Discussion Paper No. 15-078

Economic Downturn and Volunteering: Do Economic Crises Affect Content Generation on Wikipedia?

Michael Kummer, Olga Slivko, and Michael Zhang

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 15-078

Economic Downturn and Volunteering: Do Economic Crises Affect Content Generation on Wikipedia?

Michael Kummer, Olga Slivko, and Michael Zhang

First version: November 2015
Second version: February 2016
This version: September 2017

Economic Downturn and Volunteering: Do Economic Crises Affect Content Generation on Wikipedia?

Michael Kummer*, Olga Slivko†, Michael Zhang‡

First version: November 2015
Second version: February 2016
This version: September 2017

September 19, 2017

Abstract

Economic restructuring provides an opportunity for socially valuable product and process innovation. However, it's not clear whether displaced workers engage in *socially useful* activities as well. This question is particularly pressing when the advancement of computer technology creates the destructive force to replace many jobs with robotic labor and artificial intelligence. Leveraging German district-level and European country-level data we analyze the relationship between the economic crisis in 2008-2010 and a *socially useful online* activity – contributions to the largest online encyclopedia Wikipedia. We find increased contributions to the socially valuable activity in the form of knowledge acquisition and contributions to Wikipedia. We analyze both content generation at the European Country level and at the German district level. For German districts, we observe an increase in the rate of content generation on Wikipedia more severely affected districts. These effects get even stronger in our analysis at the European country level.

Keywords: online platform, Wikipedia, public goods, unemployment, user generated content.

JEL Classification Numbers: D29, D80, H41, J60, L17.

*Address: 221, Bobby Dodd Way, #222, Atlanta, GA, 30308, U.S.A, telephone: +1 678 559 5018, e-mail: *michael.kummer@econ.gatech.edu*

†Address: L7, 1 68161 Mannheim, telephone: +49(0)621123-358, e-mail: *slivko@zew.de*

‡Address: Room 918, CYT Building CUHK Business School No. 12 Chak Cheung Street Shatin, N.T. Hong Kong, telephone: +852 6081-2330, e-mail: *zhang@ust.hk*.

1 Introduction

Economic restructuring provides opportunities for product and process innovation. New production units and job functions replace old ones and the society moves forward in this incessant cycle. Schumpeter (1942) considers economic restructuring as "the essential fact about capitalism," and coined the term "creative destruction." While the cycles of destruction and creation have been considered a necessary driving force for the society as a whole, they can severely affect individuals' lives in the form of job displacement, which forces those who lost their jobs to reorganize their lives or at least their daily routines as they are faced with the threat of social decline.

But how exactly do individuals reallocate their time and, do they engage in more socially valuable or wasteful activities when they are confronted with increased levels of job displacement? We examine this question through the lens of change in the patterns of digital public goods provision during a period of significant social and economic changes that followed from the 2008-2010 economic crisis in Europe.

This question is particularly pressing when economists have just started to worry about massive job displacements in the future as a result of rapid advances in computer technology, which range from natural language processing and robotics, to automated trading of financial assets. In this new wave of economic restructuring, very important changes in the labor market are expected. When the adoption of powerful new technologies permeates more and more traditional industries, many jobs are likely to be replaced by machines (Brynjolfsson and McAfee (2014); Ford (2015)). While a possible and pessimistic outcome could be stagnation of median income and the growth of inequality, a more optimistic view is that individuals will be freed from simple and repetitive tasks and will be able to focus on more demanding and socially valuable activities. As much as this debate is important and interesting, a meaningful prediction is challenging: It is hard to associate job displacement with individuals' activities.

Job displacement during the economic crisis around 2008-2010 offers a glimpse into the potential outcome of this trend. One may certainly argue that the technology-induced job displacement, or "technological unemployment" is different from the job losses during economic crisis. However, there are several reasons why our results from the recent economic crisis can shed light on the effects of technological unemployment. First, in both cases, job displacement is fundamentally a result of broader economic restructuring. In the "old" economy, jobs are also destroyed and created by technological advancements. In his book *Politics*, Aristotle writes: "the servant is himself an instrument which takes precedence of all other instruments. For if every instrument could accomplish its own work, obeying or anticipating the will of others, [. . . , if] the shuttle would weave and the plectrum touch the lyre without a hand to guide them, chief workmen would not want servants, nor masters slaves." In other words, workers are no longer needed when machines are sophisticated enough (Campa, 2014). Later, 19-century Luddites fought against job displacement by destroying manufacturing machines. Each economic restructuring unavoidably brings the need for workers with old skills to be replaced. Second, in both cases, relatively lower skilled workers are replaced first. In an economic crisis, they lose to more efficient workers, and in technological unemployment, they are the first to be replaced by the machines. Also higher skilled workers are not safe in both cases. According to Farber (2015), 3.2% of long-tenured workers displaced between January 2001 and December 2003 could be considered to be high-skilled, working in managerial, professional, or related positions.[1] In the near future, highly paid and high-skilled workers in such industries as financial services, law, medicine and education are also threatened.[2] Third, similar in both cases, job displacement does not only increase unemployment, but those still in the labor force are likely to be influenced too. Shifting the focus from just unemployed workers to the general work force, we examine the overall allocation of activities of the population. Finally, since the economic crisis of 2008-2010 was largely unexpected, it works as a natural experiment: participants of the economy took various actions to react to the increased level of job displacement and we are able to use a difference-in-differences (DID) framework to tease out the resulting effects. While other macro- and micro-economic parameters such as outsourcing or deregulation can also lead to continuous and gradual job displacement, it is more plausible to attribute observed short-term changes in online activities to job losses in the aftermath of the financial and economic crisis.

To analyze the effect of increased unemployment on 'digital public goods provision,' we assembled a data set which covers the 402 German districts (in what follows also "Kreise" or "Landkreise"). We leverage the differential intensity with which the economic crisis in 2009 affected different districts and identify how individuals in each district changed their online activity on Wikipedia. We thus observe how much time they allocate to a potentially socially useful online activity. Thanks to Wikipedia's policy to report the IP addresses of anonymous contributions, we are able to define the location of anonymous contributors, while for registered contributors we can observe only self-declared location on their profile pages. We also analyze whether contributions are made during work or leisure hours. Finally, we extend our analysis to a data set on 20 European countries and their Wikipedia language editions, thus testing the robustness of our analysis and generalizing our results.

Contributions to Wikipedia can serve as a good proxy for socially valuable activities. The effect of an economic crisis on digital public goods provision is not straightforward ex ante. On the one hand, the observed shift in time allocation towards more computer use and increased civic engagement might lead to *increased* provision of public goods, thus more contributions to Wikipedia. Previous contributors might be able to allocate more time when displaced from their jobs and contribute their time to the public information good. People who were not aware of Wikipedia might begin searching for information on the Internet and discover Wikipedia. Consequently, they might become interested in volunteering. Even those individuals who still hold their jobs might be doing more online search for useful information

[1] cf. http://www.nber.org/papers/w21216.pdf on page 11.

[2] for examples, see McKinsey Quarterly report http://www.mckinsey.com/business-functions/business-technology/our-insights/where-machines-could-replace-humans-and-where-they-cant-yet

and end up contributing to Wikipedia. On the other hand, the crisis may lead to *reduced* contributions to Wikipedia, because contributing to public goods is clearly not an obvious reaction when people's jobs are threatened. Employed and unemployed individuals may feel threatened by social decline and might find it difficult to contribute to Wikipedia during the period of large scale job displacement as the opportunity cost of their time is higher. Our contribution is to shed light on these questions by empirically analyzing how job displacement affects online knowledge generation, and, specifically, contributions to Wikipedia.

We find that job displacement in German districts leads to higher participation of volunteers and increased content generation on Wikipedia. The number of anonymous edits to Wikipedia articles increases, and there is slightly weaker evidence for an increased overall content growth. We replicate our analysis with an EU country-level dataset. The results for European countries are consistent with those obtained from the district-level in Germany. We find even stronger effects on content generation due to the higher variation in unemployment in Europe during the crisis.

The remainder of the paper is structured as follows. The next section reviews the literature. Section 2 describes the dataset and Section 3 discusses the empirical approach. Sections 4 and 5 report results on contributions to Wikipedia for Germany and at the country level, respectively. Section 6 discusses and contextualizes the findings and limitations of our study, and Section 7 concludes.

1.1 Related Research

The traditional literature on the private provision of public goods suggests that people contribute because they are altruistic. But Olson (1965) suggests: "people are sometimes motivated by a desire to win prestige, respect, friendship, and other social and psychological objectives." The previous theoretical and empirical studies analyzed the private incentives for voluntary public goods provision from the perspective of the interplay between free-riding incentives and social effects. This literature established that social pressure, guilt, or sympathy, may play important roles in the decisions (Andreoni, 1988, 1989, 1990, 2007). For example, in Andreoni (2007) the provision of public goods is shown to be congestible. That is, an increase in the number of recipients increases the total public goods provision but at a decreasing rate. This finding received empirical support in the context of online public goods, such as open-source software and online peer productive communities (Algan et al., 2013; Comino et al., 2007; Kandel and Lazear, 1992). Comino et al. (2007) find that the size of the "community of developers" in open-source projects increases the chances of progress but this effect decreases as the community gets larger. Zhang and Zhu (2011) show the importance of the recipient group size for individual incentives for knowledge provision using exogenous variation, a block by the Chinese government, in the recipient group size on Wikipedia.[3]

We contribute to the literature on public goods provision by highlighting an additional channel

[3] In addition, since the late 1980s, researchers have increasingly contrasted theoretical models with experimental studies in the lab. The main insights of this extensive literature have been surveyed by Vesterlund (2006).

through which private motivations can foster contributions to an online public good. We show that individuals change their time allocation to digital public goods due to the economic downturn. Specifically, we use the recent economic crisis in Europe in the aftermath of the US-american financial crisis in 2008 as an exogenous shock to the time spent online. As unemployment rises during the economic crisis, people who become unemployed have to reduce their working time. When they increase their allocation of time to online activities, additional learning activities on Wikipedia lead to more contributions to this public information good. The resulting "doing by learning" during the economic crisis highlights a new driving force of digital public goods provision. Note that a similar mechanism could be at play on other digital content platforms such as github, stackoverflow, or in android mobile applications.

Previous studies have looked at how the unemployed allocate their time, considering a range of potentially beneficial and wasteful time uses (Aguiar and Hurst, 2007; Aguiar et al., 2012, 2013; Knabe et al., 2010; Krueger and Mueller, 2012). Although unemployed people have more time to be spent on leisure, they are less satisfied with life and specific activities (Knabe et al., 2010). This finding might stem from the threat of permanent loss of skills and subsequent social decline that the recently unemployed are faced with as time progresses (Pissarides, 1992). Krueger and Mueller (2012) find that previously unemployed sharply decrease time devoted to leisure activities at the time of reemployment (by 35 per cent of the time now allocated to working). In their paper, leisure includes computer and Internet use. Aguiar et al. (2012) use the American Time Use Survey (ATUS) to analyze trends in time allocation. They state that since the 1960s, individuals spend more time on leisure. This category includes personal use of computer by definition as well as other activities such as watching television or engaging in sports. By analyzing time diary data from four different countries, Burda and Hamermesh (2010) conclude that only a small share of the additional time is used for home production, also indicating that unemployed spend more time on other activities such as computer use. The ATUS analysis of Aguiar et al. (2013) focuses on the period of the global recession in the late 2000s, and confirms earlier findings. They find that more than 50 per cent of the additional time is spent on leisure activities, yet two-thirds are absorbed by watching TV and sleeping. More interestingly, roughly two percent of the foregone market hours are allocated to civic and religious engagement.

Our paper adds to this discussion by distinguishing between useful and wasteful Internet uses and focusing specifically on time spent for providing online contributions to the largest online encyclopedia, Wikipedia. This encyclopedia is produced collaboratively and is accessible to anyone with an Internet access. Wikipedia can thus be regarded as a digital public good by definition, since it is non-excludable and non-rival (Hess and Ostrom, 2003). Wikipedia is becoming a standard source that is referenced by many online search engines, and the popularity of Wikipedia (6th most visited website) is a clear indication that lots of individuals worldwide are among its readers and even contributors.

Online leisure time was shown to be a substitute for work since most of the time spent online is spent on social networks, online games, email and portals (Wallsten, 2013). Moreover, young people spend more time online. These findings are complemented by Goldfarb and Prince (2008) who show that, conditional on having Internet access, poorer people spend more time online as they have a lower opportunity cost of time than wealthier people. At times of economic crisis both these groups, younger and poorer people, can be threatened by increased unemployment rates or decreased salaries. Taken together, the existing research leads us to expect that people, who experience a sudden increase in available time, will reallocate parts of it towards online activities.

An economic downturn may also lead to a decrease in online content generation. For civic public goods, unemployment has been shown to be negatively correlated with both religious as well as secular volunteering (Freeman, 1997; Uslaner, 2002). However, Uslaner (2002) uses cross-sectional data from the U.S. and Canada and thus gives no information about effects of rising unemployment over time. Freeman (1997) also finds that volunteers predominantly have access to "higher potential earnings or greater demands on their time: the employed, married persons, those with larger families, persons in the 35-54 peak earnings ages, the more highly educated, professionals and managers." Moreover, among men, working more hours is even positively correlated with participation in volunteering. This is also in line with Taniguchi (2006), who studies the effect of gender differences and employment on volunteering using the National Survey of Midlife Development in the United States (MIDUS) 1995-1996. His results suggest that unemployment has a negative effect on men's volunteering, which is not the case for women. Moreover, working part-time and working full time makes no difference in men's efforts in volunteering.

We add a new lens to this literature by focusing on online volunteering and by specifically highlighting the role of general and youth unemployment for online contributions. In addition, we analyze observational data on voluntary contributions to a digital public good, and studying both the local level in Germany and aggregate contributions on the European level. We thus add a valuable type of data that complements existing findings based on surveys.

To summarize our contributions, our study (i) contributes to the literature on public goods by highlighting a new "doing by learning" channel that fosters contributions to an important digital public good. (ii) These contributions highlight a socially valuable online activity, that results from higher unemployment. (iii) Our observations are based on observational data from German Wikipedia, and (iv) our findings highlight the role of youth unemployment in this context of online volunteering.

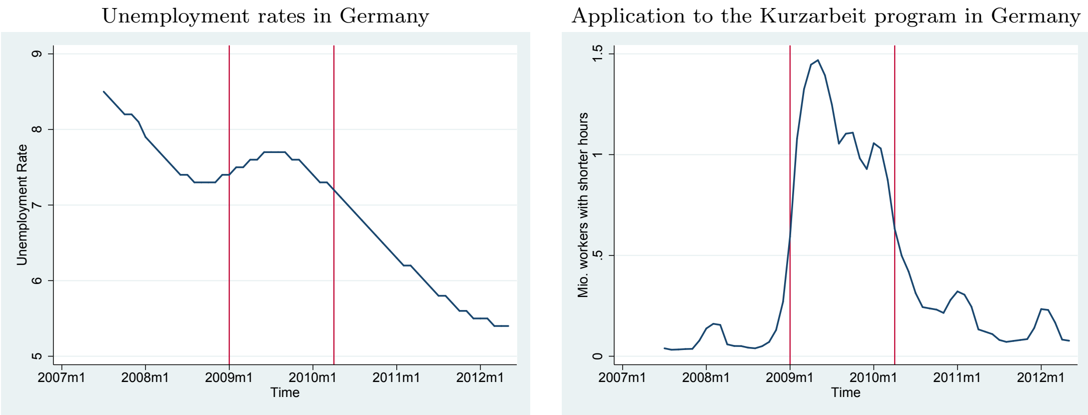
2 Data

Our main analysis of the relationship between unemployment and contributions of content is based on German data that we collected at the district (NUTS 3) level. Despite the fact that the German economy was relatively robust to the economic crisis, there was considerable variation in how different districts were affected by the rise in unemployment as well as by the applications of the "Kurzarbeit" (temporary reduced working hours) program. For our analysis, we combine economic indicators of unemployment and reduced working hours at the district level with data on online activity and contributions to German Wikipedia in the districts.

2.1 Unemployment and Reduced Working Hours in Germany in 2009-2010

In January 2009 the German government announced the necessity to combat the crisis. The German unemployment rate started rising in the months around the beginning of 2009, and, in addition, many companies applied the extended the Kurzarbeit program. As a result the government proposed to address the crisis by massively expanding the pre-existing Kurzarbeit program. According to the rules of the program, employers experiencing a negative demand shock could activate reducing working hours to their employees. They would keep paying to the employees according to hours worked and the government would compensate to the workers about 60% of the foregone income. In January 2009, this program was extended from 6-12 to 18, and later 24 months, and its scope was broadened to cover a much larger number of industries.[4] Overall, the seasonally adjusted unemployment rate rose by 1 percentage point and 300,000 people participated in Kurzarbeit.[5] We thus define January 2009 as the onset of the great economic crisis and the moment when the crisis becomes significant for the German economy.

Figure 1: Unemployment and reduced working hours in Germany in 2009-2010



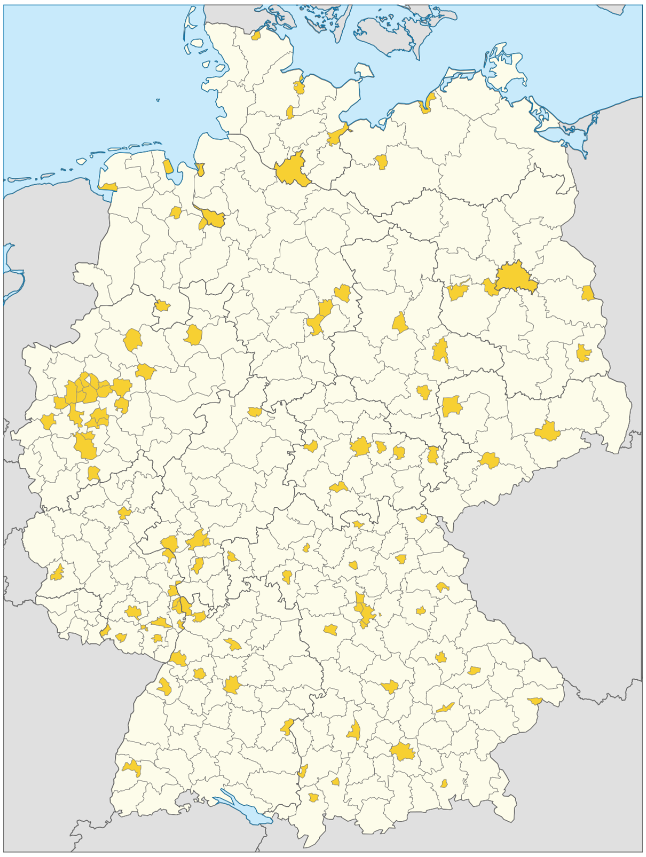
NOTES: Combined, the trends show how the rise in unemployment between 2009 and 2010 in Germany corresponded to a massive application of the Kurzarbeit program in the same period. Source: *Bundesagentur fuer Arbeit.*

We obtained monthly data on the number of unemployed, unemployment rates and participants in the Kurzarbeit programs at the district-level, and generated a district-level dataset. Across all 16

[4]The Kurzarbeit program existed even before the official announcement of the financial crisis. According to general terms, the period of application is 6 months. However, under exceptional economic situations the program can be extended. As a measure of combating the economic crisis, German government varied this extension period. Thus, from January to June 2007 employers could use this program to retain their important employees for up to 15 month in case the company faces a temporal reduction in the demand. Then, this period was reduced to 12 months, but as Germany officially entered into recession, in January 2009, this period was extended to 18 months. Six months later, this period was extended to 24 months for those employees who joined the program during the first six months of the recession. As a result, the workers of industries experiencing the negative shock could apply reduced working hours up to two years starting from January 2009.

[5]http://www.handelsblatt.com/politik/deutschland/2009-kurzarbeit-rettete-mehr-als-300000-arbeitsplaetze/3336760.html

Figure 2: German administrative units on the district level



NOTES: The figure shows the German districts. Darker districts are densely populated districts in cities.

German federal states, we observe 402 districts, which are shown in Figure 2. Our main estimation data is based on the 402 German Kreise, which are similar to mid-sized U.S. counties.[6] Table 1 summarizes the monthly panel data on the Kreis-level. We observe 402 Kreise over the 6 months before and after the shock. The average unemployment rate is 7.5%, and unemployment increased on average by 1 percent. The smallest change in unemployment was a decrease by 1.1% and the maximum increase was 3.4%. In the same table we also show our variables that allow us to analyze content generation on Wikipedia.[7]

Whether the districts are affected or unaffected by the economic crisis is defined based on changes in their unemployment rate after the crisis. To have sufficient variation between affected and unaffected districts, we rank the districts in terms of the change of unemployment rate and define the top 30% as affected by the crisis. The 30% of districts with the lowest, sometimes even negative, changes are defined as unaffected and used as the control group for our estimation. Note that the magnitude of the treatment was considerable, because unemployment increased by approx. 1.1 percentage points more *on*

[6] Georgia with roughly 10 million residents has more than 150 counties, while California with more than 40 million residents has only 58 counties. Germany with 80 million citizens has 402 Kreise so the average Kreis counts approximately 200,000 residents.

[7] Specifically we observe the number of edits and the amount of content (in kilobytes) and we distinguish edits from registered users and anonymous IP addresses.

Table 1: Summary statistics: Main Variables in the Regression Dataset

	mean	sd	min	p10	p50	p90	max
GDP/Capita (1000 EUR)	29	12	13	18	25	44	94
Inhabitants (1000)	200	232	0	70	144	349	3443
Total Hours Worked (million)	142	185	26	50	96	247	2397
# Unemployed (1000)	8.3	14	.71	1.9	5.3	16	245
Unemp. rate,%	7.5	3.7	1.4	3.5	6.7	13	20
Youth Unemp. rate (age$\leq$ 25),%	7.3	3.7	.67	3.1	6.6	13	22
# Businesses using Kurzarbeit.	78	106	1	5	45	183	1371
# Part time employees	1761	2862	1	34	693	4658	25431
Δ Unemp. rate	1.1	.68	-1.1	.24	1	1.9	3.4
Dummy: Affected	.5	.5	0	0	.5	1	1
# Registered Users	3.7	13	0	0	1	8	219
# Revisions to delete	.96	2.4	0	0	0	3	53
# Edits from district	464	1127	1	38	166	1038	20206
# Anonymous edits	210	457	1	34	119	378	8519
# Registered edits	254	808	0	0	4	631	11989
Total added content (KB)	59	224	-1569	-32	9.9	170	4757
Total anononymous content (KB)	4.3	99	-799	-47	3.7	41	3104
Total registered content (KB)	54	193	-1555	0	.066	140	3106
# Local Edits (about district)	631	1082	0	97	273	1429	13623
# Anonymous local edits	109	196	0	14	49	240	2274
# Registered local edits	523	900	0	76	223	1197	11355
Total added local content (KB)	9.7	20	0	.83	3.4	22	269
Total anon. local content (KB)	1.7	3.1	0	.13	.64	3.8	32
Total regist. local content (KB)	8.1	17	0	.64	2.7	18	266

NOTES: The table shows the distribution of the main variables. The unit of observations is Kreis i in month t. The time variable is normalized and runs from -6 to 6. The upper panel shows the districts' economic outcomes, and the panel in the middle shows content generation from the districts. No. of Obs = 5226; No. of Kreise = 402; No. of States = 16. The lowest panel shows the distribution of the variables that measure local content generation (*about* districts; No. of Obs = 3120)

Table 2: Unemployment Indicators in 16 German States

	# Unemployed (1000)	Unemp. Rate,%	Youth Unemp. Rate ,%	Δ Unempl. Rate,%	Share of affected districts
Baden-Wuerttemberg	254.1	4.69	4.13	1.15	0.62
Bavaria	291.2	4.46	4.08	1.18	0.73
Berlin	232.6	13.80	15.36	1.03	0.00
Brandenburg	169.6	13.21	13.27	0.87	0.55
Bremen	37.1	13.23	11.83	0.61	0.50
Hamburg	75.2	8.33	7.72	0.79	0.00
Hessen	209.2	6.71	6.78	0.78	0.18
Lower Saxony	304.3	7.85	7.91	0.68	0.11
Mecklenburg-Western Pomerania	120.1	13.74	12.41	1.05	0.83
North Rhine-Westphalia	774.7	8.47	8.11	0.90	0.26
Rhineland-Palatinate	120.4	6.18	6.71	1.01	0.42
Saarland	37.5	6.62	6.47	1.01	0.00
Saxony	277.2	12.80	12.59	1.61	0.82
Saxony-Anhalt	169.7	13.72	13.32	1.40	0.80
Schleswig-Holstein	108.4	8.10	8.62	0.72	0.10
Thuringia	135.3	11.35	10.53	1.76	0.76

This table shows mean values of unemployment indicators, the number of unemployed and the rates, as well as difference in the unemployment rate before and after the shock and the share of districts affected by the shock for each German state.

average in the affected districts.[8]

In Table 2 we aggregate the unemployment rates to the state level to show how unemployment varied across the 16 states. The table also shows the shares of affected districts per German State. Our definition of crisis based on change in unemployment rate implicitly controls for the baseline of economic status of the states. As a result, the highest shares of affected districts can be observed in traditionally economically strong industrial German states, such as Bavaria or Baden-Wuerttemberg. Weaker states like Thuringia also had a large share of affected districts.

2.2 Contributions to Wikipedia

Anonymous Contributions to German Wikipedia at the District Level: Based on a large data set that contains the revision history of all articles of German Wikipedia, we aggregate individual monthly contributions and compute total contributions by districts. For this aggregation, we map the IP-addresses associated with edits to the corresponding German districts.[9]

In terms of overall editing activities on German Wikipedia, anonymous edits represent about 16% of all edits during the period of our analysis (2008-2010). While we do not suppose that anonymous edits are representative for all editing activities, we deem it highly relevant for our research question, because anonymous edits are typically made by very occasional or unexperienced editors. Thus, our measures of anonymous contributions to Wikipedia at the district level account either for contributions by newcomers, or for occasional and relatively small contributions in terms of content generated.

Registered Contributions at the German District Level: In addition to anonymous editing activities, we collect information on the location of registered Wikipedia contributors, whenever they reveal it publicly on their user talk/profile pages. We thus match almost 25% of the registered edits by users who edited Wikipedia under their username, to a district of origin. Given that these registered edits were made by editors with a well developed user talk/profile page, we consider them to be representative of edits by very active Wikipedia users, thus covering the other side of the spectrum.

The center panel of Table 1 shows both anonymous and registered monthly edits from a district to German Wikipedia, together with the number of registered users, and the number of reverted edits. Since both types of edits cover approximately 20% of all contributions, the number of registered edits that we could match only slightly exceeds the number of anonymous edits. However, because registered users are very active, they generate a much larger amount of content than anonymous users do. We also show the sum of all edits that we could match to each district via one of the two approaches. Together, we could thus associate about 35% of all activities on German Wikipedia to a district of origin.

[8] With a std. error of less than 0.05, results not shown, but available upon request.

[9] Contributions associated with IP-addresses are made only by contributors who skipped the log-in procedure, that is, only by "anonymous" contributors.

We verify that the large uncovered percentage of edits does not distort our results, by analyzing edits about districts (see next part). In this analysis of district specific content we can match the edits of all users who edit in categories that can be associated to the local interest of a district.[10]

Content *About* Districts To examine the full activity on Wikipedia at the level of districts, we look deeper into the topical content of contributions, and identify content that is specific to a particular district. To that end we exploit the "category tree," which is Wikipedia's extensive categorization tool. For attributing articles on Wikipedia to German districts we used the fact that contributors assign each article on Wikipedia to different categories. We match the articles to the districts we use Wiki-topics (categories) and Wikipedia's hierarchical category tree. To do so we extract the ID of all pages that belong to a district's local interest category and recompute all our measures of monthly district-specific activities on Wikipedia only based on the local interest pages. We are thus able to track contributions which refer to our specific districts by the category of each article.

Table 3 shows an example of the categories which contain the name of the district 'Verden' (Lower Saxony), in the north-west of Germany. Among others, these categories include the district's geography, famous people, sports, construction churches and water sources. In our data, we find 82 categories per district on average. Once we identified the district-specific categories, we can use the set of Wikipedia articles in these categories to analyze district-specific content generation. This approach allows us to consider all contributions to these district-specific articles, including all contributions by authors who could not be matched to a region.

Table 3: Local categories describing Wikipedia content about German district Verden

3361	Verden	Naturschutzgebiet_im_Landkreis_Verden
3361	Verden	Ehemalige_Gemeinde_(Landkreis_Verden)
3361	Verden	Gewässer_im_Landkreis_Verden
3361	Verden	Unternehmen_(Landkreis_Verden)
3361	Verden	Kirchengebäude_im_Landkreis_Verden
3361	Verden	Verkehr_(Landkreis_Verden)
3361	Verden	Ort_im_Landkreis_Verden
3361	Verden	Geographie_(Landkreis_Verden)
3361	Verden	Person_(Verden_(Aller))
3361	Verden	Bauwerk_im_Landkreis_Verden
3361	Verden	Verden_(Aller)
3361	Verden	Achim_(Landkreis_Verden)
3361	Verden	Benutzer:aus_dem_Landkreis_Verden
3361	Verden	Domherr_(Verden)
3361	Verden	Blender_(Landkreis_Verden)
3361	Verden	Sport_(Landkreis_Verden)
3361	Verden	Bischof_von_Verden
3361	Verden	Landkreis_Verden
3361	Verden	Person_(Landkreis_Verden)
3361	Verden	Fürstbischof_(Verden)
3361	Verden	Landkreis_Verden_nach_Gemeinde
3361	Verden	Benutzer:aus_Verden

NOTES: The table shows shows an example of how we identified subcategories that consisted of articles with content *about* a district (here: the district "Verden").

The resulting measures of contributions of district-specific knowledge are shown in the lower panel of Table 1. Again, we show both anonymous and registered monthly edits *about* a district to German

[10]Moreover, we analyze data from country-level models that also consider all contributions by registered users, and the results are even stronger. We report these results in section 5.

Wikipedia. Since these measures include the edits of all users, the ratio of anonymous to registered edits is about 1:6. However, the registered share of local content is smaller than for our data on all contributions in the center panel of Table 1. The variable 'Total contributions' simply measures the sum of both types of edits about a district.

European Country Level Analysis In Section 5, we present an analysis of data on the level of European countries. The analysis supports the findings based on German districts. Since these data originated from a different data source, we defer the discussion of the European data to that section.

2.3 Descriptive Evidence

Before describing our empirical approach, we present a descriptive visualization of our data on the German districts (Kreise) in Figure 3. The left panel shows anonymous edits *from* the districts, while the right panel shows anonymous edits *about* these districts over time. We show the district-specific anonymous editing behavior six months before and after the onset of the crisis (highlighted by the thin red line at $t = 0$). The two figures in the upper row show the median values of the normalized number of edits, while the lower row shows the absolute difference between affected and unaffected districts.

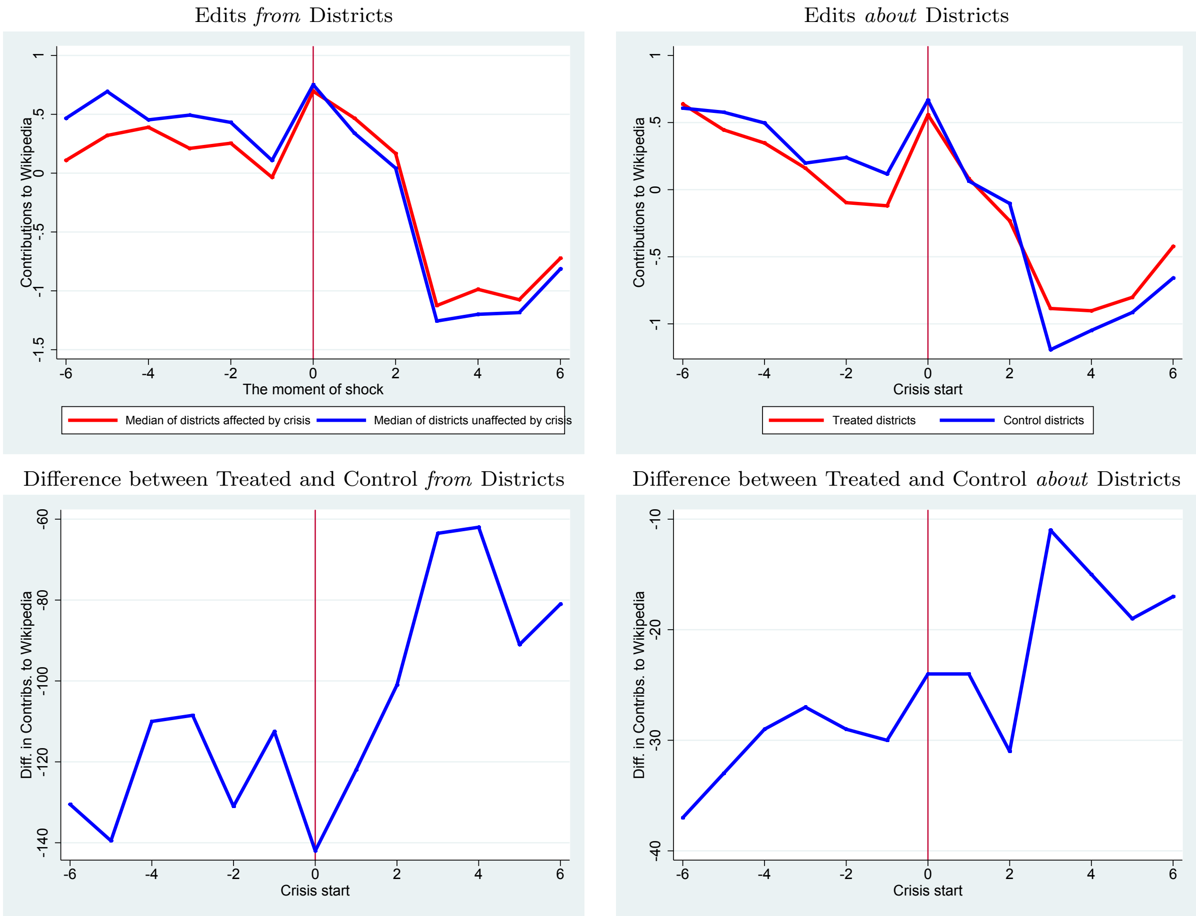
Contributions *from* and *about* affected districts (i.e., treatment group) increased relaive to those from and about unaffected districts (i.e., control group). A comparison of the normalized trends reveals that both anonymous contributions (from districts) and registered contributions (about districts) experienced a relative increase in the districts that were more affected by the crisis.

3 Empirical Approach

We analyze the relationship between the economic crisis and the voluntary contributions of online knowledge to Wikipedia in two frameworks. First, we check whether the variation in the unemployment rates can explain the variation in content generation. We use fixed-effects OLS regressions, and analyze whether unemployment is a key driver of increased online knowledge contributions.

Second, in our main specification, we rely on a difference-in-differences approach. The shock to unemployment is used as a source of exogenous variation to "disposable time" in the economic system of Germany, and we compare content generation in German districts where the crisis was felt stronger with the other districts. In Section 5, we adopt this approach to compare severely affected European countries with countries that experienced only a moderate increase in unemployment. Compared to the country-level analysis, the analysis of German districts in this Section allows us to focus on a framework with many (almost 400) units of observations in a very homogeneous institutional context. The advantage of the country-level analysis lies in the considerably larger variation in European unemployment rates.

Figure 3: Development of Main Outcomes for Content at the District Level



NOTES: The figure shows the median values of the normalized number of edits *from* users in the district (left) and median values of the normalized number of edits *to the district's* "local interest" content (right). The upper panel graphs show the median of the normalized number of edits for the affected and unaffected districts separately. The figure is based on monthly data 6 months before and after the crisis. The lower panel graphs show the absolute difference in the number of edits by the two user groups.

3.1 Fixed-Effects Panel Regressions

We first check the relationship between the variation in unemployment and online knowledge contributions to Wikipedia. For that, we rely on our panel data and use fixed effects OLS-regressions, which analyze edits and size of contributions in bytes by anonymous and registered users. We regress each of these variables on the unemployment rates in the districts. The regression equation is given by:

$$Contributions_{it} = \beta \ Unemployment \ Rate_{it} + \gamma + \mu_i + \nu_t + \epsilon_{it}, \tag{1}$$

where i refers to the districts, t is the time period (month) and γ is a vector of parameters, each corresponding to one of the control variables. Monthly dummies as well as fixed effects are included to rule out time-trend effects and individual unobserved heterogeneities.

3.2 Difference-in-Differences Framework

In our main specification we use a difference-in-differences (DID) approach to the data from German administrative districts (Kreise). The first difference compares content generation on Wikipedia before and after the shock, and the second difference compares content generation in affected German districts to content generation in relatively unaffected districts. This identification strategy allows us to measure the impact of job displacement on contributions to Wikipedia over a given time interval while controlling for all other possible sources of influence. The central assumption we need to make for the DID framework is that the changes in the readership and contribution activities are indeed due to the crisis and not due to some unobservable factors that correlate with the timing of the crisis. Moreover, treated and untreated districts have to share their pre-trend dynamics, which we consider plausible, given the visual evidence.

The unit of observation in our data is a district in Germany with all corresponding statistics (e.g., unemployment rates, etc.) and aggregated contributions to Wikipedia observed every month before and after the official beginning of the economic crisis in Germany, which was announced in January 2009. The estimated equation is as follows:

$$Contributions_{it} = \alpha\ AfterT_t + \beta\ (AfterT_t \times\ Treated_i) + \mu_i + \nu_t +\ \epsilon_{it}, \tag{2}$$

$AfterT_t$ and $Treated_i$ are dummy variables. $Treated_i$ separates districts that were affected by the economic crises from those unaffected. $AfterT_t$ equals one if the time period is after t_0. The coefficient of interest is the coefficient for the interaction term between these two dummies, β, which measures the difference-in-differences after the shock to unemployment. In all regressions, we include time (month) dummies as well as fixed effects of German districts to rule out common time trends and district-specific unobserved heterogeneity.

The contribution of district i in month t are measured in several ways. As we were able to match anonymous contributions by IP-addresses of edits to districts and to match registered contributors who reveal their location to districts we can construct various measures of contributions to Wikipedia overall for all contributors as well as distinguishing between anonymous and registered contributions. For all contributions, as well as by the type of contribution, anonymous or registered, we compute the number of edits and the number of bytes. Furthermore, based on information on registered users we can proxy the arrival of new editors to Wikipedia with the number of new registered editors that appear in a given district every month. Finally, we compute the number of edits made within the districts which represent previous edit reverts. In Wikipedia, a revert means that the edit with which the content was added before got deleted. This allows us to measure how valuable the added content is. Many reverts could indicate that the content of articles experiences vandalism or editing wars so that it is important to control for reverts. All dependent variables used in the regressions are transformed into natural logarithms.

4 Analysis of the Regional Level for Germany

In this section we present the results of our analysis at the district level. The subsequent section (Section 5) repeats the analogous evidence at the European country level.

4.1 Results for German Districts

In this subsection we present the results for all content generated by users within the district. Overall total contributions to German Wikipedia fall after the crisis, but our findings suggest that in districts with higher unemployment the negative overall trend is slower. In the next subsection we present the results for content *about* a district.

4.1.1 Fixed Effects Estimation Results

Our first set of results on the relationship between the unemployment rate and contributions to Wikipedia is based on a linear fixed effects panel regression framework. Table 4 shows the results for regressing several measures of content generation on the unemployment rate and a district fixed effects.[11] The observed outcomes in Columns 1-6 are content by all users (Cols. 1-2), and we also distinguish content from anonymous users (Cols. 3-4) and from registered users (Cols. 5-6), and we differentiate between activity (# Edits, Cols. 1,3 & 5) and the amount of content (in Kbytes, Cols. 2,4 & 6). Beyond content we analyze the number of registered users as a measure of increased user commitment (Col. 7), and the number of reverts (Col. 8) to control for vandalism. All dependent variables are in logs, and we control for monthly dummies to capture country-wide temporal dynamics. The results suggest that an increase in the regional unemployment rate is strongly related to the variation in the anonymous contributions to Wikipedia. A one per cent increase in unemployment rate is associated with a 3 per cent increase in anonymous edits and 10 per cent increase in anonymous bytes added. Based on average editing activity German Wikipedia in 2009 and 2010, that corresponds to almost 5880 additional edits and 401 Kbytes over the 6 month period after the shock we observe.

4.1.2 Difference-in-Differences Analysis

Our main results are based on a difference-in-differences specification for German districts, and they are displayed in Table 5. As before, each column presents estimation results for one of our main measures of editing activity and contributions to Wikipedia: (1) all editing activity (2) the amount of all content (3) the number of anonymous edits, (4) the total anonymous contribution length in Kbytes, the same indicators for registered users in Columns 5 and 6. Column 7 shows the number of registered contributors who reveal their geographical location on user pages, and column 8 the number of reverted edits, signaling

[11] The fixed effect essentially covers all available control variables, since macroeconomic indicators like the population structure or internet penetration do not vary month on month.

Table 4: OLS regression for the relationship between unemployment and activity on Wikipedia (range: 6 months before and after crisis start)

	(1) Log All KB	(2) Log # All Edits	(3) Log Anon.KB	(4) Log # Anon.Edits	(5) Log Reg.KB	(6) Log # Reg.Edits	(7) Log reg.editors	(8) Log reverts
Unemployment, %	0.033	0.018	0.103**	0.027**	-0.003	-0.021	-0.001	-0.012
	(0.041)	(0.013)	(0.042)	(0.011)	(0.027)	(0.030)	(0.007)	(0.014)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3806	5226	3415	5226	5135	5226	5226	5226

NOTES: The table shows the results of a fixed-effects OLS analysis that directly relates Wikipedia contributions to the unemployment rate of German districts around the onset of the European Financial Crisis in 2009. The columns contain different measures of contribution activity to Wikipedia: (1) the number of edits (revisions) by anonymous editors, (2) the total contribution *length of 'anonymous edits'* (in KB) (3) the number of edits by registered editors (4) the length of 'registered edits' (in KB) (5) the number of registered *editors* from the district (6) the number of reverted edits (which is a combined measure of unsuccessful attempts and vandalism). The independent variable of interest is *Unemployment rate* in period t for each district. All specifications include time period (month-year) dummies. Standard errors, clustered by districts, in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5: DID Regression for German Districts (range: 6 months before and after crisis start)

	(1) Log All KB	(2) Log # All Edits	(3) Log Anon.KB	(4) Log # Anon.Edits	(5) Log Reg.KB	(6) Log # Reg.Edits	(7) Log reg.editors	(8) Log reverts
After Treatment	0.036	-0.214***	-0.143	-0.307***	0.012	-0.012	-0.020	-0.130***
	(0.116)	(0.042)	(0.107)	(0.028)	(0.086)	(0.094)	(0.023)	(0.046)
Treated districts after T	0.063	0.006	0.284***	0.061**	0.002	-0.041	0.026	0.042
	(0.102)	(0.041)	(0.096)	(0.027)	(0.090)	(0.102)	(0.020)	(0.034)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2270	3120	2071	3120	3058	3120	3120	3120

NOTES: The table shows the results of our main difference-in-difference analysis that contrasts affected and unaffected German districts around the onset of the European Financial Crisis in 2009. The columns contain different measures of contribution activity to Wikipedia: (1) the number of edits (revisions) by anonymous editors, (2) the total contribution *length of 'anonymous edits'* (in KB) (3) the number of edits by registered editors (4) the length of 'registered edits' (in KB) (5) the number of registered *editors* from the district (6) the number of reverted edits (which is a combined measure of unsuccessful attempts and vandalism). The variable of interest, which captures the treatment effect, *Treated districts after T*(reatment), is an interaction term between dummies for the districts that are affected by the crisis with the time dummy indicating the period after the crisis. All specifications include time period (month-year) dummies. Standard errors, clustered by districts, in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: DID Regression for German Districts (range: 6 months before and after crisis start)

	(1) Log All KB	(2) Log # All Edits	(3) Log Anon.KB	(4) Log # Anon.Edits	(5) Log Reg.KB	(6) Log # Reg.Edits	(7) Log reg.editors	(8) Log reverts
After Treatment	0.136	-0.190***	-0.077	-0.278***	0.032	-0.002	-0.044	-0.138**
	(0.127)	(0.049)	(0.116)	(0.035)	(0.104)	(0.108)	(0.028)	(0.057)
Treated 3-1 p bef. T	-0.012	0.039	-0.220	0.051*	0.020	-0.040	0.005	0.011
	(0.132)	(0.038)	(0.141)	(0.029)	(0.086)	(0.093)	(0.026)	(0.052)
Treated in p0	0.352*	0.065	0.345*	0.092**	0.053	-0.029	0.006	-0.056
	(0.211)	(0.052)	(0.207)	(0.038)	(0.139)	(0.145)	(0.036)	(0.077)
Treated in p1	0.206	0.067	0.297	0.096**	0.026	-0.108	0.019	-0.072
	(0.203)	(0.052)	(0.223)	(0.038)	(0.121)	(0.140)	(0.035)	(0.073)
Treated in p2	-0.108	0.013	0.545**	0.106***	-0.185	-0.125	-0.028	0.044
	(0.242)	(0.054)	(0.242)	(0.038)	(0.132)	(0.148)	(0.034)	(0.070)
Treated in p3	0.078	0.024	0.077	0.098*	0.002	-0.067	-0.007	0.070
	(0.196)	(0.072)	(0.189)	(0.052)	(0.126)	(0.139)	(0.033)	(0.056)
Treated in p4	0.025	0.036	0.161	0.128**	0.014	-0.017	0.076**	0.151***
	(0.169)	(0.067)	(0.168)	(0.050)	(0.139)	(0.145)	(0.035)	(0.056)
Treated in p5	0.071	-0.027	0.176	0.032	0.190	0.016	0.054	0.127*
	(0.180)	(0.070)	(0.162)	(0.050)	(0.158)	(0.137)	(0.034)	(0.069)
Treated in p6	-0.145	-0.003	-0.056	0.054	-0.019	-0.101	0.077**	0.067
	(0.161)	(0.065)	(0.146)	(0.053)	(0.137)	(0.150)	(0.034)	(0.067)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2270	3120	2071	3120	3058	3120	3120	3120

NOTES: The table shows the per-period results of our main difference-in-difference analysis that contrasts affected and unaffected German districts around the onset of the European Financial Crisis in 2009. The columns contain different measures of contribution activity to Wikipedia: (1) the number of edits (revisions) by anonymous editors, (2) the total contribution *length of 'anonymous edits'* (in KB) (3) the number of edits by registered editors (4) the length of 'registered edits' (in KB) (5) the number of registered *editors* from the district (6) the number of reverted edits (which is a combined measure of unsuccessful attempts and vandalism). The variables of interest that decompose the treatment effect are *Treated in p1* to *Treated in p6*. These variables are interactions between a dummy for districts that are affected by the crisis with time dummies that correspond to the respective period after the crisis. All specifications include time period (month-year) dummies. Standard errors, clustered by districts, in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

vandalism or edit wars. The coefficient of interest, which measures the treatment effect, belongs to the cross-term *Treated districts after T*. It is the interaction term between dummies for the districts that are affected by the crisis with the time dummy indicating the period after the crisis.

The results for a 12-month interval (6 months before and 6 months after the shock) suggest several interesting patterns. First we note that there is no increase in overall contributions. Instead, there is an insignificant overall decrease in contributions, and especially anonymous edits in Germany fall significantly after the crisis begins in January 2009 (by 31 percent). However, the negative trend is mitigated by additional activity on Wikipedia in affected districts that experienced an increase in unemployment. Our difference-in-differences estimates suggest that this effect is about 6 percent in edits and 28 percent in Kbytes. Moreover, for registered edits (and overall), registrations and reverts we do not find such a mitigating effect.

The image becomes slightly more nuanced once we decompose the treatment effect after the shock into effects by months after treatment (see Table 6). In this table we can see that the positive effect on the number of anonymous edits is the strongest over the first four months after the shock. As for the size, in the month of the shock the amount of contributed Kbytes rises by about 35 percent, while in the second month by almost 55 percent. However, as the effect on anonymous edits fades in the fourth month after the shock, we observe a significant increase in the number of registered users in German Wikipedia (almost 8 percent in months 4 and 6). Over the same months, more edits (by 13-15 percent) got reverted. Altogether the findings could be suggestive of a pattern by which increased number of new potential editors turn to Wikipedia while searching for knowledge after the shock. As time progresses they make their first edits anonymously without registration and then some of them register in the subsequent months. Moreover, as time elapses and the users get more experience on Wikipedia they could be engaging into cleaning Wikipedia from vandalism.

Overall we find that total contributions to German Wikipedia fall after the crisis, but additional activity by new users mitigates the negative overall trend in districts with increased unemployment.

4.1.3 Robustness check: Controlling for District-Specific Pre-treatment Trends

We perform a check where we control for any district-specific patterns in the contributions to Wikipedia before the the crisis. Such district-specific patterns, if systematic, could interfere with the assumptions of our natural experiment setting. To implement this check we compute each district's individual pre-treatment trend on the dependent variable. We then extrapolate each district's pre-treatment trend into the period after the treatment and include the new variable in the estimation. We show the corresponding results in Table 7. The results on our main dependent variables of interest remain unchanged when controlling for district-specific pre-treatment trends in the contributions to Wikipedia.

4.1.4 Robustness check: Source of Change in Contributions

To shed light on the underlying mechanism that drive the changes in contributions in affected districts we explore time stamps of edits in our sample. Based on the time of contribution, we aggregate edits made during the working time, from Monday to Friday in the interval from 9 a.m. to 6 p.m., and in the remaining leisure hours of the day, including the weekend. Table 13 shows the results of this analysis. The coefficients are slightly larger for both outcomes but only the number of anonymous edits made during leisure time remains significant indicating that the positive effect of unemployment on contributions to Wikipedia is mostly driven by anonymous edits during leisure time. The analogous DiD regressions for contributions during working hours show no significant effect. Based on this evidence, we cannot discern whether the additional content is generated by unemployed users who spent more time online, or by working users who generate the content in their leisure time. Content generation by anonymous users during leisure time would be in line with previous research from labour economics, which has shown that recently unemployed workers shift household production into the working hours and engage in additional activities during leisure time (Aguiar et al., 2013; Burda and Hamermesh, 2010). Alternatively the finding might simply suggest that employed workers generate the edits, so that our finding offers a path for further research.

In our next strategy to attribute contributions to a "Kreis," we focus on Wikipedia pages *about* district-specific knowledge. Examples of such pages would be local points of interest, hiking tracks, rail tracks or regional soccer clubs etc. The results from this approach are presented in the next subsection.

4.2 Results for Local Knowledge Contributions to German Wikipedia

In this subsection we present our results when we focus on edits *about* a district. Specifically we identify pages about institutions, points of interest and local infrastructures such as bicycle paths, which are clearly specific to each district. To identify these pages and match them to the districts we use Wiki-topics (categories) and Wikipedia's hierarchical category tree. Among these categories, we have identified those related to our observed German districts if they mention the names of the German districts in their titles and recompute all variables based on the resulting set of "local interest pages." We then evaluate if we can observe similar patterns of more activities on these pages as the jobmarket worsened at the onset of the crisis, in Jan 2009. As before, we use changes in the unemployment rate to distinguish *affected and unaffected* districts.[12]

The results of this estimation are shown in Table 8. The table shows the difference-in-differences approach from equation 2. The observed outcomes in Columns 1-6 are the same as in the first six columns of the previous tables: Content from all users (Cols. 1-2), content from anonymous users (Cols.

[12] The magnitude of the treatment was considerable, because unemployment increased by approx. 1.1 percentage points more *on average* in the affected districts ($se < 0.05$, results not shown).

Table 7: DID Regression for German Districts (range: 6 months before and after crisis start)

	(1) Log All KB	(2) Log # All Edits	(3) Log Anon.KB	(4) Log # Anon.Edits	(5) Log Reg.KB	(6) Log # Reg.Edits	(7) Log reg.editors	(8) Log reverts
After Treatment	0.095	-0.190***	-0.057	-0.289***	0.066	-0.005	-0.020	-0.141***
	(0.116)	(0.045)	(0.111)	(0.030)	(0.091)	(0.100)	(0.025)	(0.047)
Treated districts after T	0.070	-0.016	0.293***	0.046*	-0.027	-0.034	0.027	0.043
	(0.100)	(0.040)	(0.098)	(0.027)	(0.083)	(0.094)	(0.020)	(0.035)
Pretrend	0.062***	0.178***	0.121***	0.122***	0.263***	0.247***	0.115***	0.090***
	(0.015)	(0.036)	(0.025)	(0.031)	(0.046)	(0.046)	(0.028)	(0.028)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2270	3120	2071	3120	3058	3120	3120	3120

NOTES: The table shows the results of our main difference-in-difference analysis that contrasts affected and unaffected German districts around the onset of the European Financial Crisis in 2009. The columns contain different measures of contribution activity to Wikipedia: (1) the number of edits (revisions) by anonymous editors, (2) the total contribution *length of 'anonymous edits'* (in KB) (3) the number of edits by registered editors (4) the length of 'registered edits' (in KB) (5) the number of registered *editors* from the district (6) the number of reverted edits (which is a combined measure of unsuccessful attempts and vandalism). The variable of interest, which captures the treatment effect, *Treated districts after T*(reatment), is an interaction term between dummies for the districts that are affected by the crisis with the time dummy indicating the period after the crisis. All specifications include time period (month-year) dummies. Standard errors, clustered by districts, in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: DID for local knowledge edits from German Districts (based on : 6 months before/after the crisis start)

	(1) Log All Contribs. (KB)	(2) Log # All Edits	(3) Log Anon.Contribs (KB)	(4) Log # Anonym.Edits	(5) Log Reg. Contribs (KB)	(6) Log # Reg. Edits
After treatment (start)	-0.141***	-0.052	-0.258***	-0.268***	-0.103**	-0.002
	(0.039)	(0.035)	(0.046)	(0.038)	(0.043)	(0.037)
Treated districts after T	0.064*	-0.001	0.100***	0.046*	0.046	-0.014
	(0.037)	(0.030)	(0.036)	(0.028)	(0.040)	(0.033)
Period FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3094	3094	3094	3094	3094	3094

NOTES: The table shows the results of our main difference-in-difference analysis that contrasts affected and unaffected German districts around the onset of the European Financial Crisis in 2009. The columns contain different measures of contribution activity to Wikipedia: (1) the number of edits (revisions) by anonymous editors, (2) the total contribution *length of 'anonymous edits'* (in KB) (3) the number of edits by registered editors (4) the length of 'registered edits' (in KB) (5) the number of registered *editors* from the district (6) the number of reverted edits (which is a combined measure of unsuccessful attempts and vandalism). The variable of interest, which captures the treatment effect, *Treated districts after T*(reatment), is an interaction term between dummies for the districts that are affected by the crisis with the time dummy indicating the period after the crisis. All specifications include time period (month-year) dummies. Standard errors, clustered by districts, in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3-4) and content from registered users (Cols. 5-6). Note however, that we capture the edits of *all* registered users now, since the edits are already matched to each district via the set of (local interest) articles under study. As before, we differentiate between activity (# Edits, Cols. 1, 3 and 5) and the amount of content (in KB, Cols. 2, 4 and 6), and the variable of interest is the cross term.

The findings corroborate the analysis in the previous subsection. A Wikipedia wide decrease in content generation is mitigated in affected districts with a greater increase in unemployment. As in the main specification, the relative increase of content is driven by an increase in anonymous editing, with our DID-estimates suggesting that anonymous edits went up by almost 5 percent and the additional amount of content (in Kbytes) went up by approximately 10 percent. Unlike in the previous approach, this effect translates to a significant overall effect on all edits, and like before, the results on anonymous editing continue to hold when we control for district-specific pre-treatment trends (in Table 14).

5 Analysis for European Countries

To examine the robustness and the potential to generalize our results from the German dataset, we perform an additional analysis for a sample of European countries. We focus on countries that were affected by the economic crises during the period 2008-2009 and compare them to countries relatively less affected by the economic crises or those where successful policies were implemented to prevent the deepening of the crises. This allows us to examine the relationship of interest in a setting where the variation in the intensity of shocks to unemployment is much higher than across German districts. Another advantage of this analysis is that we can consider total edits to Wikipedia and, hence, we have more and better measures of Wikipedia content, such as views, new words, or active users with various frequency of contributions. We can even observe average edits per article, and the average number of hyperlinks set between articles within Wikipedia and to external web-sites. Different from the previous analysis, we can also leverage the well-established differential timing of different countries entering crisis.

Before presenting the emprical specification and the results (in subsections 5.2 and 5.3), we briefly discuss our data which covers twenty European countries and Wikipedia editions in languages spoken predominantly in those countries.

5.1 Country-level Data

To analyze content generation on the European country level, we combine data on European countries' labour markets with aggregate contributions to various versions of Wikipedia of the corresponding countries. Contributions to Wikipedia come from Wikipedia's monthly statistics for different language editions of Wikipedia provided by the Wikimedia Foundation. These statistics include the number of Wikipedians, the number of articles in Wikipedia, database sizes, number of words, and readership statistics for all language versions of Wikipedia. To study the relationship between country level unemployment and Wikipedia, we need to focus on countries that have a unique language. For example, some of the most heavily affected countries, such as the United Kingdom, Spain and Portugal, had to be excluded since their languages are spoken not exclusively in these countries, but all over the world. Therefore, our efforts to measure the effect of unemployment on the activity on Wikipedia in those countries would be distorted by contributions from e.g. Latin America (or the United States/Australia and other countries with many speakers of English).

Table 10 shows the Wikipedia language versions that we could use in this paper. The share of language speakers who live in the corresponding country of origin varies from 71 percent to 99 percent (see column 1). To substitute for the Spanish Wikipedia, we add the Catalan version, which is also actively promoted by the Catalan population. We exclude another Spanish region, the Basque Country, because of the elevated activity of automated scripts, or "bots", in the Basque Wikipedia. According to the Wikimedia Foundation, 75% of all edits and 50% of all new articles in the Basque Wikipedia were created by bots in 2009. Bots are active in other Wikipedia editions as well, but not at such a high level.[13] We also exclude Ireland because people in Ireland mostly speak English and only 45% of the speakers of Irish live within the country. The final sample consists of 20 Wikipedia language editions. In addition to the largest European versions of Wikipedia, we include the small ones such as the one for Iceland, which is a country that was heavily affected by the European economic crisis. Beyond European languages, we also include Japanese version of Wikipedia, because Japan is the only country for which we were able to find monthly total working hours in the economy and monthly unemployment rates.

From Wikipedia's language statistics we retrieve seven relevant indicators of user activity: (1) aggregate views per month, (2) the number of active Wikipedians with a modest number of monthly edits ranging from 5 to 100, (3) the number of active Wikipedians with more than 100 monthly edits, (4) average edits per article, and (5) the content growth of a corresponding language edition of Wikipedia in terms of words, (6) the number of hyperlinks between the articles in Wikipedia, and (7) the number of references from Wikipedia to external websites. Having several measures for contributions allows us to analyze different aspects of the increased contributions to Wikipedia to both the quantity and quality of the content of the online encyclopedia. For example, the growth in the number of words would indicate more content on Wikipedia, while more edits per article could mean that with increased participation, articles on average got more attention for further improvement.

Table 9 gives an overview over the countries in the sample. It also clarifies which countries we consider affected or unaffected by the crisis. We consulted the European Central Bank reports about the crisis 2008-2009 specifically to find information about whether a country was affected by drastically increased unemployment or reductions in hours worked, and also when the crisis started. Countries were considered

[13] `http://stats.wikimedia.org/EN/BotActivityMatrixCreates.htm`

Table 9: Crisis Indicators: unemployment rates (%)

	Affected by crisis	Crisis start	Unemp.rate,%	Change in Unempl.,%
Bulgarian	yes	Oct 2008	6.1	0.6
Catalan	yes	Oct 2008	11.5	7.8
Czech	yes	Oct 2008	5.3	2.7
Danish	no	.	4.4	3.2
Dutch	no	.	3.3	0.6
Finnish	no	.	7.1	2.1
German	no	.	7.7	0.1
Greek	yes	June 2009	9.5	2.7
Hungarian	yes	March 2009	9.3	2.1
Icelandic	yes	Oct 2008	4.6	6.9
Italian	yes	May 2009	7.6	1.2
Japanese	no	.	4.4	1.4
Norwegian	no	.	2.8	0.8
Polish	no	.	7.4	-0.2
Romanian	yes	Oct 2008	6.2	0.8
Russian	yes	Oct 2008	7.0	2.3
Slovakian	yes	Oct 2008	10.5	2.9
Slovene	yes	Oct 2008	5.0	1.1
Swedish	no	.	6.9	1.7
Turkish	yes	Oct 2008	10.8	3.7

NOTES: This Table shows how countries' unemployment rates were affected during the crisis. Affected countries were identified either by a sharp increase in unemployment or a decrease in the hours worked in the economy.

Table 10: Wikipedia key variables within the period of 12 months before and 12 months after the crisis

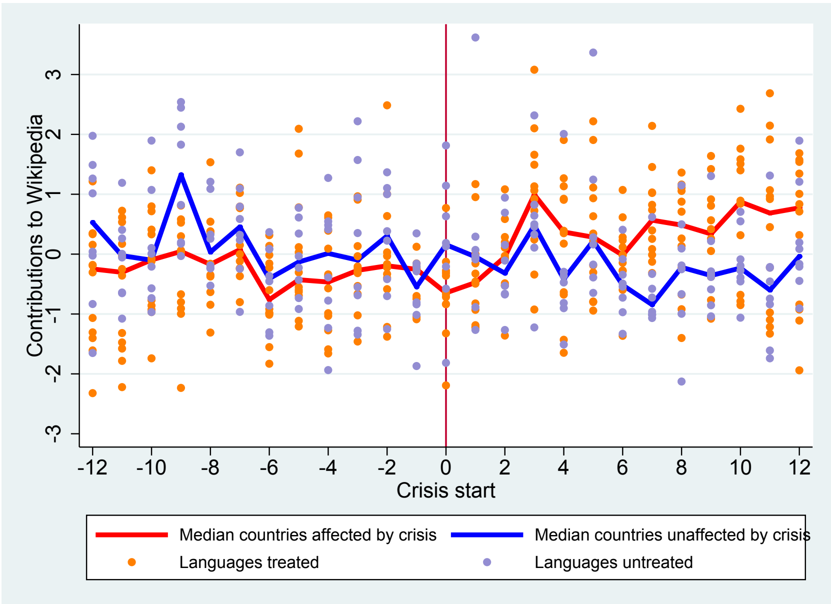
	Language speakers (m)	In main country, %	Views per speaker	Wikipedians, %	Active 5-100 edits, %	Active > 100 edits, %
Bulgarian	8.16	86.05	2	0.02	10.7	3.5
Catalan	4.08	.	3	0.06	13.3	3.9
Czech	10.62	97.93	4	0.04	12.3	2.8
Danish	5.52	97.42	3	0.06	9.9	2.3
Dutch	21.94	71.54	6	0.06	9.6	2.2
Finnish	5.39	94.58	10	0.13	9.7	2.3
German	78.25	89.21	11	0.11	8.5	1.4
Greek	13.43	79.65	1	0.02	6.0	1.8
Hungarian	12.61	78.06	2	0.04	12.8	3.3
Icelandic	0.24	94.32	9	0.16	11.8	5.4
Italian	63.66	90.64	5	0.04	10.2	2.2
Japanese	122.06	99.13	8	0.03	12.7	1.5
Norwegian	4.74	97.85	6	0.13	9.7	2.1
Polish	38.66	94.66	8	0.04	10.8	2.3
Romanian	23.78	83.67	1	0.01	12.1	2.9
Russian	167.33	81.87	1	0.01	17.0	3.4
Slovakian	5.19	91.56	2	0.03	11.5	3.6
Slovene	2.09	91.60	4	0.07	12.6	2.8
Swedish	9.20	96.12	7	0.09	11.0	2.4
Turkish	70.81	93.92	1	0.01	10.6	1.9
Total	33.39	89.99	5	0.06	11.1	2.7

Columns (3)-(6) are means of the interval 12 months before to 12 months after crisis

Sources: stats.wikimedia.org

to be affected by the crisis, if they experienced a significant decrease in hours worked, or an increase in unemployment. For the beginning of the shock we looked at the months of 2008 or 2009 mentioned in the reports and also at the country level statistics on hours worked. In the data we would see when there is a significant decrease in hours worked and we would take the second month of a sustained decrease in hours worked as crisis onset month. If no clear point in time was found from this procedure, we used October 2008 (the first month of Q4.2008), which is generally considered the point in time when most countries were officially in recession. Furthermore, we exclude the month in which we estimate the crisis begin from estimations to make sure our classification procedure does not drive our estimated effects.

Figure 4: Monthly Development of Words contributed



NOTES: The figure shows monthly content growth measured in words added. The median values for affected and unaffected countries across the 20 language versions of Wikipedia in our sample are shown as the two lines. The time spans 12 months before and after the crisis.

Figure 4 gives a descriptive account of one of the key outcomes in Wikipedia: monthly growth measured in words added. We separately show the medians of the affected and unaffected language editions of Wikipedia in our sample 12 months before and after the crisis. We show the medians together with scatter plot of different language versions of Wikipedia. The graph illustrates that, before the crisis, countries that would be affected grew slower than the unaffected countries, whereas after the crisis content growth in the affected countries was faster than in unaffected ones. The patterns are similar for views, edits per article and active Wikipedians, but not for occasional editors. For this variable we see a difference in the trends, that must be accounted for in the regression analysis.

One of the main concerns about the country level data above is the fact that the countries are quite heterogenous both culturally and economically. While we cannot easily deal with this issue at the country level, the figure above with the sharp behavioral change gives some confidence that economic crisis may indeed played a role in changing people's incentives to contribute to Wikipedia.

5.2 Empirical approach

At the country level, similarly to district level in Germany, we estimate the difference-in-differences model. The regression equation is given by:

$$Contributions_{it} = \alpha\ After_t + \beta\ (After_t \times\ Affected_i) + \nu_t + \mu_i +\ \epsilon_{it} \tag{3}$$

The unit of observation is country i (and its Wikipedia) in month t. The dependent variable $Contributions_{it}$ measures contributions to Wikipedia. We take logarithm transformation on all our measures of contributions to Wikipedia.[14] $After_t$ and $Affected_i$ are dummy variables. $Affected_i$ now distinguishes between countries that were affected by the economic crises and unaffected ones. $After_t$ equals one after the month of the shock t_0. As the variable $Affected_i$ does not vary over time, it drops out in the fixed-effects specification. The coefficient of interest is β for the interaction term of these two dummies, which measures the treatment effect of interest. The country fixed effects μ_i, and time fixed effects ν_t are also included in the regressions. Since we have only twenty units of observation on the country level, we use a 24-month interval, which covers 12 months before and 12 months after the onset of the crisis.

5.3 Results

The results of the baseline difference-in-differences estimation are shown in Table 11. Each column shows the results for one of our seven dependent variables measuring activity on Wikipedia: (i) Article Views, (ii) Active editors with 5-100 edits (iii) Active editors with more than 100 edits, (iv) edits per article, (v) growth of the total data in Wikipedia, (vi) new articles per day, (vii) internal links and (vii) external links. The coefficients of interest *Treated countries after T* suggest that after the shock there is a 14 percent increase in the number of active users with few monthly contributions and a 13 percent increase in active users who heavily edit Wikipedia, contributing more than 100 edits per month. Contributions of new words to language editions of Wikipedia grow by 13 percent. Articles got on average 6 percent more edits and 14 percent more links to external sources of information on the web.

In Table 12, we shed first light on the the role of viewership as a mediating factor by analyzing the relationship betweens viewership and content growth. The table shows the results when using a fixed effects panel analysis in which we regress activity on Wikipedia on views over the 24-month period, 12 months before and 12 months after the onset of crisis. Again, in each column of Table 12 we show our

[14] Our findings do not change if we normalize with respect to their mean and standard deviation values such that the coefficients represent the changes in the dependent variables in standard deviations.

Table 11: Country Level: DID Regression for the period of 12 months before and 12 months after the crisis

	(1) Log Views	(2) Log Active 5-100e.	(3) Log Active more 100e.	(4) Log Edits p.article	(5) Log Words growth	(6) Log Int. links	(7) Log Ext. links
After treatment	0.163***	-0.003	-0.009	0.095***	0.039	0.249***	0.309***
	(0.030)	(0.032)	(0.038)	(0.014)	(0.047)	(0.015)	(0.027)
Treated countries after T	0.146**	0.142***	0.127***	0.056**	0.133**	0.057	0.141**
	(0.068)	(0.044)	(0.042)	(0.025)	(0.058)	(0.037)	(0.053)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	432	480	480	480	480	480	480

NOTES: The table shows the effect of increased unemployment on contributions to the corresponding country's Wikipedia in Europe for the following indicators: (1) views of Wikipedia, (2) the number of active Wikipedians (with at least 5 edits), (3) the number of very active Wikipedians (with more than 100 edits), (4) the average number of edits per article, (5) the new words added, (6) hyperlinks to Wikipedia articles, (7) hyperlinks to external sources. All measures of contributions to Wikipedia are in logs, and the month of the estimated crisis onset was omitted from the regressions. The variable of interest, which represents the treatment effect, *Treated countries after T*, is an interaction term between dummies for the countries that are affected by the crisis with the time dummy indicating the period after the crisis. All specifications include time (month - year) fixed effects. Standard errors, clustered by countries, are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 12: Country Level: OLS Regression for the effect of views during the period of 12 months before and 12 months after the crisis

	(1) Log Active 5-100e.	(2) Log Active more 100e.	(3) Log Edits p.article	(4) Log Words growth	(5) Log Int. links	(6) Log Ext. links
Log Views	0.306***	0.184**	0.020	0.203***	0.150***	0.158***
	(0.103)	(0.086)	(0.026)	(0.058)	(0.049)	(0.050)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	432	432	432	432	432	432

NOTES: The table shows the relationship between a country's Wikipedia views and different measures of content contribution to Wikipedia. The independent variable of interest is *Log Views*. In each column we show a different measure of contribution activity: (1) the number of active Wikipedians (with at least 5 edits), (2) the number of very active Wikipedians (with more than 100 edits), (3) the average number of edits per article, (4) the new words added, (5) hyperlinks to Wikipedia articles, (6) links to external sources. All indicators of contributions to Wikipedia are in logs. All specifications include time (month - year) fixed effects, and exclude the estimated period of the crisis onset. Standard errors, clustered by countries in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

six different measures of contributions to Wikipedia.[15]

The results in Table 12 confirm that views are a crucial predictor for edit related outcomes except the number of edits per article. An increase in views by one percent is associated with more active editors (0.31 and 0.18 percent) and more words in Wikipedia articles (0.2 percent). Moreover, views are positively related to our measures of content quality, the number of internal links set between Wikipedia articles and external links to information sources. These findings show the role of views as a key mediating factor for additional content generation.

To illustrate the magnitude of the effects we found, we consider the example of Italy, which was one of the countries that was most severely affected by the rise in unemployment. The average unemployment rate in Italy over the observed period was 7.6 per cent, and it increased by 1.2 percentage points, which is equivalent to approximately 300,000 additional unemployed people.[16] After the shock, the number of editors with few edits grows by 14 per cent, as suggested by our results in Table 11, we would observe 0.14 * 2440 = 342 additional editors.

5.4 Specification Tests

We ran several tests to check the validity of our specification. First, we decompose the interaction term of interest $After_t \times\ Affected_i$ into $\sum I(Year\_Month = t) \times\ Affected_i$, which allows us to run the difference-in-differences analysis period by period. The results in Table 15 show the cross-terms between and indicator for treatment and an indicator that takes the value one if the observation is in a given period. We use monthly intervals for the three months before and 4 months after treatment and 2-month intervals thereafter. The reference period ranges from 4 to 12 months prior to treatment. All coefficients compare the difference of treated and control observations in a given month to the same difference in the reference period. The first three coefficients measure the effect before treatment actually begins, and are expected to be nonsignificant. The subsequent coefficients show how the difference between the treated and the controls develops as the treatment takes effect after the shock.

The results suggest that after the shock in the 12 subsequent months the viewership of Wikipedia in the treated countries increased by 8 - 20 percent faster than in unaffected countries, albeit the effect becomes signifcant with a lag. Similarly, the growth of Wikipedia content, as measured in additional words, grew substantially faster, but clearly stronger only in the later months after the shock, 12 - 22 percent. The number of active users with few monthly edits increased faster in treated countries starting from month 2, and the difference in differences varies approximately 14 - 23 percent for affected authors than those unaffected. For these measures the coefficients referring to the months before the shock indicate that the treated observations were not different from the controls before the shock, suggesting

[15](1) number of active Wikipedians (with at least 5 edits), (2) number of very active Wikipedians (with more than 100 edits), (3) average number of edits per article, (4) new words added, (5) number of internal hyperlinks between articles on Wikipedia, and (6) number of references from Wikipedia articles to an external website.

[16]Based on a workforce of 24.5 millions. See http://data.worldbank.org/indicator/SL.TLF.TOTL.IN?locations=IT

proper identification. For very active authors with more than 100 edits, they contribute more before the onset of the crisis. These results also confirm positive effects for both active users who contribute few and many edits, and also suggest a pattern by which users tried editing when the crisis began, and became active editors a few months later, creating effects that become larger in later months after the shock. In contrast, for edits per article and links to outside references, we find a systematically increasing difference over time, and thus we cannot reject that they are simply on different trends. Moreover, the coefficients for the differences before the crisis suggest that the treated countries' Wikipedia editions have systematically fewer users, especially for active users with few monthly edits, than the control ones.

In addition, a natural concern arises regarding the-crisis development of content in the languages of countries we analyze. To account for that, we include country-specific trends in various measures of Wikipedia content development takes as before the shock (Table 16). Even after accounting for country-specific trends our results hold on the number of new active users with small monthly contributions, for monthly views, and for contributed text ("words growth").

In the next specification test, we verify that the unemployment rate was not positively correlated with contributions already *before* the crisis. This is important, because the crisis is likely going to hit weaker economies harder. Hence, if contributions to Wikipedia were correlated with unemployment before the crisis, then we could not exploit the European economic crisis to study how an increase in unemployment affects contributions to Wikipedia. We would simply capture the preexisting correlation and erroneously attribute it to the crisis.

Hence, in Table 17, we show an OLS fixed effects panel regression of contributions on unemployment 12 months prior to the crisis. The table contains the regression coefficients of the independent variable of interest, *Unemployment*, on our measures of contributions to Wikipedia. All specifications include time dummies. The coefficient of the variable of interest, the unemployment rate, is not significantly different from zero. We consider this to be offering evidence of no correlation between unemployment and contributions to Wikipedia before the shock.

Finally, we check the robustness of our OLS approach by using the figures of unemployment among young people (15-24 years old) as an explanatory variable. One would expect, young people are more likely to use the Internet and, consequently, to contribute to online public goods than the elder generations. The results (see Table 18) suggest that the magnitude and significance of the unemployment effect is larger for the young poplulation.

Our analysis at the country level has several potential drawbacks. Countries are very heterogenous in many aspects which we are not able to capture with our specifications. Therefore, it is hard to argue that European countries can serve as valid counterfactuals for each other. Moreover, all the country level analysis is performed for only 20 units of observation and hence constrains our flexibility in testing analysis on subsets of the data. Both of these concerns are addressed above in the above German district level analysis. However, the country level analysis adds value to our baseline results from German districts, because it allows us to exploit greater variation in the changes in economic hardship and proves that the results we find hold not only for Germany but can be viewed as generally valid across many countries.

6 Discussion, Limitations and Further Research

In this paper, we analyze the relationship of unemployment and public goods provision online during the economic crisis in Europe in 2008-2009. In a nutshell, we find that higher unemployment is associated with higher participation of volunteers in Wikipedia and an increased rate of content generation. Our findings are based on the comparison across German districts and across countries, and we exploit that some districts/countries were affected by relatively large increases in unemployment while others were not. With higher unemployment, articles were read more frequently and the number of highly active users increased, following an initial increase of anonymous and/or casual editors ("beginners"). This pattern suggests that, over time, new users started editing and that existing participating editors also increased their activities. Moreover, we find evidence for increased content growth.

Our main analysis is based on a comparison of German districts and on a country-level analysis of European economies. At the German district level, districts with higher increases in unemployment had relatively more contributions than less affected districts, when controlling for the overall downward trend in total contributions to German Wikipedia during the crisis. At the European level we find faster contribution growth where the crisis hit harder. We would like to stress several aspects of our findings: First, the effects are consistently found for edits *from* German districts (Table 5), for edits *about* German districts (Table 8) and on the European level (Table 11). Second, the pattern of contributions in Germany and on the European level are completely in line, since we find that a downward trend in generally less affected Germany was partially mitigated in districts with higher unemployment. Finally, the effects are sizeable and the estimated increases in editing activities typically range from 5 to 20%. For example, our country-level analysis suggests that the number of casual editors (with 5 to 100 edits/month) grew by 9.5 to 14 per cent, as suggested by our results in Tables 11 and 16. For the Italian Wikipedia, this would imply about 300 additional editors with 5 to 100 edits every month. Thus, the overall effect suggests that the threat of job displacement is associated with increased online public goods contribution. Since Wikipedia functions as an important knowledge base for the economy, our results document a new "doing by learning" process, that created a valuable side effect of the economic crisis. Our findings suggest that a share of the population from regions that faced job displacement would first increase their use of online knowledge repository, Wikipedia, and afterwards begin contributing to the public knowledge good.

Our findings open up a large array of further questions. Potentially, higher unemployment may be associated with greater volunteering activity and productive time usage. Yet we cannot fully answer how this mechanism works. Particularly, it seems that new editors begin to acquire new capabilities and devote their time to contributing to online public goods. As more new articles are created on Wikipedia every day, the increased participation is focused on adding to the existing knowledge as well as providing new topics. However, we found much weaker effects in the number of hyperlinks that are placed between Wikipedia articles and to external sources. When controlling for pretrends, the coefficient for Wikipedia-links even turns negative and the coefficient for external links shows nonsignificant effects. These findings might indicate that the observed increase in content generated is due to the activity of inexperienced contributors or minor edits. Unfortunately, fewer links in the content might also suggest a decrease in content quality, which is then a negative consequence of more activity.[17]

The question whether unemployment can result in increased provision of public (online) goods and private learning is crucial, given that we observe accelerating labor substitution due to digitization. Especially, if a part of the liberated capacity can result in increased knowledge documentation and generation, this may be a positive surprise. We can provide some evidence which suggests that the effect is due to increased unemployment after the crisis (and not before), and it seems that the content generation comes from edits that are made during leisure time.

While we are able to test our hypotheses from several angles and to show the robustness of our findings, several limitations cannot easily be overcome. For example, we exploit the economic crisis as a source of exogenous variation in the economic state and the unemployment rates. This strategy is based on the following identification assumptions: First, contributions to Wikipedia should not be correlated with how likely the countries would enter the crisis. While this is a plausible assumption, we cannot fully relax it with the data we have.

Second, using districts and countries as controls requires that the various Wikipedia editions are sufficiently similar, and that the districts/countries are somewhat homogeneous with respect to economic and societal developments in the period of observation. The qualitatively similar findings in both the country level and the district level analyses offer some confidence. Clearly the institutional, the macroeconomic and the political setup is more homogeneous for German districts than for European countries. Also assuming similar Wikipedia editions is no problem for German districts, since the Wikipedia under study is the *same*. However, two concerns remain for this analysis, because the regional analysis is based on the IP-addresses of anonymous contributions. First, using IP-addresses of anonymous contributors allowed only for a restricted set of available dependent variables. Specifically, we can only determine two measures of the efforts in any given region, i.e. we have to focus on the *number of edits* and the *length of edits* in bytes. This is because computing statistics like the *number of active editors*, or *edits per article* becomes meaningless when we can only observe a part of the edits from (anon. editors

[17] Wikipedia encourages all contributions are based on verifiable knowledge. As a result, the number of external links indicates reliable sources behind contributions.

and editors with self-reported location.) Second, the use of IP-addresses implies that we can only look at a specific set of contributions. These contributions come most likely from new or occasional users, because experienced users typically edit under their user name. In our research, we mitigate this problem by examining edits by registered users who reveal their (self-reported) location.

In addition, even if our assumptions for identification are satisfied, we can only provide indicative evidence on who drives the additional content generation, employed or unemployed individuals? While this question cannot be answered at all for the country level, for the German district level data we could provide an indication that additional edits come rather from leisure time.[18] Still, it remains unclear, whether the employed users increase activity during leisure hours, or whether unemployed prefer to contribute during in the hours which we classified as leisure time.

Another fruitful avenue for further research could investigate what is actually written. This question has to remain unanswered at the current stage of our research. Maybe people simply write *about* the crisis? This seems unlikely, given the overall growth that we observe. However, a smaller or larger fraction of the additional readership and content generation in the affected countries might be a direct increase in demand for economic information or the consequence of updating the encyclopedia with current events. Alternatively, increased editing activity might be dedicated to improving the overall quality of articles or individual users might contribute to their favorite topic of interest, which they also find enjoyable to write about.

To answer who makes the edits we would need user and editor level data, and to see what they write we would have to analyze articles on their content level. Further research could analyze the nature of contributions and which type of articles are edited. Also to what extent district specific articles are being improved or whether articles related to affected professions get edited would be very interesting. This is beyond the scope of this paper and, especially on the article level, this analysis is computationally intensive, but might lead to interesting additional insights from further research. On the country level this is almost unthinkable though, since the data available are too highly aggregated.

More fine grained data, on the user level ideally, would allow us to look at what information is being searched and which edits are made. Beyond that, we could contrast Wikipedia editing activities with other ways on how newly unemployed use their additional time.

7 Conclusion

In this paper, we study how individuals reallocate their time to online public goods provision when faced with increased rates of job displacement. We observe a moderate increase in socially valuable volunteering in the form of contributions to Wikipedia. The findings are consistent on both the European

[18] In line with models from labour economics, household production could be shifted into the working hours and additional activities would of recently unemployed would thus be observed during leisure times.

and the German district level. The uncovered patterns are suggestive of a creative and constructive potential that is freed as a positive side-effect of job displacement.

Human society has gone through many occasions of social and economic advancement in history. Each such social and economic restructuring brings new ways of production and consumption, but one unfortunate consequence of the progress is that workers with old skills are displaced from their jobs. Such job displacement took place when steam power incited a similar debate over two centuries ago. As it turned out, machines did not replace human beings in previous industrial revolutions. Our results give support to the relatively more optimistic view that human beings respond to such shifts by reallocating their time to both pleasure- and production-related activities.

It is important to point out that, even though robust and consistent, we find relatively small effects. Decision makers should consider how this valuable potential could be leveraged more systematically.

References

Aguiar, Mark and Erik Hurst, "Measuring trends in leisure: the allocation of time over five decades," *The Quarterly Journal of Economics*, 2007, *122* (3), 969–1006.

_ , _ , and Loukas Karabarbounis, "Recent developments in the economics of time use," *Annu. Rev. Econ.*, 2012, *4* (1), 373–397.

_ , _ , and _ , "Time use during the great recession," *The American Economic Review*, 2013, *103* (5), 1664–1696.

Algan, Yann, Yochai Benkler, Mayo Fuster Morell, and Jérôme Hergueux, "Cooperation in a Peer Production Economy Experimental Evidence from Wikipedia," in "Workshop on Information Systems and Economics, Milan, Italy" 2013, pp. 1–31.

Andreoni, James, "Privately provided public goods in a large economy: the limits of altruism," *Journal of Public Economics*, 1988, *35* (1), 57–73.

_ , "Giving with impure altruism: applications to charity and Ricardian equivalence," *The Journal of Political Economy*, 1989, pp. 1447–1458.

_ , "Impure altruism and donations to public goods: a theory of warm-glow giving," *The economic journal*, 1990, pp. 464–477.

_ , "Giving gifts to groups: How altruism depends on the number of recipients," *Journal of Public Economics*, 2007, *91* (9), 1731–1749.

Brynjolfsson, Erik and Andrew McAfee, *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*, WW Norton & Company, 2014.

Burda, Michael C and Daniel S Hamermesh, "Unemployment, market work and household production," *Economics Letters*, 2010, *107* (2), 131–133.

Campa, Riccardo, "Technological growth and unemployment: a global scenario analysis," *Journal of evolution and technology*, 2014, *24* (1), 86–103.

Comino, Stefano, Fabio M Manenti, and Maria Laura Parisi, "From planning to mature: On the success of open source projects," *Research Policy*, 2007, *36* (10), 1575–1586.

Farber, Henry S, "Job loss in the Great Recession and its aftermath: US evidence from the displaced workers survey," Technical Report, National Bureau of Economic Research 2015.

Ford, Martin, *The rise of the robots: Technology and the threat of mass unemployment*, Oneworld Publications, 2015.

Freeman, Kassie, "Increasing African Americans' participation in higher education: African American high-school students' perspectives," *Journal of Higher Education*, 1997, pp. 523–550.

Goldfarb, Avi and Jeff Prince, "Internet adoption and usage patterns are different: Implications for the digital divide," *Information Economics and Policy*, 2008, *20* (1), 2–15.

Hess, Charlotte and Elinor Ostrom, "Ideas, artifacts, and facilities: information as a common-pool resource," *Law and contemporary problems*, 2003, pp. 111–145.

Kandel, Eugene and Edward P Lazear, "Peer pressure and partnerships," *Journal of political Economy*, 1992, pp. 801–817.

Knabe, Andreas, Steffen Rätzel, Ronnie Schöb, and Joachim Weimann, "Dissatisfied with Life but Having a Good Day: Time-use and Well-being of the Unemployed*," *The Economic Journal*, 2010, *120* (547), 867–889.

Krueger, Alan B and Andreas I Mueller, "Time use, emotional well-being, and unemployment: Evidence from longitudinal data," *The American Economic Review*, 2012, *102* (3), 594–599.

Pissarides, Christopher A, "Loss of skill during unemployment and the persistence of employment shocks," *The Quarterly Journal of Economics*, 1992, *107* (4), 1371–1391.

Schumpeter, Joseph, "Creative destruction," *Capitalism, socialism and democracy*, 1942, *825.*

Taniguchi, Hiromi, "Mens and Womens Volunteering: Gender Differences in the Effects of Employment and Family Characteristics," *Nonprofit and Voluntary Sector Quarterly*, 2006, *35* (1), 83–101.

Uslaner, Eric M, "Religion and civic engagement in Canada and the United States," *Journal for the Scientific Study of Religion*, 2002, *41* (2), 239–254.

Vesterlund, Lise, "Why do people give," *The nonprofit sector: A research handbook*, 2006, *2*, 168–190.

Wallsten, Scott, "What are we not doing when we're online," Technical Report, National Bureau of Economic Research 2013.

Zhang, Xiaoquan Michael and Feng Zhu, "Group size and incentives to contribute: A natural experiment at Chinese Wikipedia," *American Economic Review*, 2011, *101* (4), 1601–1615.

A Appendix

A.1 Robustness of Regional analysis for Germany

A.1.1 Edits during Working vs. Leisure Time Hours

Table 13: DID Regression for German Districts for contributions made in working time and leisure time of the day (6pm-9am and weekends)

	Working time				Leisure time			
	(1) A. Edits	(2) A. Bytes	(3) R. Edits	(4) R. Bytes	(5) A. Edits	(6) A. Bytes	(7) R. Edits	(8) R. Bytes
After Treatment	-0.291***	-0.150	0.031	-0.014	-0.312***	-0.212*	-0.006	-0.082
	(0.037)	(0.138)	(0.092)	(0.178)	(0.032)	(0.113)	(0.089)	(0.186)
Treated x After	0.035	0.106	-0.013	0.106	0.082***	0.091	-0.068	-0.067
	(0.031)	(0.116)	(0.092)	(0.177)	(0.029)	(0.102)	(0.099)	(0.194)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3120	2018	3120	2993	3120	2229	3120	2983

NOTES: The table shows the results of our main difference-in-difference analysis that contrasts affected and unaffected German districts around the onset of the European Financial Crisis in 2009. The columns contain different measures of contribution activity to Wikipedia during working and leisure time: first, in the panel of results for working time, the number of (1) edits and (2) bytes by anonymous editors, (3) and (4) edits and bytes by registered users, then for the leisure time the number of anonymous (5) edits and (6) bytes contributed, and the number of registered (7) edits and (8) bytes. The variable of interest, which captures the treatment effect, *Treated districts after T*(reatment), is an interaction term between dummies for the districts that are affected by the crisis with the time dummy indicating the period after the crisis. All specifications include time period (month-year) dummies. Standard errors, clustered by districts, in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A.2 Robustness of Local Interest ('District-specific') Content Analysis)

A.2.1 Controlling for District-Specific pre-treatment Trends

Table 14: DID for local knowledge edits from German Districts (based on : 6 months before/after the crisis start)

	(1) Log All Contribs. (KB)	(2) Log # All Edits	(3) Log Anon.Contribs (KB)	(4) Log # Anonym.Edits	(5) Log Reg. Contribs (KB)	(6) Log # Reg. Edits
After treatment (start)	-0.139***	-0.042	-0.244***	-0.257***	-0.104**	0.003
	(0.040)	(0.035)	(0.048)	(0.038)	(0.044)	(0.038)
Treated districts after T	0.060	0.003	0.105***	0.052*	0.039	-0.010
	(0.037)	(0.030)	(0.037)	(0.029)	(0.041)	(0.033)
District pre-treatment trend	0.076**	0.117***	0.085***	0.052**	0.090***	0.143***
	(0.030)	(0.036)	(0.025)	(0.025)	(0.030)	(0.035)
Period FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3094	3094	3094	3094	3094	3094

NOTES: The table shows the results of our 'local content' analysis, when controlling for district-specific pre-treatment-trends. It shows a difference-in-difference analysis that contrasts edits about locally relevant knowledge in affected and unaffected German districts around the onset of the European Financial Crisis in 2009. Column (1) documents the strength of the treatment effect. The subsequent columns contain different measures of contribution activity to Wikipedia: (2) the length of 'all edits' (in KB) (3) the length of 'anonymous edits' (in KB) (4) the length of 'registered edits' (in KB) (5) the number of all edit (revisions) (6) the number of edits (revisions) by anonymous editors, (7) the number of edits by registered editors. The variable of interest, which captures the treatment effect, *Treated districts after T*(reatment), is an interaction term between dummies for the districts that are affected by the crisis with the time dummy indicating the period after the crisis. All dependent variables are in logs. Districts with an unemployment increase close to the median district were excluded from the regression (central 40%). All specifications include monthly dummies. Standard errors, clustered by districts, in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A.3 Country Level Analysis

A.3.1 Specification Tests and Alternative Specifications

Table 15: Country Level: DID Regression for the period of 12 months before and 12 months after the crisis

	(1) Log Views	(2) Log Active 5-100e.	(3) Log Active more 100e.	(4) Log Edits p.article	(5) Log Words growth	(6) Log Int. links	(7) Log Ext. links
Treated 3 months before	0.042	0.078	0.113**	0.032**	0.008	0.019	0.061**
	(0.079)	(0.056)	(0.046)	(0.012)	(0.073)	(0.016)	(0.028)
Treated 2 months before	-0.014	0.044	0.090*	0.034**	0.026	0.023	0.079**
	(0.086)	(0.057)	(0.048)	(0.014)	(0.087)	(0.019)	(0.032)
Treated 1 month before	-0.043	0.070	0.056	0.038**	0.061	0.027	0.091**
	(0.065)	(0.052)	(0.049)	(0.016)	(0.063)	(0.021)	(0.035)
Treated in month 0	-0.003	0.033	0.025	0.044**	-0.120	0.028	0.099**
	(0.079)	(0.061)	(0.053)	(0.018)	(0.090)	(0.024)	(0.039)
Treated in month 1	0.017	0.038	0.053	0.051**	-0.088	0.029	0.107**
	(0.082)	(0.051)	(0.041)	(0.020)	(0.114)	(0.028)	(0.043)
Treated in month 2	0.091	0.167***	0.080	0.055**	0.078	0.033	0.117**
	(0.071)	(0.056)	(0.054)	(0.022)	(0.071)	(0.030)	(0.046)
Treated in month 3	0.086	0.158**	0.133*	0.060**	0.201**	0.041	0.128**
	(0.094)	(0.072)	(0.064)	(0.024)	(0.090)	(0.033)	(0.050)
Treated in month 4	0.120	0.150**	0.143**	0.063**	0.178*	0.049	0.137**
	(0.086)	(0.061)	(0.063)	(0.025)	(0.099)	(0.036)	(0.054)
Treated 5-6 months later	0.186	0.151**	0.146**	0.064**	0.109	0.058	0.147**
	(0.109)	(0.059)	(0.055)	(0.028)	(0.069)	(0.040)	(0.061)
Treated 7-8 months later	0.190**	0.196***	0.170***	0.066**	0.213**	0.069	0.172**
	(0.085)	(0.059)	(0.059)	(0.031)	(0.079)	(0.044)	(0.065)
Treated 9-10 months later	0.164	0.171**	0.188***	0.069*	0.183*	0.081	0.189**
	(0.102)	(0.074)	(0.053)	(0.034)	(0.092)	(0.049)	(0.070)
Treated 11-12 months later	0.170*	0.176**	0.184***	0.075*	0.157	0.090	0.207**
	(0.089)	(0.064)	(0.056)	(0.037)	(0.106)	(0.052)	(0.077)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	452	500	500	500	500	500	500

NOTES: The table shows the relationship of the European economic crisis with different measures of contributions to Wikipedia (over time). In each column we show a different measure of activity on Wikipedia: (1) views of Wikipedia, (2) the number of active Wikipedians (with at least 5 edits), (3) the number of very active Wikipedians (with more than 100 edits), (4) the average number of edits per article, (5) the new words added, (6) hyperlinks to Wikipedia articles, (7) links to external sources. All measures of contributions to Wikipedia are in logs. The variable of interest is an interaction terms between an indicator for countries that are affected by the crisis with the time dummy indicating the period after the crisis. The excluded periods are the months 4-12 prior to treatment. All specifications include time (month - year) fixed effects. Standard errors, clustered by countries, are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 16: Country Level: DID Regression for the period of 12 months before and 12 months after the crisis

	(1) Log Views	(2) Log Active 5-100e.	(3) Log Active more 100e.	(4) Log Edits p.article	(5) Log Words growth	(6) Log Int. links	(7) Log Ext. links
After treatment	0.129	0.017	-0.005	-0.005	0.054	-0.076**	-0.008
	(0.095)	(0.034)	(0.042)	(0.017)	(0.051)	(0.035)	(0.059)
Treated countries after T	0.153*	0.097**	0.041	-0.001	0.115*	0.006	0.020
	(0.078)	(0.034)	(0.036)	(0.012)	(0.058)	(0.012)	(0.019)
Pretrend	0.007	0.369**	0.386***	0.718***	0.210**	0.923***	0.755***
	(0.018)	(0.131)	(0.091)	(0.081)	(0.088)	(0.093)	(0.113)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	432	480	480	480	480	480	480

NOTES: The table contains different measures of contributions to Wikipedia in each column: (1) views of Wikipedia, (2) the number of active Wikipedians (with at least 5 edits), (3) the number of very active Wikipedians (with more than 100 edits), (4) the average number of edits per article, (5) the new words added, (6) hyperlinks to Wikipedia articles, (7) hyperlinks to external sources. All measures of contributions to Wikipedia are in logs, and the month of the estimated crisis onset was omitted from the regressions. The variable of interest, which represents the treatment effect, *Treated countries after T*, is an interaction term between dummies for the countries that are affected by the crisis with the time dummy indicating the period after the crisis. All specifications include time (month - year) fixed effects. Standard errors, clustered by countries, are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 17: Country Level: OLS Regression for the period of 12 months before the crisis

	(1) Log Views	(2) Log Active 5-100e.	(3) Log Active more 100e.	(4) Log Edits p.article	(5) Log Words growth	(6) Log Int. links	(7) Log Ext. links
Unemployment	0.030	-0.005	-0.014	-0.000	0.005	0.009	0.014
	(0.028)	(0.019)	(0.017)	(0.004)	(0.026)	(0.008)	(0.016)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	192	240	240	240	240	240	240

NOTES: This table shows the relationship between unemployment and contributions to Wikipedia before the moment the economic crisis hits the country. The table contains different measures of contributions to Wikipedia in each column: (1) views of Wikipedia, (2) the number of active Wikipedians (with at least 5 edits), (3) the number of very active Wikipedians (with more than 100 edits), (4) the average number of edits per article, (5) the new words added, (6) hyperlinks to Wikipedia articles, (7) links to external sources. All measures of contributions to Wikipedia are in logs. The independent variable of interest, *Unemployment*, is the monthly unemployment rate. All specifications include time (month - year) fixed effects. Standard errors, clustered by countries are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 18: Country Level and Youth Unemployment: OLS Regression for the period of 12 months before and 12 months after the crisis

	(1) Log Views	(2) Log Active 5-100e.	(3) Log Active more 100e.	(4) Log Edits p.article	(5) Log Words growth	(6) Log Int. links	(7) Log Ext. links
Youth unemployment	0.010**	0.007***	0.010***	0.001	0.022***	0.012***	0.019***
	(0.004)	(0.002)	(0.002)	(0.002)	(0.006)	(0.003)	(0.006)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	411	456	456	456	456	456	456

NOTES: The table contains different measures of contributions to Wikipedia in each column: (1) views of Wikipedia, (2) the number of active Wikipedians (with at least 5 edits), (3) the number of very active Wikipedians (with more than 100 edits), (4) the average number of edits per article, (5) the new words added, (6) hyperlinks to Wikipedia articles, (7) links to external sources. All measures of contributions to Wikipedia are in logs. The independent variable of interest, *Youth unemployment*, is the monthly unemployment rate. All specifications include time (month - year) fixed effects and exclude the estimated period of the crisis onset. Standard errors, clustered by countries are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.