Index Numbers and Information Networks*

Robert Ackland[†] Lingfei Wu[‡]

July 29, 2013

Abstract

This paper introduces a new approach for studying behaviour in information networks. Index number theory, which has been developed in the context of comparing income and prices across countries, regions or over time, is used to derive indexes of information consumption. In particular, we construct indexes of real attention, which account for the fact that different information consumers might have different costs (in terms of attention) of consuming the same information source. An application using clickstream data for nearly 1000 websites is provided.

Keywords: index number theory, information network, attention, WWW clickstream

1 Introduction

We define an *information network* as network that consists of two types of behaviour: information production and information consumption. Information production involves very low costs of distribution; this has been referred to as joint-consumption technology (Rosen, 1981), where the cost of providing the information does not rise proportionally with the number of people consuming the information. Examples of information producers are: website authors (including bloggers), microbloggers (e.g. users of Twitter) and producers of academic research papers.

Information consumption involves, for example, reading web pages, blogs, tweets and academic papers. The *traces* of this consumption behaviour can be collected unobtrusively, and include: hyperlinks between websites and blogsites (blogger i hyperlinking to a post by blogger j indicates that i has read the post that is being linked to), retweets (we can assume a person

*Preliminary research findings were presented as a poster at the *Workshop on Information in Networks (WIN)*, New York University Stern School of Business, 30 September - 1 October 2011. The first version of this paper was completed in February 2013 and was accepted as a full paper for *Web Science 2013* (we subsequently withdrew the paper and hence it does not appear in the proceedings).

[†]Corresponding author. Australian National University. email: robert.ackland@anu.edu.au

[‡]Baidu and Web Mining Lab, City University of Hong Kong. email: wulingfei@baidu.com

would not retweet a tweet they hadn't read), and citations in academic papers. Clicksteams showing how people move or 'surf' between websites is also a digital trace of information consumption, and this is the data source used in the present paper.

In the definition of an information network used in this paper, an actor can be both a consumer and producer of information ('prosumption') (Ritzer and Jurgenson, 2010), however in this paper, we do not model the joint behaviour of producing and consuming information, and we focus only on information consumption behaviour.

The behaviour described above does not involve market transactions, or if there are market transactions, they are not observable to the researcher. This is why we refer to the behaviour as occurring within an information network, and not an information market. However, just because we are not observing market transactions, this does not mean there are no transactions between information consumers and producers.

In particular, information consumers exchange their *attention* for information that they consume and the aim of this paper is to use economic decision theory to construct new measures of attention in information networks. Our proposed information consumption index is in fact an attention index.

2 Information network consumer theory

In this section, standard results in consumer theory are adapted to studying behaviour in an information network. A comparison is made with a 'traditional' application of consumer theory, international comparisons research, where index number theory is used to make cross-country comparisons of real GDP.

2.1 Prices and quantities in information networks

The application of index number theory to market behaviour involves data on quantities of K goods and services consumed by N consumers, and the prices paid for those goods and services. In the international comparisons literature, the consumers are countries, the quantities are per capita consumption and the prices are averages of prices within the country.

In applying index number theory to the study of information networks, we have N prosumers (e.g. bloggers, tweeters) who make a decision to consume a bundle of N types of information. So with information networks, the 'goods' that are consumed are different types information and we have data on the quantities of each type of information consumed by each consumer/producer (e.g. number of hyperlinks pointing from website i to website j, number of times tweeter i retweeted tweets authored by tweeter j, number of times author i cited author j). We therefore assume that each prosumer produces a unique type of information, but it is also possible to consider an information network with $K < N$ types of information, for example, where technology bloggers are assumed to produce one type of information, and lifestyle bloggers another type.

The bundle of information goods consumed by actor i is represented by the $N \times 1$ vector $\mathbf{q}^i$, where element q^i_j is the quantity of information type j consumed by actor i. In choosing information goods bundle $\mathbf{q}^i$, actor i faces a set of prices $\mathbf{p}^i$ (a $N \times 1$ vector).

In index number theory as applied to international comparisons, prices are measured in units of the currency relevant to the consumer, e.g. if the average American consumer pays \$3 for a kilogram of apples, the price vector for the US representative consumer includes \$3/kilogram for apples. The utility or well-being a consumer gains from consuming a bundle of goods is not directly observable, but economic theory shows that utility can equally be measured by the monetary value of the bundle of goods consumed, leading to the concept of money-metric utility.

As noted above, the prices in information networks are not market prices, but are inferred prices measured in units of *attention* that reflect the relative cost for i to consume a unit of each type of information. We assume that a unit of attention paid by consumer i has the same value as a unit of attention paid by consumer j. Thus, we assume away the problem of 'different currencies' that international comparisons research has to deal with i.e. US consumers spend dollars, while Chinese consumers spend yuan. However, crucially, we do not assume that consumers i and j 'pay' the same amount of attention to consume a unit of information from producer k. If we *did* assume that i and j pay the same price to consume information produced by k, then our attention-metric utility measure (introduced below) would be equivalent to outdegree centrality.

2.2 Attention-metric utility

In order to model behaviour in an information network as utility maximising behaviour, we need to introduce the concept of the *representative consumer*. In the context of international comparisons, see for example Ackland (2008); Ackland, Dowrick, and Freyens (2013), the representative consumer can be thought of as an imaginary person who moves from one country to another, in each each country purchasing a bundle of goods by maximising utility subject to the relative prices and income of that country. So the representative consumer purchases a bundle of goods in Australia by maximising utility subject to Australian average prices and income, and then the *same* consumer then does the same exercise in China. A central aspect of international comparisons is testing for the existence of a representative consumer (testing for *common preferences*) i.e. is it the case that world consumption patters can be explained as utility maximising behaviour of a single agent, or do we have to allow for the fact that the preferences of the Chinese representative consumer are different from that of, for example, the Australian representative consumer.

The Generalized Axiom of Revealed Preference (GARP) (Afriat, 1967; Diewert, 1973; Varian,

1982) is used to test whether a given set of demand data (prices and quantities from each country) could have been generated by a representative consumer. If we cannot reject the hypothesis the hypothesis of common preferences, then we can use the *money-metric Allen welfare index* to compare the living standards of countries i and j, which compares the minimum expenditures required to achieve utilities U^j and U^i at some reference price vector.

In the context of information networks, GARP can be analogously used to test for the existence of a representative information consumer. The *attention-metric Allen welfare index* is a straight-forward adaptation of the money-metric Allen welfare index, and is defined as the ratio of the minimum attention expenditure required to achieve utilities U^j and U^i at some reference price vector, $\mathbf{p}^r$:

$$A_r^{j:i} = \frac{e[U(\mathbf{q}^j), \mathbf{p}^r]}{e[U(\mathbf{q}^i), \mathbf{p}^r]},$$

where $U(\mathbf{q})$ is a utility function. The Allen bilateral index is dependent on the choice of reference price vector, unless preferences are homothetic (the utility function is a monotonic transformation of a function which is homogeneous of degree 1). If the price and quantity data are consistent with common homothetic preferences, we can use theorems from Afriat (1967, 1984) to determine tight bounds on the Allen bilateral index.

The Paasche and Laspeyres quantity indexes are defined, respectively:

$$Q_P^{ij} = \frac{\mathbf{p}^i.\mathbf{q}^i}{\mathbf{p}^i.\mathbf{q}^j}; \quad Q_L^{ij} = \frac{\mathbf{p}^j.\mathbf{q}^i}{\mathbf{p}^j.\mathbf{q}^j}$$

The test for common homothetic preferences requires that there exists a true multilateral index, $\mathbf{a}$, such that the ratio a_i/a_j, lies between the Paasche and Laspeyres quantity indexes for every pair of observations i and j (Afriat, 1984).[1] The Homothetic Axiom of Revealed Preference (HARP) test involves using Warshall's algorithm to construct the minimum path matrix using the matrix of bilateral Laspeyres quantity indexes as input (Varian, 1983). However satisfying the test of common homothetic preferences does not imply a unique Afriat index. Rather, there are well-defined bounds within which there is an irreducible indeterminacy resulting from the fact that we do not observe utility.

2.3 Information network prices

The application of index number theory to information networks hinges on what prices are used. Our aim is to calculate a cost or price (in terms of attention paid) *per unit* of information consumed, and in this paper we propose that these prices can be derived from the structure

[1] A multilateral index satisfies circularity. In the context of information networks this means that the real attention of i relative to j is the same whether the two are compared directly or via an intermediate third consumer k.

of the 1.5 degree egonetworks of information consumers.[2]

It has long been recognised that egonetwork structure can impact on the flow of novel information, and the important role of 'weak ties' (connections to people with whom you do not share many ties in common) spanning 'structural holes' (gaps in the social structure of communication) has been emphasised (see, e.g., Granovetter, 1973; Burt, 1992).[3]

In our simplified model of an information network, we assume that the price of consuming a unit of information from a strong tie is lower than the cost of consuming a unit of information from a weak tie. This cost is related to 'transfer problem' discussed by Hansen (1999) in the context of intra-organisational knowledge transfer–when an information consumer has searched for and identified useful information, this information must be transferred to the consumer and be made use of. Hansen (1999) highlighted the role of strong ties in facilitating transfer of complex knowledge, and Aral and Alstyne (2011, p.97) suggest that such ties means that ego is "...better able to comprehend and thus receive novel information from alter".

In particular, we calculate p^i_j as:

$$p^i_j = \min_{k \in N^i} DC^i_k + |DC^i_j - \max_{k \in N^i} DC^i_k|$$

where N^i is the number of nodes in i's 1.5 degree egonetwork, DC^i_j is total degree centrality for j in i's 1.5 degree egonetwork, and is the sum of j's indegree centrality (IDC^i_j) and outdegree centrality (IDC^i_j) in i's 1.5 degree egonetwork:

$$DC^i_j = IDC^i_j + ODC^i_j = \sum_{k \in N^i} A^i_{kj} + \sum_{k \in N^i} A^i_{jk},$$

and $\mathbf{A}^i$ is the $N^i \times N^i$ adjacency matrix for i's 1.5 degree egonetwork.

Thus in the 1.5 degree egonetwork for node i, the node with the highest total degree centrality will have the lowest price, and the node with the lowest total degree centrality will have the highest price. Since we need to have a price for all nodes in the network (even those who are not in i's 1.5 degree network), we arbitrarily set this price equal to one plus the maximum price found for that 1.5 degree egonetwork.

As an illustration of the calculation of network prices, in the egonetwork for node 3 in Figure 1 node 1 has a total degree centrality of 1 (the minimum for this egonetwork), while node 2

[2]A 1.0 degree egonetwork consists of a root node ('ego') and the nodes directly connected to ego ('alters'), but does not show the connections between the alters. A 1.5 degree egonetwork is the same as the 1.0 degree egonetwork, except alter-alter ties are also shown. Finally, a 2.0 degree egonetwork includes the nodes that connect to the alters, but do not connect to ego, i.e., 'friends-of-friends'.

[3]Aral and Alstyne (2011) provide a comprehensive summary of previous research on the impact of network structure on the flow of information. They also show that while 'structurally diverse' egonetworks–measured by, for example, network density (the number of ties as a proportion of the maximum number of ties)–are a source of innovative information, this might be out-weighed by the greater information flow across the 'strong ties' (connections with friends with whom you have many fiends in common) that exist in 'structurally cohesive' networks, leading to a 'diversity-bandwidth trade-off' with regard to information and network structure.

has a total degree centrality of 5 (the maximum of this egonetwork). The network prices are $p_1^3 = 1 + |1 - 5| = 5$ and $p_2^3 = 1 + |5 - 5| = 1$.

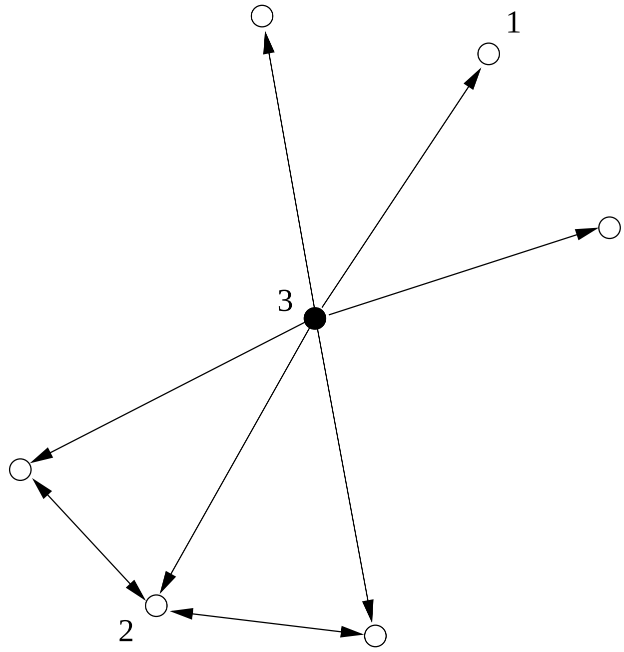

Figure 1: Example 1.5 degree egonetwork

Importantly, the price vector faced by consumer i is exogenously determined in that while the consumer makes choices over quantities of information consumed from each provider (and hence who is in his egonetwork), he does not have direct influence over how the alters in his egonetwork are connected to one another (i.e. how they consume information from one another), and it is these alter-to-alter connections that are used to derive the prices faced by consumer i. If ego *did* have some influence over the tie formation behaviour of the alters, then this would violate our assumption that the network prices (thus derived) are exogenous. For example, it might be reasonable to assume that network ties exhibit reciprocity (ego consuming information from alter increases the likelihood that alter consumes information from ego) or triadic closure (ego consuming information from both alters j and k increases the likelihood that j consumes information provided by k and/or vice-versa). But such interdependency between ties would belie our original assumption that this is an *information network*, rather than a *social network*. We assume that ties in information networks are *independent* from one another rather than interdependent, and hence our network prices are exogenous (from the perspective of ego).

For our constructed price to be acceptable for use in the construction of index number, we would expect our prices and quantities to interact in some broadly familiar way. For example, do consumers consume greater (lesser) quantities of information from a particular source when its relative price is higher (lower)? i.e. is the *own-price elasticity of demand* less than zero? The above research suggests that this should be the case, but it is an empirical question that is not addressed in this paper.

For our application of economic decision theory to information networks to be valid, do we similarly need to be able to determine that some types of information are economic *necessities* in that as expenditure rises their budget share decreases (the bread and milk of information providers), while others are economic *luxuries* (the plasma televisions of information providers) with their budget share increasing with expenditure? Again, this is not addressed in the present paper.

2.4 Multilateral indexes of real attention

We defined above the attention-metric Allen welfare index, which provides a bilateral comparison of the real attention of information consumers. Since the utility function is unobservable, the attention-metric Allen welfare index is similarly unobservable. Furthermore, unless preferences are homothetic, it is dependent on the reference price vector used in the comparison. This is problematic since it means the utility comparison will be different depending on which price vector is being used, and there is no reason to choose one price vector over another (in international comparisons research, do we compare China and Australia's real GDP using China's prices or Australia's, or some other reference price vector?).

We also summarised results that show that in the case where preferences are homothetic, a unique Allen index that is independent of prices exists (this is referred to as the 'true index'). Furthermore, tight bounds to this true index can be derived from the minimum path matrix, which is constructed from the matrix of bilateral Laspeyres indexes. Dowrick and Quiggin (1997) show that the elements of the minimum path matrix can be used to construct the lower and upper bounds of true indexes relative to the geometric mean (which they denote the Ideal Afriat Index).

Constructing multilateral true indexes of attention involves identifying subsets of information consumers who share common homothetic preferences, and then constructing the Ideal Afriat Index from elements of the minimum path matrix. This is beyond the scope of the present paper, but we can identify a potential challenge to using this approach in the context of information networks. The assumption of homotheticity is restrictive and based on our experience in international comparisons research (Ackland, Dowrick, and Freyens, 2013), the set of observations sharing common homothetic preferences is typically significantly less than the total number of observations. This makes it impossible to construct a multilateral true index for the entire dataset. It is an empirical question whether actors in information networks are found to share common homothetic preferences.

2.4.1 The EKS index

The EKS quantity index (Eltetö and Köves, 1964; Szulc, 1964) is a multilateral index commonly used in international comparisons.[4] The EKS quantity index (hereafter, the EKS index) is defined as:

$$EKS^{ij} = \prod_{n=1}^{N} \left(\frac{F^{in}}{F^{jn}} \right)^{1/N},$$

where F^{ij} is the bilateral Fisher quantity index, constructed as the geometric mean of the Paasche and Laspeyres quantity indexes defined above:

$$F^{ij} = (Q_P^{ij} Q_L^{ij})^{0.5}.$$

The main advantage of the EKS index is the fact that it is a 'superlative' index in that it is equal to the ratio of a 'flexible' utility function (a utility function that is a second-order approximation to an arbitrary twice-differentiable utility function) (Diewert, 1976, 1978)–this is also known as the property of 'exactness'. Thus while the use of the EKS index does not involve the testing of whether a set of demand data is consistent with common preferences, the fact that the EKS index is superlative means that if the demand data *were* generated by a representative consumer, then the EKS is exact for any arbitrary homogeneous quadratic utility function that could have generated the demand data.

While it is possible to generate the EKS measure of real attention in an information network, the present paper index advocates another approach for constructing multilateral indexes that involves the use of minimum spanning trees (MST).

2.4.2 Minimum spanning trees

The chaining of index numbers in the context of temporal comparisons of prices (e.g. the consumer price index used in calculating inflation) enable the construction of temporally-consistent indexes, even when the reference bundle of goods is changing over time (e.g. to reflect new consumer goods on the market). The chaining of an index involves a choice about the ordering of the observations in the chain. In the case of temporal indexes, this is straightforward since time period $t + 1$ must follow time period t in the chain.

In the case of international comparisons, however, there is no natural ordering of observations (countries), but Hill (1999) has proposed that the minimum spanning trees computed using various measures of the consumption differences between countries as the edge weights can provide an empirically-driven ordering of countries for use in a chain index.[5] Hill argues that

[4] The OECD uses the EKS method for comparisons of incomes between their member countries, and this approach is also used by the World Bank in global poverty measurement.

[5] A minimum spanning tree is an acyclic network that connects all nodes and has the smallest possible sum

the MST chaining approach for constructing multilateral indexes can improve international comparisons in two ways. First, the approach reduces the amount of data collection that needs to be undertaken by the statistical agencies of individual countries (each country would only need to collect data on prices and quantities of goods consumed by neighbouring countries in the MST). Second, the MST chaining approach improves the *characteristicity* of the multilateral index, which is the property that any multilateral comparison between observations i and j should preserve as much as possible of the variation between the two observations that is observed in the bilateral index.

In the case of the information network studied in the present paper (website clickstreams), there is huge variation in consumption behaviour of different actors. This means that a multilateral index such as the EKS index, which involves the geometric mean of all possible bilateral comparisons, is likely to have very poor characteristicity. It doesn't make sense to make a bilateral comparison between, for example, a US website and a Chinese website, when there is no overlap in the sites to which they send clicks (they consume completely different bundles of information goods).

Hill proposes several potential ways of computing the edge weight for the construction of a MST for international comparisons. The edge weight we use here is the Paasche-Laspeyres spread:

$$PLS^{ij} = \left| \ln \left(\frac{Q_L^{ij}}{Q_P^{ij}} \right) \right|$$

3 Application: WWW Clickstreams

We now apply the above framework to the study of clickstream behaviour on the WWW.

We assume that there exists a representative web user i for each of N websites. We do not test for the existence of a representative web user at the level of the website–if we were conducting revealed preference tests, it would be at the level of the clickstream network as a whole.[6]

Web user i is currently visiting website i and makes a decision as to what web content will be consumed (what websites visited) next. The clickstream flows from website i are conceptualized as the purchase by the representative web user i of a bundle of website content or information.

We assume that web users have perfect information about available website content, but are faced with prices of consuming content. These prices can be thought of as cognitive, cultural and linguistic costs associated with reading and understanding content from different websites. So, while an Australian web user knows about the content on a Chinese website, it is costly to

of edge weights.

[6]In international comparisons research the existence of a representative consumer for each country is assumed, even though econometric evidence suggests that this is not valid.

consume (relative to Australian web content) because of the requirement of using translating software, for example. Similarly, the humanities academic web user knows about the physics website but consuming that content is costly because it would involve learning mathematics. The prices are derived from the 1.5 degree egonetwork for each web user, in the manner described above.

3.1 Clickstream data

As documented in Wu and Ackland (2012), the top 1,000 websites were selected according to Google's traffic statistics in November 2010 (`http://www.google.com/adplanner/static/top1000/`). Alexa (`http://www.alexa.com`) was then used to retrieve the daily *clickstreams* between the sites (averaged over three months). The clickstream network contains 12,008 directed and weighted edges, where an edge between websites i and j indicates the percentage of global Web users who visited j immediately after visiting i. Twenty sites were dropped due to a lack of data and thus our analysis is for the remaining 980 sites. Further, as Alexa reports a maximum of ten largest inbound and outbound clickstreams for each website, the clickstream network necessarily does not include all of the clickstreams between these 980 sites. Note that our data do not allow us to know exactly how a person navigates from website i to j: navigation may occur either through a hyperlink, a search engine, or the user typing the URL into the browser (or equivalently, following a bookmark).

Nine sites were dropped because they did not have any outbound clickstreams, so the final dataset consists of 971 sites.

3.2 MST index of real attention given

The Laspeyres and Paasche quantity indexes were constructed, and the Paasche-Laspeyres spread was used as the edge weight for the MST, which is displayed in Figure 2. As discussed in Wu and Ackland (2012), the websites were classified into the following geographic-linguistic groups: Euro-American, Chinese, Japanese, Russian, Korean and Polish. The clustering of the websites in the MST clearly reflects these human-defined groups. The implication is that the multilateral index constructed from the MST will have strong characteristicity.

Next, the MST real attention index was constructed by choosing google.com as the MST root node (the choice of root node is arbitrary, but google.com was chosen as it is the website with the highest outbound click traffic) and then using chained bilateral Fisher quantity indexes to compute a real attention measure.

As an example of the calculation made, google.com has a total outbound click traffic of 45.26. One of the sites that connects to google.com in the MST is t-mobile.com. The log bilateral Fisher quantity index comparing t-mobile.com to google.com is -6.913 (real attention of t-mobile.com is 0.0995% of that of google.com), implying an MST real attention index

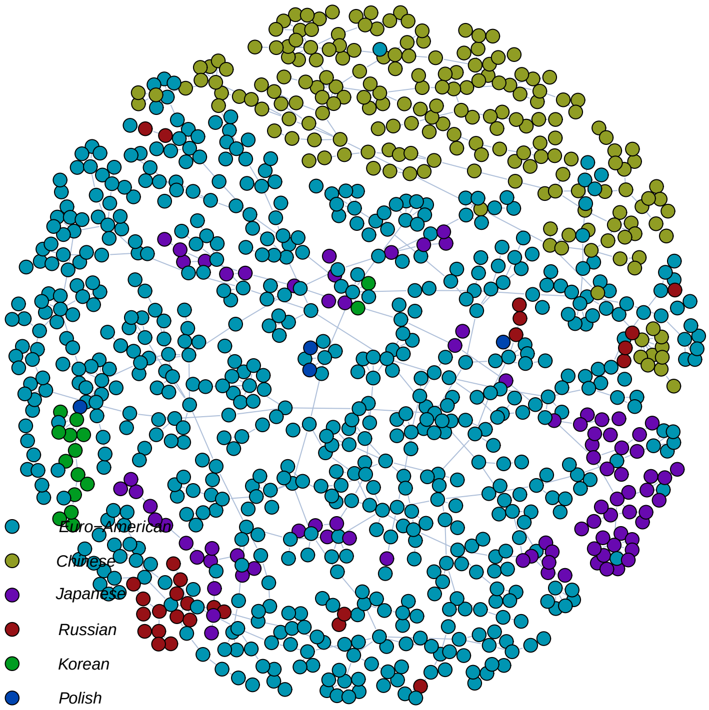

Figure 2: MST for clickstream network

(normalised against google.com) of 0.045.

Total outbound clickstreams and the MST measure of real attention are shown in columns 2 and 3 of Table 1 for the top-20 sites, where the top-20 ranking is based on total outbound clickstreams. The rankings show marked differences between the two measures. For example, jp.msn.com is ranked 9 on the basis of outbound clickstreams, but is outside the top-10 (14th) on the basis of the MST real attention measure. linkedin.com is ranked 20 on the basis of outbound clickstreams, but 26 on the basis of real attention.

3.3 Inequality of real attention given

An important aspect of international comparisons research is calculating measures of inequality of real income across countries, since information on whether global income is converging or diverging over time is important for understanding the relative success (or failure) of economic growth policies.

One of the important results of index number research is that the use of a rich country's prices to conduct a cross-country income comparison will tend to overstate poorer countries' relative income, and vice versa. This is referred to as the *Gershenkron effect*, after Gerschenkron (1951), and results in an under-estimate of the level of global income inequality.

Table 1: Information consumption (real attention), top-20 websites

Site	Q	MST	rank Q	rank MST
google.com	45.263	45.263	1	1
facebook.com	26.461	28.208	2	2
yahoo.com	15.689	19.649	3	3
live.com	11.110	15.583	4	5
baidu.com	11.026	16.292	5	4
blogger.com	10.666	10.442	6	7
youtube.com	10.622	12.002	7	6
msn.com	4.713	5.506	8	9
jp.msn.com	4.393	3.693	9	14
yahoo.co.jp	4.345	5.226	10	11
wikipedia.org	4.149	4.279	11	13
twitter.com	4.022	4.316	12	12
qq.com	3.958	7.240	13	8
alexa.com	2.743	5.264	14	10
sina.com.cn	2.059	3.372	15	15
taobao.com	2.017	3.297	16	16
amazon.com	1.894	1.818	17	20
bing.com	1.864	1.787	18	21
microsoft.com	1.555	1.434	19	24
linkedin.com	1.533	1.242	20	26
variance (N=971)	3.674	4.298		

As noted above, the sum of total outbound clickstreams is effectively a fixed-price quantity index, where the price vector consists only of ones. Table 1 shows that the variance of total outbound clickstreams (3.674) is lower than that of MST real attention (4.298). This suggests that the use of a fixed-price quantity index such as total outbound clickstreams is similarly underestimating the level of inequality of attention on the Web.

3.4 What about real attention received?

The above has focused on real attention given by websites, but what about real attention *received*? This is arguably of greater interest in this example of WWW clickstreams, since it provides a new measure of website popularity. While it is beyond the scope of the present paper to delve into this further, one way of constructing measures of real attention received would be to calculate real attention given by each website i to each other website j, and then sum up the real flows of attention to each website.

4 Conclusions

This paper recasts actors in information networks as consumers who make rational choices regarding the consumption of bundles of information goods (or equivalently, choices regarding tie formation in the information network), subject to their attention budgets and prices. This allows us to use standard tools of economic decision theory to provide new insights into behaviour within an information network. Thus, we use techniques that are typically used to study market behaviour (economic decision theory) to study behaviour in networks. Markets and networks involve different types of governance mechanisms (Demil and Lecoq, 2006), but we explore the applicability of using tools from market analysis to the study of networks.

The use of economic decision theory to analyse behaviour in information networks is a major departure from search theory, which is typically used in the economic analysis of networks. It is also significantly different to approaches that involve the application of network science to study information behaviour. The approach allows the construction of an index of information consumption, which we call a real attention index.

We construct the real attention index using clickstream data for nearly 1000 websites and find that it provides markedly different rankings of websites (on the basis of outbound clickstreams) compared with the simple summation of outbound clicks.

While this paper has focused on measuring attention *given* in information networks, future work will extend the approach to constructing measures of attention *received*, thus providing new ways of quantifying popularity or influence in information networks. Future work will also focus on conducting test of revealed preference to identify subsets of actors in information networks who share common preferences for consuming information.

References

Ackland, R. (2008): "Incorporating Minimum Subsistence Consumption into International Comparisons of Real Income," *Review of Economics and Statistics*, 90(4), 702–712.

Ackland, R., S. Dowrick, and B. Freyens (2013): "Measuring Global Poverty: Why PPP Methods Matter," *Review of Economics and Statistics*, Accepted: February 2012.

Afriat, S. (1967): "The Construction of a Utility Function from Expenditure Data," *International Economic Review*, 8(1), 67–77.

——— (1984): "The True Index," in *Demand, Equilibrium, and Trade: Essays in Honor of Ivor F Pearce*, ed. by A. Ingham. St. Martin's Press, New York.

Aral, S., and M. V. Alstyne (2011): "The Diversity-Bandwidth Trade-off," *American Journal of Sociology*, 117(1), 90–171.

Burt, R. (1992): *Structural Holes: The Social Structure of Competition.* Harvard University Press, Cambridge, Mass.

Demil, B., and X. Lecoq (2006): “Neither Market nor Hierarchy nor Network: The Emergence of Bazaar Governance,” *Organization Studies*, 27(10), 1447–66.

Diewert, W. (1973): “Afriat and Revealed Preference Theory,” *Review of Economic Studies*, 40(3), 419–26.

——— (1976): “Exact and Superlative Index Numbers,” *Journal of Econometrics*, 4, 115–145.

——— (1978): “Superlative Index Numbers and Consistency in Aggregation,” *Econometrica*, 47, 883–900.

Dowrick, S., and J. Quiggin (1997): “True Measures of GDP and Convergence,” *American Economic Review*, 87(1), 41–64.

Eltetö, O., and P. Köves (1964): “On a Problem of Index Number Computation Relating to International Comparisons,” *Statisztikai Szemle*, 42, 507–518, [in Hungarian].

Gerschenkron, A. (1951): “A Dollar Index fo Soviet Machinery Output, 1927-28 to 1938,” Rand Corporation, Santa Monica, CA.

Granovetter, M. (1973): “The Strength of Weak Ties,” *American Journal of Sociology*, 78(6), 1360–80.

Hansen, M. (1999): “The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits,” *Administrative Science Quarterly*, 44(1), 82–111.

Hill, R. (1999): “Comparing Price Levels Across Countries using Minimum Spanning Trees,” *Review of Economics and Statistics*, 81(1), 135–42.

Ritzer, G., and N. Jurgenson (2010): “Production, consumption, prosumption,” *Journal of Consumer Culture*, 10(1), 13.

Rosen, S. (1981): “The Economics of Superstars,” *American Economic Review*, 71(5), 845–858.

Szulc, B. (1964): “Index Numbers for Multilateral Regional Comparisons,” *Przeglad Statystyczny*, 3, 239–254, [in Polish].

Varian, H. (1982): “The Nonparametric Approach to Demand Analysis,” *Econometrica*, 50(4), 945–73.

——— (1983): “Non-Parametric Tests of Consumer Behaviour,” *Review of Economic Studies*, 50(1), 99–110.

Wu, L., and R. Ackland (2012): "How Web 1.0 Fails: The Mismatch Between Hyperlinks and Clickstreams," arXiv preprint arXiv:1201.6095.