

Gallus, Jana

Working Paper

Fostering Voluntary Contributions to a Public Good A Large-Scale Natural Field Experiment at Wikipedia

CREMA Working Paper, No. 2015-05

Provided in Cooperation with:

CREMA - Center for Research in Economics, Management and the Arts, Zürich

Suggested Citation: Gallus, Jana (2015) : Fostering Voluntary Contributions to a Public Good A Large-Scale Natural Field Experiment at Wikipedia, CREMA Working Paper, No. 2015-05, Center for Research in Economics, Management and the Arts (CREMA), Zürich

This Version is available at:

<https://hdl.handle.net/10419/214570>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Center for Research in Economics, Management and the Arts

Fostering Voluntary Contributions
to a Public Good
A Large-Scale Natural Field Experiment
at Wikipedia

Working Paper No. 2015-05

Fostering Voluntary Contributions to a Public Good

A Large-Scale Natural Field Experiment at Wikipedia

By JANA GALLUS*

This natural field experiment tests the effects of purely symbolic awards on volunteer retention in a public goods context. The experiment is conducted at Wikipedia, which faces declining editor retention rates. Randomization assures that receipt of the award is orthogonal to previous performance. The analysis reveals that awards have a sizeable and statistically significant effect on retention. The findings are noteworthy firstly for showing that symbolic awards with no career-related implications can positively impact behavior. Secondly, they indicate that the awards' motivating effect goes beyond serving reputational concerns and even positively affects recipients' private identity. JEL codes: C93, M52, H41.

* Dr. Jana Gallus, Department of Economics, University of Zurich. Hottingerstrasse 10, 8032 Zurich, Switzerland.
Phone: 0041 44 634 56 59. jana.gallus@econ.uzh.ch.

This is a revision of an earlier paper with the same title, dating from 15 October 2013, circulated as a mimeo, CREMA – Center for Research in Economics, Management, and the Arts, Switzerland.

Acknowledgments: I am most grateful to Bruno S. Frey for his thoughtful guidance and advice, and to Gregor Martynus for outstanding programming support. I also wish to thank Christoph Engel, Kai Konrad, and Andrew Oswald for their most helpful comments. The author acknowledges financial support from the Schweizerische Gemeinnützige Gesellschaft (SGG), and from the Center for Research in Economics, Management and the Arts (CREMA), Switzerland.

Wikipedia is an undisputed success story. It is the largest online encyclopedia, its content is provided by voluntary contributors, and its quality is similar to that of the *Encyclopedia Britannica* (Giles 2005). As of February 2014, more than 44 million registered editors have voluntarily contributed to this public good.¹ Wikipedia is among the ten most frequently visited websites worldwide² and comprises more than 30 million encyclopedic articles¹ in over 280 different languages³ (see also Greenstein and Zhu 2012).

Yet, from the point of view of standard economic theory committed to the self-interested *homo oeconomicus*, such contributions should not have been observed because they are not remunerated financially. They do not even entail reputational benefits that might translate into higher future earnings. The editors of Wikipedia are anonymous⁴ since they most often use pseudonyms. Moreover, articles are the result of common work, making it difficult, if not impossible, to trace the author of a specific part. Readers can therefore not see who wrote an article or helped to improve it. Image motivation, or non-virtual prestige benefits (Harbaugh 1998), and potential material advantages can thus be ruled out as drivers of this form of voluntary behavior. As a consequence, standard economics would predict that rational individuals do not contribute to Wikipedia.

The conflict between theoretical prediction and empirical observation has been identified in the private provision of public goods more generally. The puzzle that the existence of such public goods⁵ poses to theory has intrigued economists for a long time (see Olson 1965, Ledyard 1995, Cornes and Sandler 1996, Zelmer 2003, Chaudhuri 2011); many analyses have been

¹ https://meta.wikimedia.org/wiki/List_of_Wikipedias#Grand_Total (accessed Feb 6, 2014).

² <http://www.alexa.com/siteinfo/wikipedia.org?range=5y&size=large&y=t> (Feb 6, 2014).

³ https://meta.wikimedia.org/wiki/List_of_Wikipedias (Aug 14, 2014).

⁴ See, for instance, Soetevent (2005) on the role of anonymity in giving.

⁵ Other well-known examples of online public goods are summarized in Zhang and Zhu (2011, 1601), including products emanating from open source software development (e.g. Mozilla Firefox, Linux) and content sharing platforms (e.g. YouTube).

devoted to the motivations underlying such pro-social behavior (e.g., Fehr and Schmidt 2002, 2006, Cooper and Kagel 2013).⁶

This paper focuses on a different but related question, namely how to foster and uphold the behavior already undertaken by individuals. Voluntary commitments have become increasingly fleeting because fewer and fewer volunteers are willing or able to commit to an organization for an extended period of time (Macduff 2005). Nonprofit organizations, however, are dependent on sustained involvement of volunteers (see, e.g., Snyder and Omoto 2008). Even a project such as Wikipedia, which has seen millions of people contribute, is severely threatened by declining retention rates among its newer members (Wikimedia 2011a).

Voluntary organizations are confronted with the challenge of fostering the behavior of their members without, however, crowding out their intrinsic motivation with the rewards they choose to employ (see Gneezy, Meier, and Rey-Biel 2011 for a discussion of when incentives do and do not work to motivate pro-social behavior). These organizations have to strike a delicate balance between the nature and salience of the rewards, and the self-determinedness and self-perception of volunteers (Frey and Goette 1999, Falk and Kosfeld 2006, Tirole and Bénabou 2006, Ariely, Bracha, and Meier 2009). Voluntary organizations can neither recur to monetary incentives as for-profit firms do to incentivize their employees, nor can they bind their members with contractual agreements. Non-monetary alternatives have to be employed to motivate volunteers to stay.

The existing literature is largely silent on the effects and potential of non-monetary rewards, in particular awards, for pro-social behavior. However, it can be observed that throughout history, the provision of honor to volunteers, for instance with orders of merit, has played an important role in most—if not

⁶ Studies focusing on motivations to contribute to online public goods (many of which focus on open source software development) are, for example, Lerner and Tirole (2002), Cifforilli (2003), Lakhani and von Hippel (2003), Lakhani and Wolf (2005), Nov (2007), and Schroer and Hertel (2009).

all—societies (Frey and Neckermann 2009). Awards provide the possibility to "repay" volunteers in a non-monetary currency that may cater to their initial motivations (e.g., honor) while reducing the risk of crowding-out. Whereas states have retained the monopoly over official orders (e.g., the *President's Volunteer Service Award* in the US), non-state entities have created other forms of awards (e.g., journals bestowing service awards to a select group of Associate Editors and reviewers to recognize their voluntary contributions). There are nowadays plenty of honors given for pro-social behavior, ranging from the prestigious *Florence Nightingale Medal* of the International Committee of the Red Cross to well-respected awards bestowed by private organizations, such as the *Eagle Palms* awarded by Boy Scouts of America.

The analysis of the potential of awards in furthering the provision of public goods is confronted with a major obstacle. Since awards are by definition always given for extraordinary behavior, causality cannot be established. Omitted variables and endogeneity bias cannot be ruled out. As it is "the best" who are awarded, it is not surprising that award recipients also perform better than other persons in the future. Such superior performance may simply reflect the fact that award winners are more able and more motivated than non-winners. Observing superior performance of award recipients thus does not establish whether the award conferral as such raises subsequent performance.

Various methods can help identify an award's causal effects. One is to compare the performance of award winners with that of a closely comparable group, in particular, with the set of award nominees (Redelmeier and Singh 2001a,b, Ginsburgh 2003). It is also possible to construct a comparison group using the synthetic control method (Chan et al. 2014), or to perform a difference-in-differences analysis where awards are given for a task uncorrelated with the outcome measure of performance (Neckermann, Cueni, and Frey 2014).

An ideal identification strategy uses randomization in a natural field environment (e.g., Gneezy and List 2006 and Kube, Maréchal, and Puppe

2012 testing gift-exchange in the field, and Chen et al. 2010 analyzing the effect of personalized social information in an online community). This method has been approximated by an experiment where students were hired for a one-time data entry job of two hours and subjects in the award treatment were promised a symbolic congratulatory card in addition to the fixed wage (Kosfeld and Neckermann 2011). Notwithstanding these recent advances in the literature on awards,⁷ however, various limitations remain. In some cases the identification of a causal relationship remains uncertain (internal validity); in other cases the setting is artificial and focuses on a short time window, so that external validity and especially the extension to pro-social behavior are doubtful. True randomization is difficult to institute in the field because award-giving institutions are reluctant to relinquish control of the selection process.

To circumvent this problem, a new award scheme can be implemented, with a committee of senior practitioners to establish the award's reputation. The latter approach was chosen for this experiment. Each month, a fixed number of individuals are randomly allocated into the treatment group. An award is posted on their personal page and their name is announced on the official award page, embedded in a national Wikipedia portal. The experimental design based on a random treatment allocation allows for a straightforward identification of causality by basic mean-comparison tests.

To the best of my knowledge, this paper presents the first natural field experiment using randomization to identify the causal effects of awards on voluntary contributions to a public good. In particular, I examine two questions. The first is whether awards that are purely symbolic and of no present or future material value can motivate new contributors to uphold their engagement, thus increasing newcomer retention. The second question to be analyzed is whether awards have a motivational effect that extends beyond the

⁷ Other economic analyses of awards include the early work by Hansen and Weisbrod (1972), as well as Frey (2005), Malmendier and Tate (2009), and Siming (2012).

enhanced social reputation (public image) they usually entail. Since only the award recipients know that they are the individuals behind the pseudonyms that were awarded, the experimental context makes it possible to exclude subjects' public identity (e.g., social esteem, real-life reputation and status). I exploit this setup to identify whether awards also have an impact on individuals' intrinsic, private identity.

The experiment is conducted at the German language edition of Wikipedia, which is the second-oldest one after the English Wikipedia and ranks among the largest Wikipedia language versions⁸ in terms of article numbers (more than 1.5 million as of May 2013⁹), contributors (more than 1.6 million registered accounts¹⁰) and usage (over 1.2 million views per hour¹⁰). The experiment spans more than eight months and comprises observations on 3,066 individuals. With its focus on newcomer retention, the study addresses one of the most important challenges Wikipedia faces; for although its readership has been increasing steadily (Greenstein and Zhu 2012), Wikipedia is struggling with declining retention rates among its contributors, especially among new ones (Suh et al. 2009, Wikimedia 2011a, Halfaker et al. 2013).

The field experiment provides empirical evidence that awards do indeed increase retention among newcomers by motivating them to uphold their contributions to a public good such as Wikipedia. Activity is considered in three fundamental dimensions: General activity, direct work on articles, and behind-the-scenes coordination work. In all three cases, the positive effect of the award is not only statistically significant; it is also sizeable and of practical importance given the retention problem Wikipedia and other volunteer organizations face.

The two key findings derived from the analysis are, first, that awards positively impact individual behavior even if they cannot entail any material

⁸ A comparison of the different language versions can be viewed under https://meta.wikimedia.org/wiki/List_of_Wikipedias (accessed May 16, 2013).

⁹ <https://de.wikipedia.org/w/index.php?title=Spezial:Statistik&action=raw> (May 16, 2013).

¹⁰ <http://stats.wikimedia.org/DE/> (May 2013).

or career-related benefits. The purely symbolic award increases the share of editors remaining active in the following month by 25 percent ($p = 0.000$). It also raises the share of authors who continue to directly contribute content to articles by 19 percent ($p = 0.002$). In contrast to previous results on gift exchange involving money (Gneezy and List 2006), the symbolic award's effects persist even over longer time periods. Second, not even the enhancement of recipients' public reputation and status (public identity) is required for awards to have a motivating effect. The experimental setting shields individuals' public identity because they operate under pseudonyms which they have only recently adopted. And yet, subjects in the treatment group are more likely to remain active and contribute to articles. Moreover, the results on behind-the-scenes coordination work suggest that their self-identification as community members is strengthened. The award raises the fraction of editors who interact with others on community pages by 62 percent ($p = 0.000$). This suggests that, rather than remaining at the extrinsic level of reputational concerns, awards have an effect that goes deeper and reaches the recipients' intrinsic, private identity. The importance of awards thus goes above and beyond providing status and attention. Intrinsically motivating factors such as identifying oneself with the community, receiving attention and experiencing one's competency seem to be of great importance.

The paper proceeds as follows. Section I describes the experimental setup with background information on Wikipedia, the experimental design and the outcome variables of interest. Section II presents the theoretical background, related literature and predictions to be tested. Section III reports the results, including those from randomization and robustness checks. Section IV concludes.

I. Experimental Setup

A. Background on Wikipedia

Over the twelve years since its launch in 2001, Wikipedia has become an established online information repository.¹¹ The English- and German-language versions of Wikipedia draw more than 10.7 million and 1.2 million views per hour, respectively.¹² Articles are only one part of the project, however. Wikipedia is composed of several areas of activity (called “namespaces”) that categorize the different contributions editors can make; grouping for instance all article edits, all file uploads or all contributions to Wikipedia-related meta issues (such as policy design). Every single page has its discussion room (“talk” page) on which editors (called “users”) can exchange comments. Contributors also have a personal “user page” where they can introduce themselves to the community, as well as a corresponding talk page that serves as a platform for communicating with others. Contributors’ activities can thus take many different forms.

To improve quality and fight vandalism, a multitude of standards and rules have gradually been established (see, e.g., Viégas, Wattenberg, and McKeon 2007, Butler, Joyce, and Pike 2008). New editors wanting to contribute to Wikipedia therefore have to comply with a regulatory framework increasingly difficult to penetrate. Non-compliance with established standards often leads to quick deletion of one’s work, be it by other editors or even by automated tools (Morgan et al. 2013). Today, newcomers are more likely to be greeted with a warning rather than a welcome message (Pinchuk 2011, Halfaker 2012), harsh criticism being no exception (Kraut et al. 2012). While reverts (deletions) help preserve quality, they substantially decrease newcomers’

¹¹ Collaborations between Wikipedia and prestigious research and state entities are no exception (Butler 2008). Already in 2008, *The Economist* argued: “IT IS the biggest encyclopedia in history and the most successful example of “user-generated content” on the internet, with over 9m articles in 250 languages contributed by volunteers collaborating online.” Available at <http://www.economist.com/node/10789354> (accessed June 16, 2013).

¹² <http://stats.wikimedia.org/EN/Sitemap.htm> (accessed July 24, 2013).

willingness to further contribute to the project (Halfaker, Kittur, and Riedl 2011). Given a diminishing number of topics that do not yet have a Wikipedia entry, new editors often have to fit their contribution into an existing article and defend it in front of more tenured editors keeping watch over their field of expertise (Suh et al. 2009, Wikimedia Deutschland 2011, Jensen 2012).

In the German-language Wikipedia, newcomers' contributions even have to be screened and confirmed by established editors before being publicly visible. As a result, it can take days until new editors see the product of their work appear on Wikipedia.

B. The Experimental Design

"[W]orking closely with practitioners" is listed as one of the three central advantages of field experiments by List and Rasul (2010, 105). The present study lends support to this assertion and moreover emphasizes that practitioners' endorsement is most likely to be vital for any such endeavor. The backing and trust of several highly reputable community members stands at the heart of this experiment. These contacts were established via telephone calls, which were followed up by regular roundtable meetings with a group of editors willing to tackle the retention problem with the help of the experiment. They became official founding members of the project, which was thus established under the umbrella of the Swiss national portal,¹³ providing the awards with considerable repute and an official character. Given the high involvement of Wikipedians from the beginning on, the project did not have to be declared an experiment and could thus preserve the advantages of a natural field experiment (Harrison and List 2004), i.e. that subjects remain in their natural environment and that their behavior is not distorted by their being aware of the experiment.

The experiment proceeds in four steps (see *Table A1* in the Online Appendix). First, on the 6th of every month, I obtain a data dump of the

¹³ Wikipedia portals are pages grouping articles on specific topics and areas.

German Wikipedia, where contributors are identified by an Internet Protocol (IP) address at the time they connect to the Internet, as well as their username and ID if they register an account on Wikipedia. I am responsible for making the decision about whom to bestow with the first level award (“Edelweiss with Star”), while the core project members evaluate the previous winners to select the recipients for the second and third level awards (2 Star and 3 Star). I use a computer script to identify all new editors who have made their first edit to an article in a given month (ca. 3000) and submit them to a basic screening, whereby algorithms single out those editors who are not blocked¹⁴ and who have contributed at least twice, with a minimum of five days lying between their first and last edit. This increases the chances that editors actually return to their account and see that they have received an award (if they belong to the treatment group).¹⁵ Only registered editors are considered for the experiment;¹⁶ “bots” (automated tools) are excluded.

Second, the remaining editors (ca. 500) are examined one by one to exclude vandals, advertisers, secondary accounts (“sockpuppets”), group accounts (including those created by organizations) and accounts of Wikimedia employees according to a rulebook developed for this purpose.¹⁷ For this step, an algorithm was developed that flags an editor if specific keywords are found on his or her user page. Only editors are retained who have made at least one contribution to an article that is still visible at the day of the screening, i.e. that

¹⁴ See, e.g., Choi et al. (2010, 110): “Some users were blocked by Wikipedia because of their vandalistic edits (...). We excluded those users.”

¹⁵ This condition drastically reduces the size of the subject pool since, as shown by Panciera, Halfaker, and Terveen (2009, 55) for the English Wikipedia, roughly “60% of registered users never make another edit after their first 24 hours.”

¹⁶ IP addresses are often not permanent. They identify a specific device whereas the same person may contribute via various devices (different computers and mobile devices). Moreover, they can map to multiple computers in a network (e.g., in an office space) or to a single but public computer, which is used by different persons (e.g., library computers). Other studies follow the same approach (e.g., Zhang and Zhu 2011).

¹⁷ This is in line e.g. with Walling and Taraborelli (2012), who also “exclud[ed] to the best of [their] knowledge sockpuppets and other categories of spurious accounts.” <http://blog.wikimedia.org/tag/experiments/> (accessed June 17, 2013).

has not been deleted, whereby only articles are considered that are not tagged for deletion (as, e.g., in Zhu et al. 2013).

From the pool of remaining editors (ca. 380), 150 award recipients are randomly selected (treatment group). In a fourth and last step, on the morning of the 12th of the given month, the list of winners is posted on the award's page and a text accompanied by a graphic award is placed on the respective editor's talk page (see Online Appendix B). 15 editors (1 percent) from the treatment group and 28 editors (1 percent) from the control group have been blocked or deleted after the awarding date and are therefore not included in the data set. Treatment and control groups thus comprise 1,185 and 1,881 editors, respectively (see *Table A1* in the Online Appendix for a break-up).

Wikipedia keeps the history of every edit made by each contributor, including the timestamps. This allows me to observe the entire range of activities that contributors engage in (e.g., correcting or writing articles, uploading files), including the correspondence among editors on talk pages.¹⁸ Thus, the accurateness of the performance measurement is close to that in lab experiments, but the scope of action is not artificially imposed and even social interaction is allowed for and taken into account. At the same time, the Internet context makes it possible to treat each award recipient equally, whereas face-to-face interactions could be subject to variability of emotional expressions and inadvertent signaling by the researcher.

C. Relevant Outcome Variables

The dependent variable of interest is *retention*, i.e., whether a newcomer becomes active again in the month following the awarding date—the "original definition of activity" being that at least one edit be made in a given month (Wikimedia 2011a, 11). As stated on the Wikimedia research pages, "[t]his

¹⁸ Only the correspondence via email is not observable. However, emails only constitute a minor fraction of the correspondence, as becomes evident e.g. by the following advice: "In general, you should not expect Wikipedians to contact you by email. Instead, check back to the talk page periodically to see if your question has been answered" https://en.wikipedia.org/wiki/Wikipedia:Researching_with_Wikipedia (May 20, 2013).

metric has commonly been used (...) when examining the "decline" of participation".¹⁹ As a robustness check, the general notion of activity can be further restricted, ignoring contributions to the project's page and to the editor's own pages so as to make sure that the results are not merely driven by verbal reactions to the award.

Retention can also be more narrowly conceived of. The most conservative measure of retention would exclusively consider article edits as a form of *direct content provision*, ignoring all other activities (as, e.g., coordination work or file uploads). A binary variable indicates the shares of the treatment and control groups that remain active in this work dimension.

Several studies have also shown the "critical importance of coordination" for online production environments such as Wikipedia (Kittur and Kraut 2008, 37). Coordination is achieved by "indirect work (...) such as conflict resolution, consensus building, or community management" (Kittur et al. 2007, 3). Thus, participation in discussions is another important measure (see, e.g., Panciera, Halfaker, and Terveen (2009); Wöhner, Köhler, and Peters (2011) include it in their reputation measure). As awards are social rewards, the study will assess whether they do indeed encourage more newcomers to develop a "community sense" (Wikimedia 2012). To operationalize the measurement of an editor's *interaction with the community*, the variable "active on talk pages" assesses whether the editor has made any contributions to the different talk pages in Wikipedia. Two other variables measure whether the newcomer has edited his or her own "user page", which is most often employed to introduce oneself to the community, or the personal "user talk page", which is usually used to respond to other editors' requests.

To assure that any treatment effect found (i.e., a higher retention rate among subjects from the treatment group) is not driven by minor editing activities, the *degree of activity* is considered as an extension. The Wikimedia research team has developed a categorization of activity levels that can be used to see

¹⁹ [http://meta.wikimedia.org/wiki/Research:Metrics/survival\(t\)](http://meta.wikimedia.org/wiki/Research:Metrics/survival(t)) (accessed May 28, 2013).

whether the award also produces editors who are more active in providing content to articles than are subjects in the control group. Authors are divided into four groups according to the number of article edits they make in the following month: Those who become inactive, those making 1 to 4 article edits, "active" editors (5-99 edits), and "very active" editors (100+ edits).²⁰

Since the field experiment focuses on newcomer retention, further performance measures are consciously avoided. In particular, the number of article edits as such is not used as a variable for the analysis, despite its favorable property of being continuous. The measure has several pitfalls (see, e.g., Adler et al. 2008). Most importantly, the experimental treatment may introduce an asymmetry between treatment and control groups that inhibits the use of this measure for the analysis. A person can prepare a whole article and put it on Wikipedia in one edit, while another person may correct minor mistakes in an article and save each single change, thus generating many more edits. Each time the save button is hit, one edit is registered. If the distortions of the edit count measure were distributed equally among treatment and control groups, the measure would be flawed but the comparative analysis could still be revealing. However, receiving an award may prompt newcomers to read and abide by the community conventions, which ask each editor to reduce the number of edits made to a single article by collecting and previewing all the changes before saving them.²¹ Thus, the award may in fact lead to a decline in the number of edits made by award recipients. The analysis' focus on retention (both broadly and narrowly defined) allows circumventing the problems related to performance measurement.

²⁰ See <http://meta.wikimedia.org/wiki/Research:Metrics> (accessed June 7, 2013).

²¹ The following illustrates the convention: "It is strongly recommended that you use [the Show preview button] before saving (...) Saving the same article several times in quick succession makes it harder for people to check what changed, and clogs up the page history." https://en.wikipedia.org/wiki/Help:Show_preview (June 18, 2013).

II. Theoretical Background and Predictions

Awards such as medals of merit or badges of honor are extrinsic, non-monetary motivators. In the terminology of Bénabou and Tirole (2003: 504), these awards correspond to "discretionary" or "ex post" rewards, as opposed to "promised" or "ex ante" contingent incentives (e.g., prize contests, as studied by Casas-Arce and Martínez-Jerez 2009). Awards differ from monetary rewards mainly in that they impose little material cost on the giver, they do not require an exact definition and measurement of performance, are less likely to crowd out the recipient's intrinsic motivation, and they are more suitable for establishing special ties of loyalty and respect between the giver and the recipient (Frey 2007). The publicity as well as the selective and official nature of their bestowal, often involving an award committee, differentiate awards from positive feedback, personal praise, and gift exchange (see Deci, Koestner and Ryan 1999 on positive feedback, Gneezy and List 2006 on gift exchange in the field, and Zhu et al. 2013 on feedback in Wikipedia).

The study of awards has only recently gained momentum in economics (e.g., Ginsburgh 2003, Frey 2005, Rablen and Oswald 2008, Malmendier and Tate 2009, Kosfeld and Neckermann 2011, Chan et al. 2014, Neckermann, Cueni, and Frey 2014). It is closely related to and draws upon the literatures on social recognition (Brennan and Pettit 2004), attention (Ellingsen and Johannesson 2007, Dur 2009), experience of competency (Deci 1975, Bénabou and Tirole 2003), relative performance feedback (Blanes i Vidal and Nossol 2011, Kuhnen and Tymula 2012), self-confidence (Bénabou and Tirole 2002), and identity (Akerlof and Kranton 2000).

Studies on awards in the field are confronted with two main problems: the identification of causality, as discussed in the introduction, and the isolation of the purely non-monetary dimensions of awards.

The context of Wikipedia makes it possible to examine the effects of awards while ruling out any monetary and career-related benefits that awards

might entail; individuals operate anonymously under online pseudonyms (Wikimedia 2011b, Jensen 2012). Most studies on awards in the field involve monetary pay, even if only to remunerate subjects for their participation in the experiment. However, the mere presence of financial rewards may interact with the award's effects (Heyman and Ariely 2004). The award given to members of the treatment group on top of their regular pay may for instance add meaning to the latter. Since the present experiment does not involve any transfers of money, such possible interactions of awards with money are prevented. It is not clear *a priori* whether awards retain a meaningful effect if strictly limited to the symbolic provision of recognition. In practice, prize money is frequently added and signals the respective award's seriousness. Else, as some observers contend, awards are often "just pieces of ribbon" (Besley 2005). In this experiment, not even ribbons are used; the award is just a graphic digital symbol. And yet, awards may be valued by their recipients because they publicly confer others' esteem. The following proposition is therefore set forth:

P1. An award can be an effective motivator (increasing retention) even when carrying no material value, i.e., having no career-related implications and involving no prize money or regular pay that the award could add meaning to.

The randomized treatment allocation and the anonymous context of Wikipedia, devoid of monetary payments and excluding career-related implications, provide a unique setting for examining whether awards have a causal impact even when no material benefits may ensue. To test the prediction, the analysis considers whether subjects subsequently become active again. It looks at general activity, but also more narrowly at direct contributions to articles only.

The experiment goes further and exploits the anonymous setting of Wikipedia to analyze the relation between awards and identity. Drawing on literature from different fields, mainly from psychology and identity

economics (Akerlof and Kranton 2000, 2005), the study distinguishes two types of identity: public identity, which relates to a natural person's public image (Jenkins 2000) and is implicated in face-to-face interactions, and private identity, involving a person's sense of self (Loewenstein 1999), self-image (Lea and Webley 1997, Bénabou and Tirole 2011), self-esteem (Pyszczynski et al. 2004, Kuhnen and Tymula 2012), and self-concept (Mazar, Amir, and Ariely 2008). Private identity is intrinsic while public identity is directly related to the extrinsic status and reputation concerns of individuals (Huberman, Loch, and Öncüler 2004, Auriol and Renault 2008, Ellingsen and Johannesson 2008, Bénabou and Tirole 2009). Related to the distinction between private and public identity are the distinctions between private and public selves (Goffman 1959, Schlenker 1986), between self-signaling and social signaling (Gneezy et al. 2012), and between self-reputation and social reputational concerns (Bénabou and Tirole 2009, 462). As Kuhnen and Tymula (2013: 94) recently pointed out, "there are no empirical or experimental accounts [and no existing theory models] of behavior in such settings [where the benefit of being among the most productive players is simply ego utility, or self-esteem]."

Since awards are usually publicly bestowed in face-to-face interactions, it is difficult to examine whether they merely enhance the recipients' public identity, or whether they also affect their private identity. The experimental setup has two particular properties that, combined, shield individuals' public identity. First, in the anonymous context of Wikipedia, only the award recipients themselves know that they are the ones who have been publicly honored. The use of online pseudonyms keeps the individuals' e-reputation separated from their real-life reputation.²² Yet, individuals may have invested heavily in their e-reputation, so that they identify with it just as others do with their natural, real-life identity. This is why the second property of the setup is

²² As argued by Brennan and Pettit (2004), "a good e-reputation is an object of desire for real agents" (p. 139) and identity-integration in the virtual and non-virtual worlds is not the best strategy to maximize one's esteem. See also Anthony, Smith, and Williamson (2009, 287).

crucial. In particular, the experimental subjects are newcomers who have not yet established such an online identity. They have just created their account. This allows me to examine the following proposition:

P2. Awards affect their recipients' private identity, involving concepts such as the sense of self, self-esteem, and self-image.

The proposition is based on the premise, demonstrated in Akerlof and Kranton (2000), that individuals have "identity-related payoffs from their own actions."²³ The utility gained from contributing to Wikipedia "depends in part on its effect on identity" (p. 721). For award recipients, continuing to contribute strengthens their identity as honorable Wikipedians and may thus lead to a utility gain not enjoyed by members of the control group, who have not been awarded. By editing Wikipedia again and by interacting with the community, former award recipients conform to the behavioral prescriptions informally linked to the award and may thus "affirm [their] self-image, or identity" (p. 716). If individuals react to receiving the award although they are anonymous and their material payoffs are not positively affected (on the contrary, they incur opportunity costs of time), their concern about contributing to Wikipedia can then only be a "concern about how they feel about themselves, about their own sense of identity" (Akerlof and Kranton 2000, 725). A sense of belonging to the community of Wikipedians can strengthen their willingness to continue contributing to the common project (see, e.g., Goette, Huffman, and Meier 2006 for field evidence on the motivational benefits of social group membership).

To test proposition P2, the award's effect on the newcomers' engagement with the community will be considered—besides general activity and direct work on articles. If the award prompts its recipients to identify themselves as members of the community of Wikipedians, they are not only more likely to

²³ Bodner and Prelec (2002) use the term "moral placebo effect", according to which "a boost in self-image positively affects actions even though it leaves true preferences unchanged." The authors examine what the present paper calls private identity when looking at "pure self-signaling, separate from any desire to be regarded well by others" (p.1).

become active again and edit articles; they are also more likely to engage in discussions and behind-the-scenes coordination work where they interact with other community members.

There is one study whose results could be interpreted in the light of proposition P2, suggesting that awards could indeed have an impact on behavior even when restricted to individuals' private identity. Restivo and van de Rijt (2012) conducted a randomized field experiment at the English Wikipedia, showing that receiving an award (called "barnstar") from another editor has a significant effect, increasing productivity (i.e., article editing activity) by 60%. Since the setting of the experiment is Wikipedia, public identity is shielded off. However, only a very small share of Wikipedia editors is considered, namely the 1% most productive contributors, who have moreover never before received an award. Given that barnstars are frequent in the English Wikipedia (see the article on "Wikipedia:Barnstaritis"), never having received one although being highly active can be particularly frustrating. The considerable treatment effect the authors find may in part be due to the particular sample of highly active contributors who finally receive their long-awaited recognition.

In contrast, the present field experiment studies newcomers who have just created their pseudonym and are not necessarily among the top performers. Many of them are not even aware that awards exist in Wikipedia. If an award changes these individuals' behavior, it is because it has introduced Wikipedia as an argument in their private identity and utility function. Moreover, this experiment also allows considering interaction with others as a proxy for self-identification with the community.

In sum, the present field experiment adds insights to the literature on motivation and rewards in that (i) it is able to establish the causal effect of awards on individual behavior (i.e., retention); (ii) monetary and career related benefits from receiving the award are ruled out; (iii) the effect on recipients' private identity (self-image, self-esteem) can be examined because the use of

only recently adopted pseudonyms keeps their public identity (public image, reputation) unchanged.

Moreover, this is the first field experiment the author is aware of, which mirrors award schemes as they can often be observed in practice, with regular monthly intervals and a fixed number of recipients. A proper award jury has been established, where senior editors lend their reputation to the award in the name of a national portal.²⁴ Together with the hall of fame-like award pages, this gives the award an official nature. The existing studies on awards mostly consider one-time interventions, not veritable award schemes put in place for a long time horizon. Other studies, in particular the few that look at Wikipedia barnstars, are moreover limited to horizontal award bestowals, where one peer gives an award to another, making the transaction similar to gift giving. The implications for reciprocation are vastly different (Frey and Gallus 2014). In the latter case, the award recipient may reciprocate by giving back a gift (or barnstar). In the former case, as in this experiment, where giver and recipient are on opposite ends of a vertical hierarchy, the recipient cannot reciprocate by bestowing an award upon the giver. The possibilities for reciprocation are channeled towards conforming to the giver's expectations, e.g., by continuing to edit Wikipedia.

III. Empirical Analysis

A. Randomization

Randomly bestowing awards at first sight seems to be an impossible task because awards are designed to be given to the persons excelling in their tasks. However, the experiment shows that it can succeed if two important conditions are fulfilled. Firstly, a basic pre-selection has to exclude obviously

²⁴ In the study by Restivo and van de Rijt 2012, a single and self-declared "not very active" editor handed out the barnstars.
https://en.wikipedia.org/wiki/Wikipedia:Administrators%27_noticeboard/IncidentArchive691#IP_handing_out_random_barnstars (last accessed October 15, 2013).

undeserving candidates, like vandals (see also, e.g., Glewwe, Kremer, and Moulin (2009), who randomly attributed treatment to 25 schools from a list of 100 deserving schools). Secondly, subjects who—by chance—do not receive the award should be an unidentifiable group who are ideally ignorant of the award's existence. This is why the higher levels of the award could not be randomly bestowed. Non-receipt of the second-level award risks being demotivating for someone belonging to the identifiable group of winners of the first-level award who has made an effort to be honored again but fails to win. Such a decision can hardly be randomized.

Randomization has advantages going beyond the identification of a causal effect (Zeitoun, Osterloh, and Frey 2014). Most importantly, it prevents biased decisions (e.g., based on a jury member's political hue) and discourages strategic manipulation (e.g., awards given for work on a particular issue). Randomization allows a wider variety of subjects to be honored and to receive attention, which is in line with Wikipedia's concern for diversity.²⁵

As a randomization check, Panel A of *Table 1* displays the t-tests of all the variables a jury actively selecting award winners might take as performance criteria (i.e., number of general edits before awarding date, number of article edits, number of different articles edited, number of talk edits). Panel B includes Chi-square tests on binary variables that might predict responsiveness to an award, i.e., if editors had previously created their own user page, responded to messages on their talk page, or were actively participating in discussions on talk pages more generally (for a similar reasoning, see Zhang and Zhu 2011, 1609).

²⁵ "Wikipedia also needs more *different* Wikipedians [so as to] increase the quality and completeness of the encyclopedia [as well as] the likelihood that any new member of the community will find like-minded collaborators and feel like they fit in" (Morgan et al. 2013, 8).

TABLE 1—RANDOMIZATION CHECK

	Treatment (1)	Control (2)	Difference (3)
<i>A. Performance measures</i>			
total # of edits (ex ante)	25.81	26.25	-0.44 (-0.17)
# article edits (ex ante)	18.24	18.57	-0.34 (-0.16)
# different articles edited (ex ante)	8.58	7.65	0.93 (0.57)
# talk edits (ex ante)	3.26	2.78	0.47 (1.28)
<i>B. Responsiveness predictors</i>			
user page self-edited (ex ante)	0.21	0.19	0.01 (0.46)
user talk self-edited (ex ante)	0.18	0.16	0.02 (0.14)
active on talk pages (ex ante)	0.43	0.40	0.03 (0.14)
N	1185	1881	3066

Notes: Values are rounded to 2 decimal places. *t*-values are reported in parentheses in Panel A. Differences were estimated with Chi-square tests in Panel B, where *p*-values are indicated in parentheses. For the number of observations (N) in treatment and control groups, see *Table A1* in the Online Appendix.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

As expected, before the intervention, the differences are all negligible, point into different directions, and are never even marginally statistically significant.²⁶ The binary outcome variables—whether an editor stays active at all, or on article pages only—are not included since they are per definition positive given that the pre-selection only considers newcomers who have made their first article edit in the previous month.

However, the ordinal variable based on the Wikimedia Foundation's categorization of activity levels, which will be included as an outcome

²⁶ A power analysis indicates that even a small effect of .2 (see Cohen 1988) would be detected, if it existed, with a probability of .99991 (one-tailed) given the sample size of the two groups and an alpha error probability of .05.

variable, can be used as a further randomization check (the levels of activity being 0 article edits, 1-4 article edits, 5-99 article edits, and 100+ article edits). A Wilcoxon-Mann-Whitney test (see *Table C1*, Online Appendix C) produces statistically insignificant results ($z = 1.539$, $\text{Prob} > |z| = 0.1239$). This again confirms that assignment to treatment and control groups was random.

B. Results

The analysis considers eight awarding rounds, from September 2012 to April 2013. Each month, 150 newcomers have received the “Edelweiss with Star” award. 15 editors (1 percent) from the treatment group and 28 editors (1 percent) from the control group have been blocked or deleted after the awarding date and are therefore not included in the data set. Treatment and control groups thus comprise 1,185 and 1,881 individual editors, respectively (see *Table A1* in the Online Appendix for a break-up per monthly cohort).

26 award recipients from the eight awarding rounds considered have put a small template (called “Babel”) that was created by a recipient of the first wave (see Online Appendix B2) onto their user page. 16 have copied the entire text and award graphic (Online Appendix B1) to display it there (1 among them also included the small template). Hence, 41 winners (3.5 percent) have chosen to display their receipt of the award more prominently on their user page. Some of them have even created a subpage “Awards” for the graphics. Adding the 108 pure thank you messages²⁷ to this count, 12.6 percent of recipients (149 out of 1,185) have thus favorably reacted to the award as of June 2013. This section will explore how these immediate and mostly verbal reactions translate into observable contributions to the public good.

Thanks to the random assignment of the treatment, potential confounding variables are on average distributed equally between the treatment and control

²⁷ Such as: “I initially only wanted to make a few corrections every now and then, but this form of welcoming has highly motivated me! I am now working on my first article... Many heartfelt thanks in retrospect!” (author’s translation).

groups (Campbell, Stanley, and Gage 1963). Hence, basic and well interpretable mean-comparison tests can be used to see whether the award has a causal effect on retention in its different dimensions (general activity, direct content provision, indirect coordination work); and if so, what the size of the effect is. The findings on general activity and article editing activity are relevant for both, propositions P1 and P2. The results obtained on indirect community work are most relevant for proposition P2 since they are indicative of an individual's self-identification as a community member.

General activity.—Figure 1 plots the shares of editors in the control and treatment groups who become active again in the first month after awarding date. This basic bar chart indicates that the retention rate is 8.7 percentage points higher for recipients of the award. The error bars indicate that the 25 percent increase in the retention rate is statistically significant.

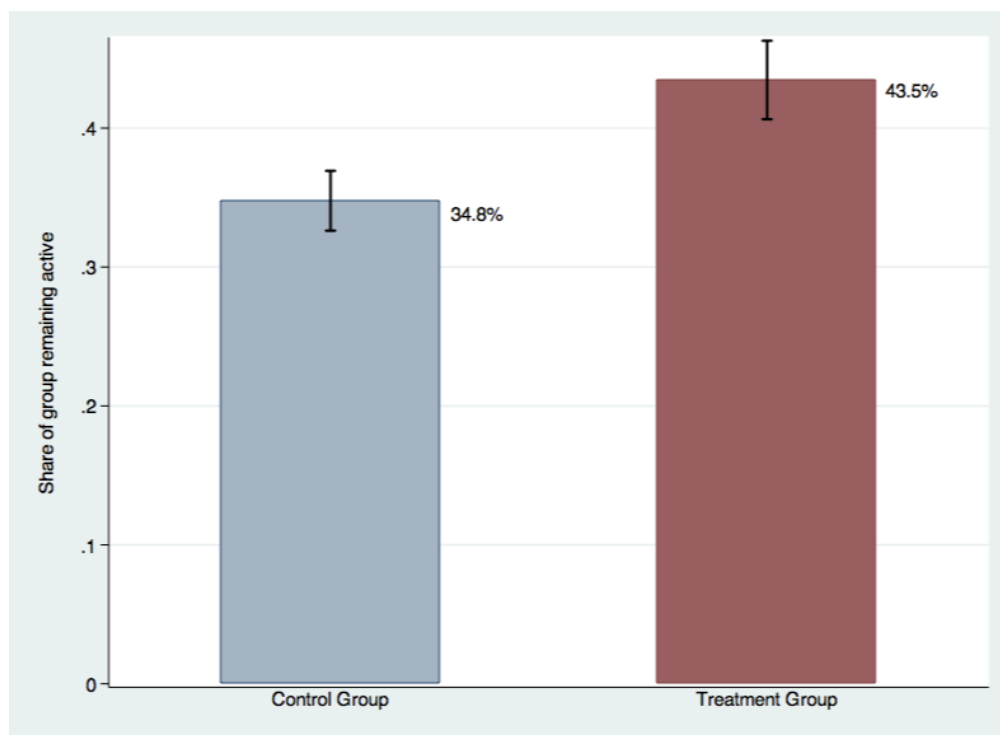


FIGURE 1. MEAN RETENTION RATES

Notes: Focusing on first month after awarding date. Error bars indicate the respective 95% confidence intervals.

Table 2 reports the retention rates in the treatment and control groups, as well as the differences between them and the p-values resulting from the Chi-square tests. Panel A focuses on the most basic measure of retention, i.e. whether any activity can be observed after the awarding date. Following up on Figure 1, the first row considers the month after awarding date and shows that the difference observed in the bar chart is indeed statistically significant at the 99.9 percent level ($\chi^2(1) = 23.28$, $p = 0.000$).

TABLE 2 — CHI-SQUARE TESTS

	Treatment (1)	Control (2)	Difference (3)
<i>A. Retention</i>			
active (1st month)	0.43	0.35	0.09*** (0.000)
active (following 2 months)	0.50	0.42	0.08*** (0.000)
<i>B. Direct content provision</i>			
active, only article edits (1st month)	0.38	0.32	0.06** (0.002)
<i>C. Interaction with community</i>			
user page self-edited (1st month)	0.08	0.04	0.04*** (0.000)
user talk self-edited (1st month)	0.10	0.05	0.05*** (0.000)
active on talk pages (1st month)	0.21	0.13	0.08*** (0.000)
N	1185	1881	3066

Notes: Average values are rounded to 2 decimal places. p -values from Chi-square tests are reported in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Row 2 in Panel A (*Table 2*) extends the period of observation to the 2 months following the intervention to account for the duration of the effects. The results in row 2 ascertain that the difference observed in row 1 is not due

to a temporal substitution effect, i.e. that award recipients do not merely advance their next period of activity to the first month instead of the second after awarding date. In the Internet context, already the first 24 hours are highly predictive of an editor's future engagement, such that the eight weeks considered are a relatively long time horizon (see, e.g., Panciera, Halfaker, and Terveen 2009, Morgan et al. 2013). Previous research on gift exchange has shown that positive effects initially found vanished after the first couple of hours (Gneezy and List 2006). The results in row 2 show that, over the 2 months following the intervention, the retention rate in the treatment group is still 8 percentage points higher than the retention rate of 42 percent observed in the control group. This treatment effect of 19 percent is again highly statistically significant ($\chi^2(1) = 17.26$, $p = 0.000$). The award's effect on retention thus persists even when the period of observation is extended.

Direct work on articles.—When restricting the definition of activity to article edits only (Table 2, Panel B), the share of award recipients who remain active in this dimension of work is 6 percentage points higher than the retention rate of 32 percent observed in the control group. The award's causal effect, raising the retention rate by 19 percent, is statistically significant at the 99-percent level ($\chi^2(1) = 9.87$, $p = 0.002$).

In a further step, the analysis of Panel B (Table 2) is extended to explore whether the award also produces more highly active authors. The Wikimedia Foundation's categorization of activity levels (focusing on article edits) is used as a basis for this part of the analysis. Figure 2 shows the distribution of subjects from the control and treatment groups in the first month after the awarding date across the different levels of article editing activity commonly considered.

As can be seen in Figure 2, the award's effect on retention is not driven by only minor article editing activity—the share of newcomers who remain active is increased on every level of activity (right side of the dashed line). The award raises the share of editors who make between 1 and 4 article edits in the

month after the awarding date by almost 14 percent (from 16.9 percent to 19.2 percent). It increases the share of editors who reach the medium level of activity (5-99 article edits) by about 18 percent (from 13.6 percent to 16 percent). The increase in the share of “very active” newcomers (100 edits and more) amounts to almost 37 percent (from 1.9 percent in the control to 2.6 percent in the treatment group). Figure 2 thus illustrates that the award not only lowers the attrition rate of new authors (as seen in Panel B of *Table 2* and on the left side of the dashed line in Figure 2); it indeed raises the share of editors on every level of activity.

The limitations of the edit count metric have been mentioned above, but transforming the measure into an ordinal variable mitigates the distortionary tendencies and gives a useful indication that helps to ascertain that the award's effect on retention found in Panel B of *Table 2* is not due to only minor article editing activity. A Wilcoxon-Mann-Whitney test (see *Table C2, Online Appendix C*) shows the statistical significance of the finding that the treatment group tends to have larger values than the control group (z -value = -3.192, $\text{Prob} > |z| = 0.0014$).²⁸

²⁸ Mann Whitney only allows to draw conclusions about the statistical significance of the test. For the effect size, ordered logit is used. The proportional odds ratio of comparing subjects of the control group with award recipients on the level of activity is 1.098703 ($\text{Prob} > |z| = 0.020$), the cut-points being 1.131517, 2.023534, and 4.133395. This means that the odds of high activity (i.e., more than 4 edits) versus the combined lower categories of activity are 1.099 times greater for subjects in the treatment group.

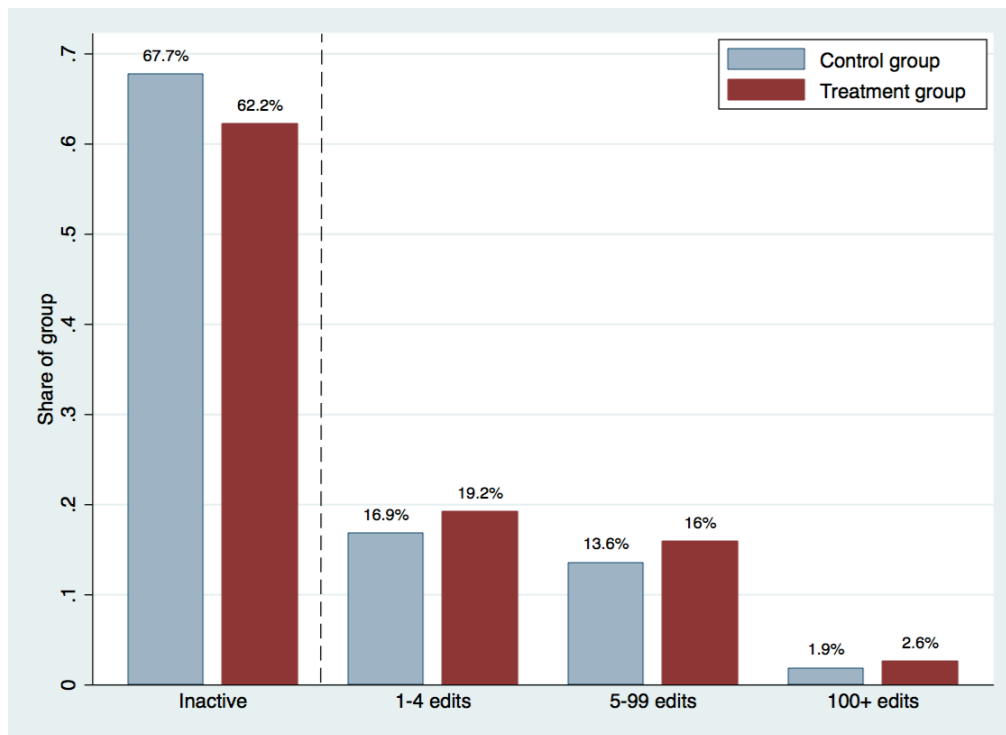


FIGURE 2.DEGREE OF ARTICLE EDITING ACTIVITY

Notes: First month after awarding date. The figure shows the degree to which those editors who continue to edit articles actually contribute (right side of dashed line). It thus extends Panel B of Table 2.

Indirect community work.—Panel C of Table 2 considers the third fundamental conception of retention: Whether editors actively engage with the community and fulfill behind-the-scenes coordination work. As expected given the social nature of awards, the intervention has a positive effect that is statistically highly significant. Compared to the control group, twice as many award recipients introduce themselves to the community on their personal user pages in the month after the intervention (8 percent vs. 4 percent; $\chi^2(1) = 23.17, p = 0.000$), thus providing others with a better possibility of contacting them (e.g., based on the fields of interest indicated on the personal page). Likewise, the award doubles the share of newcomers who respond to others' requests on their talk pages (row 2 of Panel C) and thus enter in direct contact with community members (10 percent vs. 5 percent; $\chi^2(1) = 31.26, p = 0.000$). When considering interactions on talk pages more generally (thus also including, e.g., other editors' talk pages and article talk pages), the share of

active editors is 8 percentage points higher in the award condition than in the control group (21 percent vs. 13 percent); the increase by 62 percent being highly statistically significant ($\chi^2(1) = 35.57, p = 0.000$).

C. Robustness

Two further binary activity variables are created to ascertain that the difference in retention rates between treatment and control groups found in Panel A of *Table 2*, which considers general editing activity (e.g., writing articles, discussing their content, uploading files), is not merely driven by verbal responses to the award. In line with the analysis in Panel A (*Table 2*), the two variables record any type of activity. However, the first alternative activity variable ignores contributions to the award project's page, and the second variable goes even further in that it also omits edits to the respective editor's own user and user talk pages.

Table C3 (Online Appendix C) presents the results from the Chi-square tests comparing the means of editors who stay active in the month following the awarding date. Excluding the award project's page (row 1) does not change the results from Panel A of *Table 2*. This suggests that award recipients who post a thank you note on the project's page also go on to make other contributions. Additionally excluding editors' own pages somewhat reduces the effect, but still shows a difference of 7 percentage points, which is highly statistically significant ($\chi^2(1) = 16.41, p = 0.000$). Thus, even when applying very conservative measures of activity, like that in row 2 of *Table C3* or the one that only considers direct edits to articles (*Table 2*, Panel B), a causal effect of the award on retention can be established.

IV. Concluding Remarks

Whereas research on the motivations for private contributions to public goods is extensive, little is known about the rewards and incentives that help sustain volunteers' contributions without crowding out their intrinsic

motivation. Awards such as orders and medals of valor are symbolic rewards that may foster such voluntary contributions. They are a widespread phenomenon indeed. However, any investigation into their causal effects is hindered by their juries' unwillingness to see their authority replaced by random decision-making processes.

This study presents a large-scale natural field experiment where an award scheme is implemented and randomization is employed to establish clear causal effects of awards on voluntary contributions to a public good. The experiment is conducted at the German-language version of Wikipedia and explores whether awards can be used to increase newcomer retention and thereby respond to one of the online encyclopedia's biggest challenges. Even though the awards are given in an anonymous context, where no material or non-virtual benefits such as status among one's peers can arise, the award has a considerable and statistically highly significant impact on retention rates.

The findings support the proposition that even purely symbolic awards can be effective motivators. The share of newcomers who remain active in the month after the awarding date is 9 percentage points higher for the treatment group than for the control group. This is a 25 percent increase in the retention rate ($p = 0.000$), from a share of 35 percent to one of 44 percent, which is not driven by a temporal substitution effect or by only verbal responses to the award. When considering only direct contributions to articles, the award also increases the fraction of newcomers who remain active according to this conservative definition of work, resulting in a difference of 6 percentage points between treatment and control groups. The 19 percent increase (from 32 percent to 38 percent, $p = 0.002$) is not driven by minor editing activities; the award increases the shares of authors on all three commonly considered activity levels.

The proposition relating to the award's effects on recipients' private identity is also backed by the findings. First, the above-mentioned results show that even ruling out any effects on public identity (reputation, status among peers)

does not inhibit the award from positively affecting individual behavior. Second, the award also has a substantial impact on the number of newcomers who interact with the community and engage in behind-the-scenes coordination work, suggesting that their private identity as a community member ("Wikipedian") has been strengthened. After the intervention, 8 percentage points more editors in the treatment group are active in this work dimension, a 62 percent increase from the rate of 13 percent of editors in the control group who engage in such tasks ($p = 0.000$). The importance of indirect work has been highlighted by several studies (Kittur and Kraut 2008, Wöhner, Köhler, and Peters 2011, Morgan et al. 2013), and the social ties thus established further increase the likelihood that the respective newcomer also stays within the community (McPherson, Popielarz, and Drobnic 1992).

The findings are worth noting not only because the award is costless and has no material implications for the recipients, but also because it is given to newcomers who operate under only recently adopted pseudonyms. The estimates are conservative since non-responsiveness to the intervention may be due to two causes: Indifference, but also unawareness since some recipients simply do not (or only belatedly) return to their user page and thus do not see that they won an award (see, e.g., Panciera, Halfaker, and Terveen 2009, 55, and Morgan et al. 2013, 5).

The analysis focuses on the treatment's effect on newcomer retention and leaves aside possible spill-over effects of the award. Some recipients have for instance written messages to other, more tenured editors, stating for example: "(...) I want to thank you most warmly since I would not have succeeded [in writing my first article] without your help. As such, a petal of the edelweiss belongs to you; just choose one" (author's translation).

While this suggests that the award's beneficial impact may exceed the effects found in this analysis, two limitations have to be pointed out. Firstly, the value of awards hinges on them being rare; they have to be used sparingly to prevent inflationary tendencies. Awards should therefore be integrated into a broader

reward scheme to increase retention rates. It is an interesting topic for future research to explore the relationship between the value of an award and the quantity and frequency with which it is bestowed. Varying other award parameters, such as the degree of publicity, would be of no less interest.

Secondly, the experiment is conducted in an anonymous online context. As stated by Zhang and Zhu (2011, 1613), it is an interesting question for future research to examine the findings' generalizability to non-virtual fields and other public goods contexts. At the same time, however, the results as such are already important given the increasing Internet penetration of the professional and private spheres. The finding that anonymous newcomers to an online community are motivated by social recognition, the experience of competency, self-confidence and attention is important and should receive further attention by researchers. It might inform policies dealing with anonymity and its impacts on individual behavior.

REFERENCES

- Adler, B. Thomas, Luca De Alfaro, Ian Pye, and Vishwanath Raman. 2008. "Measuring Author Contributions to the Wikipedia." *Proceedings of the 4th International Symposium on Wikis*: 1–15.
- Akerlof, George A. and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115: 715–753.
- Akerlof, George A. and Rachel E. Kranton. 2005. "Identity and the Economics of Organizations." *Journal of Economic Perspectives* 19: 9–32.
- Anthony, Denise, Sean W. Smith, and Timothy Williamson. 2009. "Reputation And Reliability In Collective Goods: The Case of the Online Encyclopedia Wikipedia." *Rationality and Society* 21: 283–306.
- Ariely, Dan, Anat Bracha, and Stephan Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review* 99: 544–555.

- Auriol, Emmanuelle and Régis Renault. 2008. "Status and Incentives." *RAND Journal of Economics* 39: 305–326.
- Bénabou, Roland and Jean Tirole. 2002. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics* 117: 871–915.
- Bénabou, Roland and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* 70: 489–520.
- Bénabou, Roland and Jean Tirole. 2009. "Over my Dead Body: Bargaining and the Price of Dignity." *American Economic Review* 99: 459–465.
- Bénabou, Roland and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." *Quarterly Journal of Economics* 126: 805–855.
- Besley, Timothy. 2005. "Notes on Honours." Mimeo, London School of Economics.
- Blanes i Vidal, Jordi and Mareike Nossol. 2011. "Tournaments without prizes: Evidence from personnel records." *Management Science* 57: 1721–1736.
- Bodner, Ronit and Drazen Prelec. 2002. "Self-signaling and diagnostic utility in everyday decision making." In *Collected Essays in Psychology and Economics*, edited by Isabelle Brocas, and Juan D. Carillo, 1–22. Oxford, UK: Oxford University Press.
- Brennan, Geoffrey and Philip Pettit. 2004. "Esteem, Identifiability and the Internet." *Analyse und Kritik* 26: 139–157.
- Bryant, Susan L., Andrea Forte, and Amy Bruckman. 2005. "Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia." *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*: 1–10.
- Butler, Brian, Elisabeth Joyce, and Jacqueline Pike. 2008. "Don't Look Now, But We've Created a Bureaucracy: The Nature and Roles of Policies and Rules in Wikipedia." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 1101–1110.

- Butler, Declan. 2008. "Publish in Wikipedia or Perish." *Nature News*. Accessed May 17, 2013. <http://www.nature.com/news/2008/081216/full/news.2008.1312.html>.
- Campbell, Donald Thomas, Julian C. Stanley, and Nathaniel Lees Gage. 1963. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin Company.
- Casas-Arce, Pablo and F. Asís Martínez-Jerez. 2009. "Relative Performance Compensation, Contests, and Dynamic Incentives." *Management Science* 55: 1306–1320.
- Chan, Ho Fai, Bruno S. Frey, Jana Gallus, and Benno Torgler. 2014. "Academic Honors and Performance." *Labour Economics*, forthcoming.
- Chaudhuri, Ananish. 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments. A Selective Survey of the Literature." *Experimental Economics* 14: 47–83.
- Chen, Yan, F. Maxwell Harper, Joseph Konstan, and Sherry Xin Li. 2010. "Social comparisons and contributions to online communities: A field experiment on MovieLens." *American Economic Review* 100: 1358–1398.
- Choi, Boreum, Kira Alexander, Robert E. Kraut, and John M. Levine. 2010. "Socialization Tactics in Wikipedia and their Effects." *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*: 107–116.
- Cifolilli, Andrea. 2003. "Phantom Authority, Self-Selective Recruitment and Retention of Members in Virtual Communities." *First Monday* 8. Accessed May 17, 2013. <http://firstmonday.org/ojs/index.php/fm/article/view/1108/1028>.
- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cooper, David J. and John H. Kagel. 2013. "Other-Regarding Preferences: A Selective Survey of Experimental Results." In *The Handbook of*

- Experimental Economics II*, edited by John H. Kagel, and Alvin E. Roth. Princeton, NJ: Princeton University Press.
- Cornes, Richard and Todd Sandler. 1996. *The Theory of Externalities, Public Goods, and Club Goods*. Cambridge, UK: Cambridge University Press.
- Deci, Edward L. 1975. *Intrinsic Motivation*. New York: Plenum Press.
- Deci, Edward L., Richard Koestner, and Richard M. Ryan. 1999. "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation." *Psychological Bulletin* 125: 627–668.
- Dur, Robert. 2009. "Gift Exchange in the Workplace: Money or Attention?" *Journal of the European Economic Association* 7: 550–560.
- Ellingsen, Tore and Magnus Johannesson. 2007. "Paying Respect." *Journal of Economic Perspectives* 21: 135–150.
- Ellingsen, Tore and Magnus Johannesson. 2008. "Pride and Prejudice: The Human Side of Incentive Theory." *American Economic Review* 98: 990–1008.
- Falk, Armin and Michael Kosfeld. 2006. "The Hidden Costs of Control." *American Economic Review* 96: 1611–1630.
- Fehr, Ernst and Klaus M. Schmidt. 2002. "Theories of Fairness and Reciprocity. Evidence and Economic Applications." In *Advances in Economics and Econometrics. Theory and Applications, Eighth World Congress, Econometric Society Monographs*, edited by Mathias Dewatripont, and Stephen J. Turnovsky, 208–257. Cambridge, UK: Cambridge University Press.
- Fehr, Ernst and Klaus M. Schmidt. 2006. "The Economics of Fairness, Reciprocity and Altruism. Experimental Evidence and New Theories." In *Handbook of the Economics of Giving, Altruism and Reciprocity*, edited by Serge-Christophe Kolm, and Jean Mercier Ythier, 615–691. Amsterdam, Oxford: North-Holland.

- Frey, Bruno S. 2005. "Knight Fever: Towards an Economics of Awards." CESifo Working Paper 1468. Accessed May 17, 2013. <http://www.econstor.eu/handle/10419/18832>.
- Frey, Bruno S. 2007. "Awards as Compensation." *European Management Review* 4: 6–14.
- Frey, Bruno S. and Jana Gallus. 2014. "Awards are a special kind of signal." CREMA Working Paper No. 2014–04.
- Frey, Bruno S. and Lorenz Goette. 1999. "Does Pay Motivate Volunteers?" IEW Working Paper 7. Accessed May 17, 2013. <http://ideas.repec.org/p/zur/iewwp/007.html>.
- Frey, Bruno S. and Susanne Neckermann. 2009. "Awards: A Disregarded Source of Motivation." In *Perspectives in Moral Science*, edited by Michael Baumann, and Bernd Lahno, 177–182. Frankfurt am Main: Frankfurt School Verlag.
- Giles, Jim. 2005. "Internet Encyclopaedias Go Head to Head." *Nature* 438: 900–901.
- Ginsburgh, Victor. 2003. "Awards, Success and Aesthetic Quality in the Arts." *Journal of Economic Perspectives* 17: 99–111.
- Gleitman, Henry. 1996. *Basic Psychology*. New York: Norton.
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics* 1: 112–135.
- Gneezy, Ayelet, Uri Gneezy, Gerhard Riener, and Leif D. Nelson. 2012. "Pay-what-you-want, identity, and self-signaling in markets." *Proceedings of the National Academy of Sciences* 109: 7236–7240.
- Gneezy, Uri and John A. List. 2006. "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments." *Econometrica* 74: 1365–1384.

- Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and why incentives (don't) work to modify behavior." *Journal of Economic Perspectives* 25: 191–209.
- Goette, Lorenz, David Huffman, and Stephan Meier. 2006. "The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups." *American Economic Review* 96: 212–216.
- Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. New York: Anchor Books.
- Greenstein, Shane and Feng Zhu. 2012. "Is Wikipedia Biased?" *American Economic Review* 102: 343–348.
- Halfaker, Aaron. 2012. "Kids These Days: The Quality of New Wikipedia Editors over Time." *The Wikimedia Community Blog*, March 27.
<http://blog.wikimedia.org/2012/03/27/analysis-of-the-quality-of-newcomers-in-wikipedia-over-time/>.
- Halfaker, Aaron, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl. 2013. "The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity is Causing its Decline." *American Behavioral Scientist* 57: 664–688.
- Halfaker, Aaron, Aniket Kittur, and John Riedl. 2011. "Don't Bite the Newbies: How Reverts Affect the Quantity and Quality of Wikipedia Work." *WikiSym – Proceedings of the 7th International Symposium on Wikis and Open Collaboration*: 163–172.
- Hansen, W. Lee and Burton Weisbrod. 1972. "Toward a General Theory of Awards, or, Do Economists Need a Hall of Fame?" *Journal of Political Economy* 80: 422–431.
- Harbaugh, William T. 1998. "The Prestige Motive for Making Charitable Transfers." *American Economic Review* 88: 277–282.
- Harrison, Glenn W. and John A. List. 2004. "Field Experiments." *Journal of Economic Literature* 42: 1009–1055.

- Heyman, James and Dan Ariely. 2004. "Effort For Payment: A Tale of Two Markets." *Psychological Science* 15: 787–793.
- Huberman, Bernardo A., Christoph H. Loch, and Ayse Önçüler. 2004. "Status as a Valued Resource." *Social Psychology Quarterly* 67: 103–114.
- Jenkins, Richard. 2000. "Categorization: Identity, Social Process and Epistemology." *Current Sociology* 48: 7–25.
- Jensen, Richard. 2012. "Military History on the Electronic Frontier: Wikipedia Fights the War of 1812." *Journal of Military History* 76: 1165–1182.
- Kittur, Aniket and Robert E. Kraut. 2008. "Harnessing the Wisdom of Crowds in Wikipedia: Quality through Coordination." *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*: 37–46.
- Kittur, Aniket, Bongwon Suh, Bryan A. Pendleton, and Ed H. Chi. 2007. "He Says, She Says: Conflict and Coordination in Wikipedia." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 453–462.
- Kosfeld, Michael and Susanne Neckermann. 2011. "Getting More Work for Nothing? Symbolic Awards and Worker Performance." *American Economic Journal: Microeconomics* 3: 86–99.
- Kraut, Robert E., Aniket Kittur, Haiyi Zhu, Jenny He, and Amy Zhang. 2012. "Effects of Feedback on Participation in Wikipedia." *Wikimedia Research*. Accessed May 17, 2013.
http://meta.wikimedia.org/wiki/Research:Effects_of_Feedback_on_Participation_in_Wikipedia.
- Kube, Sebastian, Michel André Maréchal, and Clemens Puppe. 2012. "The Currency of Reciprocity: Gift Exchange in the Workplace." *American Economic Review* 102: 1644–1662.
- Kuhnen, Camelia M. and Agnieszka Tymula. 2012. "Feedback, self-esteem, and performance in organizations." *Management Science* 58: 94–113.
- Lakhani, Karim R. and Eric von Hippel. 2003. "How Open Source Software Works: "Free" User-to-User Assistance." *Research Policy* 32: 923–943.

- Lakhani, Karim R. and Robert G. Wolf. 2005. "Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects." In *Perspectives on Free and Open Source Software*, edited by Joseph Feller, Brian Fitzgerald, Scott A. Hissam, and Karim R. Lakhani, 3–21. Cambridge, MA: MIT Press.
- Lave, Jean and Etienne Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press.
- Lea, Stephen E.G. and Paul Webley. 1997. "Pride in Economic Psychology." *Journal of Economic Psychology* 18: 323–340.
- Ledyard, John O. 1995. "Public Goods: A Survey of Experimental Research." In *The Handbook of Experimental Economics*, edited by John H. Kagel, and Alvin E. Roth, 111–194. Princeton, NJ: Princeton University Press.
- Lerner, Josh and Jean Tirole. 2002. "Some Simple Economics of Open Source." *Journal of Industrial Economics* 50: 197–234.
- List, John A. and Imran Rasul. 2010. "Field Experiments in Labor Economics." In *Handbook of Labor Economics*, edited by Orley C. Ashenfelter and David Edward Card, 103–228. San Diego, CA, and Amsterdam: Elsevier.
- Loewenstein, George. 1999. "Because it is there: The Challenge of Mountaineering... for Utility Theory." *Kyklos* 52: 315–343.
- Macduff, Nancy. 2005. "Societal Changes and the Rise of the Episodic Volunteer." In *Emerging Areas of Volunteering*, edited by Jeffrey L. Brudney, 49–62. Indianapolis: Association for Research on Nonprofit Organizations and Voluntary Associations.
- Malmendier, Ulrike and Geoffrey Tate. 2009. "Superstar CEOs." *Quarterly Journal of Economics* 124: 1593–1638.
- Mazar, Nina, On Amir, and Dan Ariely. 2008. "The Dishonesty of Honest People: A Theory of Self Concept Maintenance." *Journal of Marketing Research* 45: 633–644.

- McPherson, J. Miller, Pamela A. Popielarz, and Sonja Drobnic. 1992. "Social Networks and Organizational Dynamics." *American Sociological Review* 57: 153–170.
- Morgan, Jonathan T., Siko Bouterse, Sarah Stierch, and Heather Walls. 2013. "Tea & Sympathy: Crafting Positive New User Experiences on Wikipedia." *Computer Supported Cooperative Work*. Accessed May 17, 2013. http://dub.washington.edu/djangosite/media/papers/morgan_cscw2013_final.pdf.
- Neckermann, Susanne, Reto Cueni, and Bruno S. Frey. 2014. "Awards at Work." *Labour Economics*, forthcoming.
- Nov, Oded. 2007. "What Motivates Wikipedians?" *Communications of the ACM* 50: 60–64.
- Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.
- Panciera, Katherine, Aaron Halfaker, and Loren Terveen. 2009. "Wikipedians Are Born, Not Made: A Study of Power Editors on Wikipedia." *Proceedings of the ACM International Conference on Supporting Group Work*: 51–60.
- Pinchuk, Maryana. 2011. "The Rise of Warnings to New Editors on English Wikipedia." *The Wikimedia Community Blog*, May 2.
- Pyszczynski, Tom, Jeff Greenberg, Sheldon Solomon, Jamie Arndt, and Jeff Schimel. 2004. "Why Do People Need Self-Esteem? A Theoretical and Empirical Review." *Psychological Bulletin* 130: 435–468.
- Rablen, Matthew D. and Andrew J. Oswald. 2008. "Mortality and immortality: The Nobel Prize as an experiment into the effect of status upon longevity." *Journal of Health Economics* 27: 1462–1471.
- Redelmeier, Donald A. and Sheldon M. Singh. 2001a. "Survival in Academy Award-Winning Actors and Actresses." *Annals of Internal Medicine* 134: 955–962.

- Redelmeier, Donald A. and Sheldon M. Singh. 2001b. "Longevity of Screenwriters Who Win an Academy Award: Longitudinal Study." *British Medical Journal* 323: 1491–1496.
- Restivo, Michael and Arnout van de Rijt. 2012. "Experimental Study of Informal Rewards in Peer Production." *PloS ONE* 7: e34358.
- Schlenker, Barry R. 1986. "Self-Identification: Toward an Integration of the Private and Public Self." In *Public Self and Private Self*, edited by Roy F. Baumeister, 21–62. New York: Springer.
- Schroer, Joachim and Guido Hertel. 2009. "Voluntary Engagement in an Open Web-Based Encyclopedia: Wikipedians and Why They Do It." *Media Psychology* 12: 96–120.
- Shaw, Aaron and Benjamin Mako Hill. 2012. "Can Social Awards Create Better Wikis?" Paper presented at the annual Wikimania conference, Washington, DC, July 12–15.
- Siming, Linus. 2012. "Orders of Merit and CEO Compensation: Evidence from a Natural Experiment." Available at SSRN. Accessed May 17, 2013. <http://ssrn.com/abstract=2100499> or <http://dx.doi.org/10.2139/ssrn.2100499>.
- Snyder, Mark and Allen M. Omoto. 2008. "Volunteerism: Social Issues Perspectives and Social Policy Implications." *Social Issues and Policy Review* 2: 1–36.
- Soetevent, Adriaan R. 2005. "Anonymity in Giving in a Natural Context: A Field Experiment in 30 Churches." *Journal of Public Economics* 89: 2301–2323.
- Suh, Bongwon, Gregorio Convertino, Ed H. Chi, and Peter Pirolli. 2009. "The Singularity is Not Near: Slowing Growth of Wikipedia." *Proceedings of the 5th International Symposium on Wikis and Open Collaboration*: 8.
- Tirole, Jean and Roland J. Bénabou. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96: 1652–1678.

- Viégas, Fernanda B., Martin Wattenberg, and Matthew M. McKeon. 2007. "The Hidden Order of Wikipedia." *Online Communities and Social Computing*: 445–454.
- Walling, Steven and Dario Taraborelli. 2012. "Is This Thing On? Giving New Wikipedians Feedback Post-Edit." *The Wikimedia Blog*, September 24.
- Wikimedia. 2011a. "Editor Trends Study – Results." *Wikimedia Strategic Planning*, May 14.
http://strategy.wikimedia.org/wiki/Editor_Trends_Study/Results.
- Wikimedia. 2011b. "Semi-Annual Survey of Wikipedia Editors." *Wikimedia Foundation*, April. Accessed June 17, 2013.
http://upload.wikimedia.org/wikipedia/commons/1/10/WP_2011_Editor's_Survey_-_Topline.pdf.
- Wikimedia. 2012. "Research:Editor Milestones." *Editor Engagement Project*, September 21. http://meta.wikimedia.org/wiki/Research:Editor_milestones.
- Wikimedia Deutschland. 2011. *Alles über Wikipedia und die Menschen hinter der größten Enzyklopädie der Welt*. Hamburg: Hoffmann und Campe.
- Wöhner, Thomas, Sebastian Köhler, and Ralf Peters. 2011. "Automatic Reputation Assessment in Wikipedia." *ICIS 2011 Proceedings* Paper 5. Accessed May 17, 2013.
<http://aisel.aisnet.org/icis2011/proceedings/onlinecommunity/5>.
- Zeitoun, Hossam, Margit Osterloh, and Bruno S. Frey. 2014. "Learning From Ancient Athens: Demarchy and Corporate Governance." *Academy of Management Perspectives* 28: 1–14.
- Zelmer, Jennifer. 2003. "Linear Public Goods. A Meta-Analysis." *Experimental Economics* 6: 299–310.
- Zhang, Xiaoquan (Michael) and Feng Zhu. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia." *American Economic Review* 101: 1601–1615.
- Zhu, Haiyi, Amy Zhang, Jiping He, Robert E. Kraut, and Aniket Kittur. 2013. "Effects of Peer Feedback on Contribution: A Field Experiment in

Wikipedia." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 2253–2262.

Fostering Voluntary Contributions to a Public Good
A Large-Scale Natural Field Experiment at Wikipedia

ONLINE APPENDIX

ONLINE APPENDIX TABLE A1 – EXPERIMENTAL IMPLEMENTATION

Cohort	New accounts	1st time-contributors	1st time-contributors of article content	<i>STEP 1: AUTOMATIC FILTER</i>	Remaining	<i>STEP 2: SCREENING w/ RULEBOOK</i>	Remaining for random selection	Accounts deleted after awarding date (treat./control)
1 (Sep12)	7,187	3,202	2,662	-2,177	485	-122	363	8 (1/7)
2 (Oct12)	7,917	3,567	2,931	-2,590	341	-44	297	5 (2/3)
3 (Nov12)	9,088	3,739	3,016	-2,587	429	-39	390	9 (4/5)
4 (Dec12)	7,663	3,178	2,629	-2,194	435	-78	357	6 (4/2)
5 (Jan13)	9,147	4,225	3,525	-3,102	423	-99	324	2 (0/2)
6 (Feb13)	8,460	3,788	3,128	-2,443	685	-170	515	7 (2/5)
7 (Mar13)	8,439	3,638	3,013	-2,382	631	-170	461	5 (2/3)
8 (Apr13)	7,456	3,351	2,729	-2,165	564	-162	402	1 (0/1)
Sum							3,109	43 (15/28)

STEP 3: RANDOM ALLOCATION		
Cohort	Treatment group*	Control group*
1 (Sep12)	149	206
2 (Oct12)	148	144
3 (Nov12)	146	235
4 (Dec12)	146	205
5 (Jan13)	150	172
6 (Feb13)	148	360
7 (Mar13)	148	308
8 (Apr13)	150	251
Sum	1,185	1,881
Total		3,066

* Accounts that have been deleted after awarding date are no longer included (figures as of June 29, 2013).

ONLINE APPENDIX B — THE AWARD

The design of the award page was modeled on that of the Swiss national portal, the banner of which is prominently displayed on top of the website.¹ The award and a link to its page also figure on the front of the Portal Switzerland page. The text on the award page briefly describes the award and the idea behind it, of course without giving any details on performance criteria other than the condition that the editor has made the first contribution to the German Wikipedia in the previous month (whereby "contribution" is not defined). The aim of the award is declared, i.e. "to honor new users [editors] and their precious contributions to the German Wikipedia. They deserve our thanks and recognition." The names of some of the team members are provided for questions, suggestions or criticism. The page then displays the month's lists of recipients of the 1, 2 and 3 Star awards, as well as a link to the lists of former recipients (similar to a "hall of fame"). On the bottom of the page, a small template (called "Babel") that was created by a recipient of the first wave is offered for those interested to copy and put it onto their personal user page.²

The award that is put on recipients' talk pages resembles a medal and displays a Swiss national symbol (an edelweiss flower) and a golden star, adorned by a laurel wreath (see Appendix B1 below). It reads: "We hereby present [user name] with the award Edelweiss with Star of the Portal Switzerland for contributions to the German language Wikipedia. [With smaller font size:] Signed by The Project Edelweiss Award." The graphic is accompanied by the following text, under the heading "An Edelweiss for You": "Hello [user name], from over 4000 new authors in the month [previous month] you belong to those who have enriched Wikipedia with their work. As

¹ See <http://de.wikipedia.org/wiki/Portal:Schweiz/Edelweiss-Auszeichnung>.

² Editors frequently display such templates on their user pages to provide information on their language skills, for instance. See <https://en.wikipedia.org/wiki/Wikipedia:Babel> (accessed June 27, 2013).

a small thank-you for your contributions to the German language Wikipedia, we hereby present you with an Edelweiss from Switzerland. With best regards -- The Project Edelweiss Award in the Portal Switzerland, [date]".

Award recipients usually post their verbal reactions underneath the award on their talk page or on the project's talk page.³ Some put a babel template (Appendix B2) created by one of the award recipients on their user page, others also send an email to the official project account, for the purpose of which I had created an email account named after the award. A couple of recipients moreover write to third parties (e.g., a mentor) to show their gratitude for the help they have received.

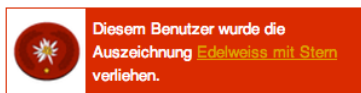
Appendix B1. Award template inserted on user talk pages

Ein Edelweiss für Dich

Hallo XYZ, von über 4000 neuen Autoren im Monat Juni gehörst du zu denen, die die Wikipedia durch ihre Mitarbeit bereits bereichert haben. Als kleines Dankeschön für deine Beiträge zur deutschsprachigen Wikipedia überreichen wir dir hiermit ein Edelweiss aus der Schweiz. Mit besten Grüßen – Das Projekt Edelweiss-Auszeichnung im Portal Schweiz, 12. Sep. 2012



Appendix B2. Babel template



³ An exemplary reaction would be, for instance, "Hello Portal Switzerland, a heartfelt Merci vielmal [Swiss German for "many thanks"] for the award, about which I was tremendously happy! I will do my best to live up to the honor. Best regards, ..." (author's translation).

ONLINE APPENDIX C

TABLE C1 – RANDOMIZATION CHECK

Ordinal variable: Level of activity.

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

award	obs	rank sum	expected
0	1881	2916095	2884513.5
1	1185	1785616	1817197.5
combined	3066	4701711	4701711

unadjusted variance 5.697e+08

adjustment for ties -1.484e+08

adjusted variance 4.213e+08

Ho: n~ea_ord(award==0) = n~ea_ord(award==1)

z = 1.539

Prob > |z| = 0.1239

TABLE C2 – EFFECT ON LEVEL OF ACTIVITY

1st month after awarding date.

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

award	obs	rank sum	expected
0	1881	2820362	2884513.5
1	1185	1881349	1817197.5
combined	3066	4701711	4701711

unadjusted variance 5.697e+08

adjustment for ties -1.657e+08

adjusted variance 4.040e+08

Ho: n~ta_ord(award==0) = n~ta_ord(award==1)

z = -3.192

Prob > |z| = 0.0014

TABLE C3 — ROBUSTNESS CHECKS

	Treatment (1)	Control (2)	Difference (3)
<i>A. Retention</i>			
active, excl. award pages (1st month)	0.43	0.35	0.09*** (0.000)
active, excl. award & own pages (1st month)	0.42	0.34	0.07*** (0.000)
N	1185	1881	3066

Notes: Average values rounded to 2 decimal places. *p*-values from Chi-square tests in parentheses.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.