

[This is the post-print version of an article forthcoming in *Telematics and Informatics* in 2007. See the journal website at: <http://www.sciencedirect.com/science/journal/07365853>]

Internet-based ‘social sharing’ as a new form of global production: The case of SETI@home

Hans-Jürgen Engelbrecht*

*Department of Applied and International Economics,
College of Business,
Massey University,
Turitea Campus,
Palmerston North,
New Zealand.*

Abstract

Benkler (‘Sharing Nicely’, *Yale Law Journal*, 2004, Vol. 114, pp. 273-358) has argued that ‘social sharing’ via Internet-based distributed computing is a new, so far under-appreciated modality of economic production. This paper presents results from an empirical study of SETI@home (the Search for Extraterrestrial Intelligence), which is the classic example of such a computing project. The aim is to explain SETI@home participation and its intensity in a cross-country setting. The data are for a sample of 172 developed and developing countries for the years 2002-2004. The results indicate that SETI@home participation and its intensity can be explained largely by the degree of ICT access (proxied by the International Telecommunication Union’s ‘Digital Access Index’), as well as GDP per capita and dummy variables for major country groups. Some other variables, such as the Human Development Index, perform less well. Although SETI@home is a global phenomenon, it is never-the-less mostly concentrated in rich countries. However, there are indications of a slowly narrowing global SETI@home digital divide.

JEL Classification: D20 Other, L31 Nonprofit institutions, L86 Information and Internet Services, Computer Software

Key Words: SETI@home, social sharing, shareable goods, economic production, distributed computing, digital divide.

* Tel: +64 6 350 5968; Fax: +64 6 350 5660; Email: H.Engelbrecht@massey.ac.nz

1. Introduction

The Internet, and Information and Communication Technologies (ICTs) in general, are ‘general purpose technologies’ (GPTs), i.e. technologies that greatly affect most sectors of the economy, and that do so increasingly over time, spawning many additional innovations.¹ One of these innovations is distributed computing and its associated technologies. For example, Foster (2002) sees the emergence of cheaper and more powerful computer networks (grids) as a new computing infrastructure and paradigm that is likely to transform the practice of science and engineering.² Demand for and supply of distributed computing technologies are re-enforcing each other, propelling the development of highly computationally intensive knowledge production.

However, the potential impacts of these technologies are even broader. Benkler (2004) has argued that the Internet has not only remarkably reduced the cost of sharing information, but that sharing via the Internet is also emerging for certain physical, rivalrous goods, and that this will increase in future, resulting in a new, as yet under-appreciated, mode of economic production, which he labels ‘social sharing’. This new mode of economic production takes place alongside market production and state-based production, sometimes substituting for, and sometimes complementing, these other modes. The physical, rivalrous and privately owned goods shared are mostly computing power and bandwidth (a non-information age example is car-pooling). They are widely distributed amongst individuals and they have large unused capacity.³ Moreover, the participating individuals are either only loosely connected or complete strangers, i.e. they can possess a low level of ‘social capital’, yet their sharing combines “to form large-scale and effective systems for provisioning goods, services, and resources“ (ibid., p. 276).⁴

Undoubtedly, current developments in computing technology and computer science are greatly increasing the technological potential for social sharing via distributed computing over the Internet. However, although the prevalence of social sharing as a form of production is sensitive to technological conditions, it is not solely determined by them. Benkler observes (ibid., p. 340): “It is important to emphasize that technology imposes *threshold* constraints on effective sharing, but it cannot unilaterally determine the level of sharing practiced in a society. At any given level of technically feasible sharing, societies may differ in their cultural practices and tastes.” Similarly, David (2004) has emphasized the importance and interplay of appropriate technical, social and legal conditions for realising a cyber-infrastructure for scientific

¹ See Helpman (1998) for a discussion of GPTs.

² Joseph and Fellenstein (2004) argue that grid computing, as well as other distributed computing technologies, grew out of a more homogenous distributed computing discipline. They provide an overview of grid computing, its emerging sub-disciplines, and the many commercial and non-commercial grid-computing initiatives underway. See also the Grid-Café at CERN (<<http://gridcafe.web.cern.ch/gridcafe/>>), and Ian Foster’s website (<<http://www-fp.mcs.anl.gov/~foster/>>).

³ Benkler (2004, p. 276/277) highlights two features of such ‘shareable goods’: They are lumpy (i.e. they come in discrete packages like PCs), and they are of ‘mid-grained’ granularity (i.e. they are widely privately owned and systematically have slack capacity).

⁴ For a delineation of ‘social sharing’ from the related and more common, but also narrower, economic concepts of reciprocity and gift exchange, see Benkler (2004).

collaboration, i.e. e-science. In important respects the non-technical elements are as critical as the technical ones, if not more so.

This paper analyses the prime example of a voluntary non-commercial Internet-based distributed computing project, i.e. the **Search for Extraterrestrial Intelligence (SETI)@home**.⁵ It is a first attempt to explore the broad determinants of SETI@home participation and its intensity in a cross-country setting. By necessity, the approach taken is to use the current state-of-the-art composite variable for measuring Internet access across the largest number of countries possible, plus other variables that mostly account for the level of development. It is recognised that further research will have to focus on smaller groups of countries for which better data are available in order to disentangle the technical and non-technical factors determining social sharing.

The SETI@home project and the SETI data used in this study are briefly discussed in section 2. Section 3 introduces the explanatory variables employed in the empirical model. Regression results are reported in Section 4. Next, a snapshot of the global SETI@home digital divide is provided. Section 6 contains concluding comments. An appendix provides information on the countries included in the data sample and on the definition of the country dummy variables used in the analysis.

2. SETI@home – a worldwide phenomenon

Through a mix of popular appeal, various incentives, good technology for the task at hand and the right economics, SETI@home has been able to attract participants from around the world.⁶ It enlists PCs and uses their spare capacity to analyse data from the Arecibo radio telescope for signs of extraterrestrial intelligence. Launched on 13 May 1999, it became the most powerful special purpose super computer in the world for part of its existence. It now provides more than 1000 years of CPU time a day. The form of distributed computing utilised by SETI@home is usually, but not always, classified as peer-to-peer (P2P) computing, i.e. a particular form of distributed computing characterised by its decentralised nature.⁷

According to Gray (2003) SETI@home ‘pays’ for using participants’ computing resources “by providing a screen saver, by appealing to people’s interest in finding extra-terrestrial intelligence, and by creating competition among teams that want to demonstrate the performance of their systems.” Moreover, its economics is right: The compute cost to network cost of SETI@home is 10,000 to 1, i.e. it is very cpu-intensive (ibid.). Gray estimates that SETI@home performed a multi-billion dollar computation for a million dollars (the cost being made up of those for network bandwidth, voluntarily donated cpu time and electricity).

⁵ See <<http://setiathome.ssl.berkeley.edu/>>.

⁶ It should be noted that the designer of SETI has also gone commercial, hoping to sell access to the world’s idle computers, or to sell the software to link idle computers within an enterprise (Foster, 2000).

⁷ For a survey of P2P computing, including SETI@home and other P2P applications, see Milojevic et al. (2002). They contrast it with ‘grid computing’ and define P2P computing as “a class of systems and applications that employ distributed resources to perform a critical function in a decentralized manner” (ibid., p. 1).

The official SETI@home website provides a number of statistics about the performance of SETI and its participants. These are updated frequently, but not always on a daily basis. I downloaded the ‘country statistics’ reported for December 10th, 2002, December 11th, 2003 and December 13th, 2004.⁸ The period covered in the analysis is one during which the number of countries connected to the Internet has been virtually constant. While in 1990 only 20 countries were connected, that increased to 115 in 1995 and surpassed the 200 mark in 2000 (ITU, 2003, Figure 1.1, p. 5). Only since the turn of the millennium has the Internet spread to virtually every corner of the globe. December 2004 SETI country statistics were available for 226 countries, i.e. SETI@home was run basically everywhere in the world, the only notable exceptions being Mauritius, Vatican City and Palestine. By December 2004, there had been more than five million SETI@home users who had contributed over 2 million years of CPU time (1000 years during the last day alone) (see Table 1). 75.4% of users were located at home, 19.4% at work, 4.1% at school, and 1.1% elsewhere.

Table 1: Total statistics at 06:44:11, 13 December 2004:

	Total	Last 24 Hours
Users	5,282,992	1,471
Results received	1,680,253,315	1,515,916
Total CPU time	2,156,513.796 years	1,091.221 years
Floating point operations	6.137173e+21	5.912072e+18 (68.43 TeraFLOPs/sec)
Average CPU time per work unit	11 hr 14 min 34.7 sec	6 hr 18 min 20.9 sec

I use two alternative dependent variables in the regression analysis: “SETI participants per capita” and “SETI results per capita”. Participants per capita data can be criticized for the same reason that ‘Number of Internet users per 100 inhabitants’ or ‘Internet host density’ has been criticised (Kirkman et al., 2002): They indicate the degree of participation, but not its intensity. It should also be noted that the ‘SETI results per capita’ variable is arguably an improvement over the often used Internet-intensity variable ‘hours of use’ (e.g. ‘average hours of Internet access per day’), because the latter does not say anything about outcomes, i.e. quality of use. In the case of SETI@home, intensity of participation is measured by *actual outcomes* (the number of data units processed).

3. Variables that might explain SETI@home participation and its intensity

Previous empirical studies of Internet and ICT diffusion and use report a host of (sometimes contradictory) findings with regard to explanatory variables (see, e.g., Kiiski and Pohjola, 2002, Chinn and Fairlie, 2004, Caselli and Coleman II, 2001, Pohjola, 2003). The relative importance of variables like GDP per capita (‘gdp’), Internet & telecommunication infrastructure and access costs, demographic and

⁸ They cover countries and some territories. Where appropriate, in the following, the term ‘countries’ covers both. The type of country statistics used in this paper seems to have become unavailable since a SETI@home software upgrade in early 2006.

institutional factors (e.g. quality of regulation) and human capital (e.g. educational attainment), tends to vary between studies. This is not surprising, given differences in country coverage and time periods, as well as differences in modelling approaches.

The approach taken in this study is somewhat different from that followed in the existing literature. I explore the explanatory power of only a few key variables likely to capture the main determinants of SETI@home participation and its intensity in order to include the largest number of countries possible. These variables are, firstly, the International Telecommunication Union's (ITU)'s *Digital Access Index* (DAI), which is the current state-of-the-art *composite* variable to measure Internet access across countries, and, secondly, gdp. I also experimented with the use of the United Nations Development Programme's (UNDP's) *Human Development Index* (HDI), another composite variable, and country group dummy variables.

3.1 *The Digital Access Index*

The DAI tries to measure “*the overall ability of individuals in a country to access and use ICTs*”, while also having as large a country coverage as possible and being transparent and intuitively understandable (ITU, 2003, p. 103/105). Introduced in the ITU's 2003 World Telecommunication Development Report, it is promoted as providing the first truly global ICT ranking. Although several international organisations have developed composite indices that rank countries' ICT capabilities, the ITU (ibid., p. 20) argues that none is completely satisfactory for measuring ICT access, which it regards as a most fundamental requisite for an inclusive information society.⁹ However, determining what to include in a composite index is a balancing act. This is nicely expressed in the ITU report (ibid., p. 6): “This report explains the different ways of measuring access to ICTs and offers a middle way between too much and too little, between relevance for the majority of countries or only for a minority, between what is achievable within existing constraints and what would require significantly increased resources.” In short, the design of such an index is always open to criticism.

The DAI uses eight indicators to capture (a) infrastructure (fixed telephone and mobile telephone subscribers), (b) affordability (internet access price), (c) knowledge (adult literacy, school enrolment), (d) quality (broadband subscribers, international internet bandwidth) and (e) actual usage of ICTs (internet users)(see Table 2). It is calculated as follows: (1) Each indicator is divided by its goalpost, i.e. the maximum value established for it. This normalizes all indicators to take values between 0 and 1. (2) The resulting values are multiplied by their weight and added to obtain a category index. (3) The overall DAI is obtained by multiplying each of the five category indices by 0.2 and adding them up. This equal weighting scheme is defended as the most transparent weighting method. Moreover, the robustness of the weighting was tested by changing the percentages accounted for by the five category indices (but still

⁹ The non-exhaustive list of other composite indices includes (see ITU, 2003, pp. 100-103): The World Economic Forum's “Network Readiness Index”, the International Data Corporation's “Information Society Index”, the Economist Intelligence Unit's “e-readiness ranking”, Mosaic Group's measures of the state of Internet diffusion (they are not reported as a composite index, but others have used them to develop one), Orbicom's “Infostate index”, the UNDP's “Technology Achievement Index”, and the United Nation Conference for Trade and Development's “ICT Diffusion Index” and their other related indices.

keeping the various percentages the same across countries). Correlation tests indicated that all the different weighting methods are statistically identical in terms of the overall rankings of the DAI (ITU, 2003, p. 110).¹⁰

Table 2: Components of the Digital Access Index (DAI), 2002

Indicator	Goal-post	Sub-Components (in italics) and their composition
Fixed telephone subscribers per 100 inhabitants ¹	60	Each has a one half weight for <i>infrastructure</i> , which proxies overall ICT network development.
Mobile subscribers per 100 inhabitants	100	
Adult literacy ²	100	Literacy has a two-third weight and enrolment a one-third weight for <i>knowledge</i> , which affects a country's ability to use new technologies.
Overall school enrolment (primary, secondary and tertiary) ²	100	
Internet access price (20 hours per month) as percent of monthly per capita income ³	1	This is subtracted from 1 to form an indicator that proxies <i>affordability</i> of Internet access. (1=free Internet)
Broadband subscribers per 100 inhabitants ⁴	30	Each has a one half weight for <i>quality</i> of access to ICTs.
International Internet bandwidth per capita	10'000	
Internet users per 100 inhabitants	85	This indicator proxies Internet <i>usage</i> .

Notes: 1. Public Switched Telephone Network (PSTN) plus Integrated Services Digital Network (ISDN) subscribers. 2. Obtained from the UNDP's Human Development Index. 3. Cheapest dial-up or broadband plan averaged over 20 hours of peak and 20 hours of off-peak usage. Annual average exchange rates from the IMF are used to convert the Internet tariffs into US dollars. GNI per capita data are from the World Bank. 4. Including Digital Subscriber Line (DSL), cable modem and other technologies faster than 128 kbit/s in at least one direction.

Source: ITU (2003, Table 5.1, p. 106, Table 5.2, p. 108).

I hypothesise that DAI is a positive and statistically significant determinant of SETI@home participation and its intensity. Should this prove correct, it would suggest that on average, the level of SETI participation and its intensity matches ICT accessibility.

However, the use of a composite index in regression analysis is controversial. A composite index is only as good as its subindices and the way they are aggregated. The ITU (2003, p. 113-116) expresses the hope that the existence of the DAI will stimulate improvements in some of the underlying data that are currently of uncertain quality (e.g. for about half the countries the data on the number of Internet users are not based on survey data, and the school enrolment variable is not quality-adjusted). Moreover, the weights for the subindices are assumed to be the same across countries, which might not be appropriate across the large sample of countries used in this study. The ITU (*ibid.*, p. 99) partially acknowledges this by qualifying the statement that one

¹⁰ For a more detailed discussion of definitional and methodological issues of the DAI see chapter 5 in ITU (2003).

of the main benefits of the DAI is its inter-country comparability: comparisons are most valuable for similar countries and for individual countries over time. This has to be kept in mind when interpreting the empirical results presented in sections 4&5 below.¹¹

3.2 *GDP per capita*

This variable has traditionally been used as a proxy for the standard of living in a country. Numerous studies exploring the determinants of Internet & ICT diffusion, use and the digital divide have included it as a key explanatory variable and found it positive and statistically significant. Moreover, it can arguably be regarded as a proxy for all the other variables positively correlated with gdp that are not included amongst the explanatory variables (Caselli and Coleman II, 2001). The gdp data in purchasing power parity (PPP) adjusted US \$ used in this study were taken from UNDP (2004, pp. 139-141). They have been transformed into \$ 000's.

3.3 *The Human Development Index*

The HDI is a composite index that “focuses on three measurable dimensions of human development: living a long and healthy life, being educated and having a decent standard of living...it combines measures of life expectancy, school enrolment, literary, and income...” (UNDP, 2004, p. 128). It is calculated as the simple average of three dimension sub-indices (the life expectancy index, the education index, and the gdp index). These are based on data for life expectancy at birth, the adult literacy rate, the (combined) gross (primary, secondary and tertiary) enrolment ratio, and gdp. To calculate the dimension sub-indices, maximum and minimum values for the underlying data are first chosen. This enables the dimension sub-indices to be expressed as values between 0 and 1. Note that the education index is a weighted average of the adult literary and the gross enrolment ratio, using the same weights as in the DAI.¹²

Although a consensus is emerging that the HDI is the preferred measure of development (Meier and Rauch, 2005, p. 2), it is recognised that it is still not comprehensive enough. The view adopted in this study, and probably in most of the literature, is that the HDI captures important dimensions of human development neglected by the level of income. It is best included in regressions as an explanatory variable alongside gdp and other variables. However, whether this improves the explanatory power of a regression is an empirical question. As in the case of DAI and gdp, I expect a positive relationship between human development and SETI@home participation and its intensity.

The same critical issues concerning composite indices that were mentioned in the discussion of DAI arise with respect to HDI. Comparing the composition of the two composite indices, there is a minor overlap with regard to adult literacy and school

¹¹ An exploration of issues concerning the validity of various related composite indices is beyond the scope of this paper. For some critical comments on composite indices see, for example, Grigorovici et al. (2002), Engelbrecht (2003), Giacomello and Picci (2003).

¹² For further details, see Technical Note 1, UNDP (2004, pp. 258/9).

enrolment rate (they are much more important in the HDI, accounting for half of the index, compared to 20% of the DAI).

3.4 Country group dummy variables

I include broad country group dummy variables amongst the explanatory variables, in particular “developed&advanced” countries versus “the rest”. The former are listed in ITU (2003, p. xi) and comprise 27 out of the 172 countries in the data set (see appendix). I further experimented with the use of six regional dummy variables similar to those employed by Caselli and Coleman II (2001). Their definitions are also provided in the appendix. Although only an analysis using time-series data can properly account for country-specific effects, the broad country group dummy variables and regional dummy variables account for some major inter-country differences that need to be taken into account in order to obtain reliable statistical estimates.

4. Regression results

For 54 countries participating in SETI@home (about 24% of the total) there were no matching data on DAI, HDI, and gdp. These countries have been deleted from the analysis, leaving 172. The main regressions are reported in Table 3. They include the dependent variables in, alternatively, 2004 levels as well as in 2002-2004 changes and DAI, gdp (both for 2002) as well as the developed&advanced country group dummy variable. Five countries (Yemen, Mali, Myanmar, Nigeria, Ethiopia) had fewer participants per capita in December 2004 than in December 2002, i.e. the change in participants per capita in logs could not be used. These observations were therefore deleted from the data used in regression (4), resulting in a slightly smaller sample size.

All estimates reported in Table 3 are highly statistically significant. In regressions (1), (2), and (4), the estimates for DAI and gdp are their elasticities. In regression (3), the elasticities for DAI and gdp evaluated at the sample means are, respectively, 0.767 and 1.127. The elasticities are all quite similar. Increasing DAI and gdp by 1% increases the dependent variables by a similar percentage. The explanatory power of the regressions is high as judged by the adjusted R^2 values. It seems to make relatively little difference whether SETI participants per capita or SETI results per capita is used as dependent variable.

All regressions, i.e. not just regression (3), were also estimated by Box Cox transformation which indicated that the double log form is greatly preferred to the linear form (in case of the latter, the residuals are highly non-normal). Normality of residuals further required the use of gdp and of the developed&advanced country group dummy variable. Multiple regional dummy variables did not improve the regressions.

HDI (for 2002) performed less well than DAI. When it was used instead of DAI as an explanatory variable alongside gdp, its coefficient estimate was either positive but not statistically significant, or positive and statistically significant but with highly non-normal residuals. Using both DAI and HDI in the same regression produced

Table 3: Regression results

	(1)	(2)	(3)	(4)
Dependent Variable:	Results per capita 2004	Participants per capita 2004	Change in results per capita, 2002-2004	Change in participants per capita, 2002-2004
Indep. Variables:				
DAI 2002	0.945 ^b (2.25)	1.065 ^a (3.54)	0.745 ^b (2.17)	0.892 ^a (3.02)
gdp 2002	1.144 ^a (4.82)	1.031 ^a (5.36)	0.938 ^a (4.96)	0.985 ^a (5.43)
Dev&adv dummy	1.121 ^a (3.52)	0.769 ^a (3.09)	1.013 ^a (2.98)	0.689 ^a (2.91)
Intercept	-4.383 ^a (-5.59)	-9.348 ^a (-15.86)	-4.457 ^a (-6.94)	-10.668 ^a (-18.65)
Adj. R ²	0.70	0.77	0.69	0.73
DW	1.76	1.57	1.73	1.73
JB	1.28	1.08	5.68	0.32
No. of obs.	172	172	172	167
λ			0.05	

^a Denotes statistical significance at the 1% level (one-sided test).

^b Denotes statistical significance at the 2.5% level (one-sided test).

All regressions were estimated by Ordinary Least Squares with White's heteroskedasticity correction and, except for regression (3), were estimated in double-log form. t-ratios are given in brackets. DW is the Durbin-Watson d test for auto-correlation. The null hypothesis of no autocorrelation is accepted at the 1% level of significance for all regressions except regression (2), for which it is rejected ($d_L=1.6$). JB is the Jarque-Bera test for normality of residuals. The critical value at the 1% level of significance is $\chi^2_{(2)} = 9.21$. λ indicates the exponent in the (extended) Box-Cox regression that was used to transform both the dependent variable and explanatory variables (except the dummy variables).

statistically insignificant and negative coefficient estimates for the latter, which is not surprising given the very high correlation between the two (0.95). The regression estimates including HDI have therefore not been reported.

To sum up, the regression results indicate that SETI@home participation and its intensity are not 'idiosyncratic', i.e. they are indeed largely explained by DAI and gdp. This seems to be a promising result for all those that support SETI@home and similar projects, and for those supporting social sharing via Internet-based distributed computing in general. However, one should be cautious in using these findings to predict the future growth of such activities. This would require a detailed analysis of the socio-institutional conditions affecting the further development of the underlying network infrastructures and incentives of participants.

5. SETI@home and the global digital divide

There is a large literature on the digital divide between (and within) rich and poor countries (see, e.g., Chinn and Fairlie, 2004, and the survey by Hargittai, 2003). A general definition of the term with regard to the Internet is ‘inequality in access and use of the medium’ (ibid., p. 824).¹³ The ITU (2003, Figure 1.3, p. 3) reports that in 2003 there were 43.7 Internet users per 100 inhabitants in high income countries, compared to 1.3 in the low income countries (i.e. in the lowest of five groups of countries ranked by income levels). However, some authors, for example Fink and Kenny (2003) and Dutta et al. (2004, p. 18/19), are challenging the view that the digital divide is widening. Instead, there seems to be some evidence of digital convergence.¹⁴

Table 4 provides a snapshot of the digital divide in terms of the SETI@home data on number of total users, total results received, and population for the groups of developed&advanced countries versus ‘the rest’. Although SETI@home is a world-wide phenomenon, it is largely concentrated in the developed&advanced countries. By December 2004 the former, accounting for just over 15% of the population of the 172 countries in the sample, had contributed over 90% of all results. However, the December 2002 to December 2004 growth rates for users and results have been faster in ‘the rest’. The picture that emerges is that of a slowly narrowing global SETI@home digital divide. Of course, the time period covered is very short. It will be interesting to see whether the observed trend continues in future.

¹³ The divide can be defined in many dimensions, for example in terms of all ICTs, or, like in this study, for a particular form of technology, or by geographic region, or by social group, etc. (Hargittai, 2003). In short, one should always be careful to specify what type of digital divide is being analysed.

¹⁴ This is similar to recent findings in the literature on broader convergence and divergence issues (Neumayer, 2003, Kenny, 2005).

Table 4:	Developed & Advanced Countries versus 'the Rest' (172 country sample)		
Table 4A:			
December 2004 totals	Developed&advanced countries:	The rest:	% the rest:
No. of total users:	4,511,609	604,821	11.8%
No. of total results:	1,507,659,459	164,968,938	9.9%
Total population:	900,395,111	4,975,626,881	84.7%
December 2002 totals:	Developed&advanced countries:	The rest:	% the rest:
No. of total users:	3,559,136	425,122	10.7%
No. of total results:	661,872,079	47,186,802	6.7%
Total population:	899,335,782	4,975,979,881	84.7%
Table 4B:	Percentage increase Dec. 2002- Dec. 2004:		
	Developed&advanced countries	The rest:	
No. of total users:	26.8%	42.3%	
No. of results:	127.8%	249.6%	
Table 4C:	Per capita and per user results:		
	Developed&advanced countries:	The rest:	The rest as % of dev&adv countries:
No. of total users per capita by Dec. 2004:	0.00501	0.00012	2.43%
No. of total users per capita by Dec. 2002:	0.00396	0.000085	2.16%
No. of total results per capita by Dec. 2004:	1.67444	0.03316	1.98%
No. of total results per capita by Dec. 2002:	0.73596	0.00948	1.29%
No. of total results per user by Dec. 2004:	334.17	272.76	81.6%
No. of total results per user by Dec. 2002:	185.96	111	59.7%

6. Concluding comments

This paper has presented a first attempt at determining the major factors explaining SETI@home participation and its intensity. Variables measuring Internet access and standard of living explain a large part of the cross-country variation, as does the general divide between rich and poor countries captured by the dummy variable used in the regressions. There is a need for studies of other social sharing projects to

establish whether SETI@home is representative of social sharing in general, and for the use of more sophisticated models than that tested in this paper.¹⁵

Suffice it to say that whether social sharing via the Internet will in future become a dominant mode of production alongside others is an interesting open question. On the one hand, the technical potential for social sharing seems great. So far, only a small fraction of the world's PCs, which by one estimate totalled over 820 million in 2004 (CIA, 2005), are participating in voluntary grid computing projects (also see Table 4C). However, distributed computing and its associated technologies also provide opportunities for commercial applications. Should technological progress render widespread commercial grid computing feasible and profitable to the extent that many PC owners can participate in it, the willingness of the public to contribute to non-commercial projects might be negatively affected. Moreover, should 'computing on demand' become reality, individuals would not have to own a PC with excess computing power usable for social sharing. In short, there are various business models potentially applicable to distributed computing that could affect the prevalence of social sharing. It remains to be seen whether commercial interests will prevail in future and reduce the extent of non-commercial sharing over the Internet.

On the other hand, as pointed out by David (2004), apart from technological developments distributed computing and its associated technologies also crucially depend on supportive socio-institutional elements, like legal and administrative regimes, that might be difficult to establish. If at all, these softer elements were only partly captured by the variable DAI in this paper. Therefore, further research should focus on narrower and more homogenous groups of countries. That should make it possible to include more explanatory variables, and also to use more sophisticated modelling approaches (like simultaneous equation modelling), that would allow the disentangling of the various factors determining future trends. In particular, more detailed models are needed to address the important issue of how social sharing and its relationship with other forms of production might inform economic policy, and how these relationships and policies might differ for countries at different stages of economic development. This goes beyond the scope of this paper, but the stakes for economic progress are potentially very high.

¹⁵ If it could be established that SETI@home participation is a representative network application, one might even be able to use it as an explanatory variable to predict network infrastructure in some countries, as suggested by one of the referees.

Appendix

Table A.1: List of countries* and definition of dummy variables:

Countries (all included in 172 observation data set):	(a) Developed & Advanced Countries ¹	(b) Developing Countries ²	(c) Regional Dummy Variables ³
1) Iceland	X		-
2) Finland	X		-
3) Denmark	X		-
4) Canada	X		-
5) Netherlands	X		-
6) United States	X		-
7) Sweden	X		-
8) United Kingdom	X		-
9) New Zealand	X		-
10) Luxembourg	X		-
11) Australia	X		-
12) Germany	X		-
13) Norway	X		-
14) Austria	X		-
15) Switzerland	X		-
16) Estonia		X	EEU
17) Belgium	X		-
18) Antigua and Barbuda		X	LAC
19) Ireland	X		-
20) Slovenia		X	EEU
21) Saint Kitts and Nevis		X	LAC
22) Portugal	X		-
23) Czech Republic		X	-
24) France	X		-
25) Hungary		X	-
26) Spain	X		-
27) Croatia		X	EEU
28) Israel	(X)		-
29) Malta		X	-
30) Poland		X	-
31) Singapore	(X)		EA
32) Greece	X		-
33) Italy	X		-
34) Cyprus	(X)		-
35) Dominica		X	LAC
36) Japan	X		-
37) Grenada		X	LAC
38) Slovakia		X	-
39) Brunei Darussalam		X	EA
40) Lithuania		X	EEU
41) Uruguay		X	LAC
42) Latvia		X	EEU
43) Chile		X	LAC
44) Bahamas		X	LAC
45) Bulgaria		X	EEU
46) Barbados		X	LAC
47) Argentina		X	LAC
48) Belize		X	LAC

49) Costa Rica		X	LAC
50) Cape Verde		(X)	SSA
51) United Arab Emirates		X	ARW
52) Seychelles		X	SSA
53) Saint Lucia		X	LAC
54) Former Yugoslav Republic of Macedonia		X	EEU
55) Bahrain		X	ARW
56) Republic of Korea	(X)		-
57) Saint Vincent and Grenadines		X	LAC
58) Romania		X	EEU
59) Panama		X	LAC
60) Malaysia		X	EA
61) Albania		X	EEU
62) Trinidad and Tobago		X	LAC
63) Mexico		X	-
64) Qatar		X	ARW
65) Venezuela		X	LAC
66) Kuwait		X	ARW
67) Brazil		X	LAC
68) Sao Tome and Principe		(X)	SSA
69) Vanuatu		(X)	ROA
70) Comoros		(X)	SSA
71) Maldives		(X)	ROA
72) South Africa		X	SSA
73) Samoa		(X)	ROA
74) Bosnia and Herzegovina		X	EEU
75) Colombia		X	LAC
76) Turkey		X	-
77) Swaziland		X	SSA
78) Guyana		X	LAC
79) Dominican Republic		X	LAC
80) Lebanon		X	ARW
81) Fiji		X	ROA
82) Jamaica		X	LAC
83) Djibouti		(X)	SSA
84) Namibia		X	SSA
85) Russian Federation		X	EEU
86) Paraguay		X	LAC
87) Suriname		X	LAC
88) Ecuador		X	LAC
89) Congo		X	SSA
90) El Salvador		X	LAC
91) Jordan		X	ARW
92) Oman		X	ARW
93) Gabon		X	SSA
94) Belarus		X	EEU
95) Bolivia		X	LAC
96) Honduras		X	LAC
97) Peru		X	LAC
98) Botswana		X	SSA
99) Equatorial Guinea		(X)	SSA
100) Armenia		X	ROA
101) Guatemala		X	LAC
102) Thailand		X	EA
103) Solomon Islands		(X)	ROA
104) Gambia		(X)	SSA
105) Georgia		X	ROA

106) Ukraine		X	EEU
107) Nicaragua		X	LAC
108) Republic of Moldova		X	EEU
109) Philippines		X	EA
110) Angola		(X)	SSA
111) Algeria		X	ARW
112) Guinea		(X)	SSA
113) Azerbaijan		X	ROA
114) Zimbabwe		X	SSA
115) China		X	ROA
116) Kazakstan		X	ROA
117) Bhutan		(X)	ROA
118) Mongolia		X	ROA
119) Tunisia		X	ARW
120) Saudi Arabia		X	ARW
121) Guinea-Bissau		(X)	SSA
122) Kyrgyzstan		X	ROA
123) Sri Lanka		X	ROA
124) Iran		X	ARW
125) Cuba		X	LAC
126) Central African Republic		(X)	SSA
127) Mauritania		(X)	SSA
128) Egypt		X	ARW
129) Cambodia		(X)	ROA
130) Haiti		(X)	LAC
131) Pakistan		X	ROA
132) Morocco		X	ARW
133) Lesotho		(X)	SSA
134) India		X	ROA
135) Burkina Faso		(X)	SSA
136) Papua New Guinea		X	ROA
137) Ghana		X	SSA
138) Eritrea		(X)	SSA
139) Zambia		(X)	SSA
140) Nepal		(X)	ROA
141) Turkmenistan		X	ROA
142) Cameroon		X	SSA
143) Tajikistan		X	ROA
144) Malawi		(X)	SSA
145) Libya		X	ARW
146) Burundi		(X)	SSA
147) Chad		(X)	SSA
148) Benin		(X)	SSA
149) Indonesia		X	EA
150) Togo		(X)	SSA
151) Sierra Leone		(X)	SSA
152) Laos		(X)	ROA
153) Senegal		(X)	SSA
154) Viet Nam		X	ROA
155) Kenya		X	SSA
156) United Rep. of Tanzania		(X)	SSA
157) Cote d'Ivoire		X	SSA
158) Syrian Arab Republic		X	ARW
159) Uzbekistan		X	ROA
160) Yemen		(X)	ARW
161) Madagascar		(X)	SSA
162) Rwanda		(X)	SSA
163) Mozambique		(X)	SSA

164) Uganda		(X)	SSA
165) Niger		(X)	SSA
166) Bangladesh		(X)	ROA
167) Mali		(X)	SSA
168) Sudan		(X)	SSA
169) Myanmar		(X)	ROA
170) Dem. Republic of the Congo		(X)	SSA
171) Nigeria		X	SSA
172) Ethiopia		(X)	SSA

Notes:

* The 54 countries deleted from the 226 countries for which SETI@home data were available are: Afghanistan, American Samoa, Andorra, Anquilla, Antarctica, Aruba, Bermuda, British Virgin Islands, Cayman Islands, Channel Islands, Cook Islands, Democratic People's Republic of Korea, East Timor, Faeroe Islands, Falkland Islands, French Guiana, French Polynesia, Gaza Strip, Gibraltar, Greenland, Guadeloupe, Guam, Iraq, Isle of Man, Kiribati, Liberia, Lichtenstein, Macau, Marshall Islands, Martinique, Micronesia, Monaco, Montserrat, Nauru, Netherlands Antilles, New Caledonia, Niue, Northern Mariana Islands, Palau, Pitcairn Islands, Puerto Rico, Reunion, Saint Pierre and Miquelon, San Marino, Serbia and Montenegro, Somalia, Taiwan, Tokelau, Tonga, Turks and Caicos Islands, Tuvalu, United States Virgin Islands, Wallis and Futuna Islands, Western Sahara.

¹ Developed and advanced countries (advanced countries in brackets), 172 country data set. These countries are listed in ITU (2003, p. xi).

² 'The Rest': Developing countries (least developed countries in brackets), 172 country data set.

³ Regional dummy variables: Sub-Saharan Africa (SSA), Latin America and Caribbean (LAC), Arab World (ARW), East Asia (EA), Rest of Asia and Pacific (ROA), Eastern Europe (EEU). The remaining countries belong mostly to the OECD. These dummy variables are similar to those used in Caselli and Coleman II (2001).

Acknowledgements

An earlier version of this paper was presented at the International Telecommunications Society Africa-Asia-Australasia Regional Conference, Perth, Western Australia, August 28-30, 2005. I thank participants of the distributed computing sessions, especially Martin Weiss, for useful comments. Discussions with Don Lambertson also helped to improve the paper. All remaining errors are naturally mine. The URLs cited were last checked on 18 August 2006.

References

- Benkler, Y., 2004. Sharing nicely: on shareable goods and the emergence of sharing as a modality of economic production. *Yale Law Journal* 114, 273-358.
- Caselli, F., Coleman II, W. J., 2001. Cross-country technology diffusion: the case of computers. *American Economic Review, Papers and Proceedings* 91(2), 328-335.
- Chinn, M., Fairlie, R., 2004. The determinants of the global digital divide: a cross-country analysis of computer and internet penetration. National Bureau of Economic Research Working Paper No. 10686, Cambridge, MA.

- Computer Industry Almanac (CIA)(2005). PCs in-use surpassed 820M in 2004, PCs in-use will top 1B in 2007. Press Release, March 9th. Available from: <<http://www.c-i-a.com/pr0305.htm>>.
- David, P., 2004. Towards a cyberinfrastructure for enhanced scientific collaboration: providing its 'soft' foundations may be the hardest part. Research Report No. 4, Oxford Internet Institute, Oxford, UK. Available from: <http://www.oii.ox.ac.uk/resources/publications/OIIRR4_200408.pdf>.
- Dutta, S., Lanvin, B., Paua, F., 2004. The Global Information Technology Report 2003-2004: Towards an Equitable Information Society. Oxford University Press, New York & Oxford.
- Engelbrecht, H.-J., 2003. Data issues in the New Economy. In: Jones, D. (Ed.), New Economy Handbook. Elsevier/Academic Press, San Diego & London, pp. 57-76.
- Fink, C., Kenny, C., 2003. W(h)ither the digital divide? INFO 5(6), 15-24.
- Foster, I., 2000. Internet computing and the emerging grid. Nature Web Matters, 7 December, 6 pp.
Available from: <<http://www.nature.com/nature/webmatters/grid/grid.html>>.
- Foster, I., 2002. The grid: a new infrastructure for 21st century science. Physics Today 55(2), 42-47.
- Giacomello, G., Picci, L., 2003. My scale or your meter? Evaluating methods of measuring the internet. Information Economics and Policy 15(3), 363-383.
- Gray, J., 2003. Distributed computing economics. Technical Report MSR-TR-2003-24, Microsoft Research, Microsoft Corporation, Redmond, WA. Available from: <http://research.microsoft.com/research/pubs/view.aspx?tr_id=655>.
- Grigorovici, D., Schement, J. R., Taylor, R., 2002. Weighing the intangible: towards a framework for information society indices. E-business Research Center Working Paper, SMEAL College of Business, Penn State University.
Available from: <http://www.smeal.psu.edu/ebrc/publications/res_papers/2002_14.pdf>.
- Hargittai, E., 2003. The digital divide and what to do about it. In: Jones, D. (Ed.), New Economy Handbook. Elsevier/Academic Press, San Diego & London, pp. 821-839.
- Helpman, E., (Ed.), 1998. General Purpose Technologies and Economic Growth. MIT Press, Cambridge, MA.
- International Telecommunication Union (ITU), 2003. World Telecommunication Development Report 2003: Access Indicators for the Information Society. Geneva.

- Joseph, J., Fellenstein, C., 2004. *Grid Computing*. Prentice Hall/Pearson Education, Upper Saddle River, NJ.
- Kenny, C., 2005. Why are we worried about income? Nearly everything that matters is converging. *World Development* 33(1), 1-19.
- Kiiski, S., Pohjola, M., 2002. Cross-country diffusion of the internet. *Information Economics and Policy* 14(2), 297-310.
- Kirkman, G., Cornelius, P., Sachs, J., Schwab, K., 2002. *The Global Information Technology Report 2001-2002: Readiness for the Networked World*. Oxford University Press, New York & Oxford.
- Meier, G., Rauch, J., 2005. *Leading Issues in Economic Development*, eighth ed. Oxford University Press, New York & Oxford.
- Milojicic, D., Kalogeraki, V., Lukose, R., Nagaraja, K., Pruyne, J., Richard, B., Rollins, S., Xu, Z., 2002. Peer-to-peer computing. HPL-2002-57R1, March 8th, HP Laboratories, Hewlett-Packard Company, Palo Alto. Available from: <<http://www.hpl.hp.com/techreports/2002/HPL-2002-57R1.pdf>>.
- Neumayer, E., 2003. Beyond income: convergence in living standards, big time. *Structural Change and Economic Dynamics*, 14(3), 275-296.
- Pohjola, M., 2003. The adoption and diffusion of information and communication technology across countries: patterns and determinants. In: Jones, D. (Ed.), *New Economy Handbook*. Elsevier/Academic Press, San Diego & London, pp. 77-100.
- United Nations Development Programme (UNDP), 2004. *Human Development Report 2004: Cultural Liberty in Today's Diverse World*. New York.