Cooperation in a Peer Production Economy
Experimental Evidence from Wikipedia*

Jérôme Hergueux† Yann Algan‡ Yochai Benkler§ Mayo Fuster Morell¶

December 7, 2015

Abstract

We experimentally elicit social motives among Wikipedia contributors to extend the literature on public goods provision beyond its current focus on explaining individual contributions. We highlight the heterogeneity of contributors, analyze the determinants of their level of collaborativeness and study the policing activities of platform administrators. Reciprocity, altruism and social image motives all strongly relate to field contributions, but differently for "casual" and "super" contributors. More reciprocal and altruistic subjects are more cooperative when contributing, while subjects who reveal social image concerns start relatively more editing conflicts. Decreased experimental trust is associated with increased policing activity among Wikipedia administrators.

JEL classification: H41, C93, D01, Z13

Keywords: Field Experiment, Public Goods, Social Preferences, Peer Production, Internet, Wikipedia

* We gratefully acknowledge financial support from the European Research Council (ERC Starting Grant) and logistical support from the Sciences Po médialab and the Wikimedia Foundation. We are grateful to Anne l'Hôte, Romain Guillebert and David Laniado for outsdanding research assistance. We are greatly indebted to Dario Taraborelli for his instrumental help in moving this project forward, Benjamin Mako Hill and Aaron Shaw for generously sharing some of their data and Timo Tijhof for coding the recruitment banner in coordination with us. This paper has benefited from invaluable conversations with Nava Ashraf, Sam Bowles, Shane Greenstein, Guillaume Fréchette, Stuart Geiger, David Laibson, Sendhil Mullainathan, Klaus Schmidt, Andrei Shleifer, Gabriel Smagghue and Marie Claire Villeval. We thank the Berkman Cooperation group and numerous seminar participants for very helpful comments. Our last thank goes to the Wikimedia community and, in particular, to the 850 Wikipedia contributors who took part in this study.

† *Job market paper.* ETH Zurich, Center for Law and Economics. e-mail: jerome.hergueux@gess.ethz.ch

‡ Sciences Po, Department of Economics. e-mail: yann.algan@sciencespo.fr

§ Harvard Law School. e-mail: ybenkler@law.harvard.edu

¶ Open University of Catalonia. e-mail: mayo.fuster@eui.eu

"The problem with Wikipedia is that it only works in practice. In theory, it can never work."
Kizor, Wikipedia administrator.[1]

1 Introduction

Wikipedia, the collaboratively edited encyclopedia, is one of the Internet's most valuable global public goods. With 37 million freely usable articles in 285 languages and 500 million unique visitors per month worldwide, its revealed informational value seems to be enormous to society.[2] Although the reliability of Wikipedia's content remains subject to some controversy, research has shown that its scientific entries are practically indistinguishable from the Britannica ones (Giles 2005)[3], while its political entries, provided that they attract a sufficient number of contributors, are actually less biased than their Britannica counterparts (Greenstein and Zhu 2014).

In this paper, we run an online experiment on the Wikipedia website to elicit the prosocial motivations of its contributor base. We then couple this information with the detailed records of contributors' activity within the site. The richness of the Wikipedia setting allows us to extend the literature on public goods provision beyond its current focus on explaining individual contributions. We stress the heterogeneity of contributors, look at *how* individuals behave while contributing and shed light on the resource's governance system (Ostrom 1990). Wikipedia is a particularly clean setting to perform those tests, as (i) it is difficult to derive monetary rewards from one's contributions to the project, (ii) those contributions carry little signaling value on the labor market, and (iii) contributors are remote strangers who interact through the site with the sole purpose of providing the public good (Reagle 2010). Another interesting feature of Wikipedia is that it is a work environment in which individuals voluntarily self-assign tasks and monitor their work, in the absence of monetary incentives and formal leadership (Benkler 2006). This feature crucially distinguishes Wikipedia from for-profit organizations, but also traditional volunteering ones, where production processes crucially relies on hierarchy to coordinate and monitor individuals' work. Understanding the nature of the work incentives at play in the Wikipedia context is therefore important, as it appears as particularly close to the ideal-type of "modern high-performance work systems" (Bartling, Fehr, and Schmidt 2012), where organizations seek to reduce monitoring costs and increase productivity by granting more autonomy to workers over the choice of their tasks and effort level, at the risk of shirking.

We extend the current literature in three distinct directions. First, Wikipedia allows to explore the relevance of the existing theory of the private provision of public goods for the population of "super contributors" – i.e. individuals who exhibit extreme patterns of involvement with the project, having contribution records that exceed several thousand edits. Indeed, while discrete examples of individuals making extreme contributions to public goods abound, the sheer size of Wikipedia's contributor base provides an opportunity to recruit a sample which allows statistical analysis. Second, we use the detailed records of contributors' activity in order to explain the level of collaborativeness of individual editors while contributing. Indeed, irrespective of the amount of content that they contribute, uncivil contributors can work against the overall provision of the public good, as their behavior increases the cost of cooperation for other contributors, which may drive them away from the project. Third, we

[1] See http://www.nytimes.com/2007/04/23/technology/23link.html?ei=5124&en=435e5b69b6b3ceac&ex=13&_r=0, accessed February 2013.
[2] See http://www.nytimes.com/2014/02/10/technology/wikipedia-vs-the-small-screen.html?_r=1, accessed November 2015.
[3] As a matter of fact, 60% of European doctors declare using Wikipedia for professional purposes. See http://www.pmlive.com/find_an_article/allarticles/categories/General/2011/june_2011/features/dr_wikipedia_will_see_you_now..._280528, accessed November 2015.

analyze how trust in strangers determines the extent to which platform administrators use their policing rights within the site. Wikipedia administrators notably use those rights in order to protect the existing resource from being vandalized by malicious contributors. Since the decision to exercise them always involves a trade-off between the necessity to protect the resource and the risk of driving well-intentioned contributors away, it is important to understand the determinants of such decisions.

Controlling for a vector of demographic variables, we find strong evidence in favor of the existing theory of the private provision of public goods. The quantity of field contributions made to Wikipedia by our subjects is strongly related their taste for reciprocal exchange (Rabin 1993; Dufwenberg and Kirchsteiger 2004; Falk and Fischbacher 2006) and to their revealed preference for social image (Holländer 1990; Bénabou and Tirole 2006; Andreoni and Bernheim 2009; Ellingsen and Johannesson 2008; Ellingsen and Johannesson 2011). Altruism also plays a significant role (Andreoni 1989; Andreoni 1990; Anderson, Goeree, and Holt 1998), but only concerns a minority of contributors.

There is much less support for the existing theory when one tries to explain the contribution patterns of super contributors, however. First, super contributors do not statistically differ from casual ones in terms of their average prosociality levels. Second, above and beyond all demographic characteristics, the only significant predictor of increased participation within this group is a revealed preference for social image within the community of editors. At the same time the super contributors who are at the very top in terms of contribution levels are actually more likely to reveal selfish preferences than reciprocal or altruistic ones in our experiment.

Turning to individual collaborativeness, we find that more reciprocal and altruistic subjects are significantly more cooperative while contributing. Strikingly, however, subjects who reveal social image concerns start relatively *more* editing conflicts on average. The fact that those subjects sometimes behave in a relatively more confrontational way is aligned with the finding that, contrary to common wisdom, the community of engaged editors seems to (marginally) reward such behavior under certain conditions. More research would be needed on this point, since our results suggest that, depending on the social incentives to which they are exposed, a trade-off may arise between the value of the contributions that editors motivated by social image make and the potential externality costs that they could impose on other contributors.

Finally, as far as governance is concerned, we find that experimental trust is a strong predictor of the behavior of the administrators in charge of enforcing Wikipedia's policy rules. Less trusting administrators are significantly more likely to block other users from editing, more likely to delete recent Wikipedia pages and dedicate a higher proportion of their working time on Wikipedia to admin related activities. Future research should therefore seek to tackle the question of the optimal level of trust that administrators should exhibit in order to efficiently fulfill their role.

Prior economic research on Wikipedia *per se* is relatively scarce. One important paper in our context is Zhang and Zhu (2011). The authors exploit a natural experiment at Chinese Wikipedia to investigate the role of group size on incentives to contribute to public goods. They show that an exogenous reduction in the size of the community of contributors led to a decrease in individual contributions. The authors therefore hypothesize that "social benefits" are likely to accrue to contributors as the size of their group grows. Our paper confirms their hypothesis and makes it precise by pointing at the prominent role of reciprocal exchange and social image motives for incentivizing sustained contributions.

Our work is closely related to a growing literature which explores the predictive power of experimental measures of social motives on field outcomes. In his seminal work, Karlan (2005) uses the Trust game to obtain individual measures of reciprocity and shows that those can be used to predict loan repayment among participants in a microcredit program in Peru. Laury and Taylor (2008) and De Oliveira et al. (2009) relate the propensity of their subjects to cooperate in a Public Goods game in the

lab to their propensity to contribute to a charitable cause in the field. Benz and Meier (2008) collect field data about their subjects' behavior in a charitable giving situation prior to conducting a charitable giving experiment in the classroom. Barr and Serneels (2009) conduct a Trust game among workers in Ghana and establish a relationship between individual measures of reciprocity and the observed labor productivity of their firm at the aggregate level. Similarly, Carpenter and Seki (2011) conduct a repeated Public Goods game among Japanese fishermen and show that fishing crews that exhibit higher levels of reciprocity are more productive. Perhaps most similar to the present study, Carpenter and Myers (2010) rely on an experimental measure of altruism and an observational measure of social image concerns within a population of volunteer firefighters and non-volunteer community members to show that both preferences predict the decision to join the volunteer fire service. Fehr and Leibbrandt (2011) and Leibbrandt (2012) conduct a Public Goods game among Brazilian shrimp catchers and sellers, respectively, and show that more prosocial shrimp catchers are less likely to engage in overextraction, while more prosocial shrimp sellers achieve higher prices for the same goods. Finally, Gneezy et al. (2015) also conduct a battery of social preferences games with Brazilian fishermen. They show that social preferences evolve in the long run as a result of natural factors, as communities which settled in an environment in which working in groups is necessary exhibit higher current levels of cooperation and trust.

By focusing on an Internet-based global public good, this paper distinguishes itself by looking at several dimensions of cooperation and governance simultaneously. As Wikipedia contributors are remote strangers who only interact for the purpose of providing the public good, our findings should be more likely to generalize to other contexts and provide broader insights into the motivations for contributing. To the best of our knowledge, our design also improves on the existing literature by allowing us to assess the relative predictive power of prosocial motives within one empirical framework.

Throughout this paper, we largely assume that social preferences are exogenous. As we document the predictive power of social motives on a variety of field behaviors, we certainly cannot fully exclude reverse causality whereby social preferences would also evolve as a result of Wikipedia participation. While this is a growing concern in the literature, we think that the resulting bias would be relatively limited in our context. First, recent research suggests that social preferences are actually stable over periods of time that span several years (Carlsson et al. 2014). Furthermore, direct evidence from the patterns of contributions of Wikipedia editors suggests that learning effects cannot tell the whole story. Indeed, in an extensive investigation of Wikipedia contributors' activity, Panciera et al. (2009) show that engaged Wikipedia contributors virtually always start to contribute intensely from the time of their very first edit, hence their title: "Wikipedians are born, not made".

The rest of the paper proceeds as follows. Section 2 provides some important knowledge background on the Wikipedia project and its community of contributors. Section 3 reports on the design and implementation of the study. Section 4, documents the association between prosocial motives and the quantity of contributions made to Wikipedia, while section 5 focuses on explaining contributors' individual collaborativeness while contributing. Finally, section 6 focuses on the sub-group of Wikipedia administrators to study the relationship between experimental trust in strangers and policing activity patterns. Section 7 concludes.

2 Background on Wikipedia

Wikipedia is a free online encyclopedia that is collaboratively edited by volunteers over the Internet. Every Wikipedia reader holding some private information of potential value to the encyclopedia faces a standard public goods dilemma. While it is individually costly to put one's knowledge in convenient shape for the general public to use, the content is immediately available for the general audience to see and use at no cost, and can easily be modified and even deleted by other contributors. According to the standard rational actor model, this should lead to no contributions being made in the first place. Importantly, the cost of contributing valuable information to Wikipedia in terms of effort and time is of a different nature – and arguably higher – than the cost of contributing to, e.g., a personal blog. As nicely stated in the Wikipedia Neutral Point of View policy, "articles must not *take* sides, but should *explain* the sides, fairly and without bias. This applies to both what you say and how you say it."[4] Obviously, Wikipedia would not be considered such a useful informational resource if it was merely a place for individuals to push their own personal views. Contributors are therefore expected to communicate knowledge in an encyclopedic format, provide reliable secondary sources for their claims, and resolve disputes through constructive discussions and consensus building. Since any contributor can easily revert the contributions of any other, this laudable goal would probably go unheeded without some shared cooperative norms and prosocial standards.

The Wikipedia project originates in Jimmy Wales and Larry Sanger's attempt at creating a traditional, extensively peer-reviewed online encyclopedia called "Nupedia" in March 2000. The goal of Nupedia was to get scholars and experts to volunteer their work and expertise to the project, with the goal of creating a free equivalent of the existing for-profit encyclopedias. Confronted with the difficulty of taking the project off the ground – Nupedia only got 21 articles finalized in its first year – Wales and Sanger eventually released Nupedia's content over the Internet in January 2001 as an open side project, called "Wikipedia", whose original purpose was to feed Nupedia with additional draft articles. Wikipedia quickly overtook Nupedia and became a multiple language popular project of its own, with over 20,000 encyclopedia articles created in its first year and an exponential growth of its content and contributor base since then.

Since 2003, Wikipedia has been operated by the Wikimedia Foundation, a small San Francisco-based non-profit organization, whose role is to pay the bandwidth bills, buy the servers and provide legal defense for the project. The Wikimedia Foundation mostly leverages the capital that it needs to perform this function through donations. It is important to note that while the Foundation is interested in developing technical and social solutions that could support volunteers' editing work, it has never been directly involved in developing Wikipedia's content or managing its community of contributors. This is a matter of principle, and the relationships between the Wikimedia Foundation and the body of engaged Wikipedia contributors have sometimes been notably tense, as some would repeatedly suspect the Foundation of covertly trying to influence the evolution of the project and direct its development.

On the technological side, Wikipedia is based on the wiki system, which allows the reader of any Wikipedia page to modify it easily and rapidly by clicking on an "edit" button. As a result, there exist no limitations *à priori* as to whom can contribute content to the encyclopedia. It is not necessary to create a Wikipedia account in the website in order to contribute, as this can be done "anonymously". Many regular contributors choose to create a Wikipedia account, however, notably because it gives them access to very useful collaborative editing tools. One prominent example of such a tool is the so-called "watchlist" system, which allows registered users to mark pages of interests and get automatic notices whenever a modification is implemented to them by another contributor. The wiki system archives each and every version of a given page in a revision history, together with the username of the registered contributor who authored the revision. (Contributions made "anonymously" are registered together

[4] See https://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view, accessed December 2013.

with the IP address of the computer from which it was performed.) It is customary for contributors to leave a brief summary of their contribution together with the reason why they implemented it upon saving their modifications. This "edit summary" can be read directly from the revision history of any page, which allows contributors to get a very quick sense of each modification and the justification behind its implementation. If necessary (for instance in cases of vandalism), the revision history allows contributors to easily revert a page to one of its previous state.

If they create a Wikipedia account, contributors automatically get a personal user page and a user talk page on the Wikipedia website. Those pages, like virtually every other on Wikipedia, can be edited by anyone. User pages are mostly edited by their owners to post some general information about themselves, their interest in Wikipedia, the articles they helped improve and the like. As collaborations between editors mostly form when they notice that they contribute to the same articles, either through its revision history or the watchlist system (as opposed to randomly scrawling contributors' personal user pages in search for an editor with matching interests), those pages are not crucial to the functioning of Wikipedia. Hence, a significant number of contributors choose to leave them blank. User talk pages, by contrast, are mainly edited by one's fellow editors. They play a very critical role on Wikipedia, as they are used as a convenient place for contributors to communicate with one another, request help, ask questions and coordinate work. Taken together, those technical features explain that even if many individuals do contribute to Wikipedia without having registered an account, the contributions made in this fashion are more likely to be one-offs and, in any case, cannot be much collaborative in terms of content.

The number of contributions made to Wikipedia by registered users follows a strong power law distribution. Skewness of participation characterizes many technology mediated peer production systems. It is not unique to Wikipedia and is also a structural feature of individual contributions to Open Source Software and participation in online message boards. As of 2011, about 200,000 individuals register an account on Wikipedia each month. About 2% of those individuals make 10 contributions or more within their first month, which certainly represents a non negligible influx of new contributors per month in absolute terms. However, only 10% of those early contributors still make one contribution or more within the following year.[5] As a result, the relatively limited body of editors who eventually become engaged and reach the threshold of 100 Wikipedia contributions was still responsible for almost 70% of all the contributions made in 2007 (Kittur et al. 2007). Even within the group of editors who become engaged with the project, individual contribution patterns are still highly heterogeneous. While the vast majority of engaged editors have a few hundred to a couple thousand contributions in total, about 5,000 of them (the so-called "super contributors") made more than 10,000 contributions and about 200 editors have contribution records ranging from 100,000 to 1,000,000 contributions. Overall, the size of the body of active experienced contributors who reached the threshold of 300 contributions is relatively stable since 2007, revolving around 20,000 individuals.

One surprising fact about Wikipedia is the ability of its community of engaged contributors to successfully synthesize into coherent and structured articles their often competing or opposed views about the topics at hand in a civil way. In this respect, it is interesting to note that among subjective topics, the most controversial ones are on average *better* treated within Wikipedia, precisely because they attract attention from a larger and more diverse pool of contributors (Greenstein and Zhu 2013).[6]

[5] See http://stats.wikimedia.org/EN/TablesWikipediaEN.htm and http://strategy.wikimedia.org/wiki/Editor_Trends_Study/Results, accessed February 2013.

[6] Controversy on Wikipedia is not limited to "hot" topics such as global warming or the Israeli-Palestinian conflict. One of many mundane examples is the controversy that arose in 2006 around the article on arachnophobia, in which one contributor added the picture of a tarantula. The question of interest was to determine whether the picture had any illustrative value or if it would simply drive potential readers away. A consensus eventually emerged around the idea of replacing the picture by a

As a result, Wikipedia articles which attract a lot of contributors are on average less biased than their Britannica counterparts (Greenstein and Zhu 2014). Reagle (2010) provides a very detailed account of how relationships within the community of engaged editors are generally driven by common behavioral norms that emerged through progressive consensus building as it faced collective action problems. One paradigmatic example of such a norm is the "neutral point of view" policy. It is remarkable that this policy doesn't state that editors should strive to be "neutral" or "objective" while contributing to a given article, but that a "fair" representation of all sides of the dispute should be sought. Conditional on being able to support one's point with reliable secondary sources, this guiding principle has the positive effect of shifting many debates from the question of whether it should be included in the article to the question of *how* it should be included. Another example is the "assume good faith" principle, which exhorts editors to approach others' contributions as being made in good faith and trying to help the project, unless there is specific evidence of malice. When direct discussion fails to resolve disputes among contributors, this is usually achieved by extending the debate to a larger audience, or seeking the mediation of a third party.

Besides the sheer number of contributions that they make to Wikipedia, the body of engaged contributors is thus key to the project, as they often make contributions across topical boundaries in order to curate the content and turn it into a comprehensive resource, help newcomers learn the behavioral norms and attitudes that will allow them to connect with others and make valuable contributions to the project and informally mediate disputes. In this sense, engaged Wikipedia contributors create the public good value of the encyclopedia, and distinguish its contributor base from a broad collection of individuals trying to push their own personal agendas within the site.

One particular class of super contributors, the Wikipedia administrators, are in charge of dealing with disruptive editors when good faith discussion and basic explanations about what the goal of the project is fail. To do so, they are entitled with special oversight rights over the encyclopedia which allow them to enforce the behavioral norms of the community, notably by blocking malicious editors, deleting pages that they think have no potential for developing as proper encyclopedic articles and protecting vandalized pages from editing. To obtain those policing rights, those engaged contributors decided to participate in a very competitive peer-review process that would require them to prove through their contribution history that they are valuable contributors who can handle heated debates and make difficult decisions.

3 Design of the study and data

In this section, we first describe our strategy for measuring social motives among our subjects. We then describe our experimental procedures before reporting on the practical implementation of the experiment.

3. 1 Measuring social motives

(i) Public Goods game. We rely on a modified version the Public Goods game to elicit subjects' reciprocity and altruism motives (see Fischbacher, Gächter, and Fehr 2001). First, we elicit subjects'

cartoon illustrating the fear of spiders, but only after several editors had spend hours on the issue, generating around 6,000 words of discussion for an article which is about 1,500 words long.

propensity to cooperate in a very standard Public Goods dilemma (see figure 1 which pictures the Public Goods game instructions screen). Subjects play in groups of four with an initial endowment of \$10 per player. Each dollar invested in the common project by a member of the group yields a return of 0.4 dollar to each group member.[7] Subjects have to decide on how much of their \$10 they want to invest in the common project. In a second step, we implement the so-called "strategy method" and ask subjects to provide their intended contribution for each possible value (on the scale of integers from 0 to 10) of the average contribution of the three other members. Subjects are told that their actual contribution to the common project will be randomly determined to be either their unconditional contribution from the standard Public Goods game or their conditional contribution decision.

[FIGURES 1 AND 2 ABOUT HERE]

We classify subjects into four exclusive cooperative types depending on their revealed preferences in the conditional Public Goods game. To do so, we compute (i) the slope of subjects' reaction functions to the possible average contributions of the other group members (i.e. reciprocity r) and (ii) the average proportion of the endowment that is conditionally contributed across all 11 conditional contributions decisions (i.e. mean contribution m). We then classify subjects according to the following rule:

- Free riders: $r < 1$ & $m \leq 0.2$
- Weak reciprocators: $r < 1$ & $0.2 < m < 0.8$
- Reciprocators: $r \geq 1$
- Altruists: $r < 1$ & $m \geq 0.8$

This classification distinguishes between "weak" and "non weak" reciprocators in order to remain consistent with the typology initially proposed by Fischbacher et al. (2001).[8] In the below analysis, we also report the results obtained from a more stringent classification rule which requires free riders to satisfy $m \leq 0.1$ as opposed to $m \leq 0.2$ (and, conversely, altruists to satisfy $m \geq 0.9$ as opposed to $m \geq 0.8$). Both classification rules allow our subjects for some range of decision error, while requiring that free riders satisfy $m = 0$ (i.e. the subject never makes a positive contribution, irrespective of the average contribution of the other members of the group) and that altruists satisfy $m = 1$ (i.e. the subject always contributes all of his endowment, irrespective of the average contribution of the other members of the group) does not change the nature of our results.

(ii) Trust game. Following the Public Goods game, we conduct a standard Trust game among our subjects. Subjects are sequentially attributed a role (according to their login order): either participant A or participant B. Each participant A is matched with a participant B, and both players receive a \$10 initial endowment. Participant A is the trustor and chooses how much of his endowment is transferred to participant B – the trustee. The trustee receives three times the amount sent by the trustor, and chooses how much is sent back to him. We elicit this decision through the strategy method: for each possible transfer from the trustor (from 1 to 10) the trustee chooses how much will be returned without knowing the trustor's actual choice.

[7] Each subject thus faces the following payoff function: $\pi_i = 10 - contrib_i + 0.4 \sum_{j=1}^{4} contrib_j$ (1)

[8] In the lab, the distinction between those two types of reciprocators turned out to be important. Groups constituted of reciprocators usually succeed in sustaining their contribution levels in repeated Public Goods experiments. On the other hand, the presence of free-riders and weak reciprocators in a group usually triggers a progressive decline in cooperation levels.

At the end of the experiment, we ask subjects some standard demographic questions about their age, gender, education and salary level, along with an experimentally validated question on risk aversion taken from Dohmen et al. (2011).

(iii) Social image motives. Social image motives are difficult to measure experimentally – even more so in a decontextualized fashion, that is, out of a given social context. As a result, we rely on the observational data available from Wikipedia in order to elicit this preference. First, we collect the size (in bytes) of the personal user pages of our subjects (recall from section 2 that personal user pages do not play an important functional role in Wikipedia and are mainly used to present oneself to the community of contributors). We then code as "social signalers" those contributors who have a personal Wikipedia user page whose size is higher than the median in the sample, and take this variable as a measure of subjects' social image motive.

In order to check for the consistency of our results, we exploit Wikipedia's main social rewarding practice – the Barnstars system – to provide an alternative measure for this motive. A Barnstar is a highly valued symbolic award within the community of contributors. It is typically constituted of an image accompanied by a personalized message acknowledging some important contribution made to the project by an editor (see figure 3 for an example).[9]

[FIGURE 3 ABOUT HERE]

Technically speaking, anyone can give or receive a Barnstar. This social practice, however, remains largely limited to the body of Wikipedia contributors who display relatively impressive contribution records. Barnstars are typically posted on a contributor's talk page. They thus appear within the flow of conversations between this contributor and the rest of the community. After some time, however, discussion threads are likely to be archived and/or become too long for anyone to easily notice that an award had been given. Some Wikipedia contributors choose to circumvent this by manually moving (some of) their Barnstars to their personal user pages (or some particular subsection of their user page generally labeled their "awards page"), so that they would be durably and prominently displayed for any other editor to see. We take such decisions as revealing a contributor's motive for social image. From the subsample of subjects who received Barnstars (about 48% of our sample, the vast majority of whom are super contributors) we therefore code as "social signalers" those who decided to display at least one of those awards on their personal user page, and take this variable as an alternative measure of subjects' social image motive.

3. 2 Experimental procedures

The implementation of online experiments typically poses a number of practical challenges and threats to internal validity. In this paper, we rely on Internet-specific experimental procedures which have been specifically developed so as to maximize their internal validity, and methodologically assessed through a rigorous lab / online comparison (Hergueux and Jacquemet 2015). The welcome page of the decision interface provides subjects with general information about the experiment, including the

[9] The Wikipedia "Barnstars" page starts as follows: "It is the custom to reward Wikipedia contributors for hard work and due diligence by awarding them a barnstar. To give the award to someone, just place the image on their talk page (or their awards page), and explain why it was given. If you are sure the barnstar is appropriate, *don't be shy!"* See http://en.wikipedia.org/wiki/Wikipedia:Barnstars, accessed February 2013.

number of sections, expected completion time (about 25 minutes) and how their earnings will be computed. In order to minimize potential demand effects and in-group biases when eliciting subjects' social motives, we take great care to present the study as non Wikipedia oriented.[10] Importantly, we make it very clear on the introductory screen that subjects will interact with a diverse pool of Internet users.[11]

Subjects are only informed of their earnings in each game at the very end of the experiment. Final payoffs are equal to the earnings from one randomly selected game plus a $10 participation fee. Subjects get paid upon completion of the experiment through an automated PayPal transfer.[12] We only require a valid e-mail address to process the payment. To strengthen the credibility of the payment procedure, we ask subjects to enter the e-mail address that is (or will be) associated with their PayPal account right after the introductory screen of the decision interface. It is important to stress that Wikipedia contributors can be very hostile to monetary rewards. In order to ensure that the experiment is equally incentive compatible for all subjects, we allow them to donate any amount taken from their final earnings to the Wikimedia Foundation and/or the International Committee of the Red Cross – a renowned and general purpose charitable organization, in anticipation of the fact that some subjects might not want to donate to the Wikimedia Foundation – upon completion of the experiment. This possibility was made clear on the welcome screen of the decision interface. It was not possible, however, to commit to donating one's final earnings prior to the study's completion.

All decisions made by our subjects are anonymous. This is because contrary to the social image motive – which is by definition a *public* social preference – the preferences that we elicit experimentally are *private* social preferences, meaning that they do not depend on the visibility of one's actions to be at work.[13] As we want to elicit social motives in isolation from strategic concerns and learning effects, each game is only played once and we match subjects in each game according to a perfect stranger procedure.

One important methodological concern with the online implementation of the experiment is to guarantee a quick and appropriate understanding of the decision problems when no interaction with the experimenter is possible. We strengthen the internal validity of the online experiment with three distinctive features of the interface. First, we include suggestive flash animations illustrating the written experimental instructions at the bottom of each game's instruction screen (see figure 2 for the example of the standard Public Goods game).[14] Second, the instructions screens are followed by a

[10] The specific language used on the welcome page was as follows: "Our goal is to better understand the dynamics of interactions and behavior in online social spaces. To do so, **we invite internet users with various profiles to fill out an interactive survey**. We very much welcome participation from Wikipedia users!" Our strategy for framing the study as non Wikipedia oriented eventually proved more effective than we had anticipated. When we presented this research project to the Wikimedia Foundation staff, their initial reaction was: "Several people expressed concerns that there was not a clear connection between the contents of the survey and data that would be strategically useful at this time to Wikimedia community members and the Foundation. […] We hope that you will find another suitable outlet to recruit participants for your study. We're happy to answer questions about this decision, and we hope in the future to be able to support other projects you may be working on that are relevant to Wikimedia."

[11] The Wikipedia subjects were matched with a traditional pool of laboratory subjects and with open-source software developers who both previously participated in a similar online experiment.

[12] Such a payment procedure guarantees a fungibility similar to that of cash transfers in lab experiments, as money transferred via PayPal can be readily used for online purchases or easily transferred to one's personal bank account at no cost.

[13] Note that the concept of "social image motive" as it currently stands in the economics literature conflates several motives (e.g. relative social status within a group or relative competence assessment) all of which crucially depend on the visibility of one's actions to be at work. We do not try to distinguish between those in this paper.

[14] The loop of concrete examples displayed in each animation was first randomly determined and then fixed for each game. The same loop is displayed to all subjects without any other numeric information than the subjects' initial endowments. We decided against displaying a purely random sequence of flash animations as it could have introduced uncontrolled and subject specific noise–through, e.g., anchoring on a particular behavior or sequence of events. Our goal with those animations was to illustrate the basic gist of each decision problem in an accessible way while avoiding to prime specific numerical examples and results in subjects' mind.

screen providing some examples of decisions, along with the detailed calculation of the resulting payoffs for each player. These examples are supplemented on the subsequent screen by earnings calculators. On this interactive page, subjects are allowed to test all the hypothetical scenarios they are interested in before making their decisions in the Public Goods and Trust games. In contrast to the illustrative flash animations, the numeric results of each scenario run by a subject in the earnings calculator screens are explicitly displayed. Last, the system provides a quick access to the instructions material at any moment during decision-making. On all screens, including decision-making ones, a "review description" button gives subjects a direct access to the instructions displayed at the beginning of the game. The system also allows participants to navigate at will from one screen to another – until a decision screen has been passed – through the "Previous" and "Next" buttons located at the bottom of each screen (see figure 2 for the example of the conditional Public Goods game decision screen).

3. 3 Implementation of the experiment

Our goal is to recruit as representative a sample of the underlying population of Wikipedia contributors as possible. We therefore need to capture the very wide heterogeneity that characterizes registered editors' contributions patterns (see section 2 for some background statistics on this structural feature of Wikipedia participation). To do so, we recruit subjects from the three following groups:

(i) The cohort of recently registered Wikipedia contributors, defined as all individuals who registered a Wikipedia account within the 30 days prior to the launch of the experiment, irrespective of the number of contributions that they made (if any). Eligible population = 190,327 subjects.
(ii) The group of active Wikipedia contributors, defined as all contributors who made at least 300 contributions to Wikipedia and are currently active (i.e. they made at least 20 contributions in the last 180 days).[15] Eligible population = 18,989 subjects.
(iii) The group of Wikipedia administrators. Those highly engaged contributors successfully decided to run for a very selective peer-review process, at the end of which they were entitled with special oversight rights over the encyclopedia in order to perform a policing role. They notably can block disruptive users, delete pages that they consider will not develop as proper encyclopedic articles and protect vandalized pages. Eligible population = 1,388 subjects.

We use the Wikimedia banner system as a convenient recruitment device for our experiment. The Wikimedia Foundation typically relies on this banner system to advertize its annual fundraising, which makes it relatively familiar even to non Wikipedia contributors. It is also used by the community of editors for purposes of extended internal communication (e.g., to advertise events and other community initiatives). As a result, the banner system is certainly the most powerful and trusted way of reaching out to a wide and diverse audience within Wikipedia. In coordination with the Wikimedia Foundation staff, we code this recruitment banner so that it would be displayed at the top of every Wikipedia page for all logged-in eligible users, until he or she decides either to click on it, or to disable it (see Figure 4, which features the recruitment banner).[16]

[15] Note that this definition of an "active contributor" corresponds to the community's criteria for being eligible and able to vote in the 2011 elections of the Wikimedia Foundation Board of Trustees. See http://meta.wikimedia.org/wiki/Board_elections/2011/en#Prerequisites_to_candidacy, accessed February 2013.
[16] This was the first (and is still the only) time in the history of the Wikimedia movement that the banner system was left to use by a third-party. Its selective display system remains Wikimedia's most sophisticated one to date.

[FIGURE 4 ABOUT HERE]

Upon clicking on the banner, eligible users were uniquely and automatically identified by the system (through their Wikipedia user id number, which then allowed us to collect their entire contribution history to Wikipedia) and redirected to the welcome screen of our experimental economics platform. Within each of the three above-defined groups, our system sequentially allocated subjects to the role of participant A or participant B according to their login order. We implemented this procedure both to ensure that we get relatively balanced samples and to randomize the allocation of participants in the role of participant A and participant B. The experiment was launched on December 8th 2011 and the banner recruited 850 subjects in 8 hours (i.e. about 2 complete answers per minute).

3. 4 Data and descriptive statistics

We start by documenting whether we succeeded in capturing in our subjects' pool the wide heterogeneity in individual contribution patterns which is typical of Wikipedia participation. For each targeted group, figure 5 contrasts the distribution of the number of Wikipedia contributions for our sub-sample of participants with that of the whole population of eligible contributors. Focusing on the groups of active Wikipedia contributors and Wikipedia administrators, we can see that the distributions of the number of Wikipedia contributions for our subjects closely mirror those of the reference groups. We do capture contributors with higher contribution records on average, however, as we can see from both distributions being slightly skewed to the right.[17] This activity bias is particularly strong within the cohort of recently registered Wikipedia contributors, in which 11% of participants are already significantly active within Wikipedia and made more than 10 (and up to 273) contributions (as opposed to 2% only in the reference group).

[FIGURE 5 ABOUT HERE]

Another way to look at how representative of the overall population of Wikipedia registered editors our sample of subjects might be is to pool them all together and compare their demographic characteristics against those of the 5,073 registered editors who took part in the 2011 Wikimedia editor survey. Designed by the Wikimedia Foundation, this survey was precisely implemented so as to get a precise picture of the profiles of Wikipedia editors.[18] Similar to the present study, it was advertized through a Wikipedia banner. The survey ran for 7 days over the whole population of registered Wikipedia editors. Table 2 compares the commonly available demographic information in both studies. It appears that demographic characteristics between both samples are very similar. Contrary to the popular perception that most Wikipedia contributors are high school students, we find that they are on average much older (33 years old with 48% of the population being above 29 in our study versus 32 years old with 47% being above 29 in the Wikimedia editor survey) and more educated (63% have finished college and 28% have a Master's or a PhD degree in our study versus 61 and 26% in the Wikimedia editor survey, respectively). Consistent with the common perception, however, we find the population of contributors to be predominantly male (90% in our study versus 89% in the Wikimedia editor survey). Taken together, we

[17] A Kolmogorov-Smirnov test of the equality of distribution functions confirms this conclusion at $p<0.001$ in all three experimental groups.

[18] See http://meta.wikimedia.org/wiki/Editor_Survey_2011, accessed February 2013.

interpret the above evidence as suggesting that our sample of Wikipedia subjects is representative of the diversity of contribution patterns and demographic profiles found on Wikipedia.

[TABLE 1 ABOUT HERE]

Table 2, provides some descriptive statistics on our dependent variables, social preferences measures and demographic variables. In this table, we start to organize the sample for the subsequent regression analysis. "All subjects" refers to all Wikipedia contributors who are not administrators within the site (i.e. 730 subjects overall). Indeed, because of their policing role within the community of contributors, Wikipedia administrators exhibit some rather specific patterns of activity on which we focus explicitly in section 7.

When explaining the quantity of contributions made by regular contributors in section 4, we will also be interested in distinguishing "casual contributors" from "super contributors", i.e. those individuals who exhibit extreme patterns of engagement with the project, having contribution records that exceed several thousand edits. According to the framework developed in Von Krogh et al. (2012), those super contributors ought to be different from casual ones, in the sense that their engagement would not be driven by short-term social rewards. Rather, those contributors would have experienced a deep personal identification with the project and its community of contributors, and developed a meaningful narrative about the role that they see themselves fulfilling through their engagement. We therefore provide a first attempt at characterizing the motives of those super contributors through the existing theoretical framework. In order to capture this distinction, we divide our sample of regular contributors according to the median number of Wikipedia contributions made. In the below median group, "casual contributors" have contribution records ranging from 0 to 1903 contributions, while in the above median group, "super contributors" have contribution records ranging from 1907 to 364,157 contributions.

[TABLE 2 ABOUT HERE]

Finally, figure 6 presents the distribution of cooperative types within the full sample of 850 Wikipedia contributors, which we derive from their revealed preferences in the conditional Public Goods game. The figure sequentially requires free riders to satisfy $m = 0$, $m \leq 0.1$ and $m \leq 0.2$ (and, conversely, altruists to satisfy $m = 1$, $m \geq 0.9$ and $m \geq 0.8$). (See section 3.1 for the full classification rule.)

The first interesting feature that emerges from this figure is that between 7 and 12% only of subjects behave as free riders, as opposed to 30% in the original paper by Fischbacher, Gächter and Fehr (2001). Even more surprising, between 9 and 13% of contributors actually behave as unconditional cooperators in the game, in the sense that they contribute a very high proportion of their endowment irrespective of the contributions of the other members of their group. Such an altruistic pattern of behavior is remarkable, and had not been identified in the original paper. Finally, 38 to 47% of subjects behave as weak reciprocators while 38% behave as perfect or strong reciprocators.

One important feature of this distribution is that, irrespective of the classification rule on which we rely, the proportions of cooperative types do not statistically differ between the populations of casual and super contributors, nor do they differ between those two groups and the group of Wikipedia administrators.[19]

[19] Depending on the classification rule, a two-tailed t-test of this difference between casual and super contributors yields p=0.91 and p=0.80 for free riders, p=0.11 and p=0.12 for weak reciprocators, p=0.32 and p=0.32 for reciprocators and p=0.33 and p=0.23 for altruists. Running the same test between those two groups and the group of Wikipedia administrators yields p=0.49 and

[FIGURE 6 ABOUT HERE]

4 Social motives and quantity of contributions

In this section, we focus on explaining the quantity of contributions made to Wikipedia by regular Wikipedia editors. First, we present estimates of the relationship between subjects' demographic characteristics and their patterns of contribution to Wikipedia. We then test for the predictive power of each elicited class of social motive – altruism, reciprocity and social image – on subjects' contributions.

4. 1 Specification and data analysis

The dependent variable of interest in this section is the total number of Wikipedia contributions made by each subject over his history with the project. A Wikipedia contribution, or "edit", is defined as the action of (i) going to a Wikipedia page (ii) hitting the "edit" tab (iii) implementing a modification and (iv) saving the modification. Our results, however, are robust to using a variety of other ways of measuring the quantity of contributions made to Wikipedia. We notably tried: (i) using the overall amount of content contributed to the project (measured in bytes), (ii) restricting the count of the number of contributions to the 30 days prior to the launch of the study and (iii) using the total number of unique Wikipedia articles edited.[20]

As expected, the pattern of contributions to Wikipedia within our sample follows a strong power law distribution (see section 2). As our dataset is characterized by heteroskedasticity (Breusch-Pagan test: $p<0.001$), we do not report OLS regression tables based on a log-transformation of our dependent variable, as this would induce some bias in our estimates (Silva and Tenreyro 2006). Indeed, log-linearizing our dependent variable to run OLS regressions yields estimates that are higher in magnitude than the ones presented in this paper (our conclusions, however, remain qualitatively similar).[21] A more cautious approach is to use the negative binomial pseudo-maximum likelihood estimator, which is not affected by this problem (Wooldridge 2010). This estimator is appealing because (i) it naturally accounts for the skewness of our data and (ii) the coefficients remain nicely interpreted as semi-elasticities.[22]

We start the analysis by estimating the effect of subjects' demographic characteristics on the number of contributions that they make to Wikipedia (see table 3).[23] Column (1) focuses on the group of regular contributors (as opposed to Wikipedia administrators). The model globally confirms our above qualitative observations: being one year older is on average associated with a 1.7% increase in the number of Wikipedia contributions, while moving from a high school education to getting a Master's degree is associated with a 26% increase. Being a female, however, is associated with a 44% decrease in the number of contributions made to Wikipedia. The coefficient on the salary level variable is very close

---

p=0.66 for free riders, p=0.31 and p=0.47 for weak reciprocators, p=0.83 and p=0.83 for reciprocators and p=0.63 and p=0.68 for altruists.

[20] Since those tables do not add any significant information, we do not report them here.

[21] Tables available from the authors upon request.

[22] An alternative estimator that has similar properties is the Poisson pseudo-maximum likelihood estimator. One limitation of this estimator, however, is that it does not allow for overdispersion (which is a feature of our data, likelihood ratio test: $p<0.001$). The negative binomial estimator is more flexible and estimates the form of the dispersion as an additional parameter.

[23] Note that we do not report the pseudo-R^2 from those regressions, since this statistic has no natural interpretation in negative binomial models.

to zero and not statistically significant. This result is surprising, as it suggests that subjects' opportunity cost of time does not have any significant impact on their willingness to contribute to Wikipedia. Finally, the effect of risk aversion seems somewhat counterintuitive: moving from generally being "unwilling to take risks" to being "fully prepared to take risks" is actually associated with a 43% *decrease* in the number of Wikipedia contributions.

Those average effects conceal an interesting underlying heterogeneity within our population of subjects, however. In columns (2) and (3) we run separate analyses for the sub-populations of "casual" versus "super" contributors. We can see that while the effect of our demographic variables remains qualitatively the same within the group of casual editors, none of those variables reliably predict the contribution patterns of the super contributors to the project. The individual trajectories of the few contributors who exhibit such extreme levels of engagement with the project therefore seem very difficult to predicts, even using those very standard covariates.

Last, column (5) of table 3 presents the impact of our demographic variables on Wikipedia participation within the group of Wikipedia administrators. In this group, being one year older is on average associated with a 1.6% increase in the number of Wikipedia contributions made. The coefficient on the salary level variable achieves statistical significance within this group: out of 9 possible revenue categories, moving from one to the next is associated with a 5.2% decrease in the number of Wikipedia contributions made. This is interesting, as it suggests that the opportunity cost of time has a negative impact on Wikipedia participation among Wikipedia administrators.

[TABLE 3 ABOUT HERE]

4. 2 Social motives and quantity of contributions to Wikipedia

Table 4 estimates the relationship between social motives and subjects' willingness to sustain their contributions to Wikipedia by adding our measures of altruism, reciprocity and social image motives in the above regressions. Panel A requires that free riders in the conditional Public Goods game satisfy $m \leq 0.2$ (and, conversely, that altruists satisfy $m \geq 0.8$), while panel B adopts a more stringent classification rule by requiring that they satisfy $m \leq 0.1$ (and, conversely, that altruists satisfy $m \geq 0.9$).

Recall that we code as "social signalers" subjects who have a personal Wikipedia user page whose size (in bytes) is higher than the median in the sample. We provide an alternative measure for this motive by focusing the analysis on the sub-sample of subjects who received social awards – or Barnstars – from other Wikipedia contributors (i.e. 347 subjects, representing 48% of our sample), and coding as "social signalers" those who decided to advertize at least one of those awards on their personal user page.[24] According to this alternative measure, 50% of Barnstars receivers reveal a preference for social image. Importantly (and almost by definition), 81% of Barnstars receivers are super contributors. As a result, one limitation of this variable is that it can only tell us about the role of social image motives within the group of highest contributors to Wikipedia. As we expect subjects who receive more Barnstars to have a higher probability of exhibiting one of them on their personal user page (at least in a statistical sense), and as the total number of Barnstars received should be highly correlated with the number of Wikipedia contributions made, we include the total number of Barnstars

[24] We also tried using the *proportion* of received Barnstars that subjects decided to manually move to their personal user pages as an alternative indicator of their social image motive. The results were unaffected (table available from the authors upon request).

received as a control in all the regression which rely on this measure to avoid potential spurious correlations.

Focusing on column (1), one may conclude that only social image motives are associated with a significant increase in Wikipedia participation. Indeed, subjects who reveal social image concerns make on average 265% more contributions to Wikipedia, while, by contrast, the coefficients on the weak reciprocator, reciprocator and altruist variables are not statistically significant. Such a conclusion would be wrong, however, as distinguishing between casual contributors and super contributors turns out, again, to be important. (In this respect, recall from section 3.4, there are no significant differences in the proportions of cooperative types between both groups.)

Focusing on the group of casual contributors (column (2)), we find that all social motives are in fact associated with a significant increase in Wikipedia participation. Depending on the classification rule, weak reciprocators are estimated to make between 64 and 88% more contributions than free riders on average, reciprocators between 86 and 116% and altruists between 97 and 151% (while social signalers make almost 5 times more contributions than non social signalers).

By contrast, within the group of super contributors (column (3)), social image is the only motive which remains associated with an increase in contributions. In this group, social signalers are estimated to make 33% more contributions on average. In fact, if one is to adopt a relatively strict definition of cooperative types in the conditional Public Goods game (see panel B of column (3)), a surprising *negative* relationship appears between reciprocity and altruism preferences and the quantity of contributions made within this group, whereby free riders are estimated to increase their participation by 43 to 50% on average. Those results are confirmed when we turn our attention to column (4), where we rely on our Barnstars data to construct an alternative indicator of taste for social image within the group of Barnstars receivers. Here, social signalers are consistently estimated to make between 30 and 34% more contributions than non social signalers, while reciprocal and altruistic subjects actually make between 35 and 50% less contributions than free riders.

[TABLE 4 ABOUT HERE]

Taken together, our results point to the important role played by social motives in incentivizing sustained contributions to global public goods. Typically, those motives are more strongly associated with contribution patterns than the basic demographic characteristics that we collected, such as gender, education level and age. While our results on the group of casual contributors strongly confirm existing theories on the private provision of public goods – i.e. reciprocity, social image and altruistic preferences all matter, even if altruism only motivates a minority of contributors – the motivations of the minority of editors who become super contributors do seem to differ substantially. Above and beyond all basic demographic characteristics, social image concerns appear as the only remaining driver of increased participation within this group, while free riding types are actually more likely to move at the very top of the distribution of contributions than any other. Given those sharp differences and the prominent role that super contributors often play in the management of public goods, we hope that our results will encourage further research on this understudied group.

5 Social motives and collaborativeness of contributors

Irrespective of the quantity of contributions that individuals choose to make to the public good, understanding the determinants of individual collaborativeness while contributing will often turn out

to be important. Indeed, uncivil behavior while contributing can be highly detrimental to the project, as it imposes negative externalities on other contributors by increasing the cost of cooperation, and can potentially drive some well-intentioned contributors away from the project. The externality cost that uncivil behavior imposes on the wider community of contributors has been widely overlooked by the existing literature. This cost can be very high, however, so much so as to completely offset the intrinsic value of one's contributions. In the Wikipedia setting, for instance, confrontational behavior by some experienced contributors trying to assert their views in a non-collaborative way has been argued to be an important driver of the Wikipedia gender gap, whereby only 9% of contributors report being female in the 2011 Wikimedia editor survey (Reagle 2012).[25]

5. 1 Data on individual collaborativeness

In order to explain the level of collaborativeness of Wikipedia contributors, we measure two well documented and highly detrimental types of uncivil behavior in Wikipedia:

(i) Deleting (i.e. "reverting") the contribution of another editor without providing an explanation. Wikipedia contributors generally provide a brief summary for every edit that they make, so that other contributors can get a quick sense of its purpose. As the "Edit Summary" Wikipedia help page reads: "It is considered good practice to provide a summary for every edit, especially when reverting (undoing) the actions of other editors or deleting existing text; otherwise, people may question your motives for the edit."[26] Wikipedia contributors typically consider reverts that do not have an edit summary as highly uncooperative and harmful to the project. This is consistent with the evidence reported by Halfaker et al. (2011), who find that new Wikipedia contributors who see their contributions reverted by more senior contributors without an explanation are very likely to stop contributing. On the other hand, providing some lightweight feedback on the reason for the revert significantly counteracts this effect.

Within our sample, we identify 535 regular contributors who reverted the contribution of another editor at least once. Consistent with the above community consensus, not providing a justification for one's reverts is rather uncommon (only 4% of the overall number of cases). *For each reverting contributor, we take the proportion of his reverts that do not feature an edit summary as a measure of his collaborativeness while contributing.*

(ii) Starting an edit conflict (i.e. an "edit war") with another contributor. It is rather frequent for Wikipedia contributors to disagree on the content of a Wikipedia article. In such instances, the contributors involved in the dispute generally resolve the conflict by opening a public discussion thread on the talk page associated with the article and seeking to reach a consensus which could then be implemented in the article itself. Sometimes, however, a contributor decides not to resolve disagreements through discussion, but rather to use edits to fight directly with another editor. Such behavior is considered highly counterproductive and uncooperative by the community of contributors. As the "Edit Warring" Wikipedia policy page reads: "An **edit war** occurs when editors who disagree about the content of a page repeatedly override each other's contributions, rather than trying to resolve the disagreement through discussion. Edit warring is unconstructive and creates animosity between

[25] This explanation is also put forward by former Wikimedia Foundation executive director Sue Gardner. See https://en.wikipedia.org/wiki/Gender_bias_on_Wikipedia, accessed September 2015.
[26] See https://en.wikipedia.org/wiki/Help:Edit_summary, accessed October 2014.

editors, making it harder to reach a consensus. Users who engage in edit wars risk being blocked or even banned. Note that an editor who repeatedly restores his or her preferred version is edit warring, whether or not the edits were justifiable: it is no defense to say "but my edits were right, so it wasn't edit warring".”[27]

Within the subsample of 578 regular contributors who made more than 2 contributions to a Wikipedia article (so that we could meaningfully single out edit warring patterns), we seek to identify the following *consecutive* sequence of events: (i) the subject reverts another contributor *C* on a Wikipedia article, (ii) *C* reverts the subject on the same article and (iii) the subject reverts *C*, therefore restoring the initial version of the article. If such a pattern is identified, we say that the subject is starting an edit war with contributor *C*. Our results, however, are robust to varying the way we define edit wars.[28] On average, subjects start 4.6 edit wars over their contribution history. The pattern of started edit wars is highly skewed, however, as 65.5% of contributors in the sample never started one, while 26.8% started more than a hundred of them. *We take the overall number of edit wars that each contributor has started with another editor as a second measure of his collaborativeness while contributing.*

5. 2 Social motives and individual collaborativeness

We start the analysis by estimating the effect of subjects' demographic characteristics on their level of collaborativeness (see table 5). We follow the same estimation strategy as in section 4. The only departure is that we present OLS estimates for the proportion of reverts that do not feature an edit summary, so that the estimates are intuitive to interpret. An important point is that both indicators of collaborativeness are likely to be highly correlated with the overall conflictuality potential of the set of Wikipedia articles that subjects choose to contribute to. For instance, all else equal, we expect an editor contributing to the article on Islam to be more likely to start an editing conflict than an editor contributing to the article on Mathematics. As a result, we also control for the average conflictuality score of the set of articles edited by each contributor in all regressions. (Details on the construction of this indicator are available in the appendix).

We can see that from all the demographic variables, only the age of the subject is significantly associated with his collaborativeness level, with the magnitude of the effect being relatively small. Being 10 years older is associated with a 1.1% decrease in the number of reverts that do not feature an explanation (recall that, on average, 4% of reverts in the sample are not justified in one way or another), as well as with a 11% decrease in the number of editing conflicts started. It is also interesting to note that the average conflictuality score of the set of articles edited by each contributor is significantly and positively associated with the number of editing conflicts started, but also significantly and *negatively*

[27] See https://en.wikipedia.org/wiki/Wikipedia:Edit_warring, accessed October 2014.

[28] It is notably important to note that all of our result are robust to considering chains of 4 consecutive reverts instead of chains of 3 consecutive reverts. In this case, we identify the following consecutive sequence of events: (i) a contributor *C* reverts the subject on a Wikipedia article, (ii) the subject reverts *C* on the same article, (iii) *C* reverts the subject in turn, and finally (iv) the subject reverts *C* once again, therefore restoring his initial contribution. In addition to being a more demanding test in terms of non cooperative behavior, this sequence of events is interesting in the sense that, in this case, the sequence starts by the subject being first reverted by contributor *C* (and not the other way around). Indeed, with chains of 3 consecutive reverts, one could potentially argue that the contributions first reverted by our subjects were clearly identifiable as of very low quality, so that it would not have been efficient from the group perspective to incur the cost of opening community-wide discussions on the associated conflicts. In this context, the subject may have chosen to incur a personal cost (that associated with violating the letter of Wikipedia's standards of conduct by starting an inter-personal conflict) in order to achieve an efficient outcome at the group level (i.e. the deletion of the low quality content at the lowest aggregate cost). The fact that our results are robust to considering chains of 4 consecutive reverts disproves such an interpretation, since the low quality contributions would now be on the part of our subjects.

related to the proportion of reverts that do not feature an explanation, probably because contributors recognize the relatively higher necessity to explain their reverts when articles are inherently more conflictual.

[TABLE 5 ABOUT HERE]

Controlling for our vector of demographic variables, we then connect subjects' cooperative types and social image concerns with their level of individual collaborativeness (see table 6). Focusing on the proportion of reverts which do not feature an explanation, we can see that being a reciprocator or an altruist is significantly associated with a 5.7 and a 5.0% decrease in the number of non justified reverts, respectively (column (1), panel A). The coefficients on the reciprocator and altruist variables do not reach statistical significance is we adopt a more stringent classification rule, however (panel B). Furthermore, being a social signaler by the Barnstars measure is also associated with 2.3% decrease in the number of non justified reverts (column (2)). Finally, it is interesting to note that the overall number of Barnstars received by our subjects is negatively associated with the proportion of non justified reverts, which tends to confirm the fact that contributors recognize the value of providing an explanation whenever deleting the contribution of another editor.

Turning our attention to the number of edit wars started (column (3)), we find that, depending on the classification rule, being a weak reciprocator is significantly associated with a 64 to 96% decrease in the number of editing conflicts triggered, being a reciprocator is associated with 52% decrease and being an altruist is associated with a 57 to 72% decrease.

Strikingly, however, we consistently find that social image motives are associated with a significant 34 to 93% *increase* in the number of editing conflicts triggered by contributors, irrespective of the indicator used to measure this preference (columns (3) and (4)). Similarly, the overall number of Barnstars received by our subjects is significantly *positively* associated with the number of edit wars started – even though the size of the coefficient is not practically very large. This last result is suggestive of the fact that, conditional on providing justifications for one's reverts, the community of engaged contributors actually seems to value personal inclinations towards imposing one's views in a relatively confrontational way. Contrary to the letter of the Wikipedia policy rules, contributors who are motivated by social image concerns might recognize this fact and be incentivized to act accordingly.

At the end of the day, while contributors who reveal a taste for social image make a lot more contributions to the project on average, they also appear to sometimes behave less collaboratively while contributing. This result, if it replicates, points to a potential trade-off between the quantity of contributions that those contributors make and the negative externality that they could impose on some other editors through their reduced level of collaborativeness. Reciprocity and altruistic preferences, on the other hand, always go in the direction of more collaborativeness while contributing.

[TABLE 6 ABOUT HERE]

6 Trusting behavior and policing activities

We end-up our analysis by studying the prosocial determinants of the use of policing rights within the group of Wikipedia administrators. This group forms a distinct high status class of engaged

contributors who successfully opted-in a very demanding and competitive peer review process[29] at the end of which they were granted with special oversight rights over the encyclopedia. Wikipedia administrators are in charge of enforcing the behavioral rule and standards of the Wikipedia community when basic communication between contributors fails at achieving a cooperative outcome. To perform their policing and curating role, Wikipedia administrators are trusted with a number of policing tools within the site. They notably have the ability to block disruptive users, delete Wikipedia pages if they deem that they will not develop as proper encyclopedic articles and prevent articles from being edited by certain groups of contributors in order to prevent vandalism. It is quite telling in this respect to note that those administrative tools are often likened to a janitor's mop, leading to admins sometimes being described within Wikipedia as being "given the mop".[30] Overall, there were 1,388 contributors holding admin rights at the time of the experiment, 120 of whom took part in the experiment.

As "system operators", Wikipedia administrators have to deal with a very large number of potentially malicious users, which often involves making many quick decisions in order to prevent threats and fix "bugs" in the system. In this section, we therefore investigate the relationship between administrators' level of trust in anonymous strangers and their policing activity patterns. To do so, we rely on the experimental data on trust that we collected as a byproduct of our Trust game based measure of reciprocity. *We take the proportion of the endowment that trustors decided to send to trustees in the Trust game as an experimental measure of subjects' level of trust towards anonymous strangers.*

We test for the predictive power of this experimental measure of generalized trust on the policing activity patterns of Wikipedia administrators in table 13. As a first step to the analysis, we verify that trust levels are not associated with the contribution patterns of regular contributors (who are the focus of the analysis in sections 4 and 5) as we have no reason to expect this to be the case in theory. We can see from columns (1) to (3) that the coefficients on trust for regular contributor are statistically insignificant and close to zero. This is true irrespective of whether we consider the whole sample of regular subjects or check for potential heterogeneous effects by looking at the below median and above median groups separately.[31]

As hypothesized, the picture is different when we focus on the sample of Wikipedia administrators. In this group, moving from full trust to no trust in strangers is significantly associated with a 107% increase in overall Wikipedia activity (see column (4)). In order to confirm this result, we directly look at the paradigmatic administrative activities of our admin subjects – number of users blocked, number of pages deleted and number of pages protected from editing – and test for the predictive power of our experimental measure of trust in strangers on the extent to which they engage in those policing activities. As we can see from columns (5) to (7), moving from full trust to no trust in the experiment is associated with a 173% increase in the total number of users blocked from editing, a 107% increase in the number of pages deleted, and a 87% increase in the number of pages protected from editing, with the effect being statistically significant in 2 out of 3 cases.

As a final piece of evidence, we returned to our subjects 6 months after the completion of the experiment (i.e. in July 2012) and asked them to tell us about the fraction of their working time on Wikipedia that they typically spent on activities that administrators only can perform (e.g. deleting and protecting pages, blocking and unblocking users etc.), as opposed to regular contribution activities. We received an answer from 58 Wikipedia administrators out of 120 in the original sample.

[29] See https://en.wikipedia.org/wiki/Wikipedia:Requests_for_adminship for more details on the nomination process, accessed October 2014.

[30] See https://en.wikipedia.org/wiki/Wikipedia:Administrators, accessed October 2014.

[31] Experimental trust is also unrelated to the individual measures of cooperativeness that we develop in section 5. Since those coefficients do not add any meaningful information, we do not report them here.

Column (8) of table 9 presents an OLS estimate of the relationship between trust in anonymous strangers and the fraction of their working time on Wikipedia that those administrators reported dedicating to administrative activities. Despite the small sample size, moving from full trust to no trust in the Trust game is significantly associated with a 3.7 points increase in the proportion of time dedicated to admin activities. Out of a 10 points scale, this estimate corresponds to a 1.34 standard deviation increase.

[TABLE 7 ABOUT HERE]

While the evidence reported in table 11 establishes that less trusting administrators are significantly more active and exercise their policing rights more intensely within the site, our data does not allow us to conclude regarding the optimal level of trust that administrators should exhibit in order to efficiently fulfill their role. In their quantitative study of the factors that discourage many new Wikipedia contributors from sustaining their contributions in the long run, Halfaker et al. (2011) insist on the concept of "newbie biting", and conclude that "the more curmudgeonly old-timers should be kept away from newcomers until they have gained some experience in the system." To be sure, low trusting administrators could be considered as archetypal examples of such surly old-timers. This interpretation echoes a vibrant debate in the Wikipedia community about whether the increasingly harsh and bureaucratic enforcement of Wikipedia's policies and guidelines vis-à-vis newcomers could be one of the factors behind Wikipedia's current difficulties at attracting contributors willing to sustain high contribution levels (Halfaker et al. 2012).[32] However, it could also be that an efficient administrator *needs* to be relatively suspicious of anonymous strangers and see newcomers as potential threats in order to successfully protect the project from the many non established users who could be inclined to hurt it. This is an important policy question, which we leave open for future research.

7 Conclusion

In this paper, we rely on Wikipedia – an environment in which monetary incentives play a negligible role in shaping individual behavior – in order to extend the existing empirical literature on public goods provision beyond its current focus on explaining individual contributions. Consistent with the existing lab-based experimental literature and a number of previous field studies, we establish that the quantity of field contributions that our subjects make to Wikipedia is strongly related to their taste for reciprocal exchange, their social image concerns and their altruistic preferences (although only about 9% of contributors can be classified as altruists in our experiment).

We then expand the scope of our analysis in 3 novel directions which matter for public goods provision and/or governance. First, we test for the relevance of the theory for those contributors who exhibit the most extreme levels of involvement with the project (i.e. the so-called "super contributors"). We show that above and beyond all demographic characteristics, a revealed preference for social image is the only factor which motivates those contributors to further increase their participation. Beyond this result, the relevance of the theory is limited by the fact that those contributors do not statistically differ from casual ones in terms of their average prosociality levels, while those super contributors who are at the very top in term of contribution levels reveal relatively more selfish preferences in our experiment.

[32] See http://strategy.wikimedia.org/wiki/Editor_Trends_Study/Results, accessed October 2014.

Second, irrespective of the amount of content actually contributed, we analyze the determinants of individual collaborativeness levels. We find that reciprocal and altruistic subjects are relatively more cooperative while contributing, as measured by the proportion of their reverts which they leave without a justification and the number of editing conflicts that they start. By contrast, contributors who are motivated by social image are consistently found to start relatively more editing conflicts on average. This finding is consistent with the fact that, contrary to the letter of the Wikipedia policy rules, the community of engaged contributors seems to marginally reward confrontational behavior (as long as the reverts have some justification).

Third, we analyze the relationship between an experimental measure of trust in strangers and the exercise by Wikipedia administrators of their policing rights within the platform. We find that less trusting administrators exercise their policing rights significantly more by blocking other users from editing and deleting recent Wikipedia pages.

This paper suggests that the investigation of the foundations of cooperation in public goods environments deserves moving beyond the analysis of the determinants of the quantity contributed in order to understand *how* individuals contribute depending on the particular *social and institutional environment* in which they evolve (Ostrom 1990). Our results point at some interesting future research directions. First, it would be interesting to investigate how social signalers' behavior responds to changes in the behavioral norms and standards of the group in which they evolve. Second, it would be useful to explicitly tackle the issue of the optimal level of trust that platform administrators (or their equivalents) should exhibit in order to efficiently protect a common resource while maximizing participation. Finally, understanding the motivations of the contributors who exhibit the most extreme levels of engagement might require some theoretical adjustments in order to properly guide further empirical investigations (Von Krogh et al. 2012). The relevance of this question is reinforced by the fact that super contributors often are in the best position to undertake *de facto* leadership roles within communities of contributors.

We are, of course, only beginning to uncover the nature of the intrinsic motives that drive individuals to voluntarily sustain cooperation in the field. These motives are likely to be diverse. Much more field work needs to be done to see if the literature will be able to identify some general underlying motives that would be systematically associated with sustained patterns of contribution to real-world public goods, irrespective of the context in which such contributions take place. It could also be, however, that the motives that drive contributions highly depend on the nature of the public good considered and the context of the contribution, which could in turn explain some of the contradicting laboratory results in the literature (see Vesterlund 2012). Although the Internet is a rather specific field of study, we suggest that there is increasing scope for learning from an online approach coupling experimental economics with big data analysis (Hergueux and Jacquemet 2015).

References

Anderson, Simon P, Jacob K Goeree, and Charles A Holt. 1998. "A Theoretical Analysis of Altruism and Decision Error in Public Goods Games." *Journal of Public Economics* 70 (2): 297–323.

Andreoni, James. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy* 97 (6): 1447–58.

———. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal* 100 (401): 464–77.

Andreoni, James, and B. Douglas Bernheim. 2009. "Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica* 77 (5): 1607–36.

Barr, Abigail, and Pieter Serneels. 2009. "Reciprocity in the Workplace." *Experimental Economics* 12 (1): 99–112.

Bartling, Björn, Ernst Fehr, and Klaus M. Schmidt. 2012. "Screening, Competition, and Job Design: Economic Origins of Good Jobs." *The American Economic Review* 102 (2): 834–64.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *The American Economic Review* 96 (5): 1652–78.

Benkler, Yochai. 2002. "Coase's Penguin, Or, Linux and 'The Nature of the Firm.'" *The Yale Law Journal* 112 (3): 369–446.

———. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press.

———. 2013. "Peer Production and Cooperation." In *Handbook on the Economics of the Internet, J. M. Bauer & M. Latzer (eds.)*. Cheltenham and Northampton, Edward Elgar.

Benz, Matthias, and Stephan Meier. 2008. "Do People Behave in Experiments as in the Field?—evidence from Donations." *Experimental Economics* 11 (3): 268–81.

Carlsson, Fredrik, Olof Johansson-Stenman, and Pham Khanh Nam. 2014. "Social Preferences Are Stable over Long Periods of Time." *Journal of Public Economics* 117 (September): 104–14.

Carpenter, Jeffrey, and Caitlin Knowles Myers. 2010. "Why Volunteer? Evidence on the Role of Altruism, Image, and Incentives." *Journal of Public Economics* 94 (11–12): 911–20.

Carpenter, Jeffrey, and Erika Seki. 2011. "Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay." *Economic Inquiry* 49 (2): 612–30.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences." *Journal of the European Economic Association* 9 (3): 522–50.

Dufwenberg, Martin, and Georg Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47 (2): 268–98.

Ellingsen, Tore, and Magnus Johannesson. 2008. "Pride and Prejudice: The Human Side of Incentive Theory." *The American Economic Review* 98 (3): 990–1008.

———. 2011. "Conspicuous Generosity." *Journal of Public Economics* 95 (9–10): 1131–43.

Falk, Armin, and Urs Fischbacher. 2006. "A Theory of Reciprocity." *Games and Economic Behavior* 54 (2): 293–315.

Fehr, Ernst, and Andreas Leibbrandt. 2011. "A Field Study on Cooperativeness and Impatience in the Tragedy of the Commons." *Journal of Public Economics* 95 (9–10): 1144–55.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr. 2001. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." *Economics Letters* 71 (3): 397–404.

Giles, Jim. 2005. "Internet Encyclopaedias Go Head to Head." *Nature* 438 (7070): 900–901.

Gneezy, Uri, Andreas Leibbrandt, and John A. List. 2015. "Ode to the Sea: Workplace Organizations and Norms of Cooperation." *The Economic Journal* forthcoming.

Greenstein, Shane, and Feng Zhu. 2013. "Collective Intelligence and Neutral Point of View: The Case of Wikipedia." *National Bureau of Economic Research Working Paper*.

———. 2014. "Do Experts or Collective Intelligence Write with More Bias? Evidence from Encyclopædia Britannica and Wikipedia." *HBS Working Paper*.

Halfaker, Aaron, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl. 2012. "The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline." *American Behavioral Scientist*.

Halfaker, Aaron, Aniket Kittur, and John Riedl. 2011. "Don'T Bite the Newbies: How Reverts Affect the Quantity and Quality of Wikipedia Work." In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 163–72. WikiSym '11. New York, NY, USA: ACM.

Hergueux, Jérôme, and Nicolas Jacquemet. 2015. "Social Preferences in the Online Laboratory: A Randomized Experiment." *Experimental Economics* 18 (2): 251–83.

Holländer, Heinz. 1990. "A Social Exchange Approach to Voluntary Cooperation." *The American Economic Review* 80 (5): 1157–67.

Karlan, Dean S. 2005. "Using Experimental Economics to Measure Social Capital and Predict Financial Decisions." *The American Economic Review* 95 (5): 1688–99.

Kittur, Aniket, Bryan A. Pendleton, Bongwon Suh, and Todd Mytkowicz. 2007. "Power of the Few vs. Wisdom of the Crowd: Wikipedia and the Rise of the Bourgeoisie." *Proceeding of the CIGCHI Conference on Human Factors in Computing Systems*.

Laury, Susan K., and Laura O. Taylor. 2008. "Altruism Spillovers: Are Behaviors in Context-Free Experiments Predictive of Altruism toward a Naturally Occurring Public Good?" *Journal of Economic Behavior & Organization* 65 (1): 9–29.

Leibbrandt, Andreas. 2012. "Are Social Preferences Related to Market Performance?" *Experimental Economics* 15 (4): 589–603.

Oliveira, Angela De, Rachel T. A. Croson, and Catherine C. Eckel. 2009. "Are Preferences Stable Across Domains? An Experimental Investigation of Social Preferences in the Field." *Working Paper*, January.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.
Panciera, Katherine, Aaron Halfaker, and Loren Terveen. 2009. "Wikipedians Are Born, Not Made: A Study of Power Editors on Wikipedia." In *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, 51–60. GROUP '09. New York, NY, USA: ACM.
Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *The American Economic Review* 83 (5): 1281–1302.
Reagle, Joseph. 2010. *Good Faith Collaboration: The Culture of Wikipedia*. MIT Press.
Reagle, Joseph. 2012. "'Free as in Sexist?' Free Culture and the Gender Gap." *First Monday* 18 (1).
Silva, J. M. C. Santos, and Silvana Tenreyro. 2006. "The Log of Gravity." *Review of Economics and Statistics* 88 (4): 641–58.
Vesterlund, Lise. 2012. "Voluntary Giving to Public Goods: Moving Beyond the Linear VCM." *Working Paper*.
Von Krogh, Georg, Stefan Haefliger, Sebastian Spaeth, and Martin W. Wallin. 2012. "Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development." *MIS Q.* 36 (2): 649–76.
Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.
Yasseri, Taha, Anselm Spoerri, Mark Graham, and Janos Kertesz. 2013. "The Most Controversial Topics in Wikipedia: A Multilingual and Geographical Analysis." SSRN Scholarly Paper ID 2269392.
Zhang, Xiaoquan, and Feng Zhu. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia." *The American Economic Review* 101 (4): 1601–15.

Tables and figures

Figure 1. The instruction screen of the Public Goods game

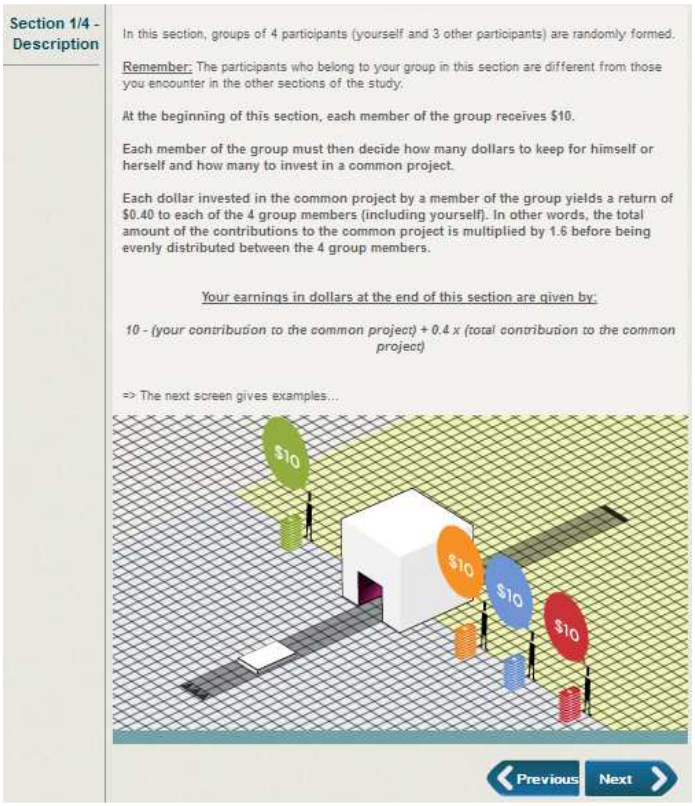


Figure 2. The decision screen of the conditional Public Goods game

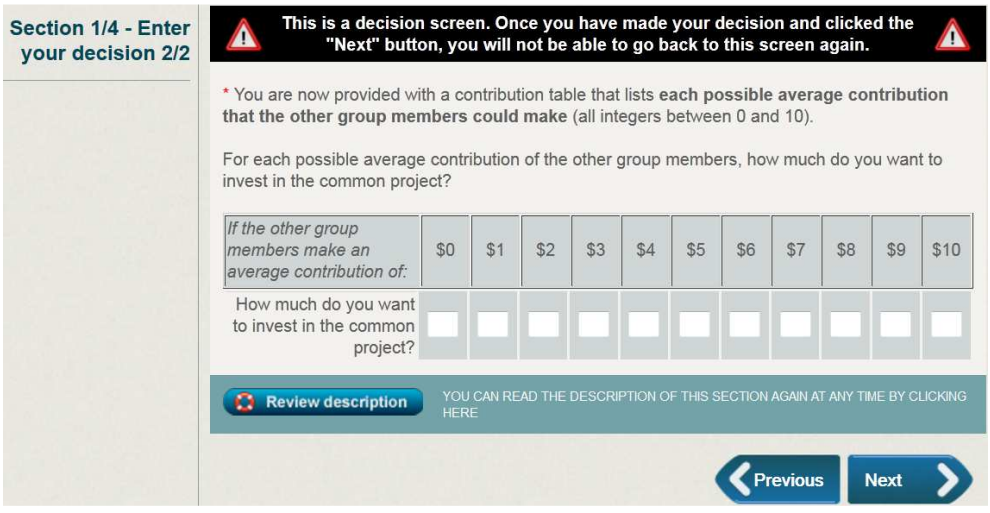


Figure 3. A typical Barnstar

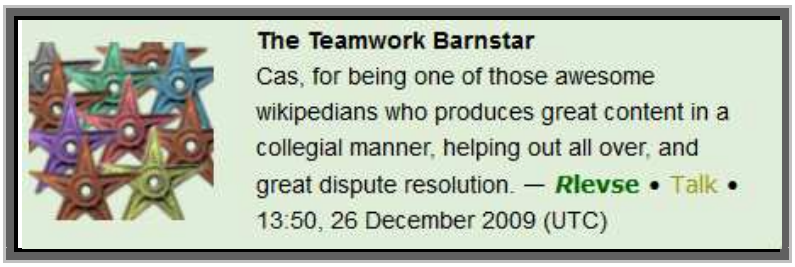
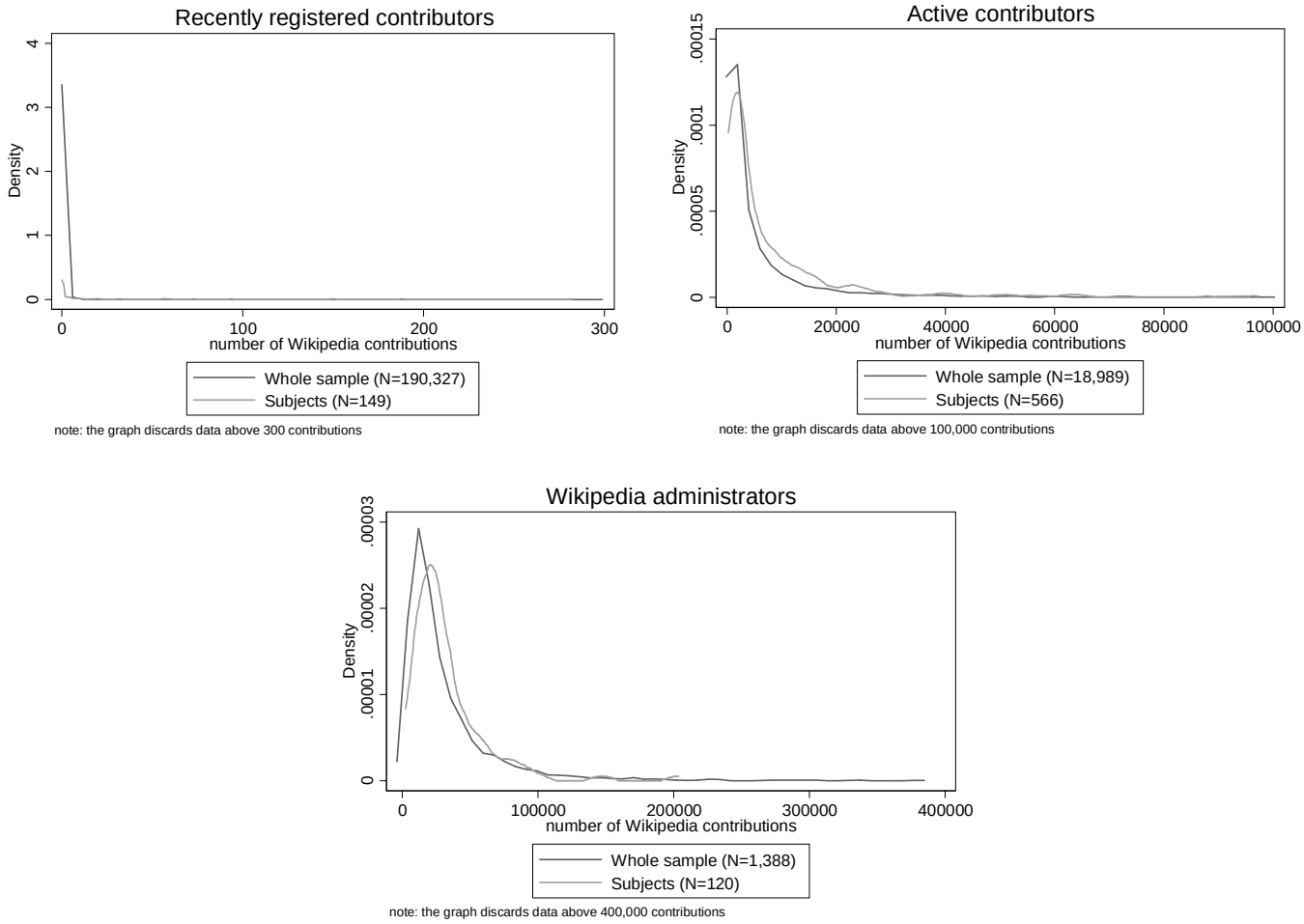


Figure 4. The Wikipedia recruitment banner

Figure 5. Distribution of the number of Wikipedia contributions:
study participants vs. eligible population



Notes: Kernel density estimates.

Figure 6. Distribution of cooperative types in the conditional Public Good game

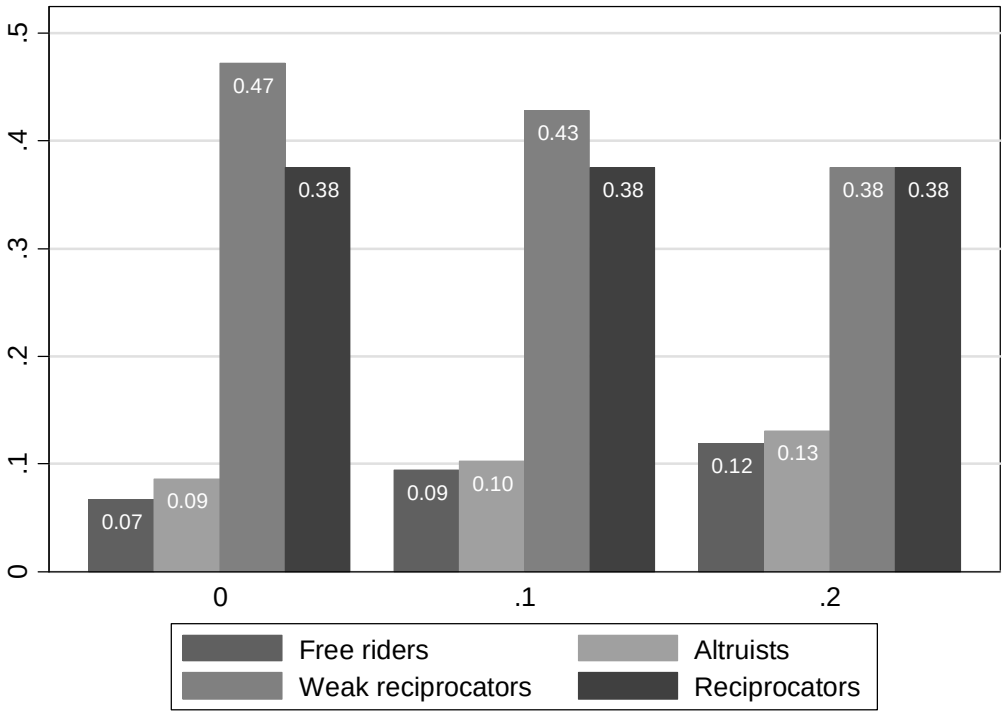


Table 1. Sample common demographic characteristics: Wikimedia editor survey vs. our study

	2011 Wikimedia editor survey	Our study
Age		
12 to 17	13%	4%
18 to 21	14%	17%
22 to 29	26%	30%
30 to 39	19%	20%
40 or more	28%	28%
Gender		
Proportion female	9%	10%
Education level		
Primary	9%	5%
Secondary	30%	31%
Bachelors / associate	35%	34%
Master's	18%	22%
PhD	8%	7%

Notes: The Wikimedia editor survey excludes respondents under 12 and over 82 from the sample. The age and gender statistics are based on the population of respondents with a positive number of Wikipedia contributions (N=4,930). The Education level statistics are based on the whole population of respondents (N=5,073). In this table, we base our own statistics on the same calculation rules.

Table 2. Descriptive statistics

	All subjects	Casual contributors	Super contributors	Administrators
Number of observations (*N*)	730	365	365	120
DEPENDENT VARIABLES				
Quantity of Wikipedia contributions				
• Mean – number of Wikipedia contributions	7549.12	545.29	14552.95	41229.24
	(21096.5)	(595.80)	(28153.43)	(86191.33)
	[0; 364157]	[0; 1903]	[1907; 364157]	[2475; 922895]
Individual collaborativeness				
• Mean – proportion of reverts w/o an explanation	0.04	0.06	0.03	0.00
	(0.11)	(0.15)	(0.07)	(0.01)
	[0; 1]	[0; 1]	[0; 1]	[0; 0.33]
• Mean – number of edit wars (3 consecutive reverts chain)	4.64	0.38	7.13	11.62
	(16.69)	(1.55)	(20.58)	(17.15)
	[0; 291]	[0; 12]	[0; 291]	[0; 94]
SOCIAL PREFERENCES				
Public goods game				
• Mean – cooperation (prop of endowment unconditionally contributed)	0.64	0.62	0.67	0.62
	(0.34)	(0.34)	(0.34)	(0.33)
Trust game				
• Mean – trust (prop of endowment sent)	0.66	0.65	0.67	0.68
	(0.35)	(0.35)	(0.34)	(0.29)
Social image data				
• Mean size of Wikipedia user page (in bytes)	7153.10	3304.35	9989.58	10055.40
	(11508.59)	(6295.71)	(13495.58)	(8887.38)
Number of Barnstars receivers	347	67	280	109
Mean – number of Barnstars received	6.08	2.13	7.03	16.8
	(8.51)	(2.31)	(9.16)	(16.00)
• Proportion signaling Barnstars	0.50	0.30	0.54	0.70
	(0.50)	(0.46)	(0.50)	(0.46)
DEMOGRAPHIC VARIABLES				
Age	33	30	36	34
	(14.36)	(12.84)	(15.15)	(12.86)
Proportion female	0.10	0.11	0.10	0.11
	(0.31)	(0.31)	(0.30)	(0.31)
Degree level	4.43	4.28	4.58	4.88
	(1.84)	(3.42)	(1.75)	(1.64)
Salary level	3.68	3.42	3.93	4.01
	(2.31)	(2.29)	(2.30)	(2.25)
Risk aversion level	5.74	5.78	5.70	5.53
	(2.35)	(2.40)	(2.30)	(2.38)

Notes: Standard errors are reported in parenthesis, minimum and maximum values are reported in brackets for the dependent variables. *Degree level*: 1 = "less than high school"; 2 = "high school"; 3 = "some college"; 4 = "2 years college degree"; 5 = "4 years college degree (BA, BS)"; 6 = "masters degree"; 7 = "professional degree (MD, JD)"; 8 = "doctoral degree". *Salary level* (monthly): 1 = "0 USD"; 2 = "less than 1000 USD"; 3 = "between 1000 and 2000 USD"; 4 = "between 2000 and 3000 USD"; 5 = "between 3000 and 4000 USD"; 6 = "between 4000 and 5000 USD"; 7 = "between 5000 and 7500 USD"; 8 = "between 7500 and 10000 USD"; 9 = "more than 10000 USD". *Risk aversion level* = whether subjects generally see themselves as fully prepared to take risks as opposed to generally trying to avoid taking risks: 0 = "unwilling to take risks" to 10 = "fully prepared to take risks".

Table 3. Contributions to Wikipedia and demographic characteristics

	(1)	(2)	(3)	(4)
Dependent variable is number of Wikipedia contributions	All subjects	Casual contributors	Super contributors	Admins
age	0.0167***	0.0136***	0.00516	0.0160***
	(0.00306)	(0.00518)	(0.00338)	(0.00541)
female	-0.365**	-0.665***	-0.0332	-0.244
	(0.147)	(0.243)	(0.152)	(0.249)
degree level	0.0582**	0.0971***	0.0165	-0.0391
	(0.0246)	(0.0362)	(0.0309)	(0.0446)
salary level	0.00282	-0.00846	-0.0133	-0.0509*
	(0.0200)	(0.0313)	(0.0225)	(0.0305)
Risk aversion	-0.0325*	-0.0640**	-0.00170	0.00533
	(0.0169)	(0.0259)	(0.0193)	(0.0279)
Constant	8.310***	5.901***	9.420***	10.46***
	(0.175)	(0.266)	(0.189)	(0.291)
N	649	325	324	113
Estimates	Neg. bin.	Neg. bin.	Neg. bin.	Neg. bin.

Notes: Standard errors are reported in parenthesis. *, ** and *** denote statistical significance at the 10, 5 and 1% levels.

Table 4. Social motives and quantity of contributions

	(1)	(2)	(3)	(4)
Dependent variable: number of Wikipedia contributions	All subjects	Casual contributors	Super contributors	All subjects
PANEL A : free riders satisfy m ≤ 0.2				
Weak reciprocator	-0.00169	0.492**	-0.200	-0.301**
	(0.138)	(0.216)	(0.152)	(0.153)
Reciprocator	0.182	0.622***	-0.201	-0.307**
	(0.135)	(0.216)	(0.146)	(0.149)
Altruist	0.143	0.677***	-0.172	-0.192
	(0.157)	(0.243)	(0.176)	(0.169)
Social signaler (user page)	1.295***	1.772***	0.287***	
	(0.0847)	(0.126)	(0.103)	
Social signaler (Barnstars)				0.289***
				(0.0975)
nb Barnstars				0.0412***
				(0.00294)
Test : Social signaler = Altruist	*p*=0.000	*p*=0.000	*p*=0.035	*p*=0.009
Test : Social signaler = Reciprocator	*p*=0.000	*p*=0.000	*p*=0.011	*p*=0.001
Test : Altruist = Reciprocator	*p*=0.737	*p*=0.745	*p*=0.833	*p*=0.369
PANEL B : free riders satisfy m ≤ 0.1				
Weak reciprocator	0.0130	0.630**	-0.405***	-0.379**
	(0.152)	(0.248)	(0.156)	(0.163)
Reciprocator	0.194	0.768***	-0.385**	-0.402**
	(0.151)	(0.249)	(0.155)	(0.160)
Altruist	0.193	0.922***	-0.355*	-0.332*
	(0.179)	(0.282)	(0.191)	(0.185)
Social signaler (user page)	1.294***	1.766***	0.289***	
	(0.0848)	(0.125)	(0.102)	
Social signaler (Barnstars)				0.265***
				(0.0980)
nb Barnstars				0.0413***
				(0.00295)
Test : Social signaler = Altruist	*p*=0.000	*p*=0.008	*p*=0.004	*p*=0.002
Test : Social signaler = Reciprocator	*p*=0.000	*p*=0.001	*p*=0.001	*p*=0.000
Test : Altruist = Reciprocator	*p*=0.995	*p*=0.407	*p*=0.838	*p*=0.607
Control variables	YES	YES	YES	YES
N	649	325	324	308
Estimates	Neg. Bin.	Neg. Bin.	Neg. Bin.	Neg. Bin.

Notes: Standard errors are reported in parenthesis. Constants not reported. *, ** and *** denote statistical significance at the 10, 5 and 1% levels.

Table 5. Collaborativeness and demographic characteristics

	(1)	(2)
Dependent variable:	Prop reverts w/o explanation	Start edit wars
age	-0.00111***	0.0107**
	(0.000421)	(0.00503)
female	-0.00945	-0.271
	(0.0116)	(0.267)
degree level	0.00163	0.0194
	(0.00305)	(0.0450)
salary level	0.00212	0.00167
	(0.00336)	(0.0329)
Risk aversion	-0.00126	0.0426
	(0.00186)	(0.0281)
ln(edited articles conflictuality score)	-0.00564**	0.205***
	(0.00286)	(0.0325)
Constant	0.0759***	0.531*
	(0.0177)	(0.282)
N	499	516
Estimates	OLS	Neg. Bin.

Notes: Standard errors are reported in parenthesis (robust standard errors are computed for the OLS estimates). *, ** and *** denote statistical significance at the 10, 5 and 1% levels.

Table 6. Social motives and collaborativeness

	(1)	(2)	(3)	(4)
Dependent variable:	Prop reverts w/o explanation	Prop reverts w/o explanation	Start edit wars	Start edit wars
PANEL A : free riders satisfy m ≤ 0.2				
Weak reciprocator	-0.0353	-0.00479	-0.493**	-0.380
	(0.0228)	(0.0219)	(0.227)	(0.250)
Reciprocator	-0.0571***	-0.0172	-0.254	-0.234
	(0.0208)	(0.0204)	(0.217)	(0.235)
Altruist	-0.0497**	-0.00928	-0.448*	-0.115
	(0.0219)	(0.0230)	(0.258)	(0.270)
Social signaler (user page)	-0.0112		0.658***	
	(0.0118)		(0.151)	
Social signaler (Barnstars)		-0.0224*		0.333**
		(0.0119)		(0.161)
nb Barnstars		-0.000744***		0.0345***
		(0.000270)		(0.00535)
ln(edited articles conflictuality score)	-0.00589**	-0.00738**	0.214***	0.201***
	(0.00291)	(0.00358)	(0.0334)	(0.0375)
Test : Social signaler = Altruist	*p*=0.129	*p*=0.978	*p*=0.000	*p*=0.126
Test : Social signaler = Reciprocator	*p*=0.052	*p*=0.915	*p*=0.001	*p*=0.040
Test : Altruist = Reciprocator	*p*=0.479	*p*=0.884	*p*=0.331	*p*=0.571
PANEL B : free riders satisfy m ≤ 0.1				
Weak reciprocator (Public Goods)	-0.00971	-0.00325	-0.673***	-0.495*
	(0.0243)	(0.0255)	(0.244)	(0.266)
Reciprocator (Public Goods)	-0.0373	-0.0172	-0.419*	-0.355
	(0.0233)	(0.0248)	(0.239)	(0.254)
Altruist (Public Goods)	-0.0316	-0.0140	-0.543*	-0.229
	(0.0249)	(0.0277)	(0.286)	(0.296)
Social signaler (user page)	-0.0123		0.641***	
	(0.0117)		(0.150)	
Social signaler (Barnstars)		-0.0228*		0.294*
		(0.0117)		(0.162)
nb Barnstars		-0.000738***		0.0344***
		(0.000277)		(0.00537)
ln(edited articles conflictuality score)	-0.00530*	-0.00722**	0.211***	0.196***
	(0.00289)	(0.00362)	(0.0335)	(0.0376)
Test : Social signaler = Altruist	*p*=0.490	*p*=0.751	*p*=0.000	*p*=0.086
Test : Social signaler = Reciprocator	*p*=0.333	*p*=0.831	*p*=0.000	*p*=0.019
Test : Altruist = Reciprocator	*p*=0.610	*p*=0.850	*p*=0.562	*p*=0.568
Control variables	YES	YES	YES	YES
N	499	302	516	305
Estimates	OLS	OLS	Neg. Bin.	Neg. Bin.

Notes: Standard errors are reported in parenthesis (robust standard errors are computed for the OLS estimates). Constants not reported. *, ** and *** denote statistical significance at the 10, 5 and 1% levels.

Table 7. Trusting behavior and policing activities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable :	number of Wikipedia contributions	number of Wikipedia contributions	number of Wikipedia contributions	number of Wikipedia contributions	number of users blocked	number of pages deleted	number of pages protected	Time spent on admin activities
	All subjects	Casual contributors	Super contributors	Admins	Admins	Admins	Admins	Admins
Trust	0.0780	-0.0265	-0.0393	-0.730**	-1.004**	-0.725*	-0.626	-3.703*
	(0.180)	(0.272)	(0.187)	(0.309)	(0.463)	(0.419)	(0.460)	(1.818)
Control variables	YES	YES	YES	YES	YES	YES	YES	YES
N	305	159	146	56	56	56	56	27
Estimates	Neg. bin.	Neg. bin.	Neg. bin.	Neg. bin.	Neg. bin.	Neg. bin.	Neg. bin.	OLS

Notes: Standard errors are reported in parenthesis (robust standard errors are computed for the OLS estimates). Constants not reported. *, ** and *** denote statistical significance at the 10, 5 and 1% levels. *Time spent on admin activities* = answer to the question: "what fraction of the time that you spend working on Wikipedia do you specifically devote to activities that admins only can perform (e.g. deleting and protecting pages, blocking and unblocking users etc.) as opposed to the regular editing activities mentioned above? Please choose one number on the following scale, where 0 means "I do not spend any of my working time on Wikipedia performing admin-related tasks" and 10 means "I spend all of my working time on Wikipedia performing admin-related tasks"."

Appendix: conflictuality score of the set of pages edited by each contributor

We build upon Yasseri et al.'s approach (2013) to compute the conflictuality score of all the articles edited by our subjects. As a first approximation for conflictuality, we identify the number of "mutual reverts" on each article. A mutual revert *m* is identified on an article when a pair of contributors *(j; k)* is identified once with *j* and once with *k* as the reverter. Since a mutual revert is more indicative of a strong conflictuality potential when it involves two experienced contributors than two inexperienced ones (or even one experienced and one inexperienced), we weight each mutual revert by the minimum of the overall number of edits N_j and N_k that both contributors made. Then, for each article, we sum-up the weights of all mutual reverts, excluding the highest weight (in order to avoid overestimating conflictuality by giving a lot of prominence to a single editorial conflict) and multiplying by the overall number of contributors *E* involved in the article. The conflictuality score C_a of an article is therefore computed as:

$$C_a = E \,.\, \sum_{m=1}^{n} \min\left[N_j;\ N_k\right] \tag{2}$$

The above measure of conflictuality is at the Wikipedia article level. In order to get a conflictuality score at the contributor level, we compute the average of the conflictuality scores over all the articles *a* that contributor C_i edited, weighted by the proportion of the overall number of contributions that the made to each article:

$$C_i = \sum_{a=1}^{n} p_a \,.\, C_a \tag{3}$$

The conflictuality control variable that we include in all the collaborativeness regressions of Tables 8 to 10 is the logarithm of the above measure of conflictuality at the contributor level: ln(edited articles conflictuality score) = ln(C_i). It has mean = 1.46; std = 1.25; min = 0 and max = 10.42. Experimenting with other ways of computing this variable such as taking the absolute value of C_i or the proportion of the articles edited by contributor C_i which have a positive conflictuality score C_a leaves our result unchanged.