CBS COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

(How) Does User-generated Content Impact Content Generated by Professionals?

Evidence From Local News

Sen, Ananya; Grad, Tom; Ferreira, Pedro ; Claussen, Jörg

Document Version
Accepted author manuscript

Published in:
Management Science

DOI:
10.1287/mnsc.2023.4962

Publication date:
2024

License
Unspecified

Citation for published version (APA):
Sen, A., Grad, T., Ferreira, P., & Claussen, J. (2024). (How) Does User-generated Content Impact Content Generated by Professionals? Evidence From Local News. *Management Science*, *70*(9), 6045-6068. https://doi.org/10.1287/mnsc.2023.4962

Link to publication in CBS Research Portal

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jul. 2025

(How) Does User Generated Content Impact Content Generated by Professionals? Evidence from Local News *

Ananya Sen
Carnegie Mellon University†

Tom Grad
Copenhagen Business School‡

Pedro Ferreira
Carnegie Mellon University§

Jörg Claussen
LMU Munich and CBS¶

March 2023

Abstract

Many platforms host User Generated Content (UGC) and content developed by professionals side by side. However, so far their impact on platform ecosystems has been mostly studied in isolation. In this paper, we use data from a network of 122 local news outlets hosted by an online news platform to study the spillover effects from UGC developed by citizen journalists to the content developed by professional journalists. We use the removal of a status index associated with citizen journalists as an exogenous shock to their supply of UGC to identify these spillover effects. We find that experienced citizen journalists reduce their production of content when this status index is removed. We then find that inexperienced professional journalists increase their output in response to this behavior. However, as a result of these changes, we find a reduction in the overall content hosted by the platform, especially in the case of local news, and in more isolated regions. We further show that this is likely to have detrimental effects for the platform. In particular, there is a decline in overall viewership and the platform may need to hire and pay salaries to more professional journalists to produce enough articles to close the gap left by the departing citizen journalists. Our work contributes to the literature on UGC and online platforms, as well as to the literature on local news.

1 Introduction

Digitization has led many online platforms to leverage User Generated Content (UGC) for tasks that professionals have historically conducted. Examples range from building knowledge pages on

*We thank Param Vir Singh, Yash Babar, Nikhil George, Grace Gu, Dominik Gutt, Alexander Staub, and Joel Waldfogel, as well as conference and seminar participants at University of Zurich, CMU, Fordham, Copenhagen Business School, Pittsburgh SITE seminar, SCECR, UW Seattle (Foster School), WISE 2020, CODE@MIT, NBER Digitization Summer Institute, Digital Economy Workshop at LMU Munich, and the Digital Economics Workshop at Tel-Aviv University for helpful suggestions.

†Carnegie Mellon University, ananyase@andrew.cmu.edu.

‡Copenhagen Business School, tg.si@cbs.dk

§Carnegie Mellon University, pedrof@cmu.edu.

¶LMU Munich, Munich School of Management (ISTO) and Copenhagen Business School, j.claussen@lmu.de.

Wikipedia (Greenstein and Zhu, 2018) to crowdsourcing online reviews on websites such as Tripadvisor and Yelp (Mayzlin et al., 2014). UGC can accelerate platform growth (Aaltonen and Seiler, 2016) and, when paired with the right platform design features, allow platforms to generate and capture additional value (Subramanian et al., 2021). Many platforms, such as app stores (Boudreau, 2019), video streaming portals (Kim, 2012), or online news websites such as *HuffPost*, host UGC and content from professionals side-by-side. In a recent analysis of 160 platforms, Subramanian et al. (2021) find that at least 23% of the platforms that primarily host firm-generated content also host UGC. At the same time, more and more platforms that traditionally focused on UGC started to host firm-generated content, such as YouTube (YouTube Originals) and Snapchat (Snap Originals).

Until now, the content and behavior of users and professionals on these platforms have been mostly studied in isolation, which leaves the potential (complex) interactions within these platforms largely unexplored (McIntyre and Srinivasan, 2017). Limited evidence on these interactions is surprising because they may have important implications for online communities and platform strategy. These interactions may be especially relevant for platforms that match users to relevant information, such as social networks or review websites (Cennamo, 2021; Tajedin et al., 2019). Moreover, the impact of UGC on the overall quality of the information available in these platforms is ambiguous. On the one hand, a larger share of UGC might lead to co-specialization as well as diversity of content, allowing these platforms to meet the heterogeneous preferences of users. On the other hand, and because of its voluntary nature, it is hard to steer or control the quality of UGC, and excess UGC might drown out high-quality content (Boudreau, 2012; Boudreau and Jeppesen, 2015).

In this paper, we study how UGC affects the *amount of content* and the *type of content* produced by professionals[1] as well as the implications that these spillovers may have for engagement with the platform.[2] The news industry is vital because, unlike standard products, it produces a good with significant information externalities. For example, there is substantial evidence that the availability of local news affects political and social outcomes (Cagé, 2020; Oberholzer-Gee and Waldfogel, 2009; Snyder and Strömberg, 2010; Drago et al., 2014; Campante and Do, 2014; Matherly and Greenwood, 2021). However, fierce competition from digital platforms and the changing patterns of news consumption challenge the traditional business models used to provision local news (Seamans and Zhu, 2014). UGC may help to address these challenges but also create uncertainty because it cannot be contractu-

[1] Note that professionals in our context are full-time employed journalists by the platform, i.e., their content can be considered firm generated content.

[2] Throughout the paper, we use the term 'spillover' to refer to the potential impact of the content production decisions of citizen journalists on the decisions of professionals journalists.

ally mandated. Instead, platforms with UGC rely on self-selection and fluid participation of interested and motivated amateurs (Bughin, 2007; Puranam et al., 2014; Levine and Prietula, 2014). We study the provision of local news by citizen and professional journalists, focusing on the complementarity and substitutability between the content provided by these two groups of journalists.

We leverage data from Meinbezirk.at, the online platform used by a network of local newspapers covering Austria, to study the spillovers between UGC and content generated by professional journalists. The platform attracts 2.2 million monthly users and consists of 122 local issues published by 77 regional offices. Articles on this platform are written by professional journalists directly employed and paid by the platform and unpaid citizen journalists who do not receive any monetary compensation for their contributions. Articles from professionals and citizens are published online in chronological order without editorial intervention.[3] Until September 2018, Meinbezirk.at used a status index aimed at recognizing the contributions of citizen journalists. This index gave them points for writing news articles and receiving comments. However, on September 26th, 2018, a third-party software supplier (Gogol) that Meinbezirk.at relies on to develop and operate their platform updated their back-end software, resulting in a new release of Meinbezirk's platform that did not include the status index. We use this platform change as an exogenous shock to the content provided by citizen journalists, which allows us to analyze the spillovers between UGC and content developed by professionals.

We find three sets of results. First, we establish that there are significant spillovers between the *amount of content* produced by citizens and professionals. In line with prior research on the motivations of crowds (see, e.g. Wiertz and de Ruyter, 2007; Wasko and Faraj, 2005; Gallus, 2017; Burtch et al., 2019; Khern-am nuai et al., 2018), the removal of the status index leads to a reduction in the content produced by citizens. This reduction is driven by "experienced" citizens, i.e., those with above-median engagement (measured by the number of clicks received) when the status index was in place. We then find that this reduction in the supply of UGC has no significant impact on the number of articles written by "experienced" professionals. However, there is a significant increase in the number of articles produced by "inexperienced" professionals (defined as those receiving less than the median number of clicks before removing the status index). We also show that the content produced by other types of journalists ("sponsored content" and content developed by "free-lance" journalists) did not change with removing the status index, which provides us with confidence that removing the

[3] In fact, the platform has recently been criticized by the Austrian Press Council for not sufficiently differentiating UGC and professional content; see (in German) https://www.derstandard.at/story/2000130435799/presserat-ruegt-meinbezirkat-wege-schlecht-gekennzeichnetuser-beitraege

status index resulted in spillover effects from experienced citizens to inexperienced professionals. We then quantify the magnitude of these spillovers using an instrumental variable approach and find that the increase in articles written by inexperienced professionals after removing the status index does not fully close the gap created by the departing experienced citizens.

Second, we analyze the heterogeneity of these spillovers across different *types of content* produced, namely local vs. national news. We use several measures, including natural language processing techniques, to do so. Interestingly, we find that local content reduces significantly relative to national content due to a sharp reduction in local content from experienced citizen journalists. Inexperienced professionals fill the gap left behind by experienced citizens by producing more local content but only partially. Moreover, we leverage regional variation to show that the reduction in content is more severe in more "geographically isolated" areas (where isolation is measured by distance to the corresponding State capital). We show that inexperienced professional journalists have a hard time filling the gap left by experienced citizen journalists in more isolated regions, which suggests that the content developed by citizen journalists may *complement* the content developed by professionals by catering to the long-tail of users with heterogeneous preferences.

Third, we analyze the implications of these spillovers between UGC and the content developed by professionals for engagement in the platform. In particular, we show a decline in engagement with local news (measured by clicks on articles) after removing the status index, which may have negative implications for the top-line revenue in an advertisement-driven business like the platform we study in this paper. We discuss how this may arise when the top experienced citizens outperform the median inexperienced professionals (in terms of clicks per article). The fact that some experienced citizens attract as many clicks per article as some inexperienced professional journalists and that the latter react when the former reduce production by increasing their output indicates a relationship of *substitutability* between UGC and content from professionals. Beyond the immediate short-term effect on engagement, the observed decline in readership can have significant long-term effects on the platform's health, given the typical transition from passive consumers to active contributors in online communities (Kane and Ransbotham, 2016). Our results also point to adverse consequences for the bottom line of the platform, which may need to hire and pay salaries to additional professional journalists to close the gap left behind by the departing unpaid citizen contributors. This consequence makes a strong case for UGC as an integral part of the strategy to develop content in online news outlets.

Our study contributes to several strands of academic literature. First, we contribute to literature that analyzes UGC (Greenstein and Zhu, 2018; Ransbotham et al., 2012; Subramanian et al., 2021; Lukyanenko et al., 2019) and especially its broader impact and spillovers (Cage et al., 2020; Reimers and Waldfogel, 2020; Tirunillai and Tellis, 2012). We leverage a natural experiment and utilize author-level micro-data to establish the existence of spillovers from UGC to the content produced by professional journalists, beyond network effects, and thus show the interdependence across different actors in the same ecosystem both in terms of the *amount* and *type* of content produced. Our findings help us understand better these dynamics in platform ecosystems (Deng et al., 2021; McIntyre and Srinivasan, 2017). In particular, our study augments the prior literature with new knowledge about when UGC and content produced by professionals are substitutes or complements. Our results show that citizen journalists, when provided the right incentives, produce UGC that substitutes for professional content still leading to similar levels of engagement (which we use in this paper as a measure of quality). Moreover, citizen journalists have an advantage when producing content relevant to specific audiences and therefore their content complements the content from professionals. This complementarity has important implications for the literature on platforms when there is significant heterogeneity in the preferences of readers for content and the quality of the information provided is of particular relevance (Cennamo, 2021; Prat and Valletti, 2021; Tajedin et al., 2019). Our study provides compelling empirical evidence for the theoretical proposition that crowdsourcing is a mechanism that translates "distant search" into "local search" (Afuah and Tucci, 2012; Tajedin et al., 2019). Many dispersed citizens with local knowledge can cater better to the heterogeneous preferences of niche audiences than a few professionals who are more concentrated in the types of content and locations that they cover.

Second, we contribute to the extant literature that studies the local news industry (Matherly and Greenwood, 2021; Campante and Do, 2014; Oberholzer-Gee and Waldfogel, 2009), especially at its intersection with the literature on digital platforms (Seamans and Zhu, 2014; Dellarocas et al., 2016; Fischer et al., 2020; Chiou and Tucker, 2017). While the prior research has mostly highlighted the detrimental effect of platforms on the provision of local news, we offer a different perspective on the role of digital platforms for local news by analyzing the fundamental role of UGC. Our paper shows that user-generated local news can complement the news produced by professionals when the right incentives are offered to contributors. Especially for isolated regions, where producing high-quality local news might be difficult for a few (remote) professionals, involving many dispersed users can be

beneficial. To the best of our knowledge, we are the first to comprehensively investigate how UGC shapes the content produced by other journalists in the same platform ecosystem and how these dynamics results in engagement. By focusing on these dimensions of platform strategy and content production decisions across different types of journalists, we complement Cage et al. (2020), who look at the spillovers of Twitter onto newsroom decisions in a more aggregated manner.

Finally, while our main interest lies in the spillovers between UGC and the content produced by professionals in the local news context, our work also speaks to the literature that analyzes the impact of non-monetary incentives on the provision of UGC (see, e.g. Burtch et al., 2019; Gallus, 2017; Goes et al., 2016). These studies focus on the effects of *receiving* a reward on the subsequent behavior of individuals. In contrast, we use a natural experiment – the *removal* of the entire incentive system for UGC – to document substantial heterogeneous effects. We find that *removing* the opportunity to accumulate status adversely affects the individuals that have accumulated the most status, even though they may have been the most active on the platform before the removal of such a system (see, e.g. Subramanian et al., 2021; Resnick et al., 2006). Still, the effect of these dynamics, if any, and who is most affected is unclear *a priori* given the potential heterogeneity across contributors and the potential behavioral "biases" involved in purely symbolic incentives. Hence, *ex-ante*, whether the removal of the status index would mirror the impact of its introduction is theoretically ambiguous.

Our work highlights the importance of understanding platforms as complex ecosystems. Given the spillovers we identify between UGC and the content developed by professionals, we show that online platforms should not treat professionals and amateurs as separate markets. Our findings suggest that it is vital for platforms to identify the interdependencies between these interacting sets of contributors to maximize platform engagement and revenue. We show that UGC can increase engagement and broaden the coverage of content on news platforms. Using non-monetary incentives can help encourage citizens to create local news content on online media platforms to foster the production of the essential public good of local news. Otherwise, (inexperienced) professionals have to be hired, which is costly and does not create as much buzz.

2 Literature & Background

In this paper we mainly draw on two streams of literature: 1) Platforms and UGC, and 2) the production of local news.

2.1 Platforms and UGC

Platforms play a fundamental role in today's digital economy (see Rietveld and Schilling, 2021, for a recent review). They increasingly rely on UGC to host relevant content (Subramanian et al., 2021), and several studies have documented the potential of amateurs to provide valuable contributions, in particular of similar quality compared to the work of professionals and experts (see, e.g. Mollick and Nanda, 2016; Reimers and Waldfogel, 2020; Poetz and Schreier, 2012; Greenstein and Zhu, 2018). Prior research has also shown that platforms can use monetary and non-monetary incentives to stimulate the production of UGC (see e.g., Burtch et al., 2019; Gallus, 2017; Khern-am nuai et al., 2018). However, the effectiveness of such incentives schemes depends highly on the underlying motivations of users to contribute (Boudreau and Jeppesen, 2015), which may differ across contributors with different levels of experience (Roberts et al., 2006), status (Levina and Arriaga, 2014; Ma and Agarwal, 2007), contribution habits (Resnick et al., 2006), and may further differ based on platform-specific capital (Subramanian et al., 2021; McAfee et al., 2010). In most cases, ensuring the quality of UGC is still notoriously difficult given the voluntary nature of the contributions from users (Bughin, 2007; Lukyanenko et al., 2019; Levine and Prietula, 2014).

Ultimately, the success of an informational platform is a direct consequence of the quality of the content offered to consumers. At first sight, the more UGC on the platform the more opportunity to match users with content that they like. However, content competes for the scarce attention from users (when at scale), and adverse network effects (e.g., too much information on the platform competing for user attention) can outweigh the potential positive effects of a more extensive contributor base (Boudreau and Jeppesen, 2015). In these platforms, curation (manual and algorithmic) and access restrictions can then become important tools to facilitate high-quality matches between users and content (Dellarocas et al., 2016; Boudreau, 2012; Rietveld and Schilling, 2021). Additionally, the prior literature has highlighted that the quality of information is a multi-dimensional construct (Wang and Strong, 1996; Lukyanenko et al., 2019). For example, Wang and Strong (1996) differentiate between intrinsic, contextual, representational, and accessibility quality. In the context of UGC and professional content, professionals will likely have an advantage in providing intrinsic and representational information given their education, training, and professional experience. However, users might have an advantage when providing contextual content because of their local embeddedness in the relevant communities and use-knowledge (Baldwin and Von Hippel, 2011). The contributions of users might be especially relevant for the highly-specialized long-tail preferences of users for content (Dellarocas

et al., 2010).

Users can learn about the quality of their contributions through platform interaction features such as likes and comments, which helps them improve the quality of their contributions over time (Riedl and Seidel, 2018). Therefore, it might be possible that a group of users becomes particularly capable of providing content of similar, or even better quality, compared to the content offered by professionals. UGC with sufficiently high quality may substitute professional content in economically beneficial ways for platforms. UGC might also be a valuable complement to professional content when professionals lack relevant contextual knowledge. In this case, UGC can broaden the scope of the content offered on the platform catering to diverse user preferences. Still, the contribution from users to online platforms is often highly skewed, with a small number of highly committed users providing a large share of the contributions (Lerner and Tirole, 2002; Van Mierlo, 2014). Also, prior research has documented the dynamics of content production within crowds (see, e.g. Aaltonen and Seiler, 2016; Li and Hitt, 2008; Le Mens et al., 2018) and from professionals to crowds (see, e.g. Kovács and Sharkey, 2014; Deng et al., 2021). However, the effect that UGC might have on content from professionals is less understood but important from a platform strategy perspective.

Platform owners need to design attractive incentive schemes to motivate user participation if they want to rely on a constant flow of high-quality UGC. When content provision from external sources drops, professionals, might need to step up to secure a) sufficient supply to avoid losing readers, and b) sufficient specialized content that allows the platform to keep its unique identity (Cennamo, 2021). The most likely candidates to step up and fill gaps in UGC are likely those professionals who are more similar to the external contributors because they face lower adjustment and opportunity costs in adjusting their contributions and content (Argyres et al., 2019, 2022).

Prior research has looked at both monetary as well as non-monetary incentives to attract UGC (see e.g. Burtch et al., 2019; Gallus, 2017; Khern-am nuai et al., 2018). For example, Gallus (2017) shows how providing symbolic awards can improve newcomer retention. However, (Goes et al., 2016) shows that these effects are often temporary.Burtch et al. (2019) study the effect of receiving peer awards on subsequent content provision and find that producing content positively affects the probability of providing future content, but content becomes less diverse over time. Similarly, Khern-am nuai et al. (2018) show that monetary rewards can increase user-generated reviews but that these reviews are of lower quality. Importantly, this literature focuses on the effects of (potentially) receiving rewards on the behavior of individuals. However, given the documented differences between gain and loss framing

(Kühberger, 1998; Tversky and Kahneman, 1991), it is not clear whether the dynamics of introduction and receiving incentives are symmetric to those that arise from removing incentives and thus studying the latter is also fundamental to understand the dynamics of platforms that rely on UGC.

2.2 Local News

Our work also contributes to the literature on local news. The availability of local news has been shown to affect important societal outcomes. For example, (Oberholzer-Gee and Waldfogel, 2009) show that local news affects not just voter turnout but also election outcomes (Drago et al., 2014). Other studies show that the presence of local news affects the efficiency of governments (Drago et al., 2014) and political accountability (Campante and Do, 2014; Snyder and Strömberg, 2010; Matherly and Greenwood, 2021).

However, the provision of local news has come under pressure from multiple angles. First, national news outlets are expanding into local markets, reducing readership for local news outlets (George and Waldfogel, 2006). Second, the emergence of online platforms, especially those focused on classifieds such as craigslist and eBay, undermined a vital revenue stream for local newspapers (Seamans and Zhu, 2014). Third, news aggregator platforms, such as Google News, promote competition for attention potentially hurting niche news outlets (Dellarocas et al., 2010; Chiou and Tucker, 2017; Meyer et al., 2022). A variety of solutions to revive local news have been experimented with, including utilizing UGC in the form of "citizen journalism", where "ordinary citizens play an active role in the process of collecting, reporting, analyzing, and disseminating news and information" (Paulussen et al., 2008). For example, in 2006, CNN started a citizen journalism initiative called *iReport* to encourage users to send in pictures, videos, and other information about the area where they live. More recently, Substack, an online platform that has become increasingly popular for professional journalists, introduced an initiative to foster the creation of local news on online platforms (see https://blog.substack.com/p/introducing-substack-local-for-a). However, professional journalists often remain critical of the quality of the content provided by amateurs (Hermida and Thurman, 2008). The feasibility of relying on users to provide local news and the role that UGC might have in providing local news by professionals is not yet well understood and our paper contributes to improve it.

3 Empirical Setting and Data Description

3.1 Empirical Setting

We have access to a unique dataset from Meinbezirk.at, a network of 122 regional news outlets in Austria. These news outlets focus mostly on local news on various topics, including local politics, health, sports, and business. "Meinbezirk", in fact, literally translates to "My District". The network has over 2 million unique users a month and reaches over 30% of Austria's online readership.[4] Meinbezirk.at is owned by Regional Medien Austria (RMA), and posts over 100 million Euros in annual revenue on average. It is considered among the country's largest and most influential news outlets.

A variety of journalists produces content at Meinbezirk.at. According to summary information from 2016 to 2018, 496 professional journalists are directly employed by Meinbezirk's news outlets. Furthermore, 12,209 citizen journalists produce UGC on various issues and events during this period. They receive no monetary compensation for the content they produce. Professional journalists on the payroll of Meinbezirk's news outlets produce output consistently. However, most citizen journalists produce content only sporadically. It is important to note that to become a professional journalist, an individual has to be formally hired by a regional office of the platform and has to be a salaried employee. Therefore, citizen journalists cannot simply 'transition' into becoming professionals by contributing more to the platform. The professional journalists (experienced and inexperienced) work on contracts with fixed monthly salaries[5], which do not depend on the number of articles written or clicks received (or other engagement metrics). However, professionals also have an interest in the platform's overall long-term success, because their jobs depend on it. The success of the overall platform depends on the ability to leverage external contributions to grow bigger more quickly than what could be achieved using internal resources alone (Subramanian et al., 2021; Rietveld and Schilling, 2021). Also, for monetization purposes, the platform relies on click-based advertising revenue, which would be affected by the amount of content offered to readers.[6]

In addition to these two groups of journalists, we have data on sponsored content produced by 379 dedicated journalists who work with companies that seek media exposure. Therefore, these sponsored-content journalists differ from the professional and citizen journalists who are the focus of our study.

[4] See https://www.regionalmedien.at/unsere-medien/meinbezirk-at/ for more information.

[5] In Austria, where the platform is based, salaries are often a result of collective bargaining agreements which is also the case for most journalists see (in German) https://www.oezv.or.at/politik-recht/kollektivvertraege/.

[6] see (in German) https://www.regionalmedien.at/wp-content/uploads/2021/12/RegionalMedienAustria_Tarif_2022.pdf for details.

Another group of journalists, the freelancers, pitch different articles or stories to the news outlets, which may pick them up on a case-by-case basis in exchange for a mutually agreed-upon monetary fee. These contributions involve an editorial process and monetary compensation, and we exclude them from our analysis.

Professionals write along three pillars of content on this platform. First, there is an annual campaign planning process where topics for the next quarter/year are broadly agreed upon. The general expectation is that professional journalists at the regional offices will then write about these topic areas. Second, there are "chronological events", i.e., newsworthy events that happen over time, such as local political events, incidents, and sporting events. Professional journalists write about these events, which account for a significant proportion of the content at Meinbezirk. Finally, professionals write on some topics that emerge cross-regionally or at the national level. For example, a report about the accessibility of public buildings for people with disabilities might arise on the national level, which is then locally adapted and locally covered. Thus, professionals at Meinbezirk receive general guidelines for what to cover based on a high-level strategy for content production. However, this guidance is far from a limiting factor. Professional journalists have significant control over the local adaptations of national campaigns and editorial oversight over the day-to-day "chronological events" to cover. This setup mirrors the strategies to produce content followed by other large news organizations (Sen and Yildirim, 2015; Beattie et al., 2021; Cage et al., 2020).

3.2 The Status Index and Its Removal

The platform uses an activity-based status index for citizen journalists to incentivize UGC, similar to that found, for example, at eBay. Citizen journalists have detailed information and receive advice about the scoring system on a dedicated platform page.[7] The accumulation of points is mostly activity driven but also involves social elements: each article written awards 5 points, each picture uploaded awards 2 points, and each comment received accrues 0.5 points.[8] The number of points accumulated is shown in the profile of each journalist, as seen in Figure 1, as well as in all articles that she writes, as seen in Figure A.1. While we do not observe this status index precisely, the granularity of our data, which we describe in further detail below, allows us to reconstruct it in great detail. A significant value set forth by this platform is to be democratic and provide a relatively leveled playing field to

[7] See https://www.meinbezirk.at/tag/regionauten-tipp for more information on this system.
[8] See https://web.archive.org/web/20170314230238/https://www.meinbezirk.at/s/hilfe for more details on the points that citizens can obtain with this status index.

new journalists relative to journalists with a long history of producing content. In line with this 'democratic stance' of the platform, articles appear on the website sequentially – as they are uploaded – with no editorial curation.

The primary source of exogenous variation in our analysis is the removal of this status index on September 26th, 2018, by the software supplier Gogol, which develops Meinbezirk's back-end. This shock can be thought of "as random" because Gogol introduced it not at the behest of Meinbezirk, as confirmed to us in our conversations with Gogol. After removing this index, the profile of citizen journalists looked different along a very significant dimension. Figure 2 shows that, after the change introduced by Gogol, the profile of citizen journalists no longer included the points (in red, as well as"Punkte"), the number of pictures ("Schnappschuesse"), and comments ("Kommentare"), among other information. Removing the status index from Meinbezirk prevented citizen journalists from accumulating points. Furthermore, as noted above, citizen journalists do not receive monetary remuneration for their content. Thus, removing the status index took away the potential for citizens to seek status and gain recognition for the content they produced. It is this variation that we intend to utilize in our empirical analysis. [9]

3.3 Descriptive Changes in the Amount of Content Produced

We observe the details of all posts at Meinbezirk from the beginning of 2016 to February 2019. For example, we see the number of words, pictures, pageviews (or clicks), comments, and the full text associated with each post. We also know each article's category (local, politics, sports), the timestamp of its publication, and the region from which the author uploaded the post. For our primary analysis, we focus on a 17 week period around the removal of the status index (in the second half of 2018), which helps us ensure that we report credible causal estimates of the effects of interest. In addition, we use data prior to the platform change to define experience measures for the journalists at Meinbezirk and train language models for text classification. A long stretch of prior data helps avoid selection issues since we can observe the journalists' content production, including entry and exit, for a prolonged time before the platform change occurred.

We begin by documenting baseline differences in the characteristics of the content produced by

[9] Similarly, in Figure A.2 in the Appendix, we can see that the total number of points accumulated by the author was removed from the article page, as opposed to what used to happen before the change introduced by Gogol, as seen in Figure A.1 in the Appendix. It is important to note that this information on an author's status index is visible only after clicking on an article written by the author. This information is not visible when the article is displayed to users on the platform before they click on it. Hence, this also implies that removing the status index did not change what the potential reader sees and does not necessarily immediately impact clicking behavior.

citizen and professional journalists. This analysis is primarily descriptive to help us understand the different dimensions of the content produced across the various news outlets on this platform. Over the period 2016-2018, we observed about 720,000 articles written, 46.9% of which were written by the 496 professional journalists employed across the various local outlets of Meinbezirk. The 12,209 unpaid citizen journalists produced 43.2% of the articles. [10] Therefore, professional journalists write more regularly. An important takeaway related to content production is that a small fraction of citizen journalists are active on the platform at any given time. For example, about 2,000 citizen journalists were active during the 16 weeks before the platform change. For our analyses in subsequent sections, we aggregate content production at the author-week level.

On the other hand, a much larger pool of citizen journalists (relative to professionals) contributes to the platform sporadically from time to time. Sponsored content journalists (379) account for 1% of the content, with the remainder (8.75%) coming from freelance journalists. Panel A of Table 1 provides additional insights.[11] Professionals receive more clicks relative to citizens at the mean article, as well as at the median and the 95th percentile. Simple OLS regressions (Table A.2 in the Appendix) indicate that even after controlling for observable characteristics, such as the number of words, pictures, and the category of the article, there is still a 24% differential in engagement between articles from professional and citizen journalists. Professionals also write longer articles with more pictures, on average, which is interesting because the platform encourages citizen journalists to take many pictures, apart from the written text, to provide information related to local community news and events. This suggests that professionals exert more effort on the different dimensions of the production process and thus receive more engagement.

The above statistics provide an overview of the platform we study in this paper. The average professional puts in more effort and outperforms the average citizen journalist. However, it is important to note that these are statements about the average. As we document below, there is significant heterogeneity within both groups of journalists in terms of experience and performance.

4 Empirical Framework

Our baseline specification aims to estimate the impact of the removal of the status index on the amount and type of content produced by different types of journalists. In particular, we use an event

[10] All articles are single-authored; that is, one author has the byline for each article.
[11] Table A.3 in the Appendix captures some correlations across variables.

study framework to estimate the following:

$$y_{it} = \alpha + \gamma_i + \theta PlatformChange_t + \rho_{st} + \epsilon_{it} \tag{1}$$

The unit of observation in our analysis is the author-week.[12] Our dependent variable y_{it} represents production, measured by the number of articles written by journalist i in week t, engagement, captured by the number of clicks received by author i in week t, and 'localness' of a news article (as explained later in section 6.2). We track journalists over time, and thus γ_i captures journalist fixed effects, which helps us account for time-invariant journalist-specific characteristics such as gender, the regional traits of where they operate from, and their innate ability, among other factors. $PlatformChange$ is a dummy variable indicating the week when the status index was removed, which happened simultaneously for all journalists. ρ_{st} are state-specific dummies indicating holiday weeks. The main coefficient of interest in this regression is θ, which measures the impact of removing the status index. We focus on eight weeks before and eight weeks after the removal of the status index as the period of our analysis. This ensures that we have long enough pre and post-event periods to observe the effects of interest, as well as a window of time that is short enough so that the pre and post-event periods are more comparable than if we had a longer time horizon. Finally, we cluster standard errors at the journalist level to account for serial correlation in their decisions to produce content over time.

The above regression captures the average effect of removing the status index. Throughout our analysis, we also provide event study evidence by explicitly showing leads and lags to ensure the absence of pre-trends and an impact only after the event (actually) occurred. In particular, we report estimates for the following regression using the number of articles as dependent variable:

$$y_{it} = \alpha + \gamma_i + \sum_{q=-8, q\neq -1}^{8} \delta_q + \rho_{st} + \epsilon_{it} \tag{2}$$

As with equation 1, we look at the eight weeks before and eight weeks after removing the status index. This specification allows us to estimate coefficients weekly to analyze the temporal dynamics associated with removing the status index. We estimate the coefficients relative to the omitted period, the one before the platform change takes place.

[12] We build a balanced panel for the entire observation period, i.e., journalists who join the platform write 0 articles before they enter, and journalists who leave the platform write 0 articles after they go. Without user-level browsing (or login) data, such an approach can account for individuals joining or utilizing the platform at a given time.

Finally, we also estimate the elasticity of substitution between the number of articles produced by experienced citizen journalists and inexperienced professional journalists using the removal of the status index as an instrumental variable. In our first stage regression, we use the removal of the status index to predict the change in the number of articles written by all experienced citizen journalists in the same state as an inexperienced professional journalist. In the second stage, we regress the number of articles produced by each professional journalist on the former, thus using the following specification:

$$ycs_{it} = \alpha + \gamma_i + \theta PlatformChange_t + \rho_{st} + \epsilon_{it} \tag{3}$$

$$yp_{it} = \beta + \delta_i + \phi \widehat{ycs}_{it} + \rho_{st} + \epsilon_{it} \tag{4}$$

In this specification, ycs_{it} is the total number of articles written by all experienced citizen journalists in the same state as inexperienced professional journalist i in week t, and yp_{it} is the number of articles written by inexperienced professional journalist i in week t. The coefficient of interest in this regression is ϕ, which measures the elasticity of substitution in the number of articles produced by an inexperienced professional journalist and the total number of articles produced by all experienced citizen journalists in the same state (the focus on experienced journalists and inexperienced professionals will become apparent later).

5 Effects on Content Produced by Citizens and Professionals

5.1 Impact on Content Produced by Citizen Journalists

We now analyze the behavior of the different types of journalists in response to the removal of the status index. Panel A of Figure 3 shows the coefficients obtained from running the regression in equation 2 for all citizen journalists. Therefore, this figure shows the change in the number of articles written by citizens during the eight weeks before and the eight weeks after removing the status index. The figure plots estimates relative to the number of articles written in the week prior to removing the status index. There are two takeaways from this figure. First, citizen journalists write fewer articles after removing the status index. The figure shows that this reduction is immediate. The

number of articles written by citizen journalists drops during the first week after removing the status index. This reduction becomes more pronounced over the next few weeks, after which it stabilizes at a new lower level. Second, we do not see significant pre-trends in the number of articles written by citizen journalists. That is, before removing the status index, the number of articles written by citizen journalists hovers around the same stable level.

The next step to analyze how the removal of the status index changed content production is to understand who among the citizen journalists decreased the number of articles produced. To this end, we introduce a new variable that measures the experience of a journalist, namely the number of articles written times the number of clicks received per article in the two years before our window of analysis begins.[13] We then split our sample of journalists at the median, with those above the median categorized as "experienced" and those below the median categorized as "inexperienced".[14] By allowing for a long time window, we ensure that we minimize the number of false positives in either category. Moreover, by considering the number of articles and the number of clicks in this measure, we ensure that we take into account different dimensions of experience and writing styles. For example, some journalists might write longer, more well-thought-out pieces at longer intervals, which may receive significant engagement through comments. Other journalists might write briefer pieces but do so at shorter time intervals. Our experience measure allows both types of content producers to accumulate experience.

Panel B in Figure 3 shows that experienced citizens write fewer articles as soon as the status index is removed. As in Panel A of Figure 3, the response is immediate and sharp, with no discernible pre-trend prior to the platform change, which is in line with the hypothesis that a valuable status signal was arguably "taken away" from exactly those who had accumulated more status over time.[15] Figure 3 Panel C shows that the removal of the status index had minimal impact on the number of articles written by inexperienced citizens. The individual coefficients for each week are statistically insignificant and significantly smaller in magnitude than those obtained for experienced citizens. These observations align with the hypothesis that these individuals had little to lose from the removal of the status index at this stage.

[13] In Table A.6 of the Appendix, we use an alternative measure of experience, which yields similar results.
[14] In Table A.5 of the Appendix we use different thresholds to spit the sample of journalists and find qualitatively similar results.
[15] This decline in content is also visible descriptively in Table A.1 in columns (2) and (4) as well as column (2) of Table 2. Moreover, we also see some movement on the extensive margin with a decrease in the number of active experienced citizens operating on the platform.

5.2 Spillover to Content Produced by Professional Journalists

The main goal of our paper is to investigate the potential spillover effects between UGC and content generated by professionals on online platforms. So far, we have documented how citizen journalists significantly reduced their content in response to removing the status index. This could have detrimental implications for the platform since UGC accounts for about 43% of all content produced. However, and *A priori*, it is not clear whether a news outlet *could*, and *would want to* fill the gap created by the reduction in UGC because citizen journalists and professional journalists may cater to different audiences. The fact that professional journalists write more articles more frequently, more lengthy articles, and receive more engagement could support such a decision. However, given the current structure of the platform, with limited editorial control, there could be an oversupply of content before the removal of the status index, so the observed reduction in UGC could actually be beneficial by making high-quality content more visible. Additionally, if professional journalists respond to the gap left behind by experienced citizen journalists, it is unclear who among them would do so.

To analyze these issues, we look at whether there was any response from professional journalists to the documented reduction in UGC. Panel A in Figure 4 shows the coefficients obtained from running the regression in equation 2 for professional journalists. Therefore, this figure shows the change in the number of articles written by professional journalists during the eight weeks before and the eight weeks after removing the status index. The figure is drawn relative to the number of articles written in the week before removing the status index. This figure suggests that removing the status index seems to have had little impact on the number of articles written by professional journalists on average. Before the removal of the status index, the number of articles written by professional journalists hovers around a stable level. After removing the index, the number of articles produced by professionals is statistically similar to what it was before, unlike what we observe for citizen journalists. Panel B in Figure 4 also shows no change in the number of articles produced by experienced professionals. However, panel C in Figure 4 shows clearly that inexperienced professionals increased the number of articles written after the reduction in UGC by experienced citizen journalists, which may suggest that inexperienced professional journalists made an effort to fill the void left by experienced citizen journalists. This increase in content is also visible descriptively in Panel B of Table 1 in columns (1) and (3) as well in regressions as seen in column (3) of Table 2.[16]

[16] Moreover, we also see movement on the extensive margin with an increase in the number of active inexperienced professionals operating on the platform in the period after the shock.

Next, we try to understand the net effect of the decline in content produced by experienced citizens and the increase in content produced by inexperienced professionals. The average number of articles produced by experienced citizens per week before removing the status index is 0.26. This statistic reduces by 0.147, as indicated in column 2 of Table 2. Therefore, the total number of articles written by experienced citizens per week before removing the status index is 12209/2*0.26=1587 (recall that there are 12209 citizen journalists in our panel, and experienced citizens are those with engagement above the median). The average number of articles written by inexperienced professionals per week before removing the status index is 0.676.[17] Therefore, the total number of articles written by inexperienced professionals per week before removing the status index is 496/2*0.676=168 (recall that there are 496 professional journalists in our panel, and inexperienced professionals are those with engagement below the median). Hence the total number of articles produced by experienced citizens and inexperienced professionals per week before removing the status index was 1755 (1587+168). The number of articles written by experienced citizens reduced by 12209/2*0.147=897 per week with the removal of the status index, while the number of articles written by inexperienced professionals increased by 496/2*0.421=105 per week.[18] Therefore, the number of articles written by experienced citizens reduced by 51% relative to the total number of articles written by citizens and professionals (897/1755), while the number of articles written by inexperienced professionals increased by 6% (105/1755). The ratio of these statistics is 6/51=11.7%, which shows clearly that the increase in the articles written by inexperienced professionals does not substitute the reduction observed for experienced citizens. In fact, the total number of articles written by experienced citizens and inexperienced professionals reduced by 792 per week, or 45% (792/1755), with the removal of the status index. To put this statistic in perspective, note that according to the results in column (1) of Table 2, removing the status index reduced the total number of articles in the platform by 18.6% (-0.0637/0.343).

An alternative approach to measuring the substitutability of content production by experienced citizen journalists and inexperienced professional journalists is to use the removal of the status index as an instrumental variable. Our approach, in this case, is to associate each inexperienced professional with the experienced citizens close to her, namely those in the same state. As documented in Table 2, removing the status index reduced the number of articles written by citizen journalists. Therefore, the number of articles written by citizen journalists at the state level should have reduced too. This constitutes the first stage of our IV regression, the results of which are reported in column (6) of Table

[17] The summary statistics for the eight weeks pre and post the platform change are in Table A.1 in the Appendix.
[18] We use the estimate 0.421 from column (3) of Table 2.

2. The average number of articles written per week by all experienced citizen journalists in a state was 252 before removing the status index. This statistic decreased by 125.[19] Column (6) in this table also shows that our first stage yields an F-statistic significantly above 10, as required to avoid a weak instrument (c.f. Stock and Yogo, 2002).

In addition, a valid instrumental variable should satisfy the exclusion restriction. In our case, this means that removing the status index from the platform (the instrument) should only affect the number of articles written by inexperienced professional journalists (our outcome of interest) through the number of articles written by experienced citizen journalists (the endogenous variable). The institutional details behind how the status index was used and removed from the platform shed light on this matter. First, based on our conversations with Gogol, the company that provides the backend software for Meinbezirk's platform, it was Gogol's unilateral decision to remove the status index. This means that decision-makers at Meinbezirk were not involved in this process. Second, following the previous point, the decision to remove the status index was made with the journalists, especially the citizen journalists, in the dark. Third, remember that professional journalists are employed and paid by the platform and therefore receive monetary compensation for their writing, unlike citizen journalists whose motivation is driven solely by non-monetary rewards such as the status index. Therefore, it is unlikely that journalists, both citizens, and professionals, anticipated the removal of the status index and that its removal affected the behavior of professional journalists in ways other than through the articles produced by citizens.

Column (7) in Table 2 shows the results obtained from our second-stage regression. The number of articles written by inexperienced professional journalists decreased by 0.0033 per unit change in the number of articles written by experienced citizen journalists in the same state. The average number of inexperienced professional journalists per state is 49. Therefore, the average number of articles written per week by inexperienced professional journalists in a state before removing the status index is 49/2*0.676=16.5. Hence, the average number of articles written per week by all inexperienced professionals and all experienced citizen journalists in the same state was 268.5 before removing the status index (252+16.5). Therefore, the increase in the number of articles written per week by an inexperienced professional journalist identified by our second stage IV regression is -0.0033*-125=0.4125 (the coefficient in the second stage of our IV regression times the change in the number of articles written by experienced citizen journalists in the same state as the inexperienced professional

[19] Note that when measuring the effect of removing the status index on the production of experienced citizen journalists at the journalist level, we obtain a similar statistic, namely 897/1587=56%, and 125/252=50% in this case.

journalist), or 0.4125*49/2=10.1 at the state level. Therefore, the change in the average number of articles written per week by all inexperienced professional journalists in a state (relative to the total number of articles written by experienced citizen journalists and inexperienced professional journalists in that state) is 10.1/268.5=3.76%. On average, the decline in the total number of articles written per week by all experienced citizen journalists in a state is 125, or 46.5% (125/268.5). The ratio of these two statistics is 3.76/46.5=8.1%. This result aligns with the one computed directly using the results in columns (2) and (3) of Table 2. It shows again that the increase in the production of inexperienced professional journalists does not make up for the decline in the production of experienced citizen journalists due to the removal of the status index.

Finally, we look at the change in the content produced by freelance journalists and by journalists that write sponsored content as an on-platform control group (see, e.g. Wu and Zhu (2018), 2018, for a similar on-platform control setup). Figure 5 shows that the number of articles these journalists wrote did not change with the removal of the status index. We show the average estimates of this result in Column (1) of Table A.8 in the Appendix. The remainder of the columns in this table shows that these journalists did not change the amount of political content written (measured by the number of articles that mention the major political parties in Austria) (column (2)), the "localness" of their articles in column (3) (which we describe more precisely in section 6.2), nor the tone of their articles (column (4); measured using the LIWC measure (Pennebaker et al., 2001)). In addition, we also see a lack of change in the total clicks received (column 5) as well as the clicks per article (column 6). The lack of change in the behavior of these journalists is important for two reasons. First, it shows that the change in the number of articles written by inexperienced professionals documented above does not seem to come from a reaction to changes in the behavior of these others journalists. Second, the fact that the behavior of these journalists did not change when removing the status index provides some evidence of the lack of concurrent changes in the platform that could drive both the behavior of experienced citizen journalists and that of inexperienced professional journalists.

5.3 Why Professional Journalists React To Citizens Output

We draw on existing theories and institutional details as well as additional statistical analyses to provide suggestive evidence for *why* (inexperienced) professionals might have incentives to respond to the decline in content produced by (experienced) citizens. .

First, we re-iterate that citizens do not receive financial compensation for their writing efforts. Like

in many other online communities and platforms, their motivation to contribute to Meinberzirk.at is driven by a mix of intrinsic and extrinsic motivation (Lakhani and Wolf, 2003; Boudreau and Jeppesen, 2015; Jeppesen and Frederiksen, 2006). While platform owners try to design their platforms in ways that appeal to intrinsic motivations, such as with challenging and fun activities, a major design component are non-monetary incentives such as points, badges and awards, that provide extrinsic motivation (see e.g. Gallus, 2017; Anderson et al., 2013). The removal of the non-monetary status index (accumulation of points for articles written and the public display of such points on the website) took away an essential part of the motivation for citizens to contribute. This should directly affect their willingness to contribute. Indeed, experienced citizens reduced their output because they would have driven the highest utility from the status they built on the platform up to when it was removed.

Next, professionals on the platform are salaried employees of the firm and therefore receive a fixed monthly compensation for writing articles and other duties within the company. Financial incentives generally have been found to have positive effects on worker output (Lazear, 2018). Professionals have an interest in the platform's overall success (besides their individual success) because their jobs depend on it. The overall success of the platform depends on the ability to leverage external contributions to grow more quickly than otherwise using only internal resources (Subramanian et al., 2021; Rietveld and Schilling, 2021). Size is a key component of platform competition alongside platform identity (Cennamo, 2021). When the provision of UGC drops, professionals might need to step up to secure a) sufficient content to not lose readers, and b) sufficient specialized content, allowing the platform to keep its unique identity (as serving Austria with local news, in this case).

Next, we compare the writing behavior of experienced citizen journalists to that of experienced and inexperienced professional journalists in order to shed additional light on why especially inexperienced professionals react to the reduction in the production of content by experienced citizens after the removal of the status index as opposed to the experienced professionals. We argue that the inexperienced professionals are the most likely candidates to step up and fill the gap left by the departing experienced citizens because they face lower adjustment and opportunity costs as they are more similar to experienced citizens in terms of the quantity and type of content provided (Argyres et al., 2019, 2022). Table 5 shows the similarity in the quantity and type of content produced by experienced and inexperienced professionals compared to the quantity and type of content produced by experienced citizens (the omitted group in these regressions).

While there are differences in absolute levels between experienced citizens and inexperienced pro-

fessionals, the content produced by the latter is significantly closer to the content produced by experienced citizens when compared to the content produced by experienced professionals. To show this, we regress several measures of output on the type of journalist using experienced citizen journalists as the omitted group. Table 5 shows the results obtained. Column (1) shows that experienced professionals write many more articles when compared to both inexperienced professionals and experienced citizens. There is no statistical and economic difference in the number of articles produced by experienced citizens and inexperienced professionals. In column (2), it is clear that experienced professionals also receive more clicks than inexperienced professionals relative to experienced citizens. This also holds for the amount of political content produced (column 3) and the localness of articles written (column 4). Columns (3) and (4) suggest that it may be harder for experienced professionals to adjust the type of content they produce, when compared to inexperienced professionals, in particular, to cater to the readers previously served by experienced citizens. We also test for equality between experienced and inexperienced professional journalists and find that they statistically differ in all these dimensions. In conclusion, in all columns of Table 5, the coefficients for inexperienced professionals are closer to zero, and thus they are more similar to experienced citizen journalists.

Finally, we also want to ensure that our results are driven by changes in quantity and not the quality of the content being produced. We conduct a formal test of the potential bias coming from unobservables. In particular, there could be a concern that the increase in the output produced by inexperienced professionals might be the result of a change in the quality of the articles written by the experienced citizens and not a result of the change in the number of articles that they write. We follow Altonji et al. (2005) to understand the potential role played by such potential unobservables in our results. This approach has been used before in Economics, Marketing, and Information Systems papers such as Petrova et al. (2021); Chae et al. (2022); Pattabhiramaiah et al. (2022). We need two entities to perform this formal test: (1) β_f, which is the coefficient of interest from the regression including a control for quality (in our case, proxied by the number of clicks) and (2) β_r, which is the coefficient of interest from the regression which includes no such control. When the unobservables are positively correlated with the observables (which is likely our case), we assess the potential bias coming from unobservables using the ratio $\frac{\beta_f}{\beta_r - \beta_f}$. As discussed at length in (Altonji et al., 2005), the intuition behind this formula is simple. First, the smaller the difference between β_r and β_f, the less the estimate of interest is affected by selection on observables, and the stronger the selection on unobservables needs to be (relative to observables) to explain the entire effect. Next, the larger the β_f,

the greater the effect that needs to be explained by selection on unobservables, and therefore the higher the ratio. This approach allows us to determine how much stronger the selection on unobservables would have to be compared to the selection on observables in order to fully explain our results.

Column (5) in Table 3 reproduces the results obtained before using instrumental variables, for sake of comparison. The results in column (6) are obtained by running the exact same regression as in column (5) but adding a control for lagged engagement, namely the number of clicks received by the articles written by experienced citizens in the previous week. The coefficients of interest in these two columns are very similar, and the ratio of interest ($\frac{\beta_f}{\beta_r - \beta_f}$) is about 58, which suggests that selection on unobserved quality would have to be 58 times the selection on observed quantity to explain all of the remaining observed effect. This magnitude is significantly higher than the thresholds discussed in prior literature (Petrova et al., 2021; Chae et al., 2022) and gives us comfort in claiming that the spillovers identified in our paper come from effects through quantities, that is, it is the change in the number of articles produced by experienced citizens that affects the number of articles produced by inexperienced professionals.

5.4 Robustness Checks

Next, we provide results from several additional analyses that show the robustness of our results:

Placebo dates We use data from exactly the same time window in 2016 and rerun our baseline specifications for experienced citizens and inexperienced professionals. The idea behind this analysis is that if the observed changes to the number of articles written by these journalists around the removal of the status index are due to other structural factors, then we should also observe them in an earlier year. We set the platform change to the last week of September 2016 as a false event date and obtain the results in Table 3. Columns (1) and (2) show that this false placebo date leads to a statistically and economically insignificant impact on content production.

Different measures of experience We carry out a check to ensure that the choice of a median cutoff does not drive our results for experience. Hence, we check whether our estimates move in line with our intuition as we use different thresholds (25th, 75th, 10th, and 90th) in Table A.5 to find that it is the most experienced citizens and the least experienced professionals who respond to the platform change. In Table A.6, we use an alternative measure of experience based on a reconstruction

of the total points an individual would have accumulated to find qualitative and quantitatively similar results.

Differences-in-differences We carry out a difference-in-differences analysis using journalists that write sponsored content as a control group. We estimate coefficients for the control group, for experienced citizen journalists and inexperienced professional journalists separately to isolate what happened to each group before and after removing the status index. The results in Columns (3) and (4) of Table 3 show that there is no statistically (and economically) significant difference in the content produced by the control group before and after this event. However, and in line with our baseline results, we find that the number of articles written by inexperienced professionals increased (Column (3)) while the number of articles written by experienced citizens decreased (Column (4)). Also, and as expected, the magnitude of the effects reported by this analysis mimics those obtained before.[20]

Non-linear models We also demonstrate that our results are robust to using non-linear models. We use two of the most commonly used non-linear models for our count variable setting (number of articles), the Poisson and the Negative Binomial. The results obtained, shown in Table A.9, are similar to the ones reported before.[21]

Other tests We use Google Trends data for the regions where articles are published in to further control for events that might heterogeneously drive the news cycle in different regions. Columns (1) and (2) in Table 4 show that our results are robust even after doing so.[22] In columns (3) and (4) of this table, we show that dropping 6 journalists who changed status between citizens and professionals from the sample leaves our results unchanged. Finally, in columns (5) and (6) of this table, we find that our results hold after controlling for lagged (weekly) success in terms of clicks.

6 Impact on Isolated Regions and Local News Content

Next, we analyze the impact of the imperfect substitution of the content written by experienced citizens by the additional content written by inexperienced professionals on important social elements, such as the coverage of news in more isolated regions and, in general, the coverage of local news.

[20] In Table A.7 in the Appendix, we also estimate a model with leads and lags to validate the difference-in-differences results further.

[21] Negative Binomial model relaxes some of the relatively stringent assumptions of the Poisson model (equidispersion) (Wooldridge, 2010).

[22] We provide the details of how we utilize the Google Trends data in the Appendix.

6.1 Impact on More Isolated Regions

There are more citizen journalists than professional journalists at Meinbezirk. Also, citizen journalists are more geographically dispersed. Therefore, the UGC at Meinbezirk can potentially come from more regions than the content written by professional journalists, and the imperfect substitution of content from experienced citizens by content from inexperienced professionals may have an effect on the representation of regions on the platform. We study this effect by analyzing what happens to citizen and professional content before and after the platform change in more isolated regions relative to less isolated regions. We follow Campante and Do (2014) and use the distance to the State capital as a measure of isolation. For that, we search for the names of all 122 regions in our dataset using the Google Maps API. We use the name of the region as the origin and the respective State capital as the destination. We then take the returned average distance and travel time (driving) to measure the region's isolation. Our focal variable of isolation is based on driving time to account for distance as well as the altitude of the regions.[23] Finally, we analyze the heterogeneity, as moderated by this measure, in the production of content at the regional level in the 8-week window before and after removing the status index.

The key results of this analysis can be seen in Table 6, where we regress the logarithm of the number of articles written in a region for experienced citizen journalists and inexperienced journalists. Column (1) shows that inexperienced professionals increase their output after removing the status index but less so in more isolated regions. At the same time, Column (2) shows that experienced citizens reduce their output the most in these regions. Overall, in column (3), we see that regions that are state capitals see a small reduction in the total number of articles produced but as the isolation increases, the number of articles written decreases. The overall effect can be seen in column (4) which shows that despite inexperienced professionals stepping in, there is a significant overall decline in content. Finally, and in line with our baseline results, there is no effect on the number of articles written by inexperienced citizens (Column (5)) and experienced professionals (Column (6)), both in total and in more isolated regions. These results indicate an important dimension of *complementarity* in the *type* of content produced by professionals and citizens. Citizens represent regions in the country better, and thus may potentially cater their content to the long-tailed preferences of readers in a way that is hard to replicate for professionals.

[23] Two regions can be close together in terms of distance but can take long to get to from one another if they are at different altitudes. This is often the case in Austria. Our results are qualitatively similar if we use a distance measure.

6.2 Impact on Local News

The fact that there are more citizen journalists than professional journalists at Meinbezirk, and that citizen journalists are more geographically dispersed, may also moderate the effect of removing the status index on the *type* of content written. In this section, we focus on the important dimension of local vs. national content, as this split has been shown to drive numerous relevant political, societal, and economic outcomes (Oberholzer-Gee and Waldfogel, 2009; Drago et al., 2014; Campante and Do, 2014; Snyder and Strömberg, 2010). We employ two different approaches to identify local and national content on the platform and analyze each of them below.

6.2.1 Measuring Localness using Word embeddings

Our main approach to identifying local and national content on the platform uses natural language processing via word embedding models. Word embedding models address shortcomings of, e.g., keyword-based approaches by identifying the location of topics in a higher dimensional vector space depending on the co-occurrence of words instead of relying on a set of pre-specified keywords (Mikolov et al., 2013a). Therefore, this method is less dependent on correctly specifying the dictionary of keywords ex-ante and has gained popularity for natural language processing applications (see, e.g. Gennaro and Ash, 2021).

We rely on a classification approach similar to Gennaro and Ash (2021). First, we train a 300-dimensional word embeddings model based on the entire corpus of news articles available on the platform since its beginning (over 2 million articles with 940,000 unique words after removing common German words) using a skip-gram specification (10-word window, 10 iterations, hierarchical softmax) of the word2vec algorithm (Mikolov et al., 2013a,b). Second, we classify all articles using this model. For this purpose, we decompose articles into words, remove common German words (i.e., stop words), obtain the vector representation of these words, and construct an article-level embedding by averaging the individual word embeddings of the article, inversely weighted by their frequency in our corpus. This results in a vector representation of each article. Third, we retrieve the vector representation of each journalist's hometown (as stated on the platform). We then calculate the cosine similarity between the article and hometown vectors. This measure, which we call localness below, leverages the locality of journalists directly and indicates how "close" the article's content is to a journalist's hometown. We then aggregate these statistics at the journalist-week level.

Columns (1) through (4) in Table 7 summarize the effect of the removal of the status index on

the localness of the content produced. Column (1) shows a statistically significant decrease in the similarity between the content of an article and the journalist's hometown for articles from experienced citizen journalists. This decline is 33% on average and thus also economically meaningful (the effect is -0.0115 out of 0.035 for the average of localness for these journalists). In line with our baseline results, column (2) shows no significant change in the localness of the articles written by inexperienced citizens. Column (3) shows a 35% increase in the localness of the articles written by inexperienced professionals (0.0842 relative to the average of 0.228 of localness for these journalists) and Column (4) shows no significant change in the localness of the articles produced by experienced professionals. In line with previous results, we see an effort from inexperienced professionals to fill the gap left by experienced citizens.

Furthermore, we classify two sets of keywords to identify the location of more local versus more national content in the 300-dimensional vector space. For local content we use the official list of 2117 municipality names in Austria, for which we retrieve the individual embeddings and average them to obtain a vector representation of localness.[24] For national content, we use the list of nine state names[25] and "Austria" as keywords, for which we again retrieve individual embeddings and average them out into one vector representation of nationalness. We then determine the cosine similarity of the article embedding vector and the local and national embedding vectors, $Sim(Local)$ and $Sim(National)$, respectively. The former indicates the extent to which an article covers local topics and the latter indicates the extent to which an article covers national topics. Lastly, we compute the ratio of $\frac{Sim(Local)+b}{Sim(National)+b}$ where b represents a smoothing parameter, which we set to 1 (c.f., Gennaro and Ash, 2021). This results in a relative measure of localness vs. nationalness for each article, where larger values indicate more local content. We then aggregate this measure to the journalist-week level.

Columns (5) through (7) of Table 7 report how the removal of the status index changed this measure. Column (5) shows that the articles written by experienced citizens shift significantly towards more national content with an 11% decrease in localness (0.134 relative to a mean of 1.20). On the other hand, Column (6) shows that the articles from inexperienced professionals shift significantly towards more local content with a 21% increase in localness (0.373 relative to a mean of 1.78).[26] Column (7) in this table shows that overall the content in the platform shifts away from local news,

[24] We obtained the names of municipalities from the statistical agency of Austria https://www.statistik.at/web_de/klassifikationen/regionale_gliederungen/gemeinden/index.html and carried out cleaning of common words and state names.

[25] Burgenland, Kärnten, Niederösterreich, Oberösterreich, Salzburg, Steiermark, Tirol, Vorarlberg, and Wien.

[26] We do not find significant changes in the localness of the articles written by inexperienced citizens or experienced professionals.

by about 6%. This negative net effect on the overall localness of the articles on the platform is a consequence of the sheer number of citizen journalists, and the decline in the localness of their articles, which still outweighs the increase in localness associated with the articles produced by the smaller number of professional journalists.

6.2.2 Measuring Localness using Keywords

We use three alternative keyword-based approaches to classify the localness and nationalness of articles. First, we mark an article as local if it includes the official name of any of the 2117 municipalities in Austria. Similarly, we mark an article as national if it includes the name of any of the 77 official political districts in the country [27]. This list of locations represents "less local" regions relative to the 2117 municipalities. Second, we still use the list of 2117 municipalities to mark articles as local, but now, recognizing that the above list of terms to categorize the nationalness of articles may not be comprehensive, we mark an article as national if it mentions "Austria" or any of the names of the 9 states in the country. This measure is more geographically aggregated than the 77 political districts and significantly more aggregated than the 2117 municipalities previously used. Third, we mark articles as local if they are tagged with the tag "Lokales" on the platform. Article tags are assigned by the journalists themselves on the platform, and this tag, which translates to "Local", is used by journalists to identify articles that focus on local content. We then regress the difference between the local dummy and the national dummy for each article on the removal of the status index for experienced citizen journalists and inexperienced professional journalists.[28] Table A.10 in the Appendix shows the results that we obtain. In all cases, we find that experienced citizen journalists reduce the localness of the articles that they produce while inexperienced professionals increase it.

7 Implications for Platform Engagement

Our analysis so far documents spillover effects in the production of content from experienced citizen journalists to inexperienced professional journalists. Our analyses also show a decline in the amount of local content available on the platform after the removal of the status index that arises from the fact that experienced citizens produce fewer local articles creating a gap that professional journalists do not fill. Still, and overall, what does this imply for engagement at the platform-level, and potentially

[27] The list of political districts was obtained from https://www.statistik.at/web_de/klassifikationen/regionale_gliederungen/politische_bezirke/index.html

[28] We use the difference instead of relative measures to account for zeros in the data at the journalist-week level.

for the click-based advertising revenue that sustains Meinbezirk? Do these dynamics between citizens and professionals increase or reduce engagement, in terms of page views?

We analyze the change in engagement associated with the removal of the status index, and thus associated with the documented spillovers between citizens and professionals, for local and national content at the article level. Several factors can affect engagement at the article level. First, there is a decline in the amount of overall content after the removal of the status index. *A priori* the effect of this change may be ambiguous, and there can be an increase in total engagement even when the overall supply of content reduces because of the potential oversupply of low-quality content before the removal of the status index (c.f. Boudreau, 2012; Boudreau and Jeppesen, 2015; Ransbotham and Kane, 2011). Indeed, and as pointed out earlier, this platform, which hosts a network of 122 news outlets, allows for complete freedom regarding who posts online, with no editorial control regarding where an article gets placed on the website. Articles are placed sequentially as they are uploaded starting from the top of the page. Given these platform features, it is unclear *ex-ante* what could happen to click-based advertising revenue after the reduction in UGC. Second, and in addition to the reduction in the amount of content, there is also a change in the composition of journalists on the platform. Experienced citizens drop off the platform, and inexperienced professionals attempt to plug the gap in content. This implies that after the removal of the status index, the population of journalists producing local news (both citizens and professionals) has also changed, which could also impact engagement. In fact, the statistics in Table 1 indicate that some experienced citizens can potentially write more engaging articles than inexperienced professionals. This table shows that the most experienced citizens (e.g., those at the 95th percentile) outperform by far the median professionals. Therefore, the substitution of top experienced citizens by median inexperienced professionals can result in a reduction of overall engagement.

We estimate article-level models regressing the logarithm of clicks per article on the platform change to measure the impact of the removal of the status index on engagement. The estimates in Panel A of Figure 6 show the effects on engagement for local news, national news, and both of them.[29] After the removal of the status index, clicks per article declined by about 9.5% for national news and 11.5% for local news. While the mean estimate for local news shows a more significant decline in clicks, it comes with the caveat that confidence intervals are wide, and we cannot reject the null hypothesis of no statistical difference between the effect on local and national news. On aggregate,

[29] We use the keyword count approach to label local and national articles.

though, we observe that clicks per article declined by about 12% (a statistically significant result at the 1% level). In Panel B of this figure, we analyze the heterogeneity of these effects along citizen and professional journalists. For citizen journalists, the reduction in engagement arises for both local and national news, though the former experience a more significant average decline. Also, there is a reduction in engagement for all local news, both from citizens and professionals.

Finally, we note that the decrease in the number of clicks per article could also arise from the fact that a shrunken journalist base may attract fewer readers, making it even less attractive to write articles. On the other hand, adverse same-side network effects, whereby articles compete for the attention of readers, could lead to an increase in the number of clicks per article after removing the status index and the number of articles on the platform reduces. We use the modeling approach in Cullen and Farronato (2021) and Boudreau and Jeppesen (2015) to test for such network effects.[30] In particular, we analyze how the total daily number of articles written affects the number of clicks per article. Furthermore, we test whether the magnitude of these network effects changed due to removing the status index. Columns (2) and (3) in Table 8 show that a 10% increase in the total number of daily articles on the platform reduces the number of clicks per article by 2.9% and 2.4% before and after removing the status index, respectively, which suggest that articles compete for attention in our setting. Column (1) combines the pre and post-period while column (4) shows explicitly that the magnitudes did not change significantly with the platform change. In sum, the presence of adverse same-side network effects, in fact, suggests that our analyses above are likely to provide a (negative) effect at the lower end because the actual effect is likely partially counteracted by the decrease in competition for attention. Moreover, the fact that these effects do not change due to the platform change suggests that there is limited interaction between the demand and supply side of the platform around the time of the shock that we analyze.

Overall, while descriptive in nature, these results shed light on the implications of UGC for platform engagement and, consequently, for its revenue. UGC can boost revenue by generating clicks, especially for "long-tail" content where UGC may have a comparative advantage. Additionally, UGC is generated without the need for news outlets to incur additional wage costs, unlike professional journalists. Although we only analyze the short-term engagement consequences in this section, these implications are likely to have lasting long-term effects due to the dynamics of consumption and contribution in

[30] As an example, in their baseline analysis, Cullen and Farronato (2021) analyze tasks per buyer as a function of the number of buyers and sellers on Task Rabbit (their partner platform) in a particular month. Boudreau and Jeppesen (2015) analyze how the platform users respond to the (lagged) number of complementors.

UGC (Kane and Ransbotham, 2016). Finally, a reduced engagement with local news content can have adverse implications for local knowledge and cohesion (Campante and Do, 2014).

8 Discussion

Today, amateurs are increasingly participating in tasks that have been the exclusive realm of professionals in the past. This often leads to a situation where UGC and content from professionals coexist side by side on digital platforms (Subramanian et al., 2021). This paper studies the substantial complementarity and substitutability between these types of content in such online platform ecosystems. In particular, we study the dynamics between UGC and professional content in terms of the *amount* and *type* of content produced in the critical domain of local news. Using data from a large online platform, we show how the decrease in the supply of UGC affects the output of professional journalists. More specifically, we show that after removing a status index used to incentivize external contributions, experienced citizen journalists decrease their contribution or leave the platform altogether. This reduction in UGC creates heterogeneous spillover effects on professional journalists. While experienced professional journalists are largely unaffected, inexperienced professionals react by increasing their output significantly. However, there is a decrease in the overall content on the platform without the status index. Using natural language processing, we show that this reduction is especially pronounced for local content and citizen journalists associated with more isolated regions. This content is likely more costly to produce for professional journalists and lends itself better to the realm of UGC production (Afuah and Tucci, 2012; Dellarocas et al., 2010). We further document that this decline has detrimental effects on engagement with the platform, impacting the top-line revenue of the platform, as well as the bottom line, given the potential need to hire more professional journalists in the medium to long run.

8.1 Implication for platforms with UGC

Our paper highlights the interdependency in content production across heterogeneous agents in multi-actor platform ecosystems. This has important practical implications for platform owners who want to benefit from using UGC. The literature and practice have highlighted the potential of using symbolic incentives to encourage individual users to contribute (see, e.g. Gallus, 2017; Burtch et al., 2019; Goes et al., 2016). Our results go beyond prior work and suggest that removing a system allowing

users to signal status affects not only the focal group, citizen journalists in our case but also other agents on the platform, professional journalists in our case, through significant spillover effects. In our context, journalists in regional offices (often comprising less than five individuals) appeared to have responded to this decline in citizen content organically without a directive from the headquarter. This echoes the finding in Hui et al. (2020), which demonstrates spillovers onto other groups on the platform when (financial) incentives are targeted to a subset of individuals on e-Bay. Platforms should account for these (general) equilibrium effects while optimizing their growth strategies. Our finding that inexperienced professionals increase output, at least partly to cover for the content experienced citizens previously produced, demonstrates that UGC and professional content can be seen as partial substitutes. Hence, platforms should aim to leverage such status signals, which bind users to the platform long-term in ways that support their contributions on par with those from professionals.

Our findings also highlight the importance of relative adjustment and opportunity costs (Argyres et al., 2022) that impact the degree to which professionals can substitute UGC. As a result, platforms built around UGC need to consider the specificity and capabilities of their resources for how they may best address fluctuations in UGC and content generated by professionals. Our study shows that there is content that cannot easily be picked up by professionals, especially long-tail content that can only be covered with specialized, i.e., "local" knowledge. For this type of content, platforms can leverage the 'crowd' to complement professional content. The findings of our study are especially relevant for information markets (Cennamo, 2021; Tajedin et al., 2019), i.e., platforms with a broad scope whose business model relies on the quality of the match between users and the information provided. Enhancing the quality of such matches is becoming increasingly important for the platform because platforms themselves are also generating content and attempting to cater to the heterogeneous preferences of their user base (e.g., YouTube Originals). Examples of platforms catering to a long tail of users with differentiated preferences include review platforms like Yelp.com or tripadvisor.com (Mayzlin et al., 2014) but also more specialized platforms in domains like health care (e.g., patient-innovation.com (Rauch and Ansari, 2022)).

8.2 Implication for Local News

We also contribute to the literature on the production of local news, or rather the lack thereof. The news industry is essential because it produces a good with significant information externalities. Prior studies have highlighted the important societal role of local news (see, e.g. Oberholzer-Gee

and Waldfogel, 2009; Snyder and Strömberg, 2010; Drago et al., 2014; Campante and Do, 2014) but also documented the adverse effects that digital platforms can have on their production (Seamans and Zhu, 2014, 2017; Matherly and Greenwood, 2021). We show that involving UGC in local news production can be an attractive solution to cover a diverse range of topics due to the large number and geographic dispersion of users that may contribute. Involving users might not be the silver bullet to fighting the societal challenges that may stem from a lack of local news, such as corruption (Matherly and Greenwood, 2021). Still, and compared to professionals, involving "local" users may be not only likely to be more cost-effective, from the platform's perspective, but also likely to produce high-quality content because users may have easier access to "local knowledge" (Afuah and Tucci, 2012), thus likely increasing overall societal welfare.

Therefore, our paper also adds a different perspective to the role of digital platforms in the production of local news. While prior research has stressed the detrimental role that platform entry, especially from news aggregators, can play for local news, we highlight a potentially positive role: By pooling diverse and dispersed users with similar interests and involving them in producing local news, digital platforms can facilitate access to local news in areas that otherwise might not get enough attention from professional journalists and would become news deserts (Campante and Do, 2014; Matherly and Greenwood, 2021).

8.3 Limitations and Future Research

Beyond these contributions, our research also demonstrates promising avenues for future research. For example, while we demonstrate that professionals react to shocks in the supply of UGC both in terms of the *amount* and *type* of content they produce, future research could further investigate the continuous co-specialization or convergence in content between users and professionals. In our empirical setting, the most relevant partitioning of content is geographical. However, it would be interesting to see whether this co-specialization can also occur in online communities split along other dimensions. Finally, it is important to note that while we follow the existing literature in our empirical approach to control for network effects, the causal identification of such effects is notoriously difficult. Therefore, future research could pick up this aspect and investigate the presence and strength of network effects between different groups of platform contributors and different types of content in more detail. Concerning local news, future research could also investigate the downstream impact of more (or less) news coverage from users and professionals. For example, it would be interesting to see

whether the composition of news coverage (e.g., only users vs. only professionals vs. mixed) impacts outcomes such as election turnout, political outcomes, or crime.

References

Aaltonen, A., and Seiler, S. (2016). "Cumulative growth in user-generated content production: evidence from wikipedia." *Management Science*, *62*(7), 2054–2069.

Afuah, A., and Tucci, C. L. (2012). "Crowdsourcing as a solution to distant search." *Academy of Management Review*, *37*(3), 355–375.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). "Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools." *Journal of Political Economy*, *113*(1), 151–184.

Anderson, A., Huttenlocher, D., Kleinberg, J., and Leskovec, J. (2013). "Steering user behavior with badges." In *Proceedings of the 22nd international conference on World Wide Web*, 95–106.

Argyres, N., Mahoney, J. T., and Nickerson, J. (2019). "Strategic responses to shocks: Comparative adjustment costs, transaction costs, and opportunity costs." *Strategic Management Journal*, *40*(3), 357–376.

Argyres, N., Nickerson, J., and Ozalp, H. (2022). "Platform competition and complementor responses: insights from combining design rules with the comparative adjustment, transaction, and opportunity cost framework." *Industrial and Corporate Change*.

Baldwin, C., and Von Hippel, E. (2011). "Modeling a paradigm shift: From producer innovation to user and open collaborative innovation." *Organization science*, *22*(6), 1399–1417.

Beattie, G., Durante, R., Knight, B., and Sen, A. (2021). "Advertising spending and media bias: Evidence from news coverage of car safety recalls." *Management Science*, *67*(2), 698–719.

Boudreau, K. (2019). "'crowds' of amateurs & professional entrepreneurs in marketplaces." *Working Paper*, *Available at: https://ssrn.com/abstract=2988308*.

Boudreau, K. J. (2012). "Let a thousand flowers bloom? an early look at large numbers of software app developers and patterns of innovation." *Organization Science*, *23*(5), 1409–1427.

Boudreau, K. J., and Jeppesen, L. B. (2015). "Unpaid crowd complementors: The platform network effect mirage." *Strategic Management Journal*, *36*(12), 1761–1777.

Bughin, J. R. (2007). "How companies can make the most of user-generated content."

Burtch, G., He, Q., Hong, Y., and Lee, D. (2019). "Peer recognition increases user content generation but reduces content novelty." *Working Paper*, *Available at: https://ssrn.com/abstract=3465879*.

Cagé, J. (2020). "Media competition, information provision and political participation: Evidence from french local newspapers and elections, 1944–2014." *Journal of Public Economics*, *185*, 104077.

Cage, J., Herve, N., and Mazoyer, B. (2020). "Social media and newsroom production decisions." *Available at: https://ssrn.com/abstract=3663899*.

Campante, F. R., and Do, Q.-A. (2014). "Isolated capital cities, accountability, and corruption: Evidence from us states." *American Economic Review*, *104*(8), 2456–81.

Cennamo, C. (2021). "Competing in digital markets: A platform-based perspective." *Academy of Management Perspectives*, *35*(2), 265–291.

Chae, I., Ha, J., and Schweidel, D. A. (2022). "Paywall suspensions and digital news subscriptions." *Marketing Science*.

Chiou, L., and Tucker, C. (2017). "Content aggregation by platforms: The case of the news media." *Journal of Economics & Management Strategy*, *26*(4), 782–805.

Cullen, Z., and Farronato, C. (2021). "Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms." *Management Science*, *67*(7), 3985–4003.

Dellarocas, C., Gao, G., and Narayan, R. (2010). "Are consumers more likely to contribute online reviews for hit or niche products?" *Journal of Management Information Systems*, *27*(2), 127–158.

Dellarocas, C., Sutanto, J., Calin, M., and Palme, E. (2016). "Attention allocation in information-rich environments: The case of news aggregators." *Management Science*, *62*(9), 2543–2562.

Deng, Y., Zheng, J., Khern-am nuai, W., and Kannan, K. (2021). "More than the quantity: The value of editorial reviews for a user-generated content platform." *Management Science*.

Drago, F., Nannicini, T., and Sobbrio, F. (2014). "Meet the press: How voters and politicians respond to newspaper entry and exit." *American Economic Journal: Applied Economics*, *6*(3), 159–88.

Fischer, S., Jaidka, K., and Lelkes, Y. (2020). "Auditing local news presence on google news." *Nature Human Behaviour*, *4*(12), 1236–1244.

Gallus, J. (2017). "Fostering public good contributions with symbolic awards: A large-scale natural field experiment at wikipedia." *Management Science*, *63*(12), 3999–4015.

Gennaro, G., and Ash, E. (2021). "Emotion and Reason in Political Language*." *The Economic Journal*, ueab104.

George, L. M., and Waldfogel, J. (2006). "The new york times and the market for local newspapers." *American Economic Review*, *96*(1), 435–447.

Goes, P. B., Guo, C., and Lin, M. (2016). "Do incentive hierarchies induce user effort? evidence from an online knowledge exchange." *Information Systems Research*, *27*(3), 497–516.

Greenstein, S., and Zhu, F. (2018). "Do experts or crowd-based models produce more bias? evidence from encyclopædia britannica and wikipedia." *MIS Quarterly*, *42*(3).

Hermida, A., and Thurman, N. (2008). "A clash of cultures: The integration of user-generated content within professional journalistic frameworks at british newspaper websites." *Journalism practice*, *2*(3), 343–356.

Hui, X., Liu, M., and Chan, T. (2020). "Targeted incentives, broad impacts: Evidence from an e-commerce platform." *Broad Impacts: Evidence from an E-commerce Platform (October 11, 2020)*.

Jeppesen, L. B., and Frederiksen, L. (2006). "Why do users contribute to firm-hosted user communities? the case of computer-controlled music instruments." *Organization Science*, *17*(1), 45–63.

Kane, G. C., and Ransbotham, S. (2016). "Content as community regulator: The recursive relationship between consumption and contribution in open collaboration communities." *Organization Science*, *27*(5), 1258–1274.

Khern-am nuai, W., Kannan, K., and Ghasemkhani, H. (2018). "Extrinsic versus intrinsic rewards for contributing reviews in an online platform." *Information Systems Research*, *29*(4), 871–892.

Kim, J. (2012). "The institutionalization of youtube: From user-generated content to professionally generated content." *Media, Culture & Society*, *34*(1), 53–67.

Kovács, B., and Sharkey, A. J. (2014). "The paradox of publicity: How awards can negatively affect the evaluation of quality." *Administrative Science Quarterly*, *59*(1), 1–33.

Kühberger, A. (1998). "The influence of framing on risky decisions: A meta-analysis." *Organizational behavior and human decision processes*, *75*(1), 23–55.

Lakhani, K. R., and Wolf, R. G. (2003). "Why hackers do what they do: Understanding motivation and effort in free/open source software projects." *Open Source Software Projects (September 2003)*.

Lazear, E. P. (2018). "Compensation and incentives in the workplace." *Journal of Economic Perspectives*, *32*(3), 195–214.

Le Mens, G., Kovács, B., Avrahami, J., and Kareev, Y. (2018). "How endogenous crowd formation undermines the wisdom of the crowd in online ratings." *Psychological Science*, *29*(9), 1475–1490.

Lerner, J., and Tirole, J. (2002). "Some simple economics of open source." *The journal of industrial economics*, *50*(2), 197–234.

Levina, N., and Arriaga, M. (2014). "Distinction and status production on user-generated content platforms: Using bourdieu's theory of cultural production to understand social dynamics in online fields." *Information Systems Research*, *25*(3), 468–488.

Levine, S. S., and Prietula, M. J. (2014). "Open collaboration for innovation: Principles and performance." *Organization Science*, *25*(5), 1414–1433.

Li, X., and Hitt, L. M. (2008). "Self-selection and information role of online product reviews." *Information Systems Research*, *19*(4), 456–474.

Lukyanenko, R., Parsons, J., Wiersma, Y. F., and Maddah, M. (2019). "Expecting the unexpected: Effects of data collection design choices on the quality of crowdsourced user-generated content." *MIS Quarterly*, *43*(2), 623–648.

Ma, M., and Agarwal, R. (2007). "Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities." *Information Systems Research*, *18*(1), 42–67.

Matherly, T., and Greenwood, B. N. (2021). "No news is bad news: Political corruption, news deserts, and the decline of the fourth estate." *Working paper*.

Mayzlin, D., Dover, Y., and Chevalier, J. (2014). "Promotional reviews: An empirical investigation of online review manipulation." *American Economic Review*, *104*(8), 2421–55.

McAfee, R. P., Mialon, H. M., and Mialon, S. H. (2010). "Do sunk costs matter?" *Economic Inquiry*, *48*(2), 323–336.

McIntyre, D. P., and Srinivasan, A. (2017). "Networks, platforms, and strategy: Emerging views and next steps." *Strategic Management Journal*, *38*(1), 141–160.

Meyer, T., Kerkhof, A., Cennamo, C., and Kretschmer, T. (2022). "Competing for attention on information platforms: The case of news outlets." *CEPR Working Paper*.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). "Efficient estimation of word representations in vector space." *arXiv preprint arXiv:1301.3781*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013b). "Distributed representations of words and phrases and their compositionality." In *Advances in neural information processing systems*, 3111–3119.

Mollick, E., and Nanda, R. (2016). "Wisdom or madness? comparing crowds with expert evaluation in funding the arts." *Management Science*, *62*(6), 1533–1553.

Oberholzer-Gee, F., and Waldfogel, J. (2009). "Media markets and localism: Does local news en espanol boost hispanic voter turnout?" *American Economic Review*, *99*(5), 2120–28.

Pattabhiramaiah, A., Overby, E., and Xu, L. (2022). "Spillovers from online engagement: how a newspaper subscriber's activation of digital paywall access affects her retention and subscription revenue." *Management Science*, *68*(5), 3528–3548.

Paulussen, S., Domingo, D., Heinonen, A., Singer, J., Quandt, T., and Vujnovic, M. (2008). "Citizen participation in online news media. an overview of current developments in four european countries and the united states." In *Journalismus Online-Partizipation oder Profession?*, 263–283, Springer.

Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). "Linguistic inquiry and word count: Liwc 2001." *Mahway: Lawrence Erlbaum Associates*, *71*(2001), 2001.

Petrova, M., Sen, A., and Yildirim, P. (2021). "Social media and political contributions: The impact of new technology on political competition." *Management Science*, *67*(5), 2997–3021.

Poetz, M. K., and Schreier, M. (2012). "The value of crowdsourcing: can users really compete with professionals in generating new product ideas?" *Journal of product innovation management*, *29*(2), 245–256.

Prat, A., and Valletti, T. M. (2021). "Attention oligopoly." *American Economic Journal: Microeconomics, Forthcoming*.

Puranam, P., Alexy, O., and Reitzig, M. (2014). "What's "new" about new forms of organizing?" *Academy of Management Review*, *39*(2), 162–180.

Ransbotham, S., Kane, G., and Lurie, N. (2012). "Network characteristics and the value of collaborative user-generated content." *Marketing Science*, *31*(3), 387–405.

Ransbotham, S., and Kane, G. C. (2011). "Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in wikipedia." *MIS Quarterly*, *35*(3), 613–627.

Rauch, M., and Ansari, S. (2022). "From 'publish or perish'to societal impact: Organizational repurposing towards responsible innovation through creating a medical platform." *Journal of Management Studies*, *59*(1), 61–91.

Reimers, I. C., and Waldfogel, J. (2020). "Digitization and pre-purchase information: The causal and welfare impacts of reviews and crowd ratings." *NBER Working Paper*, *Available at: https://www.nber.org/papers/w26776*.

Resnick, P., Zeckhauser, R., Swanson, J., and Lockwood, K. (2006). "The value of reputation on ebay: A controlled experiment." *Experimental Economics*, *9*(2), 79–101.

Riedl, C., and Seidel, V. P. (2018). "Learning from mixed signals in online innovation communities." *Organization Science*, *29*(6), 1010–1032.

Rietveld, J., and Schilling, M. A. (2021). "Platform competition: A systematic and interdisciplinary review of the literature." *Journal of Management*, *47*(6), 1528–1563.

Roberts, J. A., Hann, I.-H., and Slaughter, S. A. (2006). "Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the apache projects." *Management Science*, *52*(7), 984–999.

Seamans, R., and Zhu, F. (2014). "Responses to entry in multi-sided markets: The impact of craigslist on local newspapers." *Management Science*, *60*(2), 476–493.

Seamans, R., and Zhu, F. (2017). "Repositioning and cost-cutting: The impact of competition on platform strategies." *Strategy Science*, *2*(2), 83–99.

Sen, A., and Yildirim, P. (2015). "Clicks bias in editorial decisions: How does popularity shape online news coverage?" *Working Paper*, *Available at: https://ssrn.com/abstract=2619440*.

Snyder, J. M. J., and Strömberg, D. (2010). "Press coverage and political accountability." *Journal of Political Economy*, *118*(2), 355–408.

Stock, J. H., and Yogo, M. (2002). "Testing for weak instruments in linear iv regression."

Subramanian, H., Mitra, S., and Ransbotham, S. (2021). "Capturing value in platform business models that rely on user-generated content." *Organization Science*.

Tajedin, H., Madhok, A., and Keyhani, M. (2019). "A theory of digital firm-designed markets: Defying knowledge constraints with crowds and marketplaces." *Strategy Science*, *4*(4), 323–342.

Tirunillai, S., and Tellis, G. J. (2012). "Does chatter really matter? dynamics of user-generated content and stock performance." *Marketing Science*, *31*(2), 198–215.

Tversky, A., and Kahneman, D. (1991). "Loss aversion in riskless choice: A reference-dependent model." *The Quarterly Journal of Economics*, *106*(4), 1039–1061.

Van Mierlo, T. (2014). "The 1% rule in four digital health social networks: an observational study." *Journal of medical Internet research*, *16*(2), e2966.

Wang, R. Y., and Strong, D. M. (1996). "Beyond accuracy: What data quality means to data consumers." *Journal of management information systems*, *12*(4), 5–33.

Wasko, M. M., and Faraj, S. (2005). "Why should i share? examining social capital and knowledge contribution in electronic networks of practice." *MIS Quarterly*, *29*(1), 35–57.

Wiertz, C., and de Ruyter, K. (2007). "Beyond the call of duty: Why customers contribute to firm-hosted commercial online communities." *Organization Studies*, *28*(3), 347–376.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wu, Y., and Zhu, F. (2018). "Competition, contracts, and creativity: Evidence from novel writing in a platform market." *Working Paper*, *Available at: https://mackinstitute.wharton.upenn.edu/wp-content/uploads/2019/04/Zhu-Feng-Wu-Yanhui_Competition-Contracts-and-Creativity.-Evidence-from-Novel-Writing-in-a-Platform-Market.pdf*.

Appendix

Table 1: Characteristics of Articles Written by Professional and Citizen Journalists

Characteristic	(1) Professionals	(2) Citizens	(3) Inexperienced Professionals	(4) Experienced Citizens
Mean Clicks	284.86	137.78	202	140.14
Median Clicks	80	58	67	59
95th Percentile of Clicks	1032	427	710	434
Number of Pictures	4.49	5.67	4.55	5.77
Number of Words	180.21	98.49	173.2	97.48

The summary statistics are at the article level. The sample period is from 2016-2018.

Table 2: Impact of Platform Change on Journalistic Output

VARIABLES	Full Sample (1) Articles	Citizen Experienced (2) Articles	Professional Inexperienced (3) Articles	Citizen Experienced (4) Articles	Professional Inexperienced (5) Articles	First-Stage IV (6) Articles	Second-Stage IV (7) Articles
Platform Change	-0.0644***	-0.147***	0.421**	-0.179***	0.399**	-124.08***	
	(0.0109)	(0.0164)	(0.182)	(0.0197)	(0.185)	(3.485)	
Predicted (Citizen Content)							-0.00339***
							(0.00148)
Constant	0.344***	0.282***	1.024***	0.341***	0.947***		
	(0.00762)	(0.0104)	(0.115)	(0.0125)	(0.118)		
Journalist FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First Stage F-Stat	-	-	-	-	-	1267	-
Observations	215,883	103,530	4,182	85,986	4,199	4,182	4,182
R-squared	0.788	0.741	0.567	0.739	0.563		

The dependent variable is the number of articles written. The unit of observation is the author-week. A control for holidays in different weeks in different regions is included. Columns (4) and (5) use an alternative measure of experience based on the total number of prior articles. Column (7) uses the sub-sample of inexperienced professionals. Robust standard errors in parentheses clustered at the author level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$.

Table 3: Additional Placebos and Analysis

VARIABLES	2016 Placebo Exp. Cit. (1) Articles	2016 Placebo Inexp. Prof. (2) Articles	Prof. Trends DiD (3) Articles	Citizen Trends DiD (4) Articles	IV (5) Articles	IV with Control Lagged Engagement (6) Articles
Platform Change	0.00165	0.0244				
	(0.0200)	(0.0687)				
Platform Change x Professional			0.421**			
			(0.182)			
Platform Change x Control Group			-0.0142	-0.0142		
			(0.0169)	(0.0169)		
Platform Change x Citizen				-0.147***		
				(0.0164)		
Predicted (Citizen Content)					-0.00339**	-0.00345**
					(0.00148)	(0.00163)
Journalist FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52,853	2,839	7,395	106,743	4,182	4,182
R-squared	0.799	0.619	0.579	0.740	-	-

The unit of observation is author-week. A control for holidays in different weeks in different regions is included. Robust standard errors in parentheses clustered at the author level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$.

Table 4: Robustness Impact of Platform Change on Journalistic Output

VARIABLES	Google Trends Citizen Experienced (1) Articles	Google Trends Professional Inexperienced (2) Articles	No Switch Citizen Experienced (3) Articles	No Switch Professional Inexperienced (4) Articles	Lagged Clicks Citizen Experienced (5) Articles	Lagged Clicks Professional Inexperienced (6) Articles
Platform Change	-0.147***	0.420**	-0.147***	0.411**	-0.144***	0.388**
	(0.0164)	(0.181)	(0.0164)	(0.183)	(0.0158)	(0.177)
Constant	0.281***	1.007***	0.279***	0.960***	0.274***	0.973***
	(0.0105)	(0.119)	(0.0104)	(0.115)	(0.00993)	(0.119)
Journalist FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	103,530	4,182	103,513	4,148	103,530	4,182
R-squared	0.741	0.567	0.739	0.549	0.743	0.574

The dependent variable is the number of articles written. The unit of observation is the author-week. A control for holidays in different weeks in different regions is included. Robust standard errors in parentheses clustered at the author level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$.

Table 5: Comparing Experienced and Inexperienced Professionals to Experienced Citizens

VARIABLES	(1) Articles	(2) Engagement	(3) Political Content	(4) Localness
Experienced Professionals	8.357***	2,565***	0.128***	0.371***
	(0.577)	(196.7)	(0.0143)	(0.0287)
Inexperienced Professionals	0.0111	25.79**	0.00540**	0.0228***
	(0.0533)	(10.48)	(0.00222)	(0.00652)
Constant	0.310***	41.53***	0.00126***	0.00682***
	(0.0244)	(3.449)	(0.000161)	(0.000857)
P-val.: Exp. Prof. = Inexp. Prof.	0.00	0.00	0.02	0.00
Observations	237,024	237,024	237,024	237,024
R-squared	0.224	0.191	0.048	0.051

The unit of observation is author-week. A control for holidays in different weeks in different regions is included. Robust standard errors in parentheses clustered at the author level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$.

Table 6: Impact on Content by Level of Isolation

VARIABLES	Inexp. Prof. (1) Articles	Exp. Cit. (2) Articles	Inexp. Prof.+ Exp. Cit. (3) Articles	Inexp. Prof.+ Exp. Cit. (4) Articles	Inexp. Cit. (5) Articles	Exp. Prof. (6) Articles
Platform Change	0.415***	-0.645***	-0.227*	-0.606***	-0.093	-0.070
	(0.119)	(0.121)	(0.127)	(0.064)	(0.104)	(0.079)
Platform Change x Isolation	-0.0429***	-0.025*	-0.054***		-0.004	0.006
	(0.013)	(0.014)	(0.014)		(0.012)	(0.010)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,683	1,683	1,751	2,074	1,734	1,683
R-squared	0.651	0.685	0.663	0.659	0.495	0.867

The dependent variable is the logarithm of the number of articles written aggregated at the region level. The unit of observation is region-week. A control for holidays in different weeks in different regions is included. Robust standard errors in parentheses clustered at the region level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$.

Table 7: Impact on Local News

VARIABLES	Exp. Cit. (1) Localness	Inexp. Cit. (2) Localness	Inexp. Prof. (3) Localness	Exp. Prof. (4) Localness	Exp. Cit. (5) Local vs. National	Inexp. Prof. (6) Local vs. National	Full Sample (7) Local vs. National
Platform Change	-0.0116***	-0.00120	0.0842***	0.0337	-0.134***	0.373**	-0.0602***
	(0.00226)	(0.000892)	(0.0319)	(0.0810)	(0.0162)	(0.162)	(0.0107)
Journalist FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	103,530	103,649	4,182	4,182	103,530	4,182	215,883
R-squared	0.776	0.554	0.620	0.740	0.759	0.532	0.786

The dependent variable is the average similarity of the text of the articles written by an author in a week relative to the author's hometown based on word embeddings in columns (1)-(4) while in columns (5)-(7), the measure is the average ratio of the similarities of the article texts to the average embedding vectors of local and national keywords. The unit of observation is author-week. A control for holidays in different weeks in different regions is included. Robust standard errors in parentheses clustered at the author level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$.

Table 8: Network effects and Platform Change

VARIABLES	All Articles (1) Log(Clicks per Article)	Before (2) Log(Clicks per Article)	After (3) Log(Clicks per Article)	All Articles (4) Log(Clicks per Article)
Log(Total Daily Articles)	-0.225***	-0.299***	-0.246***	-0.287***
	(0.0333)	(0.0403)	(0.0368)	(0.0383)
Platform Change				-0.520**
				(0.214)
Log(Total Daily Articles) x Platform Change				0.0475
				(0.0328)
Journalist FE	Yes	Yes	Yes	Yes
Observations	65,665	34,231	30,784	65,665
R-squared	0.383	0.427	0.378	0.388

The dependent variable is the logarithm of the clicks per article written. The unit of observation is at the level of the article. A control for holidays in different weeks in different regions is included. Robust standard errors in parentheses clustered at the author level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$.

Figure 1: Profile of Citizen Journalist: Pre-Removal

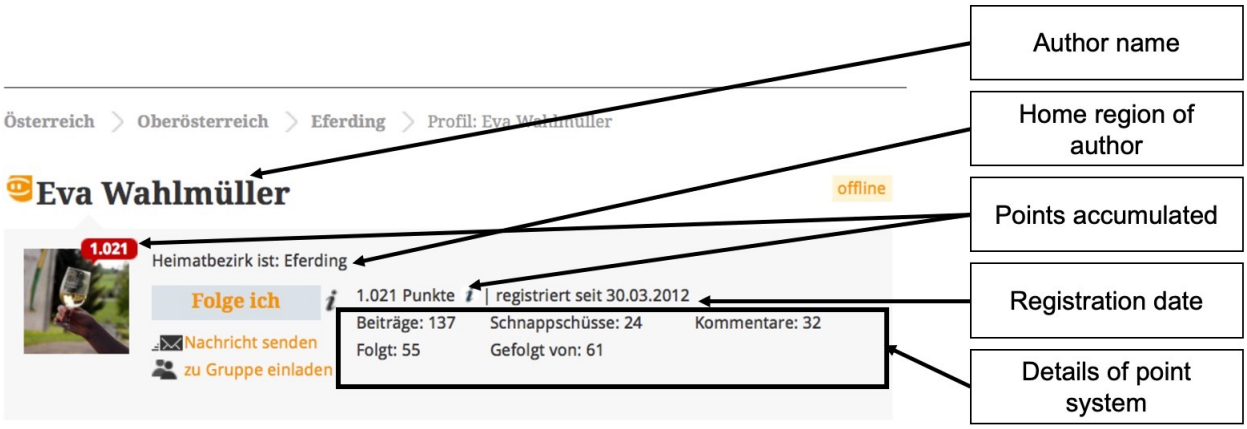


Figure 2: Profile of Citizen Journalists: Post-Removal

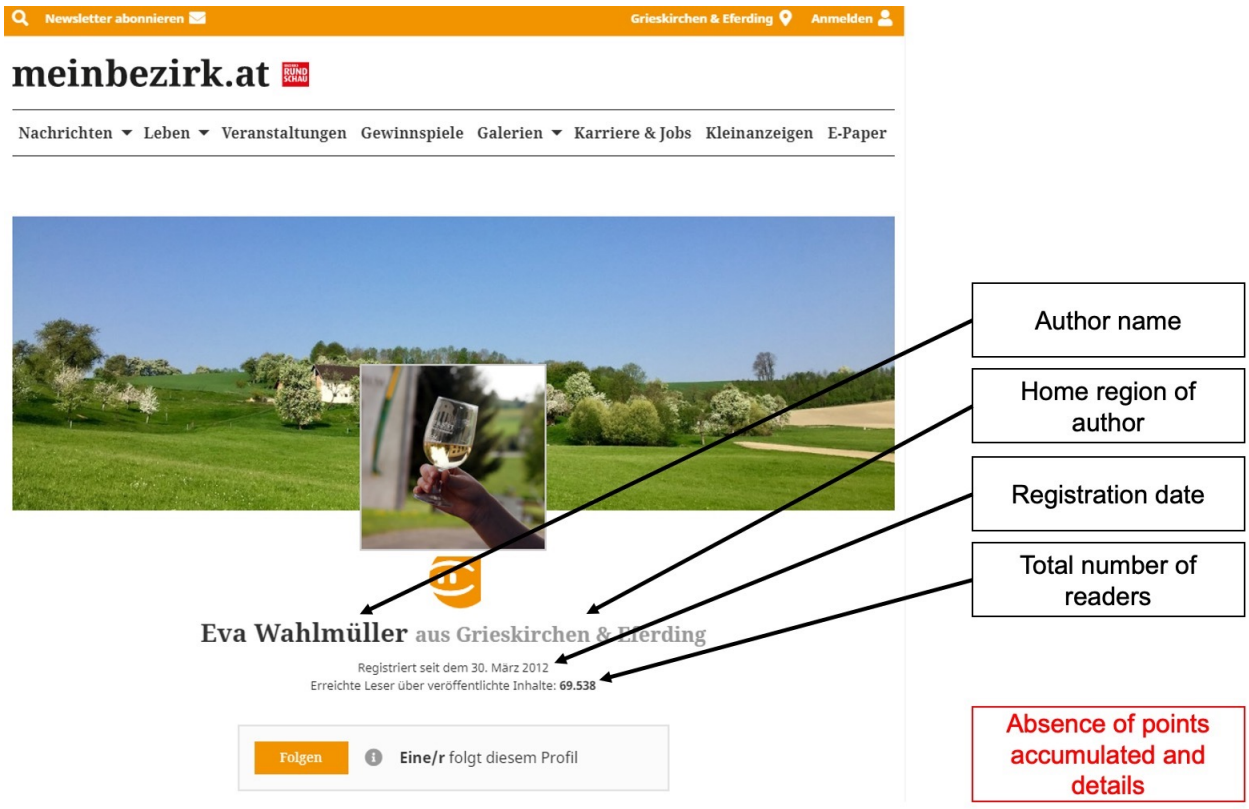
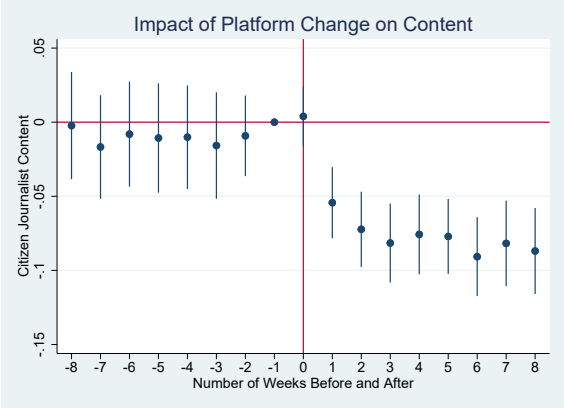
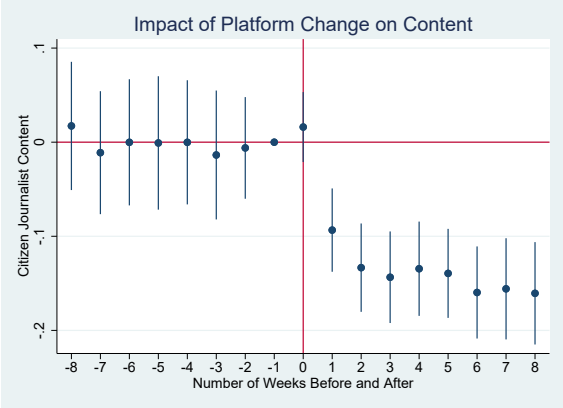


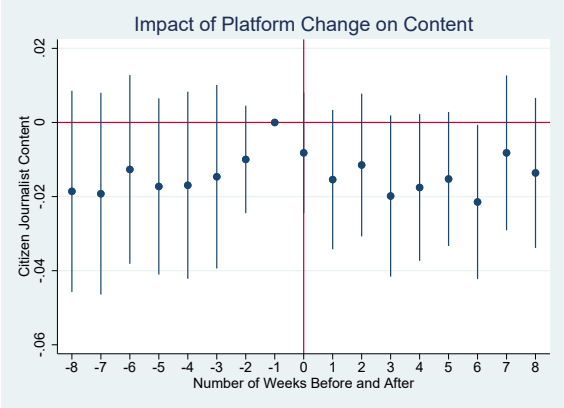
Figure 3: Impact of Platform Change on Content Production by Citizen Journalists



(a) All Citizen Journalists



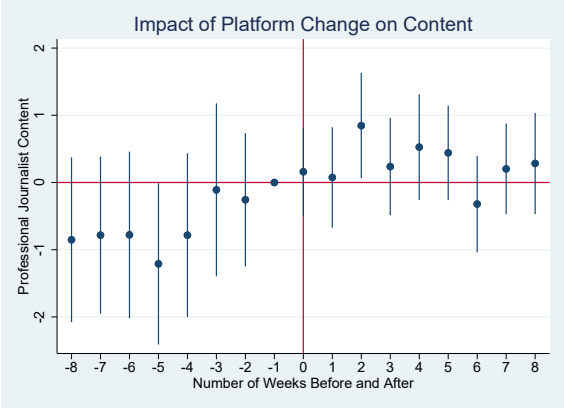
(b) Experienced Citizen Journalists



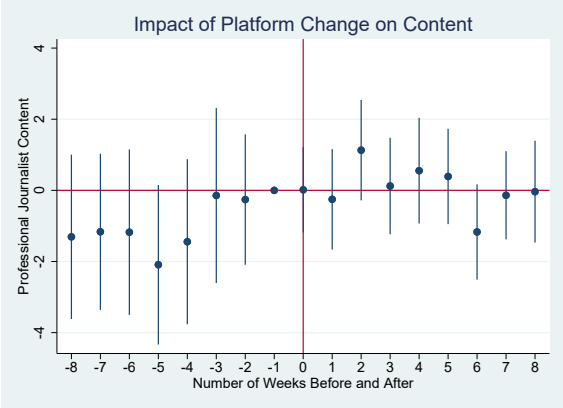
(c) Inexperienced Citizen Journalists

The figures show estimates from a regression as in equation 2. It plots the lags and leads eight weeks before and after the platform change with 99% confidence intervals. The dependent variable is the number of articles produced by an author in a week. Experience is measured based on total engagement prior to the beginning of the eight-week pre-post event study window.

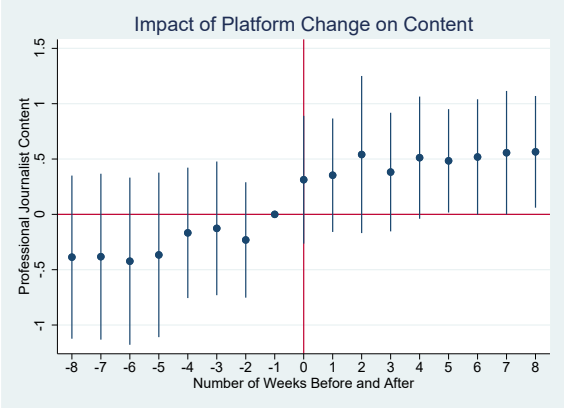
Figure 4: Impact of Platform Change on Content Production by Professional Journalists



(a) All Professional Journalists



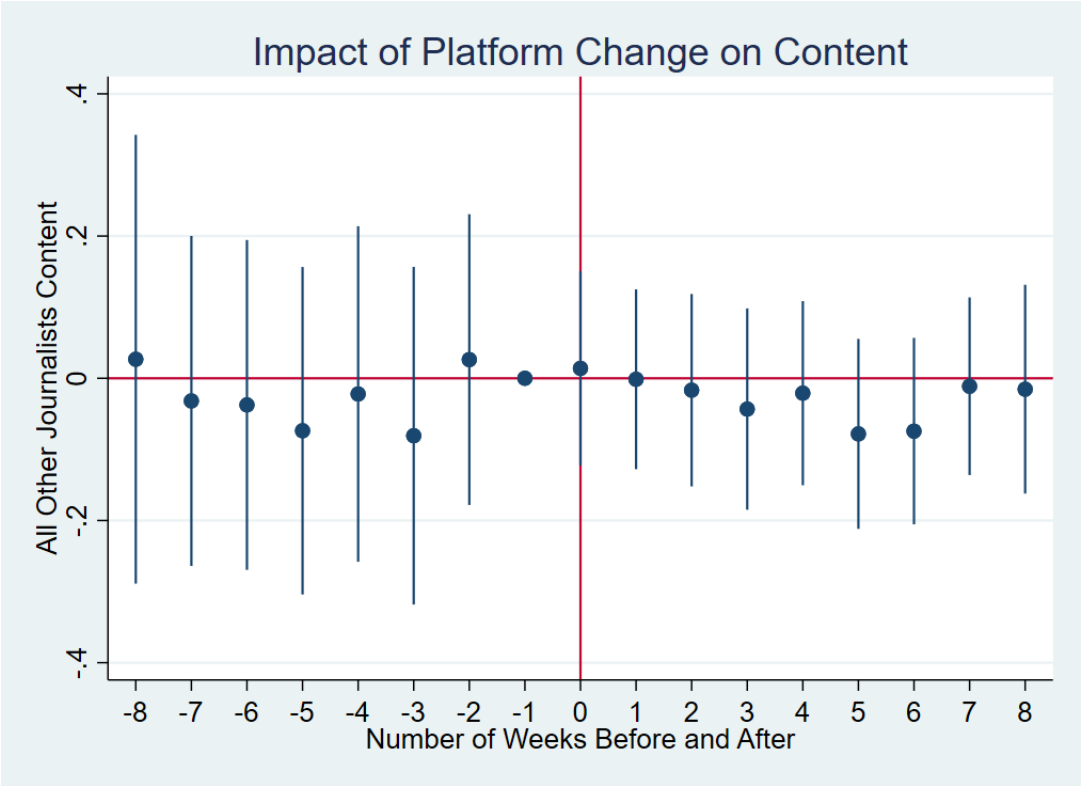
(b) Experienced Professional Journalists



(c) Inexperienced Professional Journalists

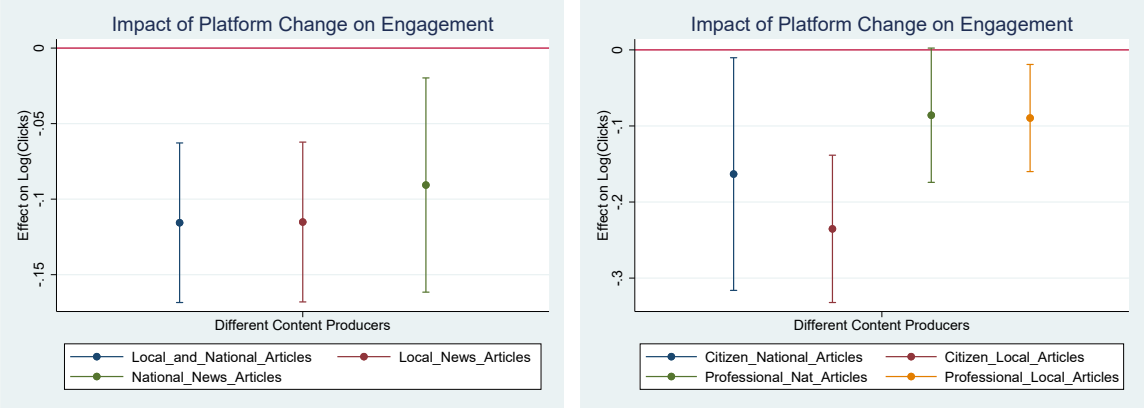
The figures show estimates from a regression as in equation 2. It plots the lags and leads eight weeks before and after the platform change with 99% confidence intervals. The dependent variable is the number of articles produced by an author in a week. Experience is measured based on total engagement prior to the beginning of the eight-week pre-post event study window.

Figure 5: Placebo: Impact of Platform Change on all other groups



The figure shows estimates from a regression as in equation 2. It plots the lags and leads eight weeks before and after the platform change with 99% confidence intervals. The dependent variable is the number of articles produced by an author in a week. Experience is measured based on total engagement prior to the beginning of the eight-week pre-post event study window.

Figure 6: Impact of platform change on engagement



(a) Aggregate Effects (b) Dissagreggated Effects

The figures show estimates from an article-level regression of engagement on the platform change. The estimates show the change (decline) in the logarithm of article-level clicks after the platform change relative to the period prior to the shock. 99% confidence intervals are presented. The dependent variable is the logarithm of the number of clicks per article.