

## AN ANALYSIS OF WIKIPEDIA

**MOHAMMAD M. RAHMAN, North South University**

School of Business, Banani C/A, Dhaka 1213, Bangladesh, Tel: +880 (181) 776 8484, Fax: +880 (2) 882 3030,  
Email: [mohammad.m.rahman@gmail.com](mailto:mohammad.m.rahman@gmail.com)

## ABSTRACT

*Wikipedia is defined by its founders as the “free encyclopedia that anyone can edit.” This property, we argue, makes Wikipedia a public good and hence subject to under-provision. A puzzling feature of Wikipedia however is its enormous size, at roughly seven times that of its commercial counterparts. What is driving this growth? And how can we assess the reliability of this giant encyclopedia arising solely from free-editing? We model contribution to Wikipedia and its reliability. We demonstrate that Wikipedia is indeed subject to free-riding, and offer a novel explanation for the mitigation of under-provision under such circumstances. We also find that the public-good feature of Wikipedia and free-riding introduce a lower-bound in the quality of Wikipedia. This finding is consistent with a previous empirical study that established Wikipedia’s surprisingly high level of quality. We identify Wikipedia as part of a general Internet phenomenon that we call the Collaborative Net, and that includes features such as citizen journalism and online reviews.*

## INTRODUCTION

Wiki, the result of an open-source effort, is a relatively new technology that allows Net surfers to freely create and edit Web page content using any Web browser. It is remarkably simple technology, and with its text syntax is equally simple to use. In the words of the Wiki community:

*“Like many simple concepts, ‘open editing’ has some profound and subtle effects on Wiki usage. Allowing everyday users to create and edit any page in a Web site is exciting in that it encourages democratic use of the Web and promotes*

*content composition by nontechnical users.”*

Perhaps the most profound, and without question the most well-known Wiki project is Wikipedia, the “free encyclopedia that anyone can edit.” It is available in several languages including Shqip and Walon; the English version started in 2001 and by December 2008 had more than 2.5 million articles. Googling exotic terms will almost certainly yield links to Wikipedia sites, which may be a sign of its quality, and is certainly a sign of its popularity. Other multilingual free-content projects include Wiktionary, Wikibooks, and Wikinews.

Didem Demirhan acted as the senior editor for this paper.

Rahman, M. M., “An Analysis of Wikipedia,” *Journal of Information Technology Theory and Application (JITTA)*, 9:3 2008, 81-98.

## MOTIVATION AND RESEARCH QUESTIONS

That Wikipedia is offered for free will not be surprising to information economists. Indeed, information goods are characterized by high up-front costs and virtually zero marginal costs of production. Since competition will drive prices to marginal costs, much of information on the Internet will be offered for “free,” (or equivalently, sold with zero price tags). Even Britannica, the well-respected commercial encyclopedia, had to reduce its price for its full volume to a fraction of what it had charged consumers for years.

But what is surprising about Wikipedia, and other free Wiki efforts is that the high up-front costs are not as significant. Wikipedia is a collaborative effort that *anyone* can contribute to. Wiki as a technology reduces publishing costs to virtually zero. A question about the reliability of Wikipedia articles then immediately follows. Can Net users gather information from Wikipedia with a reasonable degree of comfort about its reliability?

Wikipedia has received much criticism in the past about the accuracy of its contents (CNN.com, 2005). Wikipedia often responded by making changes to its design. In December 2005 for example, Wikipedia considered requiring users to register before creating or editing articles. Whereas previously, Wikipedia users did not require any formal registration process to make even the most drastic changes to an article. This was in response to a complaint in an op-ed published in USA Today by a prominent journalist, John Seigenthaler, also a former administrative assistant to Robert Kennedy. An article in Wikipedia had claimed he had been suspected in the assassination of the former attorney general, and President John F. Kennedy (Goodin, 2005). In December 2005, following the Seigenthaler controversy, the New York Times banned their reporters from using Wikipedia as a research tool (Musil 2005).

And yet increasingly, Wikipedia articles *are* being cited in many other traditionally trusted outlets such as newspaper columns (c.f. Drost, 2005). A favorite example cited by Wikipedia themselves is their use by

## CONTRIBUTION

The primary contribution of this paper is its explanation for the paradoxical growth of the public good Wikipedia, a free encyclopedia that anyone can edit. To our knowledge, our explanation is novel and contributes to the general literature on public goods. The paper also offers theoretical explanations for the surprising empirical findings on the superior quality of the giant encyclopedia arising from free-editing. This paper should be highly interesting to researchers engaged in the economic modeling of information systems projects, particularly because of the uniqueness of Wikipedia and its difference from (other) Open Source Systems. It should also be of interest to practitioners, given the increasing popularity of Wikipedia and Wiki technology.

the Parliament of Canada website as a “further reading” resource on the topic of same-sex marriage. Sreenath Srinivasan, dean and professor of Journalism at Columbia University first doubted Wikipedia but soon was surprised by its apparent reliability. Jimmy Wales, Wikipedia’s founder, intends that Wikipedia should achieve a “Britannica or better” quality. This may seem almost unrealistically ambitious, as “Encyclopaedia Britannica” is the oldest English-language general encyclopedia, first published in 1768. Recent articles in the popular press and media however indicated that Wales’ goal with Wikipedia may very well be achievable. CNN.com (2005) citing a study by Nature (Giles, 2005) claimed that “Wikipedia [was] as accurate as Britannica.” They referred to the Seigenthaler case, calling it “the exception rather than the rule.”

Wikipedia’s splendid growth is particularly surprising if it is to be assessed within a “rational” framework. That Wikipedia is free makes it a public good. While contributors may get some benefit from it, many other users will simply try to free-ride. Traditional economic theory suggests that Wikipedia as a public good will be subject to the tragedy of under-provision. Such under-

provision of public goods is usually mitigated through flavors of commercialization such as privatization, auctioning, or subsidies. The case of Wikipedia is however quite surprising. Wikipedia when compared to Britannica, its commercial counterpart, is much larger. If one measure is the number of articles, Wikipedia was already approximately 7 times larger in 2006 than Britannica which had approximately 120,000 articles then.

Wikipedia would not be the first information systems project that began as a public good but resulted in a quality product. Indeed, many successful open-source software projects, such as Unix, were created ground-up without the offering of any direct pecuniary compensation. These products today are often cited for their supremacy over related commercially-produced products. How is the case of Wikipedia then any different?

To summarize, the main research questions we ask in this paper are as follows:

- How can we explain the surprising empirical findings that an encyclopedia that anyone can edit is of a high quality?
- How can we explain the surprisingly large size of Wikipedia, when it is a public good (created by people and used by people)?
- How is the case of Wikipedia different from Open Source Software development, another information systems project that is a public good and where neither quality nor provision is compromised?

## RELATED LITERATURE

For our analysis, we invoke the solution concept of Nash (1950) equilibrium, since we model the contribution to Wikipedia as a non-cooperative game. To analyze situations involving sequential moves, we also consider Stackelberg equilibria, which in concept are the natural applications of the Nash equilibrium solution to dynamic games.

In addition, the stream of economics literature related to the private provision of public goods is relevant. Hirshleifer (1983) investigated the variation of public good provision with three different technologies. We use his seminal piece extensively as a basis for

our modeling. Cornes (1993) develops Hirshleifer's analysis further, studying the effects of changes in income allocation. Varian (1994) extends Hirshleifer's work by studying the effects on equilibria when provision involves sequential moves by the agents. The results of the literature above do not apply so readily to the case of Wikipedia since they all predict very high levels of free-riding whereas Wikipedia appears to be quite large.

There have been studies that investigate mechanisms in which free-riding is mitigated. Andreoni (1990) and Cornes and Sandler (1984) investigate the changes in results whenever consumers are allowed benefits directly from their private contribution in addition to the level of the public good. While we believe that their work is relevant, we investigate how Wiki technology in particular is able to mitigate free-riding, perhaps beyond the effects of "warm-glow" feelings suggested by Andreoni (1990) and Cornes and Sandler (1984).

Free-riding has been studied in a variety of IT application contexts. Their contexts, however, make their results too specific for insights to extend readily over to the Wikipedia case. Carlton and Chevalier (2003), for example, investigates free-riding and sales strategies for the Internet. Varian (2004), as another example, studied free-riding in the provision of (general) information systems reliability. His study has a probabilistic flavor, emphasizing the likelihood of failure of information systems, and also repeats the standard extreme free-riding result.

Two kinds of equilibria in games of incomplete information are considered in our paper. First, we consider cheap-talk equilibria (cf. Gibbons, 1992) in dynamic games of incomplete information when messaging is "cheap" or costless, and hence receivers must determine the credibility of the messages. Second, we briefly talk about signaling equilibria (cf. Vega-Redondo, 2003) when players try to signal their types using costly messaging techniques.

Wikipedia itself has been the subject in a variety of studies. Most notably, Nature's (Giles, 2005) study involved 42 articles reviewed by experts to compare the prestigious

Britannica to the free giant Wikipedia. Their results indicated that the average scientific entry in Wikipedia contained four errors or omissions, while Britannica had three. Nonetheless, Wikipedia is planning on test-launching a reviewing program. Forte and Bruckman (2005) investigate why people write for Wikipedia even when the encyclopedia does not provide bylines to credit authors for their hard work by interviewing 22 Wikipedians. They are however unable to derive economic explanations, and instead suggest ‘softer’ incentives such as engagement in desirable activities. Finally, IBM research (c.f. Wattenberg et al., 2007) uses sophisticated HCI technology to track the contribution by Wikipedians, and the growth of Wikipedia articles. They find that while most articles have been vandalized, “vandalism is usually repaired extremely quickly.”

## THE CONTRIBUTION PARADOX OF WIKIPEDIA

We will begin by modeling Wikipedia as a public good, and illustrate the standard derivation of the free-rider result, and its under-provision property. Suppose that there are two agents,  $i \in \{1, 2\}$ , and that each can choose to contribute a level  $x_i$  to Wikipedia. This is in very general terms. It may refer to a correction of an error, or the authoring of a major article. Doing so, each agent  $i$  receives a payoff of

$$u_i = v_i(x_i + x_j) - c_i(x_i) \quad (1)$$

where  $v_i$  is monotone concave function increasing in the *total* contribution, and  $c_i$  is the cost function for that agent, and  $j \in \{1, 2\}$ ,  $j \neq i$ .

Since (1) is key to our analysis, we will explain its form. First, we are defining Wikipedia to be  $x_i + x_j$ , the total contribution, as it is created entirely by Internet users. Every individual benefits from having such a website, hence the function  $v_i$ . That function is concave to reflect the standard diminishing marginal utility assumption. For each individual, the net utility depends also on the cost of his own contribution.

It is easy to recognize (1) as the utility function of agents in a game involving the private provision of a public good. And it is easy to envision Wikipedia in the same context. Every Internet citizen can potentially benefit from having Wikipedia. Contributing to Wikipedia however will involve a positive cost function. Not surprisingly, we can intuitively expect free-riding as a result. That is, more people will simply use Wikipedia than contribute to it. We shall now derive that standard result.

First, we need to determine the reaction functions and Nash equilibria for this simultaneous contribution game. A reaction function  $f_i(x_j)$  gives the optimal strategy of agent  $i$  given a choice of  $x_j$  by agent  $j$ ,  $j \in \{0, 1\}$ ,  $j \neq i$ . Let the reaction function of agent  $i$  be given by  $f_i$ . We assume that  $v_i$  and  $c_i$  are twice continuously differentiable for all  $i$ . The first-order condition is given by

$$v_i'(x_i + x_j) = c_i'(x_i).$$

If we assume that  $c_i$  is linear in  $x_i$ ,  $c_i(x_i) = \lambda_i x_i$ , where  $\lambda_i$  is a positive constant, and that  $v_i'$  has an inverse, we can define

$$\hat{x}_i = (v_i')^{-1} \lambda_i.$$

If we solve for agent  $i$ 's contribution under the first-order condition, we see that it is  $\hat{x}_i - x_j$ . To derive agent  $i$ 's reaction function  $f_i$ , we only need impose a non-negativity constraint to get

$$f_i(x_j) = \max\{\hat{x}_i - x_j, 0\}.$$

We can define and account for differences in the tastes of the agents in terms of  $\hat{x}_i$ . More formally, we say that agent  $i$  likes Wikipedia better if and only if  $\hat{x}_i > \hat{x}_j$ . This is because  $\hat{x}_i$  can be defined as agent  $i$ 's standalone contribution: the amount he contributes to Wikipedia when the other agent does not contribute at all.

We may plot the reaction functions  $f_1$  and  $f_2$  of agents 1 and 2 on the same pair of axes. This is given in figure 1. The  $x$ -axis

gives values of  $x_1$  whereas the  $y$ -axis gives values of  $x_2$ . Any point  $(x_1^*, x_2^*)$  where the two response function will intersect will correspond to a Nash equilibrium. This is because at  $(x_1^*, x_2^*)$ ,  $f_1(x_2^*) = x_1^*$ , and  $f_2(x_1^*) = x_2^*$  which satisfies the definition of a Nash equilibrium. When only one such point exists, the Nash equilibrium is unique. Looking at figure 1, the proposition that follows is then immediate.

**Paradox 1a:** If the contribution to Wikipedia is modeled as the standard private provision game where the utility functions have the form given in (1), there exists a unique total size of the encyclopedia in equilibrium. A Nash equilibrium always exists, and is unique when the two agents have different tastes.

Paradox 1a suggests that Wikipedia *cannot* be expected to grow perpetually, which clearly contradicts the continually growing characteristic of Wikipedia that can be observed. (Between January 2006 and December 2008, Wikipedia tripled in size.) While it was assumed that the number of agents was fixed, shortly we will demonstrate

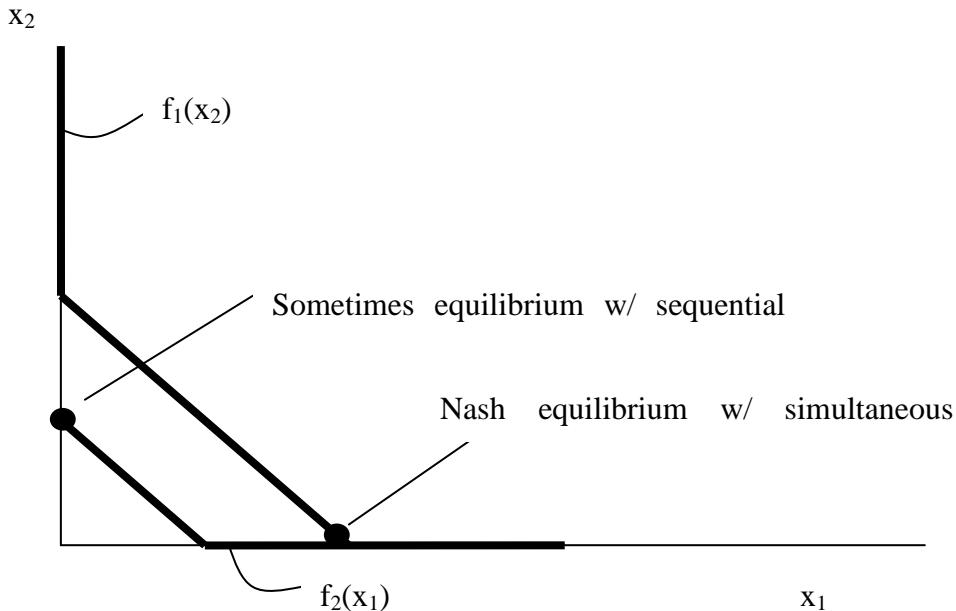
how a growth paradox persists even as the number of agents increases.

Suppose, without loss of generality, agent 2 likes Wikipedia more than agent 1, that is  $\hat{x}_2 > \hat{x}_1$ . Then in the unique Nash equilibrium, agent 2 provides all of the articles, and agent 1 free-rides. If  $\hat{x}_2 = \hat{x}_1$ , then a plethora of Nash equilibria exist as the reaction functions will coincide. But as is clear from the figure, the total size will still be unique in equilibrium.

**Paradox 1b:** In the two-agent provision game of Wikipedia where the utility functions have the form given in (1), all contributions to Wikipedia are made by the agent(s) who like it most. All other agents free-ride.

We now contrast that expected outcome against the social optimum. The social problem will solve

$$\begin{aligned} \max_{x_1, x_2} &= v_1(x_1 + x_2) + v_2(x_1 + x_2) - c_1(x_1) - c_2(x_2) \\ \text{s.t.} & \quad x_1 \geq 0, x_2 \geq 0 \end{aligned}$$



**Figure 1. Equilibria in the standard contribution game**

Without loss of generality let us suppose  $\lambda_1 < \lambda_2$ . It is clear then that the social optimum is attained when agent 1 contributes everything. The first-order condition then becomes.

$$\frac{\partial v_1(x_1 + 0)}{\partial x_1} + \frac{\partial v_2(x_1 + 0)}{\partial x_1} = \lambda_1,$$

which can be written simply as

$$v_1'(x_1) + v_2'(x_1) = \lambda_1.$$

We note that since  $v_1$  is concave,  $v_1'(x_1)$  is decreasing whereas  $\lambda_1$  is a constant. Since  $v_2'(x_1) > 0$  for all  $x_1$ , then the socially optimal level of provision,  $x_1^s$ , is

$$x_1^s = \hat{x}_1 + \delta,$$

where  $\delta > 0$ . A special case would involve  $\delta = \hat{x}_2$ .

This leads us to state formally the following:

**Paradox 2:** In the two-agent contribution game where the utility function is of the form in (1), when compared to the socially optimal case, Wikipedia will always be under-provided. Further, the wrong agent may contribute in the sense that he may like Wikipedia more, but is worse in quality.

The case of the wrong contributor  $i$  will occur whenever  $\hat{x}_i > \hat{x}_j$ , even though  $\lambda_i > \lambda_j$ .

Our findings thus far are already surprising. We expect Wikipedia to be smaller in size to the case that is optimally social, as Wikipedia can be treated as a public good. This under-provision should be lacking or at least lessened when financial incentives are introduced. Indeed, subsidies and taxes are one method of changing equilibria to points that are social optima (Landsburg, 2001). So why is it then that a purely commercial counterpart, Britannica, is much smaller in size than Wikipedia, the privately-provided public good? At the very least it warrants further analysis.

One modification that can be made is to the utility function. In our utility function for individual agents given by (1), each agent receives utility only from the total level of the public good. This may be inconsistent with real-life facts. Each agent may receive benefit not only from the level of the public good, but from his individual contribution as well. Wikipedia provides several reasons why one may contribute to the public good. Most of these are along the lines of positive feelings due to good-Samaritan deeds. For instance, Wikipedia users are asked to contribute for the poor child in Africa having no access to expensive encyclopedia. Yet there are also other direct benefits associated with contributing. Let us consider the evolution of a Wikipedian as he contributes to Wikipedia. He is increasingly recognized as a reputable volunteer, and is elevated to a higher status such as that of an ‘administrator’ and granted distinct privileges.

The case where agents benefit directly from their contribution in addition to the overall level of the public good was investigated by previous researchers. In Andreoni’s (1990) simple model, an agent’s utility function is a function of both his own contribution, and the amount of the public good. Andreoni’s (1990) model is very general, in that it captures altruism and its antithesis, egoism, as well as “impure altruism,” meant to capture concepts such as the “warm-glow” felt by agents upon their contribution to charities. Andreoni’s (1990) results however pertain to the effects of income re-distribution between agents on the contribution to charity. Cornes and Sandler’s (1984) result is more relevant. They show that when a single act of contribution is able to generate utility both directly and indirectly, as community size increases, free riding and its associated inefficiency may *decrease*. Their model is set up as a standard consumer behavior optimization problem.

An alternative explanation for Wikipedia’s size when compared to Britannica is based on the cost structure. Encyclopedia fall under a class of goods known as information goods. The production of these goods is characterized by the necessary existence of high up-front (fixed) costs, and

virtually zero marginal costs. Britannica's fixed costs would consist of two major components. First, Britannica would need to invest in signaling expenses, discussed in detail in the next section. Second, Britannica would incur costs in authoring articles for inclusion in its encyclopedia. Britannica's profits would depend on the number of articles, as the following indicates. Let  $FC$  stand for fixed cost,  $C_s$  for signaling cost,  $g$  for a convex cost function increasing in  $k$  the number of articles,  $TC$  for total cost,  $TR$  for total revenue, and let  $\Pi$  denote Britannica's profits. Then

$$FC = C_s + g(k),$$

$$TC = FC + 0,$$

and

$$\Pi = TR - TC.$$

In the first-order condition,  $k^*$  solves

$$\frac{\partial TR}{\partial k} = g'(k).$$

Assuming that it exists,  $k^*$  then is then the unique number of articles selected by Britannica for inclusion in its encyclopedia. As the above system of equations indicates, this  $k^*$  depends on  $TR$ , which in turn is of course determined by the demand and reservation values of the consumers.

A noteworthy point is that Britannica's choice of  $k^*$  will not be socially optimal, and it is straightforward to argue this. The socially efficient price necessarily equals marginal cost. But in Britannica's case, we can safely assume this to be zero. So any non-zero price will involve positive amounts of deadweight loss. Of course, there is no solution to this social inefficiency problem. Whenever a firm has decreasing average costs at the socially optimal point, as Britannica does, its profits will be negative there. There is no easy way however to compare the differences between the degrees of suboptimality of the two different cases.

Since Wikipedia use appears to be growing, it is of interest how the overall efficiency compares to the social optimum as

the number of agents using it increases. To see that, we must analyze the problem in the context of an arbitrary number of  $n$  agents. We make the simplifying assumption that the agents are all identical. In particular,  $v_i$  and  $c_i$  are all identical  $\forall i = 1, \dots, n$ . First, let us view the social optimum as a function of  $n$ . As stated above, the whole range of equilibria are possible since the tastes are identical, and since the response functions  $f_i$  will intersect in the  $n$ -dimensional space at an infinite number of points. All such equilibria satisfy the following Nash requirement, that  $\forall i$

$$x_i^* = \hat{x} - \sum_{j \neq i} x_j^*.$$

But one equilibrium may be more 'intuitive,' and somewhat informally Pareto optimal. Let us assume that in the Nash equilibrium output, all of the agents exert identical levels of effort. That is

$$x_1^* = x_2^* = \dots = x_n^* = x^*.$$

In the Nash equilibrium now,

$$x^* = \hat{x} - (n-1)x^*.$$

Solving for  $x^*$ , one sees that

$$x^* = \frac{\hat{x}}{n}.$$

On the other hand, the social optimum output is of course

$$n\hat{x}.$$

Let us define  $\alpha$ , the relative efficiency as follows:

$$\alpha = \frac{\hat{x}}{n} / n\hat{x}.$$

Then  $\alpha$  is simply the constant  $1/n^2$ , which has derivative  $-2/n^3$ . The following proposition then is immediate.

**Paradox 3:** When utility functions are of the form in (1), and the agents are all identical, the relative efficiency of Wikipedia decreases rapidly as the number of users  $n$  increases, even when all the agents are contributing.

As mentioned before, Wikipedia by nature is very dynamic. Wikipedia grows not only as new articles are created, but also as extant articles are edited. The game of contribution to Wikipedia discussed so far however has been static in the sense that all moves were simultaneous. It would be interesting to investigate the change in the equilibrium outcome if agents moved sequentially. Varian (1994) investigated the general theory of sequential contribution to public goods. Our discussion below is based on Varian (1994), albeit using a simpler version of his elegant model.

Now the two agents move sequentially, so the game of contribution is dynamic. Suppose, without loss of generality, agent 1 moves first. Choosing  $x_1$ , he receives a utility of

$$u_1 = v_1(x_1 + x_2) - c_1(x_1)$$

when agent 2 subsequently chooses  $x_2$ . Using the reaction function of 2, we can write the right-hand-side equivalently as

$$u_1 = v_1(x_1 + f_2(x_1)) - c_1(x_1)$$

or

$$u_1 = v_1(x_1 + \max\{\hat{x}_2 - x_1\}) - c_1(x_1)$$

The reaction function can be used to split agent 1's utility into the following system of two equations,

$$u_1(x_1) = v_1(x_1) - c_1(x_1) \quad \text{if } x_1 > \hat{x}_2$$

$$u_1(x_1) = v_1(\hat{x}_2) - c_1(x_1) \quad \text{if } x_1 < \hat{x}_2,$$

since for any level of contribution  $x_1$  by agent 1, agent 2 will always contribute  $x_2$  at least and just enough so that the total

$$x_1 + x_2 = \hat{x}_2.$$

Two cases will arise, depending on the difference in preferences of the two agents. In the first case,  $\hat{x}_1 < \hat{x}_2$ , that is agent 1 likes Wikipedia *less* than agent 2. In that case, agent 1 will always choose to free-ride on agent 2 and contribute nothing to Wikipedia. To see that, first note agent 1's choice between a strictly positive contribution,  $x_1 > 0$ , or free-

riding completely,  $x_1 = 0$ . For any choice  $x_1 > 0$ , he will receive a *net benefit* of

$$v_1(\hat{x}_2) - c_1(x_1),$$

as agent 2 will always provide the remainder of the positive good. In free-riding completely however, agent 1 can receive higher net benefits as doing so will cost him nothing. That is, free-riding will yield a net benefit of  $v_1(\hat{x}_2)$ .

The second case is more complicated, and more interesting. The first-mover, agent 1, likes the good more. That is  $\hat{x}_1 > \hat{x}_2$ . As in the first case, if agent 1 chooses to free-ride completely, he can ensure a net-benefit of at least  $v_1(\hat{x}_2)$ . Although this is *less* than  $v_1(\hat{x}_1)$ , agent 1 may still free-ride completely as a positive contribution by agent 1 entails costs that he will incur, and the *net benefit* he will derive is given by

$$v_1(\hat{x}_1) - c_1(\hat{x}_1).$$

If this net benefit  $v_1(\hat{x}_1) - c_1(\hat{x}_1)$  is greater than  $v_1(\hat{x}_2)$ , agent 1 will provide all of the good himself, whereas if it is smaller, agent 1 will free-ride entirely. This is quite surprising as a result when we recall that agent 1 likes the good *better*. It is interesting to highlight the cases when free-riding by agent 1 will occur, even when he likes the good better. One case may be when tastes are similar. Indeed, if  $\hat{x}_2$  is less than  $\hat{x}_1$  but  $\hat{x}_2$  is large enough, then agent 1 will free-ride.

The analysis above may explain a seemingly strange behavior by the administrators of Wikipedia. On Wikipedia's website, the administrators encourage all users to be "bold" and make changes to *any* of Wikipedia's articles, even the same article with that message itself. Where the stability and reliability of the articles is a concern, one would perhaps expect some form of mandatory registration process, or at least a message of a much different tone, such as the following perhaps:

*"Please author an article only if you are completely sure of its contents. Please do*



*not make changes to an article unless you are completely certain of its faults.”*

Our analysis shows however that when moves are sequential, it may be advantageous to have the agent who likes the good *least* move first as doing so will result in a higher level of contribution. This is because the other agent will not have any option of free-riding, and will always provide all of the good, and hence a ‘higher’ level of the good will result. When Wikipedia administrators ask users to be “bold” and make changes, it may act as a signal to the high-level contributors that their moves are all that remain. Similarity of preferences was mentioned above as one case when we will see free-riding by agent 1 even when he likes the good better. In Wikipedia, the similarity in tastes will probably feature prominently as the number of users is large, and preference may actually be a continuous variable.

We end our discussion related to the contribution to Wikipedia by considering the case of Open-Source Software (OSS), which is another category of “public-good” projects in Information Technology that also exhibits the paradoxical growth and high-quality. In doing so we shall highlight what distinguishes Wikipedia from OSS, in order to demonstrate both the novelty and the validity of our work.

Open-Source development projects as public-goods are driven in growth by the seemingly altruistic efforts of a variety of individuals. As with Wikipedia, of wonder has been the effect of voluntary provision on both the quality and amounts of such open-source efforts (cf. Lerner and Tirole, 2000). And as with Wikipedia, open-source projects are surprisingly impressive in quality, and found in abundance. A considerable amount of research has been conducted to investigate this apparent contradiction to the traditional findings related to the private provision of public goods. The theoretical explanation that seems most feasible (Lerner and Tirole, 2002) is that participation in open-source projects allows developers to effectively signal their skills to ensure higher levels of future returns. One may be tempted to suggest that contribution to Wikipedia may be explained away in a similar manner, that individuals are editing articles to

signal their higher “quality.” We contend however that such contributions to Wikipedia cannot constitute a signaling activity, and provide an intuitive explanation. Users cannot effectively signal their “quality” because reward structures based on such signals cannot be justified. Simply put, if better articles were rewarded for, then agents could easily acquire the information at nominal costs from other sources, such as Britannica, and the bases for rewards would be unfounded.

### An Explanation for Wikipedia’s Size

So far, we have been able to reproduce only the extreme free-riding result that is standard in the public goods economics literature. Clearly though, the abstraction in the model is far too rigid to be true in practice. Instead of extreme free-riding, what we can readily observe in Wikipedia is a variety of contribution levels. The empirical study by *Nature* (Giles, 2005) suggests that there is some free-riding with only about 10% of the scientists it surveyed contributing to articles. Most articles seem to be the result of generally cooperative effort with partial contributions from a large number of users. Some contributors seem to be highly excited by their topics of interest while being indifferent to other topics. Finally, some articles seem to be reactive to others and quite large, such as the article on the late scientist Lisa Meitner who was failed to be recognized for her work in collaboration with the Nobel laureate Otto Hahn.

In this section, we modify our model and offer an explanation for Wikipedia’s size. We introduce a new property of agents known as “type.” Suppose agent  $i$  has type  $t_i > 0$ . In contributing  $x_i$  the agent receives a payoff of

$$u_i = v_i(x_i + x_j(1 - (t_i - t_j)^2)) - c_i(x_i) \quad (2)$$

where  $v_i$  is again a monotone concave function, and  $c_i = \lambda_i x_i$  is the cost function for that agent. The payoff model above is straightforward and intuitive. From the other agent’s contribution, an agent receives utility that is moderated by the degree of similarity between the two agents. For example, an economist may care little about a psychology article, and at the very extreme, a liberal may

actually receive disutility from the presence of a conservative’s article on say abortion. The first-order condition is given by

$$v_i'(x_i + x_j(1 - (t_i - t_j)^2)) = \lambda_i$$

Once again, we can define  $\hat{x}_i$  as

$$\hat{x}_i = (v_i')^{-1} \lambda_i$$

assuming linear costs and the existence of the inverse of  $v_i'$ . Now suppose agent  $j$  contributes  $\hat{x}_j$ . Some algebra shows that agent  $i$ ’s best response is to provide

$$\hat{x}_i(t_i - t_j)^2 > 0$$

whenever  $t_i \neq t_j$ , and quite surprisingly,

$$\hat{x}_i(t_i - t_j)^2 > \hat{x}_i$$

whenever  $(t_i - t_j)^2 > 1$ . So even though agent  $j$  is contributing the entire amount that agent  $i$  would have contributed himself, agent  $i$  cannot simply free-ride as agent  $j$ ’s preferences are different. The results are more clear if we plot the response functions  $f_i(x_j)$ . This is given in Figure 2.

We can state the most important result:

**Proposition 1a:** In the two-agent contribution game with the utility function of the agents having the form given in (2), as the difference between the preferences of the agents  $(t_i - t_j)^2$  increases, the free-rider problem of Wikipedia as a public good is mitigated.

From the figure, we see that the exact Nash equilibrium outcome will depend on the magnitude of  $(t_i - t_j)^2$ . The dashed lines are the different response functions of the two agents with different values of  $(t_i - t_j)^2$ . The points of intersection highlighted as Nash equilibria correspond to pairs of response functions with the same values of  $(t_i - t_j)^2$ . When  $(t_i - t_j)^2$  is small enough, then complete free-riding will result, as long as one agent prefers Wikipedia more than the other. At an intermediate level of  $(t_i - t_j)^2$ , the two agents offer their standalone contributions. Now even though agent 1 is offering his original contribution, agent 2 is not free riding

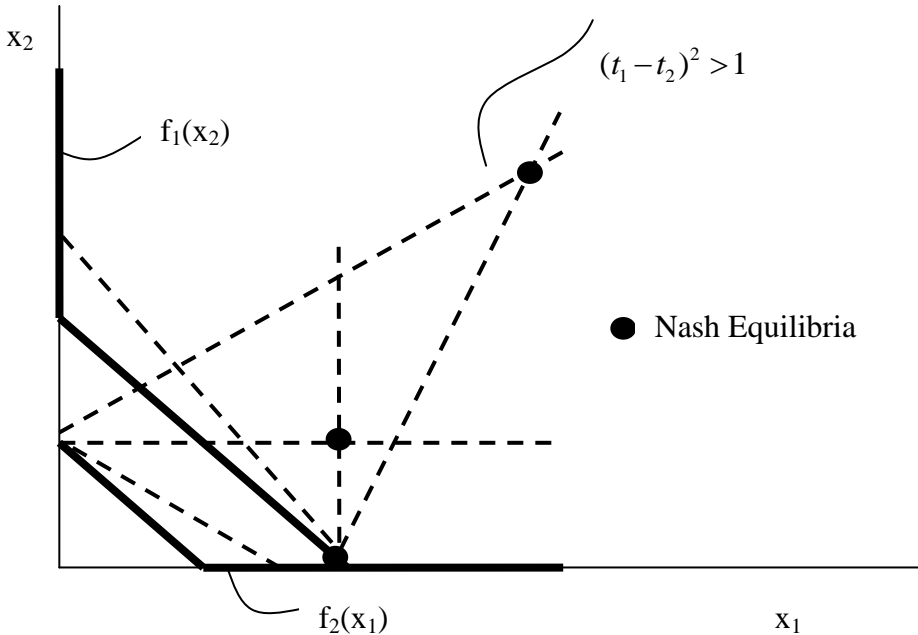


Figure 2 Equilibria in the contribution game with different preferences

anymore. Finally, if  $(t_i - t_j)^2$  is very large, then the Nash equilibrium is the outermost point, where both agents are reactively contributing very large amounts. We state that more formally below.

**Proposition 1b:** Suppose, without loss of generality,  $\hat{x}_j > \hat{x}_i$ , and that  $\delta = \hat{x}_i$ , four regions of  $(t_i - t_j)^2$  can be associated with four different kinds of equilibria:

- If  $0 < (t_i - t_j)^2 < \frac{\hat{x}_j - \hat{x}_i}{\hat{x}_j}$ , agent  $i$  will free-ride entirely on agent  $j$  as in equilibrium  $x_i = 0$ . In addition,  $x_j = \hat{x}_j$ .
- If  $\frac{\hat{x}_j - \hat{x}_i}{\hat{x}_j} < (t_i - t_j)^2 < 1$ , agent  $i$  will free-ride partially on agent  $j$  as in equilibrium  $0 < x_i < \hat{x}_i$ .
- If  $(t_i - t_j)^2 = 1$ , then agent  $i$  will not free-ride at all, being indifferent to agent  $j$ 's contribution. The outcome is  $(\hat{x}_i, \hat{x}_j)$  which can often be socially efficient.
- If  $(t_i - t_j)^2 > 1$ , then both agents contribute at levels that are high, but which can often be socially inefficient.

**Proof:** {Sketch} To derive region 1, first from figure 2 it is clear that agent  $i$  will stop free riding as soon as  $f_i(\hat{x}_j) > 0$ . We note that  $f_i(\hat{x}_j) > 0$  is given by

$$x_i = \hat{x}_i - \hat{x}_j(1 - (t_i - t_j)^2).$$

In that case, a little algebra can show that the condition on  $(t_i - t_j)^2$  is that it should be larger than  $\frac{\hat{x}_j - \hat{x}_i}{\hat{x}_j}$ . Deriving the rest of the regions is straightforward, and we omit it.

We can demonstrate the meaning of the ranges by means of an example. Suppose we have two contributors, where one agent likes

Wikipedia more. Suppose the article to be created is to profile a prominent conservative politician. If our two agents are both supporters of the politician, and so  $(t_i - t_j)^2$  is fairly small and in region 1, then the agent who likes Wikipedia more will author the article entirely. Suppose instead that one agent likes the politician's stance on opposing outsourcing labor, whereas the other agent likes his stance on the environment. Then  $(t_i - t_j)^2$  is in region 2, and agent 1 will not contribute everything. This is probably most reflective of Wikipedia's editing scenario, where free-editing allows individuals to contribute what they value most in the form of small edits. Suppose one agent supports the politician as a union member strongly opposed to outsourcing, whereas the other agent supports him as an environmental activist. Then  $(t_i - t_j)^2$  is in region 3, and each agent will write their version of the article being indifferent to each other's contribution. Finally, if one agent is liberal, then  $(t_i - t_j)^2$  is in region 4, and we may expect two different articles, one with a liberal focus and another with a conservative focus.

## RELIABILITY OF WIKIPEDIA

In this section, we extend our model of the contribution game to account for reliability. In a general setting involving  $n$  potential contributors, let the contribution by agent  $i$  be described by  $q_i \geq 0$ , where  $q_i$  is the *quality* of that contribution. The quality of Wikipedia then is a function of  $q_1, q_2, \dots, q_n$ . One particular function that is well suited as a definition of quality for Wikipedia is the geometric mean given by

$$Q = \left( \prod_{i=1}^n q_i \right)^{1/n}.$$

Although this definition may appear somewhat arbitrary, or even strange as a choice, it is well suited as it captures the idea that "weaker-links" are more significant. Weaker-links in the context of public goods have been studied by Cornes (1993), albeit within the setting of a general model. A closely related concept is that of the Cobb-Douglas function in the context of public goods (cf.

Varian, 1984). It is clear that this definition captures the essence of “weaker-links” when we take its derivative with respect to  $q_i$ ,

$$\frac{\partial Q}{\partial q_i} = \frac{Q}{nq_i}.$$

Clearly, the marginal effects on total reliability  $Q$  are higher whenever  $q_i$  is lower. In terms of the reliability of Wikipedia, “weaker-links” will determine whether or not Wikipedia can be trusted in general as a source. If it is known *a priori* only that a few articles are inaccurate, without knowing which ones exactly, then potentially all articles should be approached with caution, or at least verified using other sources. To simplify our analysis, we may consider the special case of “weaker-links,” which is the case when reliability is determined by the “weakest-link.” In that case, if it is known *a priori* only that a single article is inaccurate, without knowing exactly which one, then without verification potentially any article may be the inaccurate one and the reliability of the entire project is in question.

We may model the weakest-link concept by redefining quality as

$$Q = \min\{q_1, \dots, q_n\} \tag{3}$$

To solve for the equilibrium contribution qualities, we first need to derive the reaction functions. Suppose  $v_i$  and  $c_i$  are functions as before but accepting  $q_i$  as the input parameter instead. Also let  $n = 2$ . Agent  $i$  will receive the following in net benefit by contributing  $q_i$ :

$$u_i = v_i(\min\{q_i, q_j\}) - c_i(q_i)$$

Let us capture differences in tastes by the constants  $\hat{q}_1$  and  $\hat{q}_2$ , which are the maximum levels of qualities that agent 1 and agent 2 would care for Wikipedia to have respectively. It is then straightforward to show that for agent  $i$ , it is optimal to match the other agent’s contribution if it is less than  $\hat{q}_i$  or provide  $\hat{q}_i$  otherwise. So the reaction functions are given by

$$f_i(q_j) = \min\{q_j, \hat{q}_i\}$$

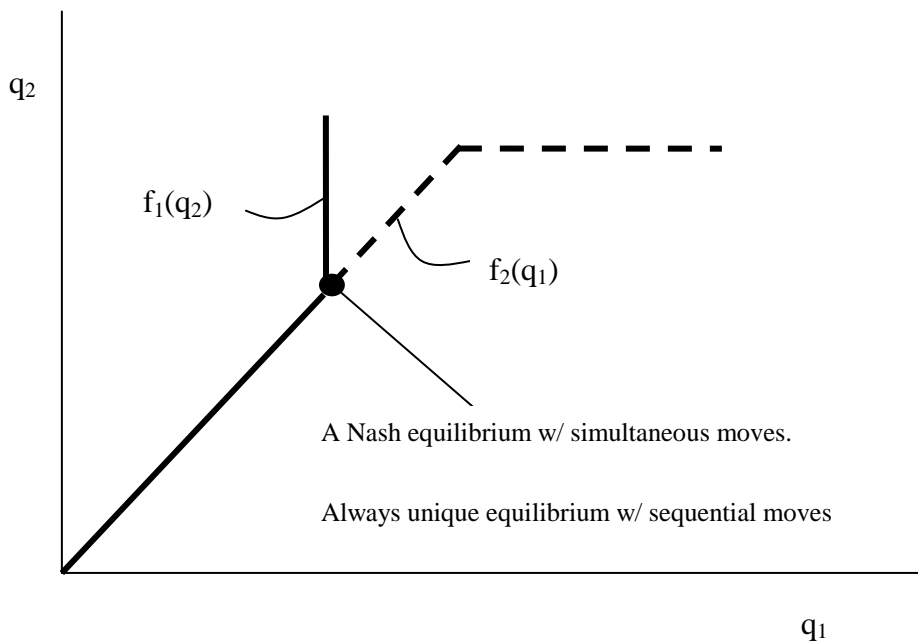


Figure 3 Reliability equilibria in the contribution game

Without loss of generality suppose  $\hat{q}_1 < \hat{q}_2$ . Plotting the reaction functions in figure 3, we can clearly see that there exists a plethora of Nash equilibria. We may select the one that Pareto dominates the other as it may perhaps be the most likely outcome. In that Nash equilibrium the reliability of Wikipedia is given by  $\min\{\hat{q}_1, \hat{q}_2\} = \hat{q}_1$ .

An interesting extension of the result above is that uniqueness of Nash equilibria can easily be established if moves are assumed to be sequential and hence the game dynamic. The uniqueness result for the general model was argued by Hirshleifer (1983) and established by Varian (2004). The argument for uniqueness is as follows. Let agent 1, without loss of generality have the first move. Agent 1 will choose  $\min\{\hat{q}_1, \hat{q}_2\}$ . Choosing any amount larger will result in forgone benefits and useless costs when the subsequent choice will be less. Choosing any amount smaller is not the best option.

**Proposition 2a:** When reliability of Wikipedia is defined as in (3) in the sequential two-agent game, the unique equilibrium level of reliability of Wikipedia is the least of the minimum levels desired by contributors.

This result above appears ‘mixed,’ in that its optimality is not clear. At first glance, it is reassuring to realize that the equilibrium reliability will most likely be at least above a bare minimum threshold. But two questions arise subsequently. How does this compare with the reliability that is socially optimal? And is such a threshold level of reliability ‘good-enough,’ per se?

To determine Wikipedia’s equilibrium reliability to the socially optimal case, we note that the social objective is to maximize the following social utility function, by choosing  $q_1$  and  $q_2$  accordingly.

$$U_s = v_1(\min\{q_1, q_2\}) + v_2(\min\{q_1, q_2\}) - c_1(q_1) - c_2(q_2)$$

In the social optimum, level, all agents need to be contributing articles of the same quality. This is easy to argue using contradiction. Suppose that  $q_1 \neq q_2$ . Without loss of generality, let  $q_1 < q_2$ . If  $q_1 < q_s$ , the

socially optimal quality level, then reliability will be compromised. If on the other hand  $q_1 \geq q_s$ , even though reliability is at the socially optimum level, social welfare can still be improved since  $c_2(\cdot)$  is increasing in  $q$  and agent 2 can increase his private benefits by compromising the quality of his contribution. So the socially optimal point is the single level of quality, say  $q_s$ , in contribution that solves

$$\max_q U_s$$

Although  $q_s$  is hard to solve, we can still draw insights from analyzing various cases. Suppose the choice of  $q_s$  is between  $\hat{q}_1$  and  $\hat{q}_2$  where  $\hat{q}_1 < \hat{q}_2$ . It is already clear that either one may be more optimal in equilibrium, depending on  $v_i$  and  $c_i$ . To allow for further analysis, let us simplify the matter and make all  $v_i$  ‘s identical, and as before  $c_i$  ‘s linear functions. Since the Nash equilibrium outcome is expected to be a choice of  $\hat{q}_1$ , let us first consider otherwise. In choosing  $\hat{q}_2$ , agent one will need to make a suboptimal choice that will decrease his net utility by

$$v(\hat{q}_1) - v(\hat{q}_2) + \lambda_1(\Delta q)$$

where  $\Delta q = \hat{q}_2 - \hat{q}_1$ . On the other hand, agent 2’s loss from agent 1 selecting  $\hat{q}_1$  is given by

$$v(\hat{q}_2) - v(\hat{q}_1) + \lambda_2(\Delta q)$$

We see that agent 2 will choose  $\hat{q}_1$  in the socially optimal case whenever

$$\Delta q > \frac{2[v(\hat{q}_2) - v(\hat{q}_1)]}{\lambda_1 + \lambda_2}$$

The above analysis leads us to state the following proposition without proof.

**Proposition 2b:** The reliability of Wikipedia, when defined by (3) in the two-agent game, will be compromised when compared to the socially optimal case if quality preferences are close enough. If quality preferences are far apart, then Wikipedia will have the socially optimal quality level.

Of course, in the analysis thus far, we assumed that no individual would be interested

in deliberately destroying the reliability of Wikipedia. That is, we considered only situations when individuals make the choice solely between benevolent contributions or free-riding. Wikipedia however is subject to “vandalism” in that perfect articles can be edited arbitrarily by individuals interested in deliberately undermining Wikipedia’s reliability. To incorporate vandalism into the model, one needs only to introduce agents with the appropriate structure of the utility function. At this point, we suggest no normative guidelines for mitigating this problem. Instead, we highlight that self-regulation will be possible insofar as there are agents who endure significant disutility from vandalism actions. At the very least, these agents can always change an article back to its original state. Obviously, at any point in time an article may be the subject of vandalism, and that fact may not be clear to the casual user. It may be worthwhile for users to check the time-stamped history of the article to determine whether they can reliably use it for references.

We can still model the concept of vandalism and demonstrate equilibrium outputs as follows. Suppose there are two agents, a *vandal* and one other who is a benefactor whom we call the *contributor*. Suppose the vandal has the following utility function

$$u_v = \begin{cases} \tilde{v} - \tilde{c}(d) & \text{if } d > x \\ -\tilde{c}(d) & \text{if } d \leq x \end{cases} \quad (4)$$

when he contributes an amount  $d$  of damage to Wikipedia, while  $x$  is the positive contribution of the benefactor.

In that function  $\tilde{v}$  is the benefit the vandal derives whenever the reliability of Wikipedia is compromised, and  $\tilde{c}$  is the cost function increasing in  $d$ . Suppose the contributor has the following utility function

$$u_c = \begin{cases} v(x) - \lambda x & \text{if } x \geq d \\ -\lambda x & \text{if } x < d \end{cases} \quad (5)$$

where  $v$  is again increasing monotone, concave in  $x$ , and  $\lambda > 0$  is a cost constant.

We now need to derive the response functions. We start with the response function

$f_v$  of the vandal, as that is easier. For each value of  $x$ , the vandal needs to choose between a positive contribution  $d > 0$  or abstaining from vandalism. A positive contribution will be better if and only if  $\tilde{v} > \tilde{c}d$  and  $d > x$ . The optimal positive contribution  $d^*$  will be infinitesimally larger than  $x$ . To derive the full response function  $f_c$  of the contributor, we note that the choice is between a positive contribution  $x > 0$  and no contribution. For  $x > 0$ , there is always a unique positive  $x^*$  that solves

$$\max_x v(x) - cx$$

So the choice is between  $x^*$  and  $x = 0$ . The condition for reliability to be maintained in equilibrium is of course that  $x^* > \tilde{v}/\tilde{c}$ . This leads us to state the following.

**Proposition 2c:** In the two-agent (a vandal and a contributor) game of Wikipedia where the utility functions are of the form in (4) and (5) articles in equilibrium will not sway back and forth between vandals and contributors. To prevent vandalism of an article in the unique equilibrium outcome, the least quality preference among those of the editors should be at least equal to the benefit derived from vandalism.

Although never due to vandalism, there is one situation when articles may sway back and forth through editing. When readers are unsure of the contents of an article, and have no way of verifying their accuracy, the article may be subject to repeated editing by polarised editors. The situation may be exacerbated by the fact that editing is not a ‘costly’ activity. Consequently, we may use the solution concept of cheap-talk equilibria (cf. Gibbons, 1992) that investigate credibility in environments sustaining costless messaging.

Suppose chance draws an editor’s type to be either a ‘lying’ or ‘honest’ individual. The editor behaves according to his type and edits an article. The users of Wikipedia then have the option of believing or disbelieving the contents of the article, but have no idea about the true type of the editor. Whether believing

or disbelieving occurs in equilibrium will depend on the fulfilment of three conditions, and eventually on the payoffs to the editor and users. The first condition is that the users' preferred actions will need to depend on the type of the editor. This condition is trivially fulfilled; users will prefer to believe honest editors and disbelieve others. The second condition is that editors must have different preferences across the users' strategies. This condition fails to be fulfilled. Both types of editors would like users to believe them. Although the third condition does not need to be stated, we do so anyway for the sake of completeness. That condition requires that the editors' and users' preferences not be completely opposed. This condition is also not fulfilled with Wikipedia. Specifically, users prefer to believe when editors are honest and disbelieve otherwise. But lying editors prefer that users believe, and hence communication cannot occur.

That many Wikipedia articles cannot or will not be believed by users is not a result that is destructive to the value of the project. Rather, we contend that it is a major strength. In particular, the polarization result of free-editing in Wikipedia will act as a perfect signalling device to users for highlighting situations when they should disbelieve articles that may be highly opinionated and less factual. This is one explanation for the common saying about Wikipedia that even amidst discourse and vandalism, the actual discourse itself may be informative. This communication feature of Wikipedia may be a great strength in the face of its commercial counterparts, such as Britannica. First, users of Britannica may not be able to discern opinionated articles that are further from truth from those that are more factually correct. Further, these articles will often be believed because the three conditions of cheap-talk equilibria highlighted above may often be fulfilled even after critique by a review committee.

In terms of the design mechanism, we may be interested in seeing what the effect will be of punishing vandals. Suppose, for example, that vandals' IP addresses are logged and then future activities blocked.

## THE COLLABORATIVE NET

Today, a striking Internet phenomenon can easily be spotted. We call this phenomenon the "Collaborative Net," and it is a culmination of an information pool created mostly by most of its users. The simplest technological introductions are enabling this phenomenon, which in turn is drastically redefining properties of the Internet. In addition to wiki, blogs are such a technological innovation and the blogosphere is a Collaborative Net feature.

One direct consequence of the dynamic blogosphere is a concept known as "citizen journalism." Other terms used to describe the same concept include "grassroots reporting," or "media-of-the-masses." Citizen journalism as a topic has been investigated in journalism studies (cf. Andrews, 2003; Blood, 2003; Regan, 2004). Lasica (2003) for example defines it as "individuals playing an active role in the process of collecting, reporting, sorting, analysing and disseminating news and information, a task once reserved almost exclusively to the news media," and contends that the transparency of blogging has contributed to increase the accessibility of news entities. While traditional periodic sources of information involved professional activities, such as television news programs or magazines and other periodicals, citizen journalism is providing Internet surfers alternative sources. Features remarkable about citizen journalism are its vast size and its provision by ordinary citizens. These features become strikingly apparent particularly when citizen journalism is compared to the mainstream media industry, as the latter can be characterized by its composition of a limited set of players. Of concern is the credibility of the information reported through citizen journalism channels; is the information provided through citizen journalism reliable? In contrast, many mainstream media players have well-established reputations, with famous journalists possessing enviable credentials. An understanding of the credibility of information reported becomes even more important with new practices by Internet news aggregators. In a recent article titled "Can You Tell Blogs From 'Real' News?" in *Forbes*, Dicarlo (2005) observes that blogs and other user-generated

content will soon show up on Yahoo!, blurring the lines between professional media and citizen journalism.

The results established in this paper have implications for the Collaborative Net feature of citizen journalism. By means of our main result, we have established that when preferences are sufficiently different, the underprovision problem of public goods resulting is mitigated as extreme free-riding is no longer feasible. This result, and our analysis in general will extend over easily to citizen journalism. The popularity of citizen journalism comprising blogs is often attributed to the failure of traditional media outlets to cover a wide enough variety of articles. In 2005 for example, bloggers championed the case of a missing woman which did not receive national media attention as that of Laci Peterson. According to the bloggers, the missing woman was poor and African American, and deemed unworthy of media attention unlike the attractive Peterson (Farivar, 2005). Our analysis showed that traditional print media (such as Britannica) will be unable to cover all of these articles. Citizen journalism however may not be as accurate as Wikipedia when readers will have no way to verify the content. We have already mentioned that in games of incomplete information, under three conditions cheap-talk equilibria are possible when information cannot be verified. Further, unlike Wikipedia, the publisher of a blog has full control over the contents of an article. Consequently, highly opinionated articles can easily prevail. An example of this was the appearance of numerous faceless blogs that successfully defended G. W. Bush during the 2004 elections when his service record was in question.

## DISCUSSION

Wikipedia is attracting both crowds and criticism. Users are drawn to Wikipedia not only by its tempting free-offering, but also because it contains articles on almost every conceivable topic. Why is Wikipedia so large, even as a public good? And can we trust articles that “anyone can edit?”

The first characteristic of Wikipedia we investigated was its sheer size. Since Wikipedia can be considered a public-good, we demonstrated how traditional theory would suggest that the standard free-rider problem would emerge. In equilibrium, only a few volunteers would contribute. But this result would obviously contradict the empirical observation that Wikipedia is roughly seven times as large as its commercial counterpart. Consequently, we investigated mechanisms in which extreme free-riding was being mitigated. Our main result in this paper was an explanation for the size of Wikipedia based on equilibrium contributions depending on the differences in types. Free-editing allows for a variety of expressions; expressions that reflect differences in type. In addition, using well-grounded principles from information economics, we explained why Wikipedia’s commercial counterpart could be much smaller in size.

The second characteristic of the Wikipedia we investigated was its reliability. This characteristic is currently a topic of much debate. Indeed, while many critics of Wikipedia are highly skeptical about the reliability of this free-encyclopedia that anyone can edit, surprising results are emerging regarding Wikipedia’s reliability. Our results were important as we are able to establish both lower and upper bounds for the reliability of Wikipedia. Qualitatively, Wikipedia’s definition as a public good, combined with free-riding and free-editing helps to maintain the reliability of Wikipedia.

We identified Wikipedia as part of a general Internet phenomenon that we call the Collaborative Net. The effect that the simple technologies enabling the Collaborative Net inspires awe. Increasingly, users are adopting more active roles. Previously, Net users were frequently called “browsers.” And they were exactly that, being unable to edit most of the information online. Today, more and more Net users are actually the contributors to a vast pool of information as bloggers or contributors to Wiki projects. This pool of information is characteristically dynamic with changes being made to it constantly.



The primary contribution of this paper is its novel explanation for the paradoxical growth of Wikipedia, the free encyclopedia that anyone can edit. To our knowledge, our explanation is novel and contributes to the general literature on public goods. The paper also offers theoretical explanations for the surprising empirical findings on the superior quality of the giant encyclopedia arising from free-editing. Our findings have implications for the much-debated topic of credibility in the new Collaborative Net environment. We also highlight the uniqueness of Wikipedia when compared to (other) Open Source Systems.

Our study features prominently as we are able to assess contribution and reliability within the traditional economic paradigm, in contrast to other studies. Benkler (2006) for example examines the contribution to blogs and wiki projects from a social-capital perspective. In addition to its theoretical appeal, our study is also of practical relevance as wiki projects will feature as options alternative to traditional knowledge pool-models. Knowledge-management, for example, can be implemented in organizations very efficiently by the use of information systems that are designed similar to Wikipedia.

The main limitation of our study was the use of two-person games to model contributions to Wikipedia when in reality Wikipedia involves the contributions of thousands of individuals. We believe however that the insights generated are still valid for at least three reasons. First, our aim was not to derive accurate numerical results, but instead to gain an understanding of the behavior. We are convinced that two-person games are

sufficient for that purpose, particularly since they are the standard in studying provisions to public goods (c.f. Varian, 1994). Second, in many cases richer games are a collection of smaller games, and in Wikipedia, the contribution game will often be reduced to the final two individuals who like it most. Third, mathematically, many of our results will still hold when considering  $n$ -dimensional vectors.

While Wikipedia offers many areas of future research such as the social and legal aspects of a free-to-edit information source, we would particularly suggest studying the specific design features of wiki-based information systems. One may investigate, for instance, incentives to remain neutral with varying levels of policing. Such studies would further highlight the suitability of Wikipedia and other wiki-based projects as reliable, unbiased information systems for use in effective knowledge-management. This area is particularly promising as commercial applications of wiki technology are emerging. Many online retailers including Amazon.com are implementing wiki-based review systems, and ShopWiki.com is an entire online business based on user-generated reviews published using wiki. A blogger recently reported that he was offered a contract by a Microsoft PR associate to edit a Wikipedia article the associate felt was biased against Microsoft technology (Jelliffe, 2007). While the associate maintained that it was only intended to be compensation for using a neutral voice, several interesting questions arise that relate to conflicts of interest, and the effects of pecuniary compensation on the overall neutrality of wiki articles.

## REFERENCES

- Andreoni, J., "Impure Altruism and Donations to Public Goods: A Theory of Warm G, low Giving." *Economic Journal*, 1990, 100, 464-477.
- Andrews, P., "Is blogging journalism?" *Harvard Neiman Reports*, 2003, 57(3), 63.
- Benkler, Y., *The Wealth of Networks*, Yale University Press, CT, 2006.
- Blood, R., "Weblogs and journalism: Do they connect?" *Harvard Neiman Reports*, 2003, 44.
- Carlton, D.W. and J.A. Chevalier, "Free riding and sales strategies for the Internet," *NBER working paper*, 2003.
- Cornes, R., "Dyke maintenance and other stories: Some neglected types of public goods," *The Quarterly Journal of Economics*, 1993, 108(1), 259-271.
- Cornes, R., and T. Sandler, "Easy riders, joint production, and public goods," *The Economic Journal*, 1984, 94, 580-598.

- CNN.com, "Journal: Wikipedia as accurate as Britannica," 2005, Online. Available: <http://www.cnn.com/2005/TECH/internet/12/15/wikipedia.ap/index.html>, last accessed 15 December 2005.
- Cooper, R. W., DeJ. Douglas, R. Forsythe, and T. Ross. "Selection criteria in coordination games: some experimental results." *American Economic Review*, 1990, LXXX, 218-33.
- DiCarlo, L., "Can you tell blogs from 'real' news?" *Forbes*, October 11 2005,
- Drost, E., "KU graduate to ascend Mount McKinley," *The University Daily Kansan*. August 8 2005.
- Farivar, C., "Bloggers champion missing woman," *Wired News*, 2005, Available: <http://www.wired.com/news/culture/0,1284,68389,00.html> last accessed January 31 2006.
- Forte, A., and A. Bruckman, "Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia," *Proceedings of 2005 international ACM conference on Supporting Groupwork*. 2005.
- Gibbons, R., *Game theory for applied economists*, Princeton University Press, Princeton, NJ, 1992.
- Giles, J., "Internet encyclopedias go head-to-head," *Nature*, December 15, 2005. 900-901.
- Goodin, D., "Online encyclopedia to regulate entries," *A.P.* December 6 2005.
- Hirshleifer, J., "From weakest-link to best-shot: the voluntary provision of public goods," *Public Choice*, 1983, 41, 371-86.
- Jelliffe, R., "An interesting offer: get paid to contribute to Wikipedia," *O'Reilly XML.com Opinions*, Online, January 22, 2007, available: [http://www.oreillynet.com/xml/blog/2007/01/an\\_interesting\\_offer.html](http://www.oreillynet.com/xml/blog/2007/01/an_interesting_offer.html)
- Landsburg, S., *Price Theory and Applications*, South-Western, 2001.
- Lasica, J.D., "Blogs and journalism need each other." *Harvard Neiman Reports*, 2003, 57(3), 70.
- Lerner, J. and J. Tirole, "The simple economics of open-source." *NBER working paper*, 2000.
- Lerner, J. and J. Tirole. "Some simple economics of open-source," *Journal of Industrial Economics*, 2002, 46(2), 125-156.
- Musil, S., "This week in wikis," *cnet News.com*, Online, December 9 2005, available:<http://news.com.com>, last accessed December 9, 2005.
- Nash, J.F. "Equilibrium points in n-person games." *Proceedings of the National Academy of Sciences*, 1950, 36, 48-49.
- Regan, T., "Weblogs threaten and inform traditional journalism," *Harvard Neiman Reports*, 2003, 57, 68.
- Van Huyck, J. B., R. Battalio, and R. Beil, "Tacit coordination games, strategic uncertainty, and coordination failure," *American Economic Review*, 1990, LXXX, 234 -48.
- Varian, H., *Microeconomic Analysis*, New York, NY: Norton, 1984.
- Varian, H., "Sequential contribution to public goods," *University of Michigan – Center for Research on Economic and Social Theory paper*, 1994.
- Varian, H., "System reliability and free-riding." *UC Berkeley working paper*, 2004.
- Vega-Redondo, F., *Economics and the Theory of Games*, Cambridge University Press. England, 2003.
- Wattenberg, M. F. Viegas, and K. Hollenbach, "Visualizing activity on Wikipedia with Chromograms," *Human-Computer Interaction – INTERACT 2007*. 2007, pp. 272-297.

## AUTHOR



**Dr. Mohammad Rahman** teaches in the School of Business at North South University in Bangladesh. Dr. Rahman's research involves the application of rigorous techniques in explaining puzzling phenomena and 'real-life' problems that arise

in business administration. Consequently his research efforts are inter-disciplinary and offer key insights to managers. Dr. Rahman teaches technology, operations and strategic management in the MBA program. He received a Ph.D. in Business Administration from the University of Kansas, with a concentration in decision sciences.