq \rho(\overline{\tau}/\overline{m})^{\varphi}$, then the R curve passes through point C, a point above the L curve.) In this case, $Q_I < 0$ for all I, so that no optimum with $I > 0$ exists.

An increase in $s \equiv \overline{\tau}/\overline{m}$ shifts the L curve outwards, so that $\partial I^0/\partial s > 0$. An increase in b or q moves the R curve outwards (because $\varphi < 1$), so that $\partial I^0/\partial b < 0$ and $\partial I^0/\partial q < 0$. For the impact of b and q on z^0, we again have $sign\left[\partial z^0/\partial b\right] = -sign\left[\partial I^0/\partial b\right]$ and $sign\left[\partial z^0/\partial q\right] = -sign\left[\partial I^0/\partial q\right]$ from restriction $z + I = \overline{\tau}/\overline{m}$.

And (A8) with $1 - \rho(1 - \varphi) < 0$ implies $sign\,[\partial z^0/\partial s] = -sign\,[\partial I^0/\partial s]$. For $1 = \rho(1 - \varphi)$, $\partial z^0/\partial s = 0$. QED.

References

Cho, J., and Roy, S. (2004), "Impact of Search Engines on Page Popularity", *Proceedings of the Thirteenth International World Wide Web Conference*, May, New York.

Cho, J., Roy, S., and Adams, R.E. (2005), "Page Quality: In Search of an Unbiased Web Ranking", *Proceedings of the 2005 ACM SIGMOD International Conference on Management and Data*, June, Baltimore, pp. 551–562.

Drèze, J., and Estevan, F. (2007), "Research and Higher Education in Economics: Can we Deliver the Lisbon Objectives?" *Journal of the European Economic Association*, vol. 52, April-May, pp. 271–304.

Falkinger, J. (2007a), "Attention economies", *Journal of Economic Theory*, vol. 133(1), pp. 266–294.

Falkinger, J. (2007b), "Limited Attention as a Scarce Resource in Information-Rich Economies", *Economic Journal*, forthcoming.

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006), "Costly information acquisition: experimental analysis of a boundedly rational model", *American Economic Review*, vol. 96, September, pp. 1043–1068.

Huberman, B.A. (2001), *The Laws of the Web. Patterns in the Ecology of Information*, Cambridge/MA: MIT Press.

Huberman, B.A., and Wu, F. (2006), “The economics of attention: maximizing user value in information-rich environments”, *Proceedings of The First International Workshop on Data Mining and Audience Intelligence for Advertising* (ADKDD’07). August 12, 2007, San Jose, California, USA.

Kahneman, D. (1973), *Attention and Effort*, New Jersey: Prentice Hall.

Klamer, A., and Van Dalen, H.P. (2002), “Attention and the art of scientific publishing”, *Journal of Economic Methodology*, vol. 9, pp. 289–315.

Lempel, R., and Moran, S. (2003), “Predictive Caching and Prefetching of Query Results in Search Engines”, *Proceedings of the Twelfth International World Wide Web Conference*, May, Budapest.

Sims, C.A. (2003), “Implications of rational inattention”, *Journal of Monetary Economics*, vol. 50, pp. 665–690.

Simon, H.A. (1971), “Designing organizations for an information-rich world”, in M. Greenberger (ed.), *Computers, Communications, and the Public Interest*, Baltimore: John Hopkins Press, pp. 38–52 (cited in H.A. Simon (1982), *Models of bounded rationality. Volume 2: Behavioral Economics and Business Organization*, Cambridge/MA: MIT Press).

CESifo Working Paper Series

for full list see www.cesifo-group.org/wp
(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

2092 Henrik Jacobsen Kleven, Claus Thustrup Kreiner and Emmanuel Saez, The Optimal Income Taxation of Couples as a Multi-Dimensional Screening Problem, September 2007

2093 Michael Rauber and Heinrich W. Ursprung, Life Cycle and Cohort Productivity in Economic Research: The Case of Germany, September 2007

2094 David B. Audretsch, Oliver Falck and Stephan Heblich, It’s All in Marshall: The Impact of External Economies on Regional Dynamics, September 2007

2095 Michael Binder and Christian J. Offermanns, International Investment Positions and Exchange Rate Dynamics: A Dynamic Panel Analysis, September 2007

2096 Louis N. Christofides and Amy Chen Peng, Real Wage Chronologies, September 2007

2097 Martin Kolmar and Andreas Wagener, Tax Competition with Formula Apportionment: The Interaction between Tax Base and Sharing Mechanism, September 2007

2098 Daniela Treutlein, What actually Happens to EU Directives in the Member States? – A Cross-Country Cross-Sector View on National Transposition Instruments, September 2007

2099 Emmanuel C. Mamatzakis, An Analysis of the Impact of Public Infrastructure on Productivity Performance of Mexican Industry, September 2007

2100 Gunther Schnabl and Andreas Hoffmann, Monetary Policy, Vagabonding Liquidity and Bursting Bubbles in New and Emerging Markets – An Overinvestment View, September 2007

2101 Panu Poutvaara, The Expansion of Higher Education and Time-Consistent Taxation, September 2007

2102 Marko Koethenbuerger and Ben Lockwood, Does Tax Competition Really Promote Growth?, September 2007

2103 M. Hashem Pesaran and Elisa Tosetti, Large Panels with Common Factors and Spatial Correlations, September 2007

2104 Laszlo Goerke and Marco Runkel, Tax Evasion and Competition, September 2007

2105 Scott Alan Carson, Slave Prices, Geography and Insolation in 19th Century African-American Stature, September 2007

2106 Wolfram F. Richter, Efficient Tax Policy Ranks Education Higher than Saving, October 2007

2107 Jarko Fidrmuc and Roman Horváth, Volatility of Exchange Rates in Selected New EU Members: Evidence from Daily Data, October 2007

2108 Torben M. Andersen and Michael Svarer, Flexicurity – Labour Market Performance in Denmark, October 2007

2109 Jonathan P. Thomas and Tim Worrall, Limited Commitment Models of the Labor Market, October 2007

2110 Carlos Pestana Barros, Guglielmo Maria Caporale and Luis A. Gil-Alana, Identification of Segments of European Banks with a Latent Class Frontier Model, October 2007

2111 Felicitas Nowak-Lehmann D., Sebastian Vollmer and Immaculada Martínez-Zarzoso, Competitiveness – A Comparison of China and Mexico, October 2007

2112 Mark Mink, Jan P.A.M. Jacobs and Jakob de Haan, Measuring Synchronicity and Co-movement of Business Cycles with an Application to the Euro Area, October 2007

2113 Ossip Hühnerbein and Tobias Seidel, Intra-regional Tax Competition and Economic Geography, October 2007

2114 Christian Keuschnigg, Exports, Foreign Direct Investment and the Costs of Corporate Taxation, October 2007

2115 Werner Bönte, Oliver Falck and Stephan Heblich, Demography and Innovative Entrepreneurship, October 2007

2116 Katrin Assenmacher-Wesche and M. Hashem Pesaran, Assessing Forecast Uncertainties in a VECX Model for Switzerland: An Exercise in Forecast Combination across Models and Observation Windows, October 2007

2117 Ben Lockwood, Voting, Lobbying, and the Decentralization Theorem, October 2007

2118 Andrea Ichino, Guido Schwerdt, Rudolf Winter-Ebmer and Josef Zweimüller, Too Old to Work, too Young to Retire?, October 2007

2119 Wolfgang Eggert, Tim Krieger and Volker Meier, Education, Unemployment and Migration, October 2007

2120 Stefan Napel and Mika Widgrén, The European Commission – Appointment, Preferences, and Institutional Relations, October 2007

2121 Bertil Holmlund and Martin Söderström, Estimating Income Responses to Tax Changes: A Dynamic Panel Data Approach, October 2007

2122 Doina Maria Radulescu, From Separate Accounting to Formula Apportionment: Analysis in a Dynamic Framework, October 2007

2123 Jelle Brouwer, Richard Paap and Jean-Marie Viaene, The Trade and FDI Effects of EMU Enlargement, October 2007

2124 Kurt R. Brekke, Luigi Siciliani and Odd Rune Straume, Competition and Waiting Times in Hospital Markets, October 2007

2125 Alexis Direr, Flexible Life Annuities, October 2007

2126 Johannes Becker and Clemens Fuest, Quality versus Quantity – The Composition Effect of Corporate Taxation on Foreign Direct Investment, October 2007

2127 Balázs Égert, Real Convergence, Price Level Convergence and Inflation Differentials in Europe, October 2007

2128 Marko Koethenbuerger, Revisiting the "Decentralization Theorem" – On the Role of Externalities, October 2007

2129 Axel Dreher, Silvia Marchesi and James Raymond Vreeland, The Politics of IMF Forecasts, October 2007

2130 Andreas Knabe and Ronnie Schöb, Subsidizing Extra Jobs: Promoting Employment by Taming the Unions, October 2007

2131 Michel Beine and Bertrand Candelon, Liberalization and Stock Market Co-Movement between Emerging Economies, October 2007

2132 Dieter M. Urban, FDI Technology Spillovers and Wages, October 2007

2133 Valentina Bosetti, Carlo Carraro, Emanuele Massetti and Massimo Tavoni, Optimal Energy Investment and R&D Strategies to Stabilise Greenhouse Gas Atmospheric Concentrations, October 2007

2134 David-Jan Jansen and Jakob de Haan, The Importance of Being Vigilant: Has ECB Communication Influenced Euro Area Inflation Expectations?, October 2007

2135 Oliver Falck, Heavyweights – The Impact of Large Businesses on Productivity Growth, October 2007

2136 Xavier Freixas and Bruno M. Parigi, Banking Regulation and Prompt Corrective Action, November 2007

2137 Jan K. Brueckner, Partial Fiscal Decentralization, November 2007

2138 Silvia Console Battilana, Uncovered Power: External Agenda Setting, Sophisticated Voting, and Transnational Lobbying, November 2007

2139 Alan J. Auerbach, Michael P. Devereux and Helen Simpson, Taxing Corporate Income, November 2007

2140 Lorenzo Cappellari, Paolo Ghinetti and Gilberto Turati, On Time and Money Donations, November 2007

2141 Roel Beetsma and Heikki Oksanen, Pension Systems, Ageing and the Stability and Growth Pact, November 2007

2142 Hikaru Ogawa and David E. Wildasin, Think Locally, Act Locally: Spillovers, Spillbacks, and Efficient Decentralized Policymaking, November 2007

2143 Alessandro Cigno, A Theoretical Analysis of the Effects of Legislation on Marriage, Fertility, Domestic Division of Labour, and the Education of Children, November 2007

2144 Kai A. Konrad, Mobile Tax Base as a Global Common, November 2007

2145 Ola Kvaløy and Trond E. Olsen, The Rise of Individual Performance Pay, November 2007

2146 Guglielmo Maria Caporale, Yannis Georgellis, Nicholas Tsitsianis and Ya Ping Yin, Income and Happiness across Europe: Do Reference Values Matter?, November 2007

2147 Dan Anderberg, Tax Credits, Income Support and Partnership Decisions, November 2007

2148 Andreas Irmen and Rainer Klump, Factor Substitution, Income Distribution, and Growth in a Generalized Neoclassical Model, November 2007

2149 Lorenz Blume, Jens Müller and Stefan Voigt, The Economic Effects of Direct Democracy – A First Global Assessment, November 2007

2150 Axel Dreher, Pierre-Guillaume Méon and Friedrich Schneider, The Devil is in the Shadow – Do Institutions Affect Income and Productivity or only Official Income and Official Productivity?, November 2007

2151 Valentina Bosetti, Carlo Carraro, Emanuele Massetti and Massimo Tavoni, International Energy R&D Spillovers and the Economics of Greenhouse Gas Atmospheric Stabilization, November 2007

2152 Balázs Égert and Dubravko Mihaljek, Determinants of House Prices in Central and Eastern Europe, November 2007

2153 Christa Hainz and Hendrik Hakenes, The Politician and his Banker, November 2007

2154 Josef Falkinger, Distribution and Use of Knowledge under the “Laws of the Web”, December 2007