Institutions, Culture and Open Source

Sebastian v. Engelhardt[*], *Andreas Freytag*[**]

- THIS IS WORK IN PROGRESS -

Abstract

So-called open source software (OSS) is marked by free access to the software and its source code. Copyright-based OSS licenses permit users to use, change, improve and redistribute the software, which is designed and developed in a public, collaborative manner. High quality OSS-products like Linux, Apache etc. are developed by thousands of volunteers, who often do not receive direct monetary reward. Thus, OSS seems to be an example of a 'private provision of a public good', and some argue that OSS is a 'new intellectual property paradigm' for the digital economy. Therefore, OSS has been in the focus of economic research for some years.

However, it is still not known which institutional and cultural factors favor OSS development. For this reason, we perform a cross-country study of how the number of OSS developers per inhabitants of a country depends on institutional and cultural factors. For this purpose we make use of data about OSS developers registered at SourceForge. Regarding country-specific factors we take into account aspects of the legal system and regulation, social capital, the openness to novelty, the degree of individualism/self-determination of a society as well as its attitude toward competition.

We find that a culture of individualism/self-determination as well as of interpersonal trust is in favor of OSS. The openness to novelty seems to be relevant only with respect to scientific progress. While the attitude towards competition was never significant, less regulated countries have more OSS (per capita). Furthermore, the protection of intellectual property rights has (if all) a positive impact.

[*] Friedrich-Schiller-University Jena, Department of Economics and Business Administration, Carl-Zeiss-Str. 3, D-07743 Jena. E-mail: Sebastian.Engelhardt@wiwi.uni-jena.de

[**] Friedrich-Schiller-University Jena, Department of Economics and Business Administration, Carl-Zeiss-Str. 3, D-07743 Jena and ECIPE, Brussels. E-mail: A.Freytag@wiwi.uni-jena.de

1 Introduction

The Software industry is characterized by innovations not only at the level of products but also at the level of market organization, i.e. *institution*s[1]. In the case of open source software (OSS), the source code—the human-readable recipe of a software program—is 'open' (disclosed). This implies general access to the software and its source code , as well as the right to read, modify, improve, redistribute and use it. This principle of openness is codified in the (copyright based) OSS licenses. OSS seems to represent a "new intellectual property paradigm" (Maurer & Scotchmer, 2006), i.e. a new type of ownership concept that leads to different allocations of intellectual property rights and different modes of organization as compared to so called proprietary software. OSS is developed by a decentralized but nevertheless well organized 'community' that consists of thousands of volunteers, who develop software, often without direct monetary reward. Nevertheless, some OSS – like the Apache Webserver software (see figure 1) – have a remarkable market share. Thus, OSS appears to be a successful example for a "Private provision of a public good" (Johnson, 2002).

--- [Figure 1 about here] ---

OSS has been in the focus of economic research for some years. For instance, economic research has analyzed the effects of OSS on *competition* (Casadesus-Masanell & Ghemawat 2003, Bitzer 2004), as well as *open innovations* (von Hippel & von Krogh 2003, von Hippel 2005) and *firm investments in OSS* (e.g. Baake & Wichmann 2004, Henkel 2006, Lerner et al. 2006, Rossi & Bonaccorsi 2006). Another important branch of the literature is dealing with the motivations of OSS developers, asking the question of "Why should thousands of top-notch programmers contribute freely to the provision of a public good?" (Lerner & Tirole 2002), i.e. is analyzing the *incentives* and extrinsic and intrinsic *motives*. There is much literature dealing with this question on the individual level, both from economics and other social sciences.[2] We are interested in the conditions for OSS activities on the level of society; in other words we take into consideration the individual motives and search for the conditions for these motives to materialize. As we do a cross-country study, this implies that the number of OSS developers

[1] We refer here to institutions in the definition of North (1990): "Institutions are the rules of the game of a society or more formally are the humanly-devised constraints that structure human interaction. They are composed of formal rules (statute law, common law, regulations), informal constraints (conventions, norms of behavior, and self imposed codes of conduct), and the enforcement characteristics of both."

[2] Schiff (2002) provides an overview of early economic contributions to this topic. A prominent explanation referring to extrinsic motivations is the acquisition of a reputation-signal (Lerner & Tirole 2000), but intrinsic motives also play a role. An empirical examination of the motives can be found e.g. in Ghosh et al. (2002). See also Rossi (2006) on this topic.

of a country is shaped by the country's institutional and cultural factors. Hence potential dependent variables are the number of OSS developers per 1000 inhabitants as well as the level of OSS activities. Making use of Williamson's analytical framework (Williamson 2000, see also figure 1), it seems obvious that economic research so far has focused on the level of resource allocation and employment (level four, the main focus of neoclassical economics) and the level of governance (level three, „the play of the game"). Hence, with respect to OSS there is still lack of knowledge regarding the levels one and two, i.e. regarding the so-called "embeddedness" (informal institutions, customs, traditions, norms, religion) and the institutional environment (formal rules of the game, esp. property). Our paper aims to fill this gap. For this purpose, we perform a cross-country study of how the per capita number of a country's OSS developers registered at SourceForge[3] depends on institutional and cultural factors. In particular, we take into account aspects of the legal system and regulation, social capital, the openness to novelty, the degree of individualism/self-determination of a society as well as its attitude toward competition.

--- [Figure 2 about here] ---

The remainder of the paper is structured as follows: In section 2, we discuss the theoretical foundations of the paper and derive the hypotheses for the empirical study. In section 3, we operationalize the variables and describe the data and its sources. This data is used to perform the empirical assessment in section 4, where the regression results are presented. In section 5 we compare and discuss the results before we end with a summary in 6.

2 Theoretical Considerations and Hypotheses

We assume that the decision to become an (active) OSS developer is shaped by institutional and cultural factors that belong to level one and two of Williamson's framework. To analyze such institutional drivers of OSS development activities in a country, it is sensible to refer to the microeconomic foundations.

Thus, in this section we derive hypotheses about the influence of institutional and cultural factors on OSS developers and their activities respectively. This is a relatively new approach.

The only study (we are aware of) linking cultural factors with the demographics of OSS de-

[3] SourceForge is an Internet platform for OSS developers to control and manage open source software development. In a sense it is a virtual center where the developers of a certain OSS project can meet, discuss, coordinate their tasks, upload new developed code, and so on. SourceForge, is seen as the world's largest site hosting OSS-projects. While access to this developer-areas needs registration, finished version of the software can be downloaded by anybody.

velopers is Ramanujam (2007). Ramanujam took Data from Ghosh (2006) and uses Hofstede's cultural indicators in order to analyze how differences in national culture affect or influence the participation in OSS. He links the geographical distribution of developers with the four dimensions of national cultures considered by Hofstede (Power Distance, Individualism, Masculinity-Femininity, and Uncertainty Avoidance). Ramanujam states a positive correlation between the share of OSS developers and individualism, whereas power distance and uncertainty avoidance are negatively correlated each. However, with respect to OSS contribution Ramanujam (2007) distinguishes only four regions, whereas the paper at hand runs regressions with data from about 70 countries, and we analyze several cultural and institutional factors including norms and attitudes.

2.1 Intrinsic motivation

One main intrinsic motive mentioned in the literature on single programmers points towards *individualism/self-determination*. For example, Hars & Ou (2002) found that "self-determination" was with about 80% agreement the strongest intrinsic motive. Other authors report that "fun" and enjoyment of programming work itself or of solving problems, and the joy of intellectual challenge are important motives for individuals to contribute to OSS (Luthiger Stoll 2007, Lakhani & Wolf 2005, Hertel et al 2003, Lakhani et al 2002). We conclude that such motives, which are very closely related to individualistic attitudes and self-determination, require a culture with a higher degree of individualism/ self-determination. Thus, our first hypothesis states that:

H1: the degree of individualism/self-determination of a society has a positive impact on the number of OSS developers as well as on the OSS activity level.

Somehow related to the enjoyment of solving problems and the joy of intellectual challenge is the next aspect: innovation and novelty. The process of (open source) software development is to some extent a search for new solutions, i.e. is an innovative process as such. Thus, OSS development is about "coordinating innovation" (Kugler, 2005). Additionally, the rise of OSS can be seen as an innovation at the level how to organize software development. Therefore, some authors discuss it as a new intellectual property paradigm (e.g. Maurer & Scotchmer, 2006), or put open source software in line with „innovation" and an „intellectual property revolution" (Pisano 2006). Furthermore, since OSS is perceived as something new it may be sensible to assume that societies that are more open to novelty are more open to the "innovation" of OSS. In particular we assume that in countries where more people are *open to new*

ideas, OSS activities are intense. In addition, OSS is a novelty of the "cyber space": without the Internet the success of OSS would not have been possible. To some extent, a positive attitude towards technological progress (like the Internet and software/computers) is connected to a positive attitude towards *scientific progress*. Assuming that individuals with a positive attitude towards scientific progress are open to technological solutions as well, we expect that a generally positive attitude towards scientific progress in society has an impact on the number of OSS developers. However, that does not necessarily imply that general openness to scientific progress has an impact on the *level* of OSS activities. Based on this, we state the following two hypotheses:

H2: A preference for new ideas has a positive impact on the number of OSS developers as well as on the OSS activity level.

H3: A positive attitude towards scientific progress has a positive impact on the number of OSS developers.

Another related factor to be taken into account, is so-called social capital. The term social capital is related to ties between people. While some refer to the number of ties only, others stress the features, strength or quality of such ties, which can also include aspects like norms and trust. Probably the most known (and widely accepted) definition of "social capital" is by Putnam (1993, 1995). Putnam states that social capital „refers to features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit." Putnam (1995, p 67). Therefore measures of social capital can take into account aspects like number of people somebody is (weakly) connected to, but also aspects like interpersonal trust and social engagement. We focus on interpersonal trust in this paper.

As stated before, the voluntary contribution to OSS can be regarded as being a "private provision of a public good" (Johnson, 2002). Hence free rider effects tend to occur, if not being wanted. Social trust can help here to lower the individually perceived probability of such free rider behavior. Furthermore it is known that reciprocity is part of the OSS community culture (Gosh et al. 2002, Lakhani et al. 2002). The literature on public goods problems indicates that interpersonal trust has a positive impact of cooperation and reciprocate behavior (Yamagishi et al. 2005, De Cremer 1999, Elinor 1998, Yamagishi & Yamagishi 1994). Thus, based on this we can derive the following hypothesis:

H4: Social Capital in terms of interpersonal trust has a positive impact on the number of OSS developers as well as on the OSS activity level.

2.2. Extrinsic motivations

Next we discuss extrinsic motives, being aware of the ambiguous nature of OSS activities. Additional to enjoying the developing, OSS can be seen as a business model, either on the part of the individual or on the level of the firm. From the perspective of individual, motives such as self-marketing, peer recognition and reputation within the community (Lakhani et al. 2002, Hars & Ou 2002) indicate a *merit principle*. Furthermore, there are motives that are directly linked to career aspects, like the improvement of programming skills, i.e. the investment in human capital, and the aim to build up reputation signals for the job market (Lakhani & Wolf 2005, Hertel et al. 2003, Gosh et al. 2002, Hars & Ou 2002, Lakhani et al. 2002, in all cases these motive were stronger than the motives related to peer recognition). This also points to the merit principle. In a boarder sense one could argue, that these motives have aspects of an attitude towards competition in the sense that personal performance should play a role and hence incentive structures are accepted that rewards personal performance. Therefore we expect the following:

H5: A culture of positive attitudes toward competition and the merit principle has a positive impact on the number of OSS developers as well as on the OSS activity level.

Although OSS is marked by free access to the source code, there exist several OSS-business models. As the OSS-code itself can not be a profit center, OSS business models are based on selling complementary products (Maurer & Scotchmer 2006, p 289, 290ff). This can be hardware like servers or mobile phones, premium versions or different kinds of service like maintenance etc. Having in mind this business character of OSS, it may be important to assess the potential size of a market, both for the supplier of OSS and the demander of OSS programmers. As the cross-country data on individuals employed in programming is so poor, we approximate the market size for both groups (individual programmers and OSS firms) with the share of the population having active access to the Internet, stating that:

H6: A high number of Internet users is beneficial both for the number of OSS developers and the OSS activity level.

In addition, we have to discuss the potential impact of formal institutions on OSS activities. First, this discussion refers to (a) the *protection of intellectual property rights* (IPRs), measured as the de facto protection, implying a negative score by using software piracy rates, and (b) to the *degree of regulation* in a country, i.e. a set of formal institutions, directly interfering in relative prices. The latter is a proxy for the degree of collectivistic control in a society.

Without deeply entering the discussion on the background philosophies of OSS, at this

point we want to give a short account on the main lines of arguments. Some authors in academics, within the OSS community and in software engineering heatedly discuss whether OSS is an expression of collectivistic or of individualistic, if not libertarian, ideas.[4] Therefore, it is in fact not easy to decide whether OSS has, in practical terms, to be associated with collectivism and/or heavy regulation or not. Although from a philosophical point of view one might interpret it as rooted in libertarian ideals, the statements of some OSS advocates and the terms they use seem it justified to claim the opposite. However, the majority of OSS developers may just favor pragmatic solution, which implies that heavy regulation of economic activities is adverse to OSS activities, in particular if interpreted as business. Our hypothesis regarding the culture of individualism/self-determination support this, as the higher the degree of regulation in a society, the lower is generally the (accepted) degree of individual initiatives and self-determination. In sum, this yields the following:

H7: A high degree of intense economic regulation has a negative impact on the number of OSS developers as well as on the OSS activity level.

At a first glance, it is not clear, whether *protection of IPRs* and OSS have a positive or negative relationship. One might interpret the OSS community as a (social) movement against intellectual property protection of software. Based on this one would expect to see at least not a positive relationship between the de facto protection of IPRs and OSS, at least because OSS can better handle situations where the protection of intellectual property rights is weak, as it is "open", or "free" anyway. Furthermore, one might expect a positive relationship between software piracy rates and OSS, as OSS developers might not respect intellectual property. Such a view can be supported by the fact that some parts of the open source community argue in an "anti intellectual property"-way. Furthermore, the "Free Software Foundation" opposes the use

[4] The discussion is about whether OSS is "Hacking Capitalism" (Söderberg 2008) and should therefore be associated with collectivism, i.e. with socialism, communitarianism, or communism, like e.g. Glass (2004) does. Several authors question whether OSS is really „Beyond capitalism" (Economist, 10 June 2004) but rather is „about capitalism, not freebies" (Asay, M. 2009). Furthermore, in his famous libertarian interpretation of OSS, Greg Perkins argues that "it is capitalism which is harmonious with Open Source, and that collectivism is incompatible" (Perkins 1998). This discussion is somehow represented by two members of the OSS community: First, Eric S. Raymond, well known for his "Cathedral and the Bazaar" and co-founder of the Open Source Initiative. He interprets OSS as being something that is de facto more close to libertarianism than to collectivism (to some extend regardless to what some OSS activist sound like). His opponent is Richard Stallman, who developed the GNU General Public License (GPL), a widely used OSS license. Mr. Stallman is more related to left-wing opinions. Not surprisingly Stallman and Raymond have had a lot of disputes about what OSS is and shall be. The clearest distinction between the both positions can be made when it comes to the question about intellectual property (IP) as such. While Stallman refuses the idea of IP, arguing that because of *moral reasons* no one should be allowed to claim property rights on information or knowledge, Raymond supports the idea of property right claims, and hence also of IPRs, but simply argues that proprietary software (in the sense of closed source software) is simply an *inefficient* way of developing software.

of the term “intellectual property”[5], and its president Richard Stallman refuses the idea of intellectual property, arguing that because of moral reasons no one should be allowed to claim property rights on information or knowledge. This view is opposed by figure like Eric S. Raymond, co-founder of the Open Source Initiative. Raymond supports the idea of property right claims, and hence also of intellectual property rights, but simply argues that proprietary software (in the sense of closed source software) is simply an inefficient way of developing software (see Weber 2004). Others like e.g. Greg Perkins also point out that “Open Source depends on the idea of the individual human right to private property” (Perkins 1998). However, there is a good argument why OSS in practice relies on protection of IPR: The best know, and most used OSS license is the GNU General Public License (GPL) developed by Stallmann. One important feature of this license is the so-called ‘Copyleft’-Principle which ensures that the licensed software stays “open”.[6] Obviously such claims are only possible if authors have intellectual property rights and therefore the right to claim under which terms and conditions their work can be used. Thus, in practical terms OSS is not software ‘without property rights’, as the OSS licenses are rely on, i.e. make use of copyright law. Based on this our hypothesis is as follows:

H8: The protection of intellectual property rights has a positive impact on the number of OSS developers or on OSS activity level.

3 The Data

We draw our data from several sources. With respect to the institutional and cultural aspects we could make use of date available from different resources. With respect to the geographical location of the OSS developers we had to collect the data by our own. Nevertheless, whereas it is a relatively new approach to link institutional and cultural factors to the phenomena of OSS, there is already a literature describing the geographic origin of OSS developers. Basically there exists two approaches: some studies are based on survey-data, for example Ghosh

[5] "Publishers and lawyers like to describe copyright as “intellectual property”—a term that also includes patents, trademarks, and other more obscure areas of law. These laws have so little in common, and differ so much, that it is ill-advised to generalize about them. It is best to talk specifically about “copyright,” or about “patents,” or about “trademarks.”
The term “intellectual property” carries a hidden assumption—that the way to think about all these disparate issues is based on an analogy with physical objects, and our ideas of physical property.
When it comes to copying, this analogy disregards the crucial difference between material objects and information: information can be copied and shared almost effortlessly, while material objects can't be.
To avoid the bias and confusion of this term, it is best to make a firm decision not to speak or even think in terms of “intellectual property”."
Source: http://www.gnu.org/philosophy/words-to-avoid.html#IntellectualProperty

[6] Basically this is achieved with a clause that says, that any further developed software as well any derived work must be licensed as a whole under the sGPL. This ensures, that OSS stays OSS.

(2006), David et al (2003), and Ghosh et al. (2002). Other work, like Gonzalez-Barahona et al. (2008), Robles & Gonzalez-Barahona (2006), and Lancashire (2001), is based on specific data drawn from code of certain OSS projects, mailing lists or platforms like SourceForge. (Robles et al. (2001) provide a combination of both types of data collection.) Regarding the data collection about the location of OSS developers, our paper belongs to the second branch.

3.1 Data about OSS Developers registered at SourceForge

SourceForge, is the largest site hosting OSS-projects.[7] We draw our data about *OSS developers* registered at sourceforge.net from the SourceForge Research Data Archive (SRDA). SRDA is offered by the University of Notre Dame under a special agreement for scientific research. The data-base consists of dumps containing some of the information stored at the SourceForge web-page. The latest dumps containing all information necessary for our analysis were those from 2006. Namely the dump of October 2006 was used to identify users. For activity level of users we could make use of the dumps from January 2006 until December 2006.

When OSS developers register at SourceForge they have to indicate a valid email address. Additionally, when registering developers can change the time zone from the default-value to their specific time zone (e.g. “Europe/Berlin”). Furthermore, the SRDA contains tables with the IP (Internet Protocol address) of the users logged in, and also information about when and whether users posted messages. We were able to identify each user by the user-ID and connect this information with the indicated email address and time zone. Furthermore, we could connect each user to his or her IP, and we were also able to assign to each user the number of posted messages. The latter information is a proxy for the activity of an OSS developer.

The original data of the 2006 sumps delivered approximately 1.4 million datasets which have been cleaned of all duplicates, fake accounts like "nobody@nowhere.com" and non reliable data. Then, we assigned to each user his or her geographical origin by making use of the email address, time-zone and IP. In particular we used the following *methods/indicators*:

- The first two indicators are assigned to the email address:
  - The first is the country coded Top-Level Domain (ccTLD) of the email address. Thus, the assumption is, that each user's ccTLD correctly indicates his or her native country (For example, “.us” for the USA, “.nl” for the Netherlands, or “.de” for Germany). A problem are so-called open ccTLDs, like “.ws” for Western Samoa. While ccTLDs limit registration to citizens or firms of their respective countries, in the case of “open” ccTLDs registration is possible (therefore “open”) to any inter-

[7] See also the footnote on page x.

ested registrant. The reason is that e.g. ".tv" (for Tuvalu) looks like "television", or ".ws" (Western Samoa) looks like "web side". Therefore one can make money by selling such accounts. However, this implies that such open ccTLDs can not be used for geographical identification. In fact, these are "de facto" generic TLDs such as ".org" or ".com". Therefore we excluded all open ccTLDs from the dataset when identifying via country coded Top-Level Domain of the email addresses.

  - For all email accounts with generic TLDs it is possible to use information from the so-called second level domain (SLD). For example in case of "xyz@yahoo.com" is "yahoo" the SLD. It is possible to identify the location of the domain server of a SLD. Therefore we manually assigned to each of the top 1000 SLDs their domain server, and therefore the country of the server. If one now assumes that the location of the domain server of the SLD of a user's email address also indicates the country the user lives in, then it is possible to assign users with generic TLDs to countries. Clearly this method can be criticized as the probability of mistakes might be high. For example a Spanish developer using an yahoo.com email account would be counted as a citizen of the USA. We will come back to this later.

- Another indicator is the time zone indicated. Time zones like "EST" sum up several countries and can therefore not used for the analysis. The same is true if `time zone' has its default value, as it is not known whether the option time zone was just ignored, or not. Thus, members with the default or an summarizing time zone can not be geographically identified via this method. But nevertheless, well-defined and unique time zones can be used to assign a country to a user. For example, if one has chosen the time zone "Europe/Berlin", then this can be assigned to Germany. Clearly the assumption standing behind this is, that users report their time zone correctly (when changing it from the default value "CET") and that this indicates their usual place of residence.
- Finally one can make use of the information offered by the saved IP we could draw from the SRDA. We used the partially available IPs of each user and calculated their actual habitation by GeoIP. GeoIP allows to identify geographic location of Internet-connected devices via their IP-range. Namely the location of servers of Internet service providers, Universities, etc, can easily be identified. Via these "providers", the geographic location of Internet users can be identified quite correctly.[8] However, if unlucky, the IP belongs to a range that is assigned to regions but not to certain countries.

Identifying the geographical origin of OSS developers via ccTLD, IP and indicated time zone seems to be quite reliable because of theoretical reasons. In order to get an impression of

[8] To try out how exact this can be, simply visit http://www.netip.de.

the reliability we cross-checked the results that ccTLD, IP and indicated time zone deliver. ccTLD and IP had a matching of 89.16%, IP and time zone delivered the same results in 87.29% of all cases, and time zone and ccTLD had 80.45%. As already mentioned above, identifying the location via SLD is from a theoretical point of view the weakest method. Thus, not surprisingly, checking IP with SLD, and time zone with SLD delivered matching rates of only 51.83%, and 56.45% respectively.

Therefore we combined all four methods in the following way: First, when possible, we identified users' geographical location via GeoIP. The remaining users were then identified via their ccTLD, if possible. The rest was then assigned to their country using the information about the time zone. The remaining 283,028 not located users were then assigned to a country using the information about the SLD. Doing so we end up with 1,315,263 users who are assigned to their countries (83,217 users, i.e. 6% could not be identified). However, as one might doubt the results using the SLD, we always run regressions with data containing the SLD-based identifications, and without.

As already mentioned we were also able also extract the information whether, and if, how often, a user posted a message in 2006. This was used as an indicator of activity. Therefore were were able to distinguish active developers (developers who had posted in 2006). Furthermore, counting the number of messages posted by users from a country delivered us data about the OSS activity that comes from a specific country.

Weighting all these information by the number of inhabitants in 2006 (source: Worldbank), we finally end up with the following country-specific informations:

- Number of OSS developers per 1,000 inhabitants
- Number of active OSS developers per 1,000 inhabitants
- Level of OSS activity (Number of posted messages per 1,000 inhabitants)

Especially as we have information about activity levels, our data offer information about global OSS activities than any other non-survey data we are aware of. A graphical impression of these results are given by the “World Map of OSS Developers” (figure 3), the “World Map of Active OSS Developers” (figure 4), and the “World Map of OSS Activities” (figure 5).

--- [Figure 3 about here] ---

--- [Figure 4 about here] ---

--- [Figure 5 about here] ---

3.2 Data on Cultural Factors and Social Capital

One main source is for our analysis are the so-called "World Values Surveys", that offer a wide range of country-specific, and hence culture-specific data, and are often used in cross-cultural research. We made use of this for our variables about *cultural factors* as well as *social capital*. We used data from the waves of 1990, 1995/1998 and 1999/2000 as one can receive them online (www.worldvaluessurvey.org, go to Online Data Analysis). However, not all questions were asked in all countries, and additionally not in all interviews. Thus we had to correct for that and eliminate all those with too little overall coverage.[9] Some of the questions had a scale for the possible answer (like e.g. 1 = strongly disagree, up to 10 = strongly agree). Although it is very common to use the mean of such answers, this is quite critical from a methodological point of view, as in such a case one treats ordinal scaled data like being on an interval scale. It is better to choose a certain threshold, i.e. for example count the percentage of answers with scale 4 or smaller. As we want to be able to distinguish groups (here countries) from each other with respect to a certain characteristics, a good way to find such a threshold is to "ask the data". Thus, we looked at the direction the answers point to, choose those of interest, and then set different plausible thresholds. In a next step we compared the variance, and choose c.p. those with the higher variance, as this is an indication by the data that we made the right cut in order to measure the difference of the respective category. (If variances were close to each other, we choose those threshold with the distribution closer to the normal distribution). However, whenever we refer to WVS data henceforth, they were treated in the way just described if necessary.

3.2.1 Social Capital: Interpersonal Trust

With respect to *interpersonal trust*, we made use of the data offered by the World Values Survey regarding social, or: interpersonal, trust. Interpersonal trust was measured by the average percentage of respondents saying 'most people can be trusted'.(The question is "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" Possible answers are either "Most people can be trusted", or "Can´t be too careful"). We labeled this variable as "IntPer_Trust".

[9] In case of questions that were not asked in 100% of all interviews in a country (but with a sufficient high percentage) we additionally had to correct the percentage of answers, as the numbers one receives from the WVS online-dataset always sum up with the "not asked"- share to 100%.

3.2.2 Cultural Aspects: Degree of Individualism/Self-Determination, and Attitudes Toward Competition/Merit Principle

As already mentioned before, a cultural dimension of interest is *individualism/self-determination.* The degree of individualism is one famous category in cross-cultural studies. Hofstede's individualism „pertains to societies in which the ties between individuals are loose: everyone is expected to look after himself or herself and his or her immediate family“ (Hofstede 2001 pX). Hofstede developed the individualism index for 50 countries based on a worldwide survey of IBM employees that was carried out during 1978-83. The questions the individualism index was build upon asked about whether the job leaves sufficient time for personal and family life, considerable freedom to adopt own approach, includes challenging work to do, offers opportunities to improve and learn new skills, etc. (Hofstede 2001, p X). Hence, based on these categories, high scores in individualism indicate the prevalence of individual interest in a society, i.e. in a sense that people would like (and can) “do what they want to do”. However, we use an individualism measure developed and used by Diener, Gohm, Oishi, Suh, and Triandis. Basically, this is an updated and further developed version of Hofstede’s measure (namely a merge of ratings provided by Triandis and Hofstede's scores, see e.g. Suh et al. 1998, p 485). We collected the values for this measure from Diener et al. (2000), Oishi (2000), and Suh et al. (1998).

Obviously individualism (in the tradition of Hofstede's definition) should correlate with leisure time preferences, preferences for independence and self responsibility, etc. Here we could again make use of the WVS data. Treating the data as described above lead to the following categories for “leisure time” and “self-responsibility”:

- Leisure time: % of all respondents of a country saying “1 Very important” (WVS-Question was asking about how important leisure time is in ones life)
- Self-responsibility: % of all respondents of a country ranging from 1 to 4. (WVS Question was asking to put oneself on a range 1 to 10 expression own opinion, with 1 = People should take more responsibility to provide for themselves, vs 10 = The government should take more responsibility to ensure that everyone is provided for.)

Additionally the WVS data delivered the percentage of all respondents of a country who mentioned that “Feeling of responsibility” is an important quality children shout learn at home. (They were given a list of qualities that children can be encouraged to learn at home. They should choose up to five they consider to be especially important.).

For our analysis we wanted to have one measure for the degree of *individualism/self-determination*. Therefore we made use of a tool that is widely used in cultural studies: the so-called principal component analysis. The idea of this multivariate statistical technique is the

following: Assumed a (or more) certain characteristic(s) (e.g. "culture") can not be measured directly, but one has several indicators for this characteristic(s). Then the principal component analysis (PCA) is a useful tool to identify the meaningful underlying variable(s) and construct this based on the data available. In other words, the PCA tries to find components that explain the maximum amount of variance. Thus the goals of a PCA is to reduce the dimension of the data and to detect structure in the relationships between variables, that is to classify variables. Therefore one takes a set of variables of whom one expects a relationship because of theoretical reasons (e.g. as one sees them as expressions of the same attitude), and then runs the PCA, that is minimize the sum of the squared perpendicular distances to the axis of the principal component. We did so, and finally constructed a PCA-component labeled "SelfDet_Indiv", that consist of the individualism scores mentioned above, "leisure time", "self-responsibility" and whether a child should learn responsibility. "SelfDet_Indiv" has a Kaiser-Meyer-Olkin measure of sampling adequacy of about 0.7 (i.e. 0.683).

As measure for the degree of positive attitudes toward *competition and/or the merit principle* we were also able to construct a PCA based component, labeled "Comp_Merit" consisting of variables from the WVS, measuring attitudes towards income differences as incentives, whether competition is perceived as good or harmful, and the importance to teach a child independence:

- Income differences as incentives: % of respondents of a country ranging from 10 to 7 regarding "Incomes should be made more equal vs We need larger income differences as incentives" (The WVS question was asking to put oneself on a range 1 to 10 expression own opinion, with 1 = Incomes should be made more equal, vs 10 = We need larger income differences as incentives.)
- Competition is good : % of respondents of a country ranging from 1 to 2 (Question asked to range oneself according to opinion about "Competition is good, it stimulates people to work hard and develop new ideas, vs. Competition is harmful, It brings the worst in people. Range: 1 = Competition is good, 10 = Competition is harmful.)
- Importance to teach a child independence: % of all respondents of a country who mentioned that "Independence" is an important quality children shout learn at home. (They were given a list of qualities that children can be encouraged to learn at home. They should choose up to five they consider to be especially important.)

However, the PCA based component "Comp_Merit" might be problematic, as it has a Kaiser-Meyer-Olkin measure of sampling adequacy of about 0.5 (i.e. 0.517). Therefore we run regression with and without "Comp_Merit".[10]

[10] We also run regression with "Competition is a good thing" only. However the results were quite similar, therefore we did not put them into this paper.

3.3 Data on Attitudes towards Novelty

Data on attitudes towards novelty also come from the 1990, 1995/1998 and 1999/2000 waves of the World Values Survey. We could make use of the following:

- "ScienAdvan_Will_help" is the % of all respondents of a country saying that scientific advances we are making will help mankind. (The question is "In the long run, do you think the scientific advances we are making will help or harm mankind?" Possible answers: 1 Will help, 2 Will harm, 3 Some of each.)
- "Pref_New_Ideas" is the % of all respondents of a country prefering new ideas over old ones by ranging from 8 to 10. (The survey question asks to rate oneself on a scale about "Ideas stood test of time better vs New ideas better", with 1 = Ideas that stood test of time are generally best, up to 10 = New ideas are generally better than old ones.)

3.4 Data on internet users, IPR Protection and Regulation

Next we look for the numbers of internet users per 100 inhabitants ("InetUsers.per_100_inhab"). The data for this come from International Telecommunication Union (2006). In order to evaluate the degree of regulation, we used the data offered with the Economic Freedom of the World Index (Gwartney et al. 2008). The report offers an inverse index of regulation, called "Area 5: Regulation of Credit, Labor, and Business". We made use of this index in order to measure the degree of regulation in a society in 2006, and denoted this variable by "Regulation2006". This index is build upon several sub-indices measuring credit market regulations, labor market regulations, and business regulations (Gwartney et al. 2008, p 189ff). With respect to IPR, we made use of one of the sub-indices of Gwartney et al. (2006) belonging to the *property right* section: the sub-index of the protection IPR ("2C Protection of intellectual property") for the year 2004, the latest IPR-data available. This IPR sub-index is based on data from the The Global Competitiveness Report of the World Economic Forum.[11] We will denote this index by "IPR_Protection2004". Another measure related to IPR are the figures about the software piracy rates in 2006, taken from the Fifth Global Software Piracy Study (Business Software Alliance 2007). This was denoted by "PiracyRates2004".

3.5 Data for the Control Variables

Obviously, the probability that a country's inhabitant becomes an OSS developer rises in the degree of economic and technical development, in the latter case mainly with the access to the Internet. Thus, we have to control for this. Hence, we take into account the GDP per capita

[11] The question was whether "Intellectual property protection in your country is 1 = weak and not enforced, up to 7 = strong and enforced".

ppp for 2006 ("GDP2006"), data source is Worldbank (2007). We also would like to control for the importance of the software industry in a country. Unfortunately data for a sufficient number of countries was only available for the whole ICT-Sector only (share of workers). But clearly we run into problems of multicollinearity with internet users (see section 3.4) as internet access must have been installed by someone working for the IT sector. Therefore we decided to leave this out. Furthermore, we control for education, because (a) previous studies indicated that OSS developers are well-educated software engineers (or IT students), and (b) in order to be able to write software code (i.e. programming) one must be able to think in abstract terms and logic. Additionally, most programming languages are based on English and the whole communication and coordination of OSS projects is done in English. As measure for education we used the combined gross enrollment ratio for primary, secondary and tertiary schools with a four-year lag (EDUC2002, i.e. of the year2002). The source is UNDP (2004).

4. Empirical Results

This section is dedicated to the presentation and discussion of the empirical results. To test hypotheses H1 through H8, we run linear regression models (OLS) based on our hypotheses. We run different models, varying the endogenous variable as well as the set of explanatory variables. The results appear quite robust and are displayed in tables 1 through 3. All three tables are structured as follows: After the control variables, the influence of the variables presenting hypotheses H1 through H8 is shown. We present the three most representative models, each with (for each table equations 1, 3 and 5 respectively) and without "Comp_Merit". Social trust, Internet users and IPR[12] is used across the board, other variables are skipped in single equations. We are able run regressions with up to 70 countries, and we are able to distinguish with respect to the level of contribution.

Table 1 presents the regression results for model 1, using the number of OSS developers per 1,000 inhabitants including those localized using the information about the SLD. We analyze three sub-models. In the Appendix the reader can find the same regressions for OSS developers data *without* those localized using the information about the SLD. In a second model, we run regressions for the active OSS developers per 1,000 inhabitants, again including those localized using the information about the SLD (for results *without* SLD see Appendix). Finally, we analyze the OSS activity level, as usual here we present the results with those located via SLD (for the other version see Appendix).

[12] We also run regression with "PiracyRates2004" instead of "IPR_Protection2004". The results were basically the same, beside the fact that the regulation measure was no longer significant, and---more important---"PiracyRates2004" itself was *never* significant.

Table 1: *Dependent Variable:* **OSS Developers per 1,000 inhabitants (with SLD)**

Equation	1.1	1.2	1.3	1.4	1.5	1.6
(Intercept)	-1,93e+003**	-1,91e+003**	-1,09e+003*	-9,59e+002*	-1,16e+003*	-1,15e+003*
	(0.00403)	(0.00363)	(0.0255)	(0.0421)	(0.026075)	(0.023850)
Control Variables:						
GDP2006	-4,21e-003	-3,89e-003	9,82e-003	8,27e-003	-4,45e-003	-4,30e-003
	(0.64485)	(0.66160)	(0.1757)	(0.2459)	(0.634357)	(0.636556)
Educ2002	1,98e+000	2,00e+000	3,12e+000	2,30e+000	5,57e-001	5,67e-001
	(0.67257)	(0.66737)	(0.4771)	(0.5955)	(0.906606)	(0.903977)
Explaining Variables:						
SelfDet_Indiv	1,17e+002*	1,15e+002*	5,73e+001	6,85e+001•	1,23e+002*	1,22e+002**
	(0.01395)	(0.01210)	(0.1631)	(0.0872)	(0.011507)	(0.009142)
Pref_New_Ideas	1,68e+002	1,52e+002			6,58e+001	5,84e+001
	(0.69915)	(0.71943)			(0.881925)	(0.891947)
ScienAdvan_Will_help	1,25e+003**	1,22e+003**			1,17e+003*	1,15e+003**
	(0.00672)	(0.00390)			(0.012603)	(0.007141)
IntPer_Trust	1,17e+003*	1,17e+003*	6,50e+002•	5,92e+002	1,16e+003*	1,17e+003*
	(0.01171)	(0.01029)	(0.0950)	(0.1250)	(0.014211)	(0.012804)
Comp_Merit	-6,76e+000		3,37e+001		-3,10e+000	
	(0.84849)		(0.2587)		(0.931688)	
InetUsers.per_100_inhab.	1,37e+001**	1,37e+001**	1,02e+001*	1,01e+001*	1,54e+001***	1,54e+001***
	(0.00213)	(0.00194)	(0.0150)	(0.0161)	(0.000673)	(0.000595)
Regulation2006	1,05e+002•	1,04e+002•	8,52e+001	8,28e+001		
	(0.06442)	(0.06286)	(0.1243)	(0.1359)		
IPR_Protection2004	3,73e+001	3,76e+001	3,68e+001	3,56e+001	5,12e+001	5,13e+001
	(0.40249)	(0.39493)	(0.3712)	(0.3881)	(0.257374)	(0.251981)
D. of Freedom	49	50	61	62	50	51
Adjusted R^2	0.8147	0.8183	0.7845	0.7835	0.8052	809

The values in brackets are the p-values. Significance levels are denoted by *** = 99.9%, ** = 99%, * = 95%, and • = 90%.

Table 2: *Dependent* Variable*:* **Active Developers per 1,000 inhabitants (with SLD)**

Equation	2.1	2.2	2.3	2.4	2.5	2.6
(Intercept)	-4,23e+002**	-4,24e+002**	-2,45e+002*	-2,14e+002*	-2,43e+002*	-2,45e+002*
	(0.00266)	(0.00213)	(0.0171)	(0.0330)	(0.02701)	(0.02271)
Control Variables:						
GDP2006	-1,36e-003	-1,37e-003	2,82e-003•	2,43e-003	-1,42e-003	-1,47e-003
	(0.47665)	(0.46059)	(0.0671)	(0.1090)	(0.47329)	(0.44481)
Educ2002	4,99e-001	4,99e-001	5,35e-001	3,30e-001	1,66e-001	1,62e-001
	(0.61053)	(0.60722)	(0.5633)	(0.7194)	(0.86826	(0.86978)
Explaining Variables:						
SelfDet_Indiv	2,70e+001**	2,71e+001**	1,28e+001	1,57e+001•	2,85e+001**	2,88e+001**
	(0.00679)	(0.00484)	(0.1389)	(0.0658)	(0.00577)	(0.00372)
Pref_New_Ideas	4,20e+001	4,27e+001			1,81e+001	2,07e+001
	(0.64402)	(0.62869)			(0.84644)	(0.81925)
ScienAdvan_Will_help	2,15e+002*	2,16e+002*			1,95e+002*	2,01e+002*
	(0.02435)	(0.01322)			(0.04553)	(0.02475)
IntPer_Trust	2,59e+002**	2,59e+002**	1,21e+002	1,07e+002	2,58e+002**	2,57e+002**
	(0.00755)	(0.00685)	(0.1383)	(0.1897)	(0.00993)	(0.00931)
Comp_Merit	2,65e-001		8,42e+000		1,12e+000	
	(0.97136)		(0.1814)		(0.88331)	
InetUsers.per_100_inhab.	2,43e+000**	2,44e+000**	1,81e+000*	1,79e+000*	2,82e+000**	2,83e+000**
	(0.00820)	(0.00752)	(0.0388)	(0.0424)	(0.00268)	(0.00237)
Regulation2006	2,45e+001*	2,45e+001*	2,02e+001•	1,96e+001•		
	(0.03925)	(0.03684)	(0.0845)	(0.0960)		
IPR_Protection2004	1,47e+001	1,47e+001	1,15e+001	1,12e+001	1,79e+001•	1,79e+001•
	(0.11857)	(0.11493)	(0.1854)	(0.1998)	(0.06289)	(0.06064)
D. of Freedom	49	50	61	62	50	51
Adjusted R^2	0.8288	0.8322	0.8019	0.7992	0.8169	0.8204

The values in brackets are the p-values. Significance levels are denoted by *** = 99.9%, ** = 99%, * = 95%, and • = 90%.

Table 3: *Dependent Variable:* **OSS Activity Level (Messages per 1,000 inhabitants, with SLD)**

(Intercept)	-2,97e+004**	-2,94e+004**	-1,65e+004*	-1,44e+004•	-1,84e+004*	-1,83e+004*
	(0.004832)	(0.004299)	(0.0361)	(0.05951)	(0.024127)	(0.021781)
Equation	3.1	3.2	3.3	3.4	3.5	3.6
Control Variables:						
GDP2006	-1,70e-001	-1,66e-001	2,54e-001*	2,29e-001•	-1,73e-001	-1,72e-001
	(0.239417)	(0.238119)	(0.0327)	(0.05040)	(0.239166)	(0.230762)
Educ2002	5,71e+001	5,73e+001	4,71e+001	3,36e+001	3,63e+001	3,64e+001
	(0.439204)	(0.433200)	(0.5088)	(0.63387)	(0.625250)	(0.620679)
Explaining Variables:						
SelfDet_Indiv	2,77e+003***	2,74e+003***	1,56e+003*	1,75e+003**	2,86e+003***	2,85e+003***
	(0.000346)	(0.000236)	(0.0209)	(0.00824)	(0.000287)	(0.000176)
Pref_New_Ideas	3,76e+003	3,55e+003			2,26e+003	2,18e+003
	(0.582851)	(0.593131)			(0.744180)	(0.745654)
ScienAdvan_Will_help	1,25e+004•	1,21e+004•			1,13e+004	1,11e+004•
	(0.077820)	(0.061658)			(0.116069)	(0.089693)
IntPer_Trust	2,66e+004***	2,67e+004***	1,35e+004*	1,25e+004*	2,65e+004***	2,66e+004***
	(0.000398)	(0.000324)	(0.0343)	(0.04728)	(0.000518)	(0.000433)
Comp_Merit	-8,63e+001		5,56e+002		-3,29e+001	
	(0.876493)		(0.2504)		(0.953662)	
InetUsers.per_100_inhab.	1,59e+002 *	1,59e+002*	1,18e+002•	1,16e+002•	1,83e+002**	1,83e+002**
	(0.020553)	(0.019484)	(0.0784)	(0.08277)	(0.007959)	(0.007356)
Regulation2006	1,53e+003•	1,53e+003•	1,23e+003	1,19e+003		
	(0.084356)	(0.082292)	(0.1694)	(0.18452)		
IPR_Protection2004	1,02e+003	1,02e+003	5,31e+002	5,11e+002	1,22e+003•	1,22e+003•
	(0.148537)	(0.143294)	(0.4267)	(0.44538)	(0.086485)	(0.083162)
D. of Freedom	49	50	61	62	50	51
Adjusted R^2	0.8422	0.8453	809	808	0.8356	0.8388

The values in brackets are the p-values. Significance levels are denoted by *** = 99.9%, ** = 99%, * = 95%, and • = 90%.

5. Comparison and interpretation of the results

In this section we compare and interpret the results of the different models. The control variables do not contribute to the explanation of OSS activities. The evidence for the other variables is mixed. We start by discussing the intrinsic motivations of OSS developers.

H1: the degree of individualism/self-determination of a society has a positive impact on the number of OSS developers as well as on the OSS activity level.

As for hypotheses 1, stating the degree self-determination of a society has a positive impact on the number of OSS developers as well as on the OSS activity level, the number of OSS developers at SourceForge is indeed positively correlated with the degree of individualism. Interestingly the significance level rises when it comes to the *active* developers (table 2), and even more, when one looks at the activity level (table 3). This fits our expectations. The way of software production of OSS is often characterized as collective. At the same time however, the OSS contributors are volunteers, and no one can force them (i.e. command and control like in firms) to do certain things. Hence, those individuals voluntarily participate in an OSS project when they want, and in a way they want. In other words, being an (active) OSS developer can be a way of individualistic self-fulfillment. Therefore, it is highly plausible that societies with high account of self-determination are experiencing a higher OSS activity.

H2: A preference for new ideas has a positive impact on the number of OSS developers as well as on the OSS activity level.

Surprisingly "Pref_New_Ideas" is positively but not significantly correlated with OSS activities. Thus, country-wide openness to new ideas is not encouraging participation in OSS. Interestingly, this is different with respect to the attitude towards scientific progress.

H3: A positive attitude towards scientific progress has a positive impact on the number of OSS developers.

A positive attitude towards scientific progress is clearly significant with respect to the number of developers. It is also significant with respect to active developers, and the activity level. These results point to the direction we expected although it do not fit our expectations to 100%, as we expected no impact on active developers and activity.

H4: Social Capital in terms of interpersonal trust has a positive impact on the number of OSS developers as well as on the OSS activity level.

The number of OSS developers is positively correlated the degree of interpersonal trust. Therefore H3 is not rejected. This fit our expectations. Again, despite the fact that this variable is highly significant throughout all equations, it is interesting to notice that this factor is more significant when it is about the active developers, the activity level respectively. In a society generating mutual trust, the provision of public goods indeed seems more likely.

Next to the institutional and cultural setting responsible for intrinsic motivations, we briefly discuss the extrinsic motivations.

H5: A culture of positive attitudes toward competition and the merit principle has a positive impact on the number of OSS developers as well as on the OSS activity level.

The PCA component "Comp_Merit" was never significant (nor was the positive attitude towards competition solely). Therefore H2 has to be rejected. A possible explanation could be, that individualistic self-fulfillment aspects and self-determination are more important on the level of culture.

H6: A high number of Internet users is beneficial both for the number of OSS developers and the OSS activity level.

The number of Internet users is positively correlated with OSS activities. This can be interpreted from both a demand and a supply perspective. The higher the number of user, the higher the number of software producers seems to be – thus the supply channel. At the same time, a higher number of Internet users increases the number of potential "customers". However, this effect is probably less direct, as the OSS can be used beyond national borders.

H7: A high degree of intense economic regulation has a negative impact on the number of OSS developers as well as on the OSS activity level.

In some regressions the inverse measure of regulation had a positive sign and was significant. Thus H7 cannot be rejected. Hence, one could conclude that OSS activities depend on regulations, exactly as other economic activities are positively correlated with a set of reasonable regulations.

H8: The protection of intellectual property rights has a positive impact on the number of OSS developers or on OSS activity level.

In regressions without the regulation measure the number of active OSS developers and the level of activity both are positively correlated with the degree of protection of intellectual property rights. Hence, H7 cannot be rejected. It seems to be, that OSS also relies on the security of intellectual property rights. This is plausible if one remembers that, as already mentioned, OSS licenses are build upon copyright law. Thus, it seems as if OSS relies on the idea of intellectual property rights, although it uses this institution in "new" way. The deny of intellectual property rights as such might even harm OSS, as then for example Stallman's GPL could not be enforced anymore.

In sum, the results in tables 1 through 3 suggest that OSS activities have both intrinsic and extrinsic foundations.

5 Conclusions

The paper presents a cross-country study of how the (relative) number of OSS developers per inhabitants and the OSS activities of a country depend on institutional and cultural factors. There are both intrinsic and extrinsic motivations relevant; institutions and cultural aspects that foster self-determination, scientific progress and interpersonal trust, it is more likely to find (active) OSS developers. Similarly, we find no evidence that OSS developers do not respect intellectual property, obviously countries with strong IPR show more OSS activity.

References

Allik, J. & Realo, A. 2004, 'Individualism-Collectivism and Social Capital', *Journal of Cross-Cultural Psychology* **35**(1), 29–49.

Asay, M. 2009, 'Open source: It's about capitalism, not freebies.' *cnet online article,* February 9, 2009 6:02 AM PST, news.cnet.com/8301-13505_3-10159370-16.html

Baake, P. & Wichmann, T. 2004, Open source software, competition and potential entry., Berlecon Research Papers 5, Berlecon Research, Berlin.

Bessen, J. 2006, Open source software: Free provision of complex public goods, *in* J. Bitzer P. Schröder, eds, 'The Economics Of Open Source Software Development', Elsevier, p. 57–81.

Bitzer, J. 2004, 'Commercial versus open source software: the role of product heterogeneity in competition', *Economic Systems* **28**(4), 369–381.

Brand, A. & Schmid, A. 2005, Koordination in einem Open Source-projekt, Technical report.

Business Software Alliance (2007): *Fifth Annual BSA and IDC Global Software Piracy Study.*

Casadesus-Masanell, R. & Ghemawat, P. 2003, Dynamic mixed duopoly. a model motivated by linux vs. windows, IESE Research Papers 519, IESE Business School, Barcelona.

D'Antoni, M. & Rossi, M. A. 2007, Copyright vs. copyleft licencing and software development, Working Paper of the Department of Economics University of Siena 510, Department of Economics, University of Siena.

David, P & Waterman, A. & Arora, S. 2003, FLOSS-US. the free/libre/open source software survey for 2003, Technical report, Stanford Institute for Economic Policy Research.

De Cremer, D. 1999, 'Trust and fear of exploitation in a public goods dilemma', *Current Psychology,* **18**(2), p 153-163.

Diener, E. & Gohm, C. L. & Suh, E. & Oishi, S. 2000, 'Similarity of the relations between marital status and subjective well-being across cultures', *Journal of Cross-Cultural Psychology* **31**(4), 419–436.

Elinor O. 1998, 'A Behavioral Approach to the Rational Choice Theory of Collective Action', *American Political Science Review,* **92**(1), p 1-21

Gehring, R. A. 2006, 'The institutionalization of open source', *Poiesis & Praxis: International Journal of Technology Assessment and Ethics of Science* **4**(1), 54–73.

Ghosh, R. A. & Glott, R. & Krieger, B. & Robles, G. 2002, FLOSS final report, part 4: Survey

of developers, Technical report, International Institute of Infonomics, University of Maastricht.

Ghosh, R. A. 2006, Economic impact of open source software on innovation and the competitiveness of the information and communication technologies (ICT) sector in the EU, Study, European Commission.

Glass, R. L. 2004, 'A Look at the Economics of Open Source.' *Communications of the ACM*, 47(2), p25-27.

Gonzalez-Barahona, J. M. & Robles, G., Andradas-Izquierdo, R. & Ghosh, R. A. 2008, 'Geographic origin of libre software developers', *Information Economics and Policy* **20**(4), 356 – 363. Empirical Issues in Open Source Software.

Gwartney, J. & Lawson, R. & Norton, S. 2008, Economic freedom of the world: 2008 annual report, Report, The Fraser Institute. Data retrieved from www.freetheworld.com.

Hars, A. & Qu, S. 2002, `Working for Free? Motivations for Participating in Open-Source Projects', *International Journal of Electronic Commerce*, **6**(3), p 25–39.

Henkel, J. 2006, The jukebox mode of innovation - a model of commercial open source development, DRUID Working Papers 06-25, DRUID, Copenhagen Business School, Department of Industrial Economics and Strategy/Aalborg University, Department of Business Studies.

Hertel, G. & Niedner, S. & Herrmann, S. 2003, "Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel," Research Policy, Elsevier, vol. 32(7), pages 1159-1177, July.

Hofstede, G. 2001, *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*, 2nd edn, Sage Publications.

Inglehart, R. 1997, *Modernization and postmodernization: cultural, economic, and political change in 43 societies*, Princeton University Press, Princeton, N.J.

International Telecommunication Union 2006, ICT Statistics 2006, Part 4 – Internet Indicators: subscribers, users and broadband subscribers, Technical report, International Telecommunication Union. http://www.itu.int/ITU-D/ICTEYE/Reports.aspx.

Johnson, J. P. 2002, 'Open source software: Private provision of a public good', *Journal of Economics & Management Strategy* **11**(4), 637–662.

Kugler, P. 2005, Coordinating innovation: evidence from open source software development.

Lakhani, K. R. & Wolf, R. G. 2005, Why Hackers Do What They Do: Understanding Motiva-

tion Efford in Free/Open Source Software Projects. in: Feller, J., Fitzgerald, B., Hissam, S. & Lakhani, K. R. (eds.). Perspectives on Free and Open Source Software. MIT Press.

Lakhani, K., Wolf, B., Bates, J., DiBona, C., 2002, The Boston Consulting Group Hacker Survey. Available at: http://www.osdn.com/bcg.

Lancashire, D. 2001, 'Code, culture and cash: The fading altruism of open source development', *First Monday* **6**(12). http://firstmonday.org.

Lerner, J. & Pathak, P. A. & Tirole, J. 2006, 'The dynamics of open-source contributors', *The American Economic Review* **96**(2), 114–118.

Lerner, J. & Tirole, J. 2000, *The Simple Economics of Open Source.*, number 7600 *in* 'NBER Working Paper Series'.

Lerner, J. & Tirole, J. 2002, 'Some simple economics on open source', *Journal of Industrial Economics* **50**(2), 197–234.

Luthiger Stoll, B. 2007, Spass und Software-Entwicklung: Zur Motivation von Open-Source-Programmierern

Maurer, S. M. & Scotchmer, S. 2006, Open source software: the new intellectual property paradigm, *in* T. Hendershott, ed., 'Handbook of Economics and Information Systems', Elsevier, pp. 285–319.

OECD 2007*a*, Education at a glance 2007, Oecd indicators.

OECD 2007*b*, PISA 2006: Science competencies for tomorrow's world, Pisa studies.

Oishi, S. (2000). Goals as cornerstones of subjective well-being: Linking individuals and cultures. In E. Diener & E. M. Suh (Eds.), *Cross-cultural psychology of subjective well-being*. Boston: MIT Press.

Perkins, G. 1998, *Open Source and Capitalism*, online article, posted at slashdot.org, see slashdot.org/articles/980824/0854256.shtml.

Pisano, G. (2006): `Profiting from innovation and the intellectual property revolution', *Research Policy*, **35**(8).

Ramanujam, P. 2007, Who produces open source software, MPRA Paper 4253, University Library of Munich, Germany.

Robles, G. & Gonzalez-Barahona, J. M. 2006, Geographic location of developers at sourceforge, Proceedings of the mining software repositories workshop.

Robles, G. & Scheider, H. & Tretkowski, I. & Weber, N. 2001, Who is doing it? A research on libre software developers, Technical report, Technische Universitt Berlin.

Rossi, C. & Bonaccorsi, A. 2006, Intrinsic motivations and profit-oriented firms in open source software: Do firms practise what they preach? , *in* J. B. Schröder P. J. H., eds, 'The Economics Of Open Source Software Development', Elsevier, pp. 84–109.

Rossi, M. A. 2006, Decoding the free/open source software puzzle: A survey of theoretical and empirical contributions, *in* J. Bitzer P. Schröder, eds, 'The Economics Of Open Source Software Development', Elsevier, pp. 15–55.

Schiff, A. 2002, 'The economics of open source software. a survey of the early literature', *The Review of Network Economics* **1**(1), 66 – 74.

Söderberg, J. 2008,. *Hacking Capitalism: The Free and Open Source Software Movement*, Routledge.

Suh, E., Diener, E., Oishi, S., & Triandis, H. (1998). `The shifting basis of life satisfaction judgments across cultures: Emotions versus norms`, *Journal of Personality and Social Psychology*, 74, 482-493.

UNDP 2004, *Human Development Report 2004.*

von Hippel, E. & von Krogh, G. 2003, 'Open source software and the "private-collective" innovation model: Issues for organization science', *Organization Science* **14**(2), 209–223.

von Hippel, E. 2005, Open source software projects as user innovation networks - no manufacturer required, *in* J. Feller, B. Fitzgerald, S. A. Hissam K. R. Lakhani, eds, 'Perspectives On Free and Open Source Software', MIT Press, Cambridge, Mass. [u.a.].

Weber, S. 2004, *The Success Of Open Source*, Harvard University Press.

Williamson, O. E. 2000, 'The New Institutional Economics: Taking Stock, Looking Ahead' *Journal of Economic Literature,* **38**(3), p 595-613.

Yamagishi, T. & Yamagishi, M. 1994, 'Trust and commitment in the United States and Japan', *Motivation and Emotion,* **18**(2), p 129–166.

Yamagishi, T. Kanazawa, S., Mashima, R. & Terai, S. 2005, 'Seperating trust from cooperation in a dynamic relationship: Prisoner's dilemma with variable dependence', *Rationality and Society,* **17**(3), p 275–308.

Appendix

Table 4: *Dependent Variable:* **OSS Developers per 1,000 inhabitants (without SLD)**

(Intercept)	-1,84e+003**	-1,81e+003**	-1,05e+003*	-9,53e+002*	-1,10e+003*	-1,08e+003*
	(0.00781)	(0.00775)	(0.0346)	(0.0479)	(0.041180)	(0.04076)
Equation						
GDP2006	-7,77e-003	-7,15e-003	8,43e-003	7,21e-003	-8,01e-003	-7,55e-003
	(0.41305)	(0.43970)	(0.2564)	(0.3212)	(0.408647)	(0.42388)
InetUsers.per_100_inhab.	1,42e+001**	1,42e+001**	1,07e+001*	1,06e+001*	1,58e+001***	1,58e+001***
	(0.00221)	(0.00207)	(0.0127)	(0.0130)	(0.000724)	(0.00066)
Educ2002	2,32e+000	2,35e+000	3,17e+000	2,53e+000	9,38e-001	9,69e-001
	(0.63414)	(0.62717)	(0.4821)	(0.5689)	(0.848501)	(0.84203)
Explaining Variables:						
SelfDet_Indiv	1,13e+002*	1,08e+002*	4,76e+001	5,64e+001	1,19e+002*	1,15e+002*
	(0.02197)	(0.02180)	(0.2573)	(0.1664)	(0.018013)	(0.01653)
IntPer_Trust	1,18e+003*	1,20e+003*	5,87e+002	5,42e+002	1,18e+003*	1,19e+003*
	(0.01377)	(0.01183)	(0.1403)	(0.1683)	(0.016124)	(0.01414)
Comp_Merit	-1,33e+001		2,64e+001		-9,76e+000	
	(0.71750)		(0.3874)		(0.794477)	
Pref_New_Ideas	2,05e+002	1,72e+002			1,05e+002	8,23e+001
	(0.65094)	(0.69505)			(0.818084)	(0.85336)
ScienAdvan_Will_help	1,10e+003*	1,04e+003*			1,02e+003*	9,74e+002*
	(0.02026)	(0.01670)			(0.033401)	(0.02640)
IPR_Protection2004	3,17e+001	3,21e+001	3,09e+001	2,99e+001	4,51e+001	4,54e+001
	(0.49339)	(0.48347)	(0.4644)	(0.4772)	(0.333479)	(0.32669)
Regulation2006	-1,02e+002•	-1,00e+002•	8,35e+001	8,16e+001		
	(0.08370)	(0.08391)	(0.1419)	(0.1500)		
D. of Freedom	49	50	61	62	50	51
Adjusted R^2	0.7859	0.7896	0.7585	0.7596	0.7768	0.7809

The values in brackets are the p-values. Significance levels are denoted by *** = 99.9%, ** = 99%, * = 95%, and • = 90%.

Table 5: *Dependent Variable:* **Active Developers per 1,000 inhabitants (without SLD)**

(Intercept)	-4,10e+002**	-4,08e+002**	-2,40e+002*	-2,13e+002*	-2,35e+002*	-2,35e+002*
	(0.00445)	(0.00388)	(0.0226)	(0.0372)	(0.03616)	(0.03251)
Control Variables:						
GDP2006	-1,96e-003	-1,92e-003	2,60e-003	2,26e-003	-2,02e-003	-2,02e-003
	(0.31883)	(0.31629)	(0.1001)	(0.1442)	(0.31930)	(0.30704)
InetUsers.per_100_inhab.	2,52e+000**	2,52e+000**	1,90e+000*	1,88e+000*	2,90e+000**	2,90e+000**
	(0.00793)	(0.00736)	(0.0348)	(0.0368)	(0.00266)	(0.00239)
Educ2002	5,78e-001	5,80e-001	5,63e-001	3,88e-001	2,54e-001	2,54e-001
	(0.56709)	(0.56191)	(0.5542)	(0.6803)	(0.80427)	(0.80212)
Explaining Variables:						
SelfDet_Indiv	2,64e+001**	2,61e+001**	1,12e+001	1,37e+001	2,78e+001**	2,78e+001**
	(0.00997)	(0.00807)	(0.2060)	(0.1153)	(0.00831)	(0.00613)
IntPer_Trust	2,67e+002**	2,68e+002**	1,14e+002	1,02e+002	2,66e+002**	2,66e+002**
	(0.00764)	(0.00673)	(0.1734)	(0.2209)	(0.00972)	(0.00881)
Comp_Merit	-9,00e-001		7,21e+000		-6,99e-002	
	(0.90580)		(0.2643)		(0.99286)	
Pref_New_Ideas	4,95e+001	4,73e+001			2,62e+001	2,60e+001
	(0.59752)	(0.60307)			(0.78423)	(0.77940)
ScienAdvan_Will_help	1,91e+002 •	1,86e+002*			1,72e+002 •	1,71e+002 •
	(0.05063)	(0.03636)			(0.08448)	(0.05944)
IPR_Protection2004	1,35e+001	1,35e+001	1,04e+001	1,01e+001	1,67e+001 •	1,67e+001 •
	(0.16195)	(0.15672)	(0.2461)	(0.2588)	(0.09061)	(0.08734)
Regulation2006	2,38e+001	2,37e+001*	1,98e+001 •	1,93e+001		
	(0.05119)	(0.04927)	(0.1000)	(0.1095)		
D. of Freedom	49	50	61	62	50	51
Adjusted R^2	0.8288	0.8322	0.8019	0.7992	0.8169	0.8204

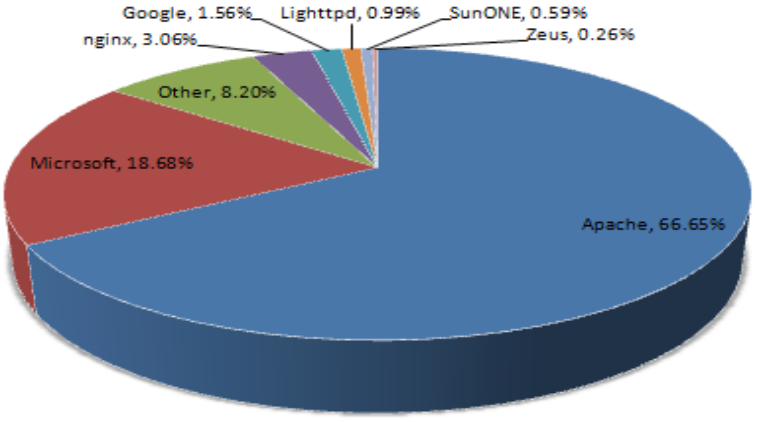
The values in brackets are the p-values. Significance levels are denoted by *** = 99.9%, ** = 99%, * = 95%, and • = 90%.

Table 6: *Dependent Variable:* **OSS Activity Level (Messages per 1,000 inhabitants) (without SLD)**

(Intercept)	-2.901e+01**	-2.867e+01**	-1.624e+01*	• -1,43e+004	-1,79e+004*	-1,78e+004*
	(0.006123)	(0.005653)	(0.0411)	(0.0634)	(0.028580)	(0.026605) *
Control Variables:						
GDP2006	-1.943e-04	-1.885e-04	2,45e-001*	• 2,21e-001	-1,98e-001	-1,94e-001
	(0.181879)	(0.183376)	(0.0419)	(0.0607)	(0.182253)	(0.177955)
InetUsers.per_100_inhab.	1.630e-01*	1.626e-01*	• 1,22e+002	• 1,21e+002	1,87e+002**	1,87e+002**
	(0.018586)	(0.017694)	(0.0718)	(0.0750)	(0.007230)	(0.006725)
Educ2002	5.897e-02	5.922e-02	4,72e+001	3,48e+001	3,85e+001	3,87e+001
	(0.428243)	(0.421764)	(0.5125)	(0.6248)	(0.606626	(0.600836)
Explaining Variables:						
SelfDet_Indiv	2.741e+00***	2.702e+00***	1,50e+003*	1,67e+003*	2,83e+003***	2,81e+003***
	(0.000423)	(0.000308)	(0.0281)	(0.0122)	(0.000348)	(0.000229)
IntPer_Trust	2.676e+01***	2.688e+01***	1,31e+004*	• 1,22e+004	2,67e+004***	2,67e+004***
	(0.000410)	(0.000329)	(0.0421)	(0.0554)	(0.000523)	(0.000431)
Comp_Merit	-1.233e-01		5,11e+002		-7,09e+001	
	(0.825601)		(0.2960)		(0.900988)	
Pref_New_Ideas	4.012e+00	3.711e+00			2,54e+003	2,37e+003
	(0.560732)	(0.579250)			(0.715449)	(0.725953)
ScienAdvan_Will_help	1.148e+01	1.087e+01			1,03e+004	9,93e+003
	(0.107685)	(0.094527)			(0.154491)	(0.131539)
IPR_Protection2004	9.801e-01	9.842e-01	4,92e+002	4,73e+002	• 1,18e+003	• 1,18e+003
	(0.167415)	(0.161472)	(0.4662)	(0.4832)	(0.099634)	(0.095864)
Regulation2006	• 1.505e+00	• 1.494e+00	1,22e+003	1,18e+003		
	(0.092253)	(0.090866)	(0.1798)	(0.1933)		
D. of Freedom	49	50	61	62	50	51
Adjusted R^2	0.8089	0.8126	0.7814	0.7805	0.7974	0.8014

The values in brackets are the p-values. Significance levels are denoted by *** = 99.9%, ** = 99%, * = 95%, and • = 90%.

Figure 1: Server Share amongst the Million Busiest Sites, March 2009



(Source: Netcraft's March 2009 Web Server Survey, www.netcraft.com)

Figure 2: Williamson's four interrelated levels of social and institutional analysis

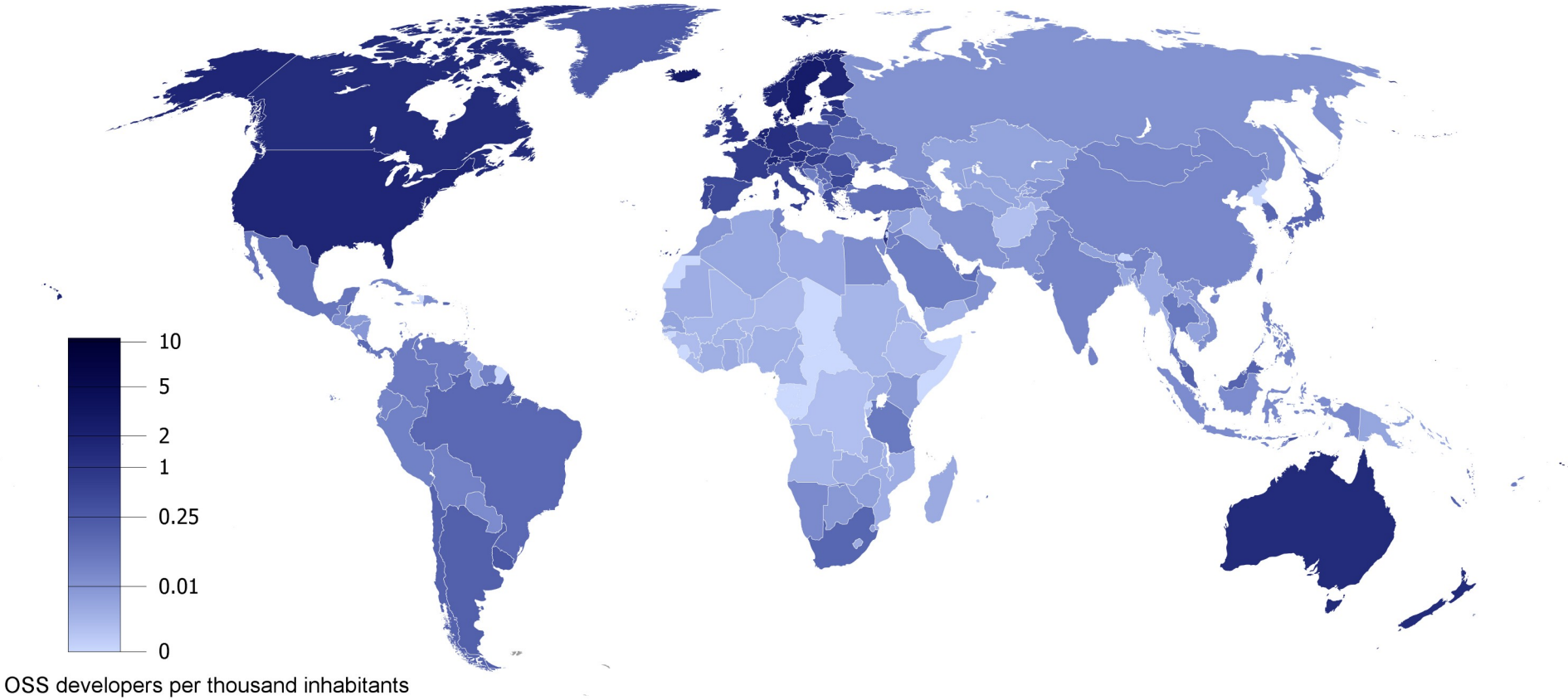
Level		Frequency (years)	Purpose
L1	Embeddedness: informal institutions, customs, traditions, norms religion	10^2 to 10^3	Often noncalculative; spontaneous (caveat: see discussion in text)
L2	Institutional environment: formal rules of the game—esp. property (polity, judiciary, bureaucracy)	10 to 10^2	Get the institutional environment right. 1st order economizing
L3	Governance: play of the game —esp. contract (aligning governance structures with transactions)	1 to 10	Get the governance structures right. 2nd order economizing
L4	Resource allocation and employment (prices and quantities; incentive alignment)	continuous	Get the marginal conditions right. 3rd order economizing

L1: social theory
L2: economics of property rights/positive political theory
L3: transaction cost economics
L4: neoclassical economics/agency theory

Figure 1. Economics of Institutions

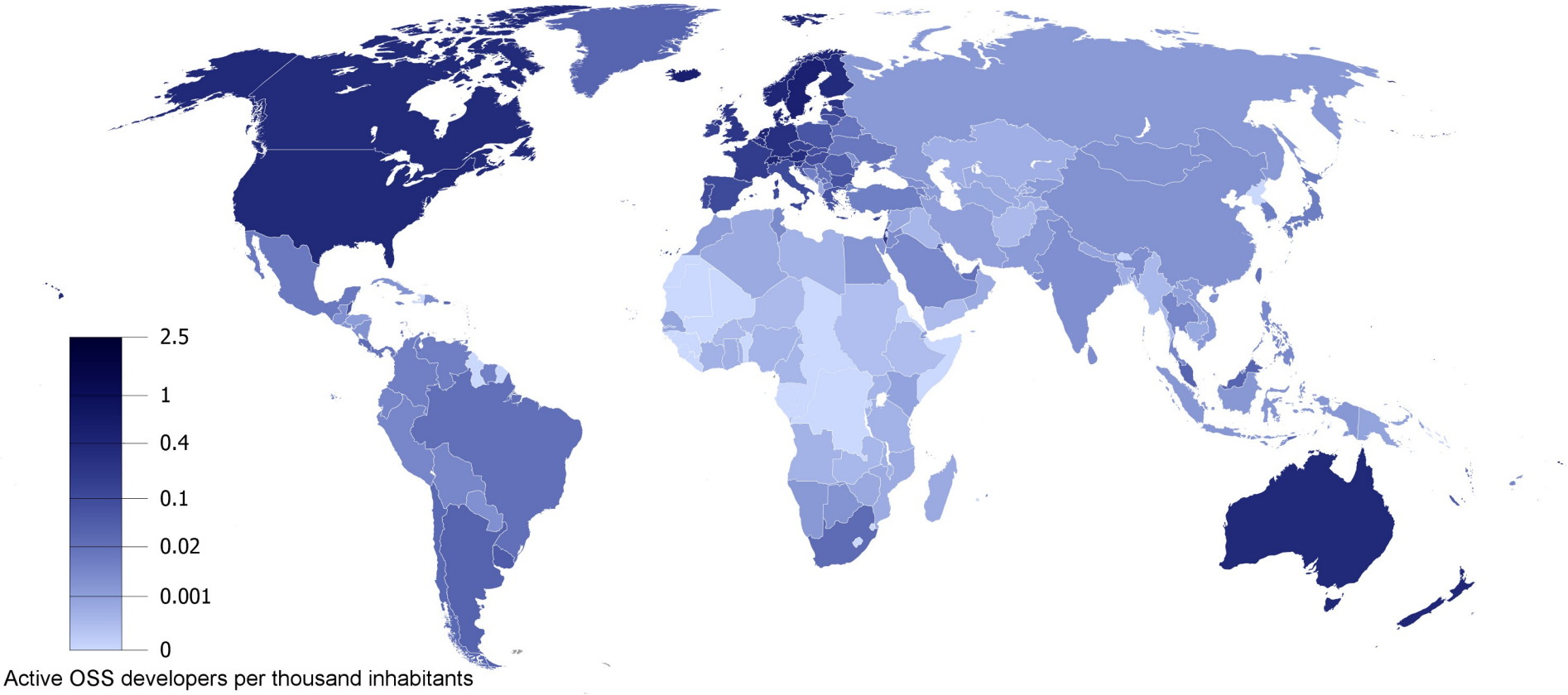
(Source: Williamson 2000, p 597)

Figure 3: World Map of OSS Developers (per thousand inhabitants)



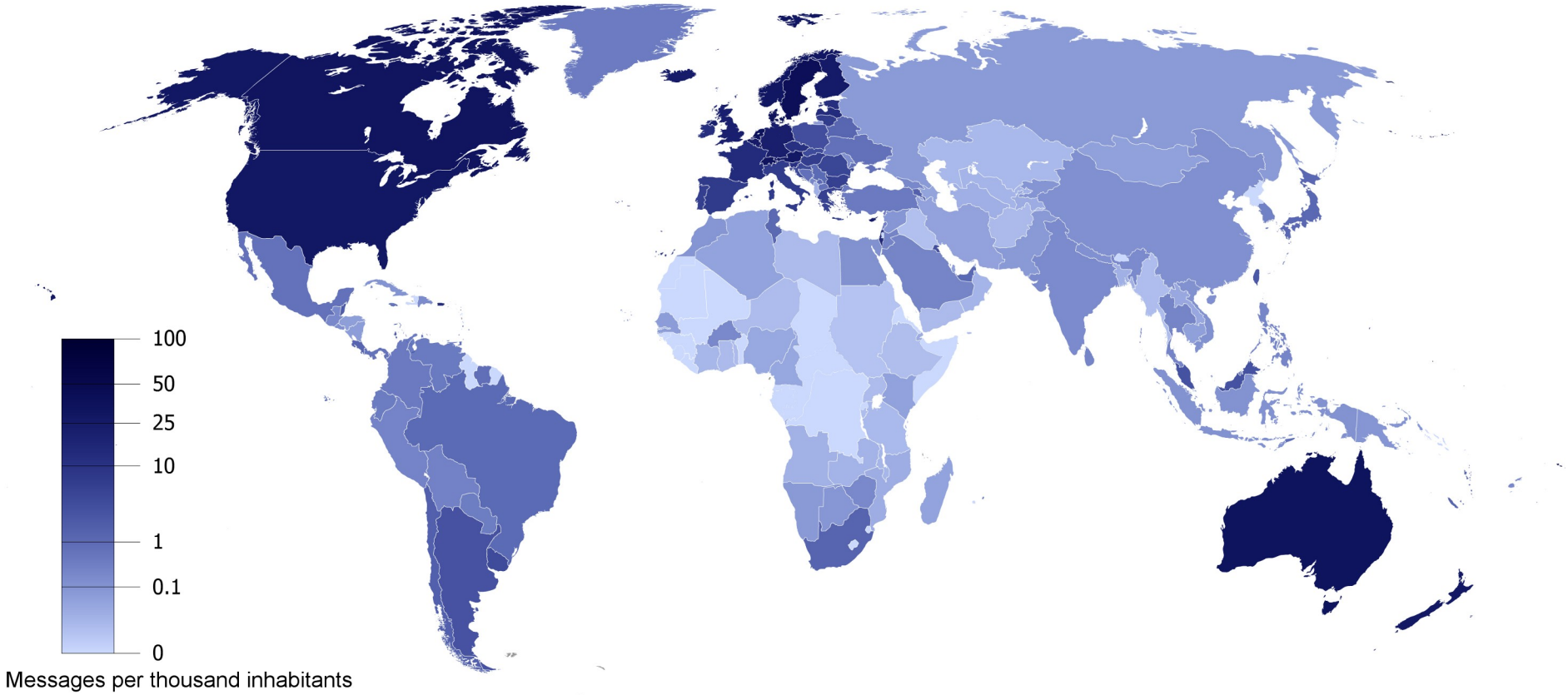
(Source: own data)

Figure 4: World Map of *Active* OSS Developers



(Source: own data)

Figure 5: World Map of OSS Activities



(Source: own data)