Incentivizing Engagement: Experimental Evidence on Journalist Performance Pay*

Ivan Balbuzanov† **Jared Gars**‡ **Mateusz Stalinski**§
Emilia Tjernström¶

October 2025

Abstract

Digital platforms increasingly compensate content creators based on engagement metrics, yet the effects of these incentives remain poorly understood. We conducted a field experiment with a Kenyan news outlet that randomly assigned writers to either pay-per-click (PPC) or piece-rate contracts to study how engagement incentives affect content production, quality, and journalist well-being in digital media. We randomly assigned writers to either pay-per-click (PPC) or piece-rate contracts. The PPC contract tripled per-article pageviews and increased daily pageviews by 107%, but reduced the number of published articles by 74%. While PPC writers earned more per article, their overall earnings fell, lowering the firm's wage bill and increasing profits. However, these gains came at a cost: PPC writers shifted content production away from local news and towards attention-grabbing political stories. PPC writers also used less positive language in both headlines and article bodies. Our results show that engagement-based pay boosts reader traffic but caution that this may come at the cost of compromised coverage diversity, local news provision, and journalist well-being.

JEL Classification: C93, J24, J33, L82, M52

Keywords: performance pay, labor productivity, media engagement, field experiment

*We would like to thank Charles Angelucci, Pascaline Dupas, Justin Holz, Eliana La Ferrara, Leah Lakdawala, Huailu Li, John List, Teresa Molina, Matthew Ridley, Devesh Rustagi, Laura Schechter, Carlo Schwarz, Alessandro Tarozzi, Maisy Wong, and numerous seminar and conference participants for helpful comments and suggestions. We also thank Jacob Coreno, Toan Le, Charlotte Wang, and Xiaolei "Adam" Wang for excellent research assistance. This study was approved by the University of Wisconsin–Madison Education and Social/Behavioral Science IRB (protocol number 2017-0527) and the Amref Ethics and Scientific Review Committee (protocol number P369-2017). We registered the experiment in the American Economic Association Registry for randomized control trials under trial number AEARCTR-0002396. This work was supported by funding from the IFREE small grants program. Jared Gars gratefully acknowledges financial support from the "Policy Design and Evaluation Research in Developing Countries" Initial Training Network (PODER, funded under the Marie Curie Actions of the EU's Seventh Framework Programme, Contract Number: 608109). The opinions expressed herein do not reflect the opinions of the sponsoring agencies.

†Department of Economics, University of Melbourne, ivan.balbuzanov@unimelb.edu.au.

‡Food and Resource Economics Department, University of Florida, jgars87@ufl.edu.

§University of Warwick and CAGE, mateusz.stalinski@warwick.ac.uk.

¶Macquarie University, emilia.tjernstrom@mq.edu.au.

1 Introduction

The digital economy has transformed how media and technology firms produce and monetize content.[1] These advertising-based business models have encouraged outlets—including news organizations, social media platforms, and streaming services—to tie creator compensation to engagement metrics such as clicks, shares, or return visits.[2] Despite growing concerns that engagement-based compensation may distort content production and news coverage, empirical evidence on their effects remains scarce (Aridor et al., 2024). Given the expansion of online news consumption, this evidence gap matters because incentive design could either improve content alignment with audience preferences or undermine broader public value.

In this paper, we investigate how engagement-based contracts affect content production and firm profits. We conducted a field experiment in partnership with a digital news outlet in Kenya, which covered national and local stories, attracting an audience of up to 6.1 million unique monthly readers. Throughout our experiment, the partner firm sourced articles from independent writers that had full discretion over coverage, contribution volume, and framing, subject only to minimal editorial checks prior to publication—similar to other digital creators whose incomes depend on audience engagement. Using the experimental variation in writers' contracts, we provide novel evidence on how incentives shape content, revealing key trade-offs for media firms and informing broader debates about platform design in the digital economy.

Our within-firm experimental design randomly assigned 146 writers to one of two contracts: the status-quo per-article piece-rate or a pay-per-click contract in which earnings depended on pageviews. We randomized an additional 71 writers to a third condition in which they could choose between the pay-per-click and control contracts.[3] Control group writers continued on their pre-intervention contract, while the firm calibrated the pay-per-click payoff function to align ex ante expected earnings with the control contract. The firm implemented the experiment simultaneously for all writers and maintained the randomized allocation for five months. We use a difference-in-differences design to estimate causal effects of pay-per-click incentives on outcomes aggregated at the writer-day level, including article supply, pageviews, content, and tone.

[1] According to Pew Research Center (2023), the share of newspaper ad revenue from digital sources has risen steadily since 2010, reaching 48% in the US by 2022. As advertising grows more important, traditional media begin to resemble social media platforms, whose business model relies almost entirely on ad revenues.

[2] Examples of engagement contracts include paying freelance contributors based on unique visitors and/or clicks (e.g. Gawker (Benton, 2009; Starkman, 2013) and Forbes (Bartlett, 2013)) and revenue-sharing programs at YouTube (Kerkhof, 2024).

[3] We provide more detail about the treatments in Section 3.3.1.

The pay-per-click contract dramatically increased article reach but sharply reduced article supply. Pageviews per article tripled in the treatment group compared to the control group. Treatment writers' articles attracted 771 more pageviews per day than their control group counterparts, representing a 107% increase. This increased engagement came without increases to the firm's total wage bill. These results show that high-powered performance pay contracts can effectively increase profits for digital firms with advertising-dependent revenue streams. However, treatment writers published an average of 0.46 fewer articles per day over the course of the experiment—a 74% decline relative to control. This effect could undermine platforms that rely on steady content volume from creators.

Pay-per-click writers earned substantially less overall despite their increased reach. While treatment writers earned 54% more per article than control writers, their overall earnings fell by 49%. This earnings drop stems partly from the payoff function structure, which did not reward per-article pageviews enough to offset treatment writers' reduced output. Treatment writers also spent 25% more time per day preparing articles compared to the control group, suggesting that writers substituted away from output volume toward effort that increased per-article engagement. The media firm captured gains from both increased traffic and reduced writer compensation costs, highlighting a conflict between firm profits and worker welfare.

Turning to article content, the engagement incentives altered the focus of news coverage across both topics and geographies. Pay-per-click writers increased the share of their writing on political topics by 33 percentage points (pp) while reducing local reporting by 19 pp. These shifts are consistent with writers responding to the new incentives by selecting topics with broad appeal. Whether the resulting decline in comprehensive and locally-relevant coverage imposes civic costs is an open question, though prior work suggests that local news provision generates positive externalities (Campante et al., 2022).

The pay-per-click contract also changed how writers framed their content. Treatment writers reduced positive words in headlines by 17% and wrote article bodies that were 58% more negative on average. We also find suggestive evidence that the treatment increased the toxicity of articles, which encompasses hateful, harassing, and profane content, as measured by a state-of-the-art machine learning classifier. Our results align with existing literature indicating that negativity and toxicity drive user engagement (Baumeister et al., 2001; Robertson et al., 2023; Beknazar-Yuzbashev et al., 2025). We demonstrate that journalists, when responding to engagement-based incentives, exploit opportunities presented by these phenomena.

Our theoretical framework contextualizes the empirical findings within a broader incentive model. We extend the linear-contract moral hazard model (e.g., Holmstrom and Milgrom, 1991, 1987) to allow agents to endogenously choose their task quantity. The model predicts that the pay-per-click contract structure increases both effort and engagement—consistent with our results. The model also reveals that sufficiently risk averse writers reduce their output under engagement-based contracts. This aligns with our finding that the pay-per-click contract led to a notable decrease in publication frequency among risk-averse individuals, while exerting little to no effect on the behavior of risk-loving writers.

This study contributes to several strands of literature. First, we advance the literature on what shapes news quality and user-generated content (e.g., Groseclose and Milyo, 2005; Baron, 2006; Gentzkow and Shapiro, 2010; Larcinese et al., 2011; Prat and Strömberg, 2013; Wu, 2017; Cagé et al., 2025). Prior work emphasized government media capture (Di Tella and Franceschelli, 2011; Szeidl and Szucs, 2021), advertiser influence (Beattie, 2020), and media market competition (Mullainathan and Shleifer, 2005; Angelucci and Cagé, 2019). We provide the first causal evidence on how individual creator incentives affect content choice and characteristics. As engagement-based pay becomes standard across both traditional and social media, understanding these incentive effects matters for both platform design and media policy.

Our result that engagement incentives reduce news coverage diversity and local reporting complements research showing similar effects from other forces like TV expansion (Angelucci et al., 2024). This represents an important trade-off given local news's demonstrated role in voter turnout, local candidates running for office, and civic participation (Oberholzer-Gee and Waldfogel, 2009; Snyder and Strömberg, 2010; Gentzkow et al., 2011; Schulhofer-Wohl and Garrido, 2013). By varying individual writers' incentives, we directly trace their impacts on content choice and framing—improving upon previous studies that rely on market-level shocks like increased competition (Angelucci and Cagé, 2019; Cagé, 2020) that limit identification of specific mechanisms.

Our study also contributes to the growing literature on the harms of digital media (e.g., Bursztyn et al., 2024; Allcott et al., 2020; Zhuravskaya et al., 2020; Aridor et al., 2024). Platforms often encourage content creators to chase engagement in ways that negatively affect users. We find that engagement contracts push journalists toward more negative and toxic content. This shift aligns with evidence that negative or toxic content drives user engagement (Baumeister et al., 2001; Robertson et al., 2023; Beknazar-Yuzbashev et al., 2024, 2025). We provide causal evidence on the supply-side mechanisms that generate harmful content in digital media markets.

As the second main contribution, we extend the large body of research on incentive schemes in firms, moving beyond piece rates.[4] We experimentally test the impact of moving from piece rates to higher-intensity contracts for content creators, defined as pay schemes that tightly link compensation to audience engagement metrics. We find substantially larger productivity gains under performance contracts than prior studies, with pageviews tripling while firm costs fell. These gains may arise because the journalists in our experiment have substantial discretion over effort allocation and hold private information about their returns to effort. Performance contracts elicit this effort while piece rates do not (Prendergast, 2002; Raith, 2008; Hong et al., 2019).

A key advantage of our setting is the direct link between article pageviews and advertising revenues, which allows us to measure how performance pay affects firm profits. Firms have used performance contracts to address principal-agent problems in various contexts (Lazear, 2000; Finan et al., 2017), but evidence on whether they improve profits remains mixed (Bandiera et al., 2011; Miller and Babiarz, 2013; Prendergast, 2015). Contracting on outcomes can crowd out intrinsic motivation (Bénabou and Tirole, 2006), reward short-term gains at the cost of long-term productivity (Bolton et al., 2006), and reduce attention to tasks that are hard to measure and therefore not contracted (Holmstrom and Milgrom, 1991). Our experiment isolates the causal effect of performance pay by randomly assigning and enforcing contracts, avoiding the selection problems that complicate earlier studies (Bandiera et al., 2007; Dohmen and Falk, 2011).

Lastly, we contribute to the literature on work incentives and participation decisions in online labor markets (Chandler and Kapelner, 2013; DellaVigna and Pope, 2018; De Quidt, 2018; List and Momeni, 2021; Butschek et al., 2022) and gig economy settings (e.g., Wu and Zhu, 2022). Our findings apply beyond news outlets to any digital platform where creators earn based on engagement, including YouTube, Twitch, TikTok, and web-based publishers. These platforms collectively employ millions of content creators who face similar engagement-based incentive structures.

Our results are robust to a battery of robustness checks. While our preferred results are estimated with difference-in-differences, we show that our results hold under an ANCOVA

[4] Most empirical studies on performance contracts examine shifts from hourly wages to piece rates. These shifts raise productivity by 20% to 50%, though increased managerial costs often dampen profit gains (e.g., Lazear, 2000; Shearer, 2004; Bandiera et al., 2005). Piece rates work well when tasks are simple, inputs are easy to monitor, workers have limited discretion, and output is not stochastic.

specification (Section 4.4). We also vary the clustering of standard errors, and show that it does not affect our conclusions. Section 4 additionally discusses results related to the Choice treatment arm, in which writers could select either the pay-per-article or pay-per-click contract at the beginning of the intervention. Around 31% of writers in this arm chose the pay-per-click contract. Similar to the main treatment, the choice group also raised firm profit, though less than the main pay-per-click arm. We also find that many writers who had the ability to generate engagement chose to stay in the pay-per-article contract, suggesting either a sub-optimal choice or high levels of risk aversion.

2 Background

2.1 Media Market in Kenya

Kenya's post-independence media landscape enjoys a broad range of media outlets, with a mix of state-owned and private outlets, most of which publish in English (see Lohner et al., 2016; Freedom House, 2017, for reviews contemporaneous with the study). The news media consists of at least four national daily newspapers, one business daily, and a handful of regional weekly newspapers. Although media coverage in Kenya has traditionally been relatively rigorous, especially in the print sector (e.g., *The Nation* and *The Standard*), editorial pressure and the political preferences of advertisers shape coverage at many outlets (Nyabuga, 2023; House, 2023). Reporters can face repercussions, including dismissal, for critical coverage.

Around the time of the experiment, human rights groups expressed concerns that press freedom was in sharp decline (Namwaya, 2018; Freedom House, 2017). In 2025, Kenya ranked 117 out of 180 countries on Reporters Without Borders' Press Freedom Index, substantially worse than its 2017 rank of 95/180.[5] Some sources associate the reduction in press freedom directly with ownership concentration and media capture. Freedom House (2017) notes that politicians or politically connected people own the majority of media outlets and that five media companies capture over 70 percent of all media consumers.[6] Simiyu (2014) surveys

[5] Among the many examples of violations of press freedom, the Kenya Communications Authority switched off a number of television stations in early 2018 for broadcasting live from the site of a political action by the opposition leader. The High Court of Kenya issued an order suspending the media ban, but the government stationed police around the relevant government offices to block court officers from serving the order (Namwaya, 2018). Reporters Without Borders (2024) provide an up-to-date Press Freedom Index for Kenya.

[6] As a benchmark, the top 5 newspaper owners in the US (Gannett, Tribune/MediaNews Group, Lee Enterprises, Adams Publishing Group, and Paxton Media Group) controlled 1,091 newspapers (Konopliov, 2024). In 2024, the US had about 5,600 newspapers, though the overall number is declining over time (Metzger, 2024). The ownership concentration of the top five companies is therefore less than 20%.

citizens and reporters after the 2013 presidential elections and finds that a stark majority of the Kenyan electorate believe the media was biased and partisan in their reporting of the election. Despite these challenges, more recent data from the Reuters Institute indicates that some print outlets remain fairly trusted by the electorate (Gicheru and Nyabuga, 2024).

2.2 The News Firm

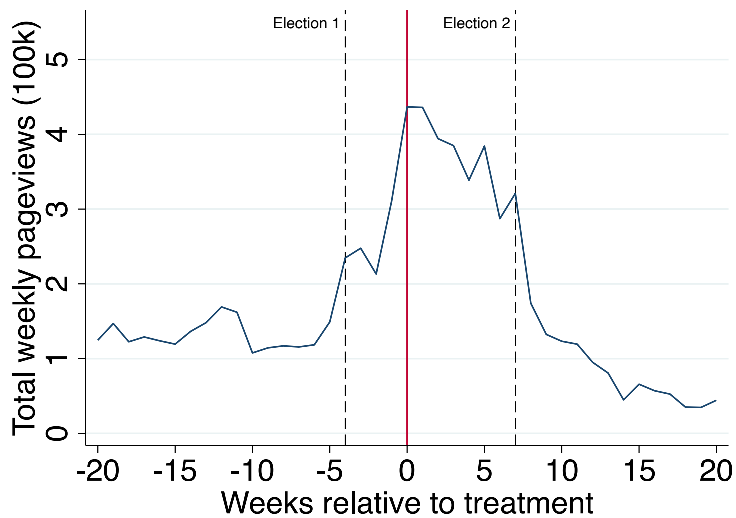
We study a digital news platform in Kenya that published both local and national news from a network of local reporters. Founded in 2014, the firm aimed to supply high-quality local news and to foster civic engagement and political participation, yet relied on an ad-based business model. Writers received a piece-rate of 100 Kenyan shillings (KES), or roughly $0.96 at the time of the experiment, for each published article. The firm accepted contributions from writers without prior qualifications and employed a Nairobi-based editorial team that managed content.

The news platform reached a large audience in the months leading up to our intervention: between May and September 2017, the site averaged over 1.5 million unique monthly readers. The rise in user engagement with the firm's content coincided with the run-up to the highly contested presidential election on August 8, 2017, and remained elevated through the legal disputes that followed.[7] Figure 1 shows that daily pageviews rose rapidly ahead of Election 1, stayed high until the launch of our experiment two weeks later, and began declining following Election 2. As a result, our study covers both a period of heightened readership and a period of subsequent decline in traffic following the electoral period.

The platform gave writers flexibility over what to write and when. Writers registered with the firm before contributing and submitted articles through a "workbench" page, where they could submit their work, edit their profiles, and track payments. Each writer had a public profile that linked their published articles, publication dates, and pageview counts. Submissions faced a 500-word limit but the platform did not suggest or control topics or content, leaving writers with full control over what stories to cover and how often to submit.

Editors reviewed articles in the order received and published those that met a minimum quality standard. The most common reasons for rejection were: being insufficiently engaging,

[7]Kenya held a presidential election on August 8, 2017 (Election 1), in which the incumbent President Uhuru Kenyatta won reelection with 54% of the vote. Opposition leader Raila Odinga challenged the results in the Supreme Court, which annulled the outcome and ordered a new election for October 17, 2017 (Election 2). On October 10, 2017, Odinga withdrew, citing concerns about the impartiality of Kenya's Independent Electoral and Boundaries Commission (IEBC) and the withdrawals of key members of his political coalition.



Note: The figure shows weekly pageviews for articles published by sample writers. The x-axis reports weeks relative to the start of the intervention. Dashed vertical lines mark the dates of Kenya's two 2017 presidential elections.

Figure 1: Weekly Pageviews

plagiarism, informal language, and excessive similarity to existing articles. Writers received payments for published articles via the mobile money system M-PESA. These were disbursed weekly until June 6, 2017, after which writers were paid every three days.

3 Experimental Design

3.1 Overview

The intervention lasted five months, from September 2, 2017 to January 31, 2018. We randomly assigned writers to one of three experimental conditions, without revealing that the contract changes were part of an experiment. In the Control (Pay-Per-Article) group, writers continued to receive the same piece rate for each published article as before the experiment. In the Pay-Per-Click (PPC) group, payment varied with the number of pageviews that a published article received. In the Choice group, writers selected either the PPC contract or the Control contract at the beginning of the intervention, and their choice remained fixed for the full five months.

We collected data on article supply, reader engagement (pageviews), and other article characteristics throughout both a baseline period and the five-month intervention. In February 2017, the firm changed how it recorded pageviews. To ensure consistent measurement, we therefore define our baseline period as starting on March 1, 2017, approximately six months before the start of the experiment, and ending on September 2, 2017, the day before the intervention.

3.2 Sample and Randomization

3.2.1 Sample Size and Composition

Our dataset contains data on 480 writers who were registered with the firm and submitted at least one article in the year preceding the intervention. At the start of the intervention, we randomized these writers to one of the three experimental conditions. Our main sample retains writers who published at least one article during the baseline period. We exclude 3 writers in the Choice condition, as the firm administrative data shows multiple values for their contract choice. Therefore, we are uncertain about whether the firm correctly implemented their contract choice. We also exclude 4 editors from our sample—although they could not process their own articles, we might worry that they could influence their articles' publication chances through their colleagues. After these restrictions, our main sample consists of 217 writers.

The writers were freelance contributors based across Kenya. They did not know each other personally before or after the experiment, making coordination or communication unlikely. Our partner platform occasionally communicated with writers using WhatsApp. They created new WhatsApp groups, which kept the three treatment groups separate for the duration of the intervention. While writers could stop submitting articles to the news platform at any time, they could not select out of the experiment.

3.2.2 Randomization

We randomly assigned writers in our sample to three conditions: the PPC contract ($n = 74$), Choice ($n = 71$), and Control ($n = 72$). We stratified the randomization on three variables measured over the 12 months preceding the intervention: above/below median pageviews, above/below median articles published per week, and publication recency. We classified recency into four groups based on when the writers last published prior to the experiment: (i) within one month, (ii) two to three months, (iii) four to six months, and (iv) more than six months earlier. This created 16 strata.

3.2.3 Descriptive Statistics

Column 1 of Table 1 reports descriptive statistics based on data from the baseline period. These include the number of articles published, pageviews, pageviews per publication, shares of articles covering national events and politics, as well as sentiment and toxicity measures. Over the six-month baseline period, the average writer published 69 articles and attracted over 113,000 pageviews—more than 1,500 views per publication. Our measure of pageviews follows the "users" metric in Google Analytics, which counts the number of *unique* IP addresses per day. This metric reduces the scope for manipulation, such as refreshing an article repeatedly to garner more views. The numbers above show both high writer activity and the platform's strong reader engagement. For the average writer, national and political news make up 35% and 37% of their publications, respectively, but are not dominant categories. This is line with the firm's mission to provide local news on a variety of topics and serve local communities.

Columns 2-5 of Table 1 compare baseline covariates across the treatment groups. We report means and standard deviations for 11 baseline period covariates.[8] We detect statistically significant differences across groups for only one variable: the share of political articles is higher in the Choice condition. Overall, the sample appears well balanced.

3.3 Treatments

We now describe the treatment contracts and the information that writers in each treatment group received. All contracts took effect on the same date, September 2, 2017, for everyone.

3.3.1 Contracts

Control (Piece rate). Writers in the Control condition continued with the status-quo contract, receiving 100 KES per published article. This contract had been in place since the site launched in May 2014. To receive payment, a writer's article had to pass the editorial review described in Section 2.2. The firm then paid writers via M-Pesa on the next scheduled payday after publication.

Treatment 1 (Pay-per-click). The PPC treatment group contract tied pay to article pageviews using the kinked and discontinuous piecewise-linear fee structure depicted in Figure

[8] We use baseline period data for all outcomes except tenure, which we define as the number of months between a writer's registration with the firm and the intervention start date. Many writers joined before March 2017, the beginning of the baseline period, so the average tenure (7.6 months) exceeds the length of the baseline period.

Table 1: Balance on Writer Characteristics

	Treatment Arm				
	All (1)	Pay per article (2)	Pay per click (3)	Choice (4)	p-value (5)
Tenure (months)	7.88	7.92	8.11	7.59	0.87
	(0.42)	(0.76)	(0.68)	(0.74)	
Articles published	69.14	58.96	74.11	74.28	0.45
	(9.01)	(12.72)	(16.42)	(17.43)	
Total views	113333.11	96425.60	121269.78	122206.73	0.50
	(15768.38)	(23598.19)	(27822.31)	(30406.94)	
Views/Published	1502.96	1524.54	1402.34	1585.94	0.72
	(116.05)	(202.40)	(169.12)	(231.36)	
Share national	0.35	0.36	0.36	0.34	0.93
	(0.03)	(0.05)	(0.04)	(0.05)	
Share politics	0.37	0.36	0.30	0.46	0.01
	(0.02)	(0.04)	(0.04)	(0.05)	
Negative (Title)	0.63	0.68	0.59	0.63	0.35
	(0.03)	(0.06)	(0.05)	(0.05)	
Positive (Title)	0.33	0.29	0.31	0.37	0.30
	(0.02)	(0.03)	(0.03)	(0.04)	
Tone (Body)	-0.33	-0.44	-0.34	-0.21	0.78
	(0.17)	(0.31)	(0.31)	(0.27)	
Toxicity	0.01	0.01	0.01	0.01	0.81
	(0.00)	(0.00)	(0.00)	(0.00)	
Toxicity (Binary)	0.02	0.03	0.01	0.02	0.28
	(0.01)	(0.02)	(0.01)	(0.01)	
Observations	217	72	74	71	

Note: This table reports means and standard deviations (in parentheses) for 11 observables, overall (column 1) and by treatment condition (columns 2-4). Column 5 summarizes balance tests based on regressions of each variable on the treatment indicators (PPC dummy and Choice dummy) with strata fixed effects. Specifically, it reports the p-values for a joint test that the two treatment dummies equal zero. The table is based on the main experimental sample of N=217 writers, using data aggregated at the writer-week level. Summary statistics cover the full baseline period, except for tenure with the company, which measures months from a writer's registration with the firm and the intervention start date. We report the following covariates: (1) tenure at the firm at the time of randomization (in months), (2) the number of articles published (winsorized at the 95th percentile), (3) total pageviews (winsorized at the 95th percentile), (4) per-article pageviews, (5) the share of articles covering national issues, (6) the share of articles covering politics, (7) the number of negative words in the title, (8) the number of positive words in the title, (9) article body tone, defined as the number of positive words less the number of negative words scaled by the overall length of the article, (10) the toxicity score of the article body, measured by Unitary's Detoxify library, (11) the proportion of articles with a toxicity score over 0.1.

2. Writers received no payment for articles with less than 400 pageviews. Between 400 and 800 views, writers earned 125 KES per 1000 pageviews, and beyond 800 pageviews, the rate dropped to 12 KES per 1,000 pageviews. There was no upper limit on the number of pageviews that could count towards payment.

The firm calibrated these contract parameters on historical traffic data. It aimed to reward high-performing articles while keeping ex ante expected earnings the same. The steep payoff slope between 400 and 800 pageviews was designed to reduce the number of "losers," defined as articles earning less than the 100 KES the status quo contract would pay. In the calibration data, 44% of articles would have earned less under PPC than the status-quo contract, and 27% of articles would have received no payment at all under the PPC. At the other end of the

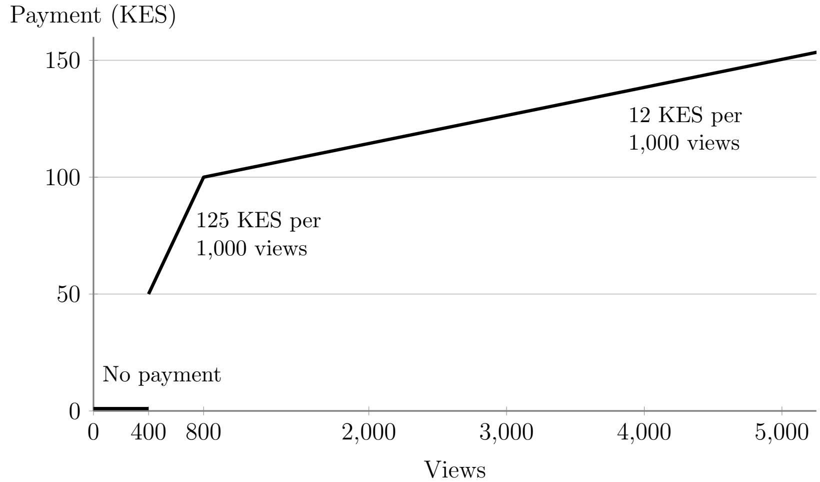


Figure 2: Payment Schedule for PPC Writers

spectrum, nearly 11% of articles would have earned at least twice as much. Thus, the PPC contract introduced substantial downside risk, but with the potential for higher returns for greater effort.

Writers received M-PESA payments on the first payday after publication, based on cumulative pageviews up to that date. If an article continued to generate engagement, the writer could receive additional payments in future pay periods. The fee structure described above applied to all pageviews after article publication and view counts did not reset in each pay period. Most of the article engagement happened quickly: 87% came within two days of publication, and nearly all within the first week. Figure B1 shows the distribution of pageviews by day since publication.

Treatment 2 (Choice). When they first logged in after the intervention launched, writers in the Choice group were invited to select either the piece-rate or the PPC contract. The firm informed them that they had to choose before they could submit another article and that their selection would be permanent. The payment rules and timing were identical to those described above.

3.3.2 Communication

The firm informed all writers of the new contracts via SMS on the day the intervention launched. Each writer also received a message from the Head Editor on their workbench highlighting the firm's recent growth and thanking them for their contributions to this success. This message served both to standardize communications across all treatment groups and to update writers on the firm's performance.

In addition to the above message, writers in the treatment group also learned that they would be paid according to a new contract going forward, effective immediately. The message to PPC writers explained that the firm was testing a new way to reward high-performing contributors for their work. The firm did not reveal the complete pay schedule. Instead, writers saw a simplified table summarizing the key kinks in the payout schedule, shown in Figure 3.

Article visitors	Your Pay
300	0
400	50
800	100
5000	150
15000	270
35000	510

Figure 3: Payment Schedule Communicated to Writers

PPC writers also received some information about how their payments would be calculated using Google Analytics data. The message noted that a closely correlated pageview measure was visible on their profile pages and explained how to find payment details for each article. Writers in the Choice group received the same message, with the added instruction to select their contract by logging into their writer page.

Editor compensation and responsibilities remained unchanged throughout the experiment. Editors earned a fixed hourly wage for their editorial work and could continue submitting articles under the piece-rate contract. Editors were blind to writers' treatment assignment, which rules out any confounding from editorial discretion.

To further reduce information gaps across treatment groups, we provided all writers with access to their articles' pageview counts six weeks before the experiment. This ensures that writers had consistent access to information about their articles' popularity across treatment groups. Writers could see these pageview counts on their public profile page and on the submission dashboard. The public profile page was easily accessible from article bylines via a writer profile link.

3.4 Outcomes

We group our outcomes into three categories: (1) article supply and writer performance, (2) writer effort and predictions, and (3) article characteristics. The first set captures how the higher-intensity contract affects output and profit. The second documents how well writers predict the link between effort and user engagement. The third illuminates outcomes that may affect the firm's reputation and long-term readership, with broader implications for the quality of journalism and political discourse. We additionally report the share of writers in the Choice group who selected the PPC contract, alongside the main results for that condition.

3.4.1 Article Supply and Writer Performance

Article Publications. We start with our first pre-registered outcome: the number of articles that writers submit. We aggregate submissions at the writer-day level, using article publication timestamps provided by the firm. We also report a binary indicator for whether a writer submitted any article on a given day. This second measure gives us insight into whether or not a writer was active.

Writer Performance. Second, we turn to writer performance, for which we have two pre-registered outcomes: pageviews and earnings. We compute the number of pageviews using Google Analytics data. Specifically, we count the number of users (IP addresses) who viewed each article at least once in a given day, and allocate the total views over the 7 days following publication to the day of publication. Because nearly all article views occurred within this seven-day window—and most within a single day (see Figure B1)—this measure effectively constituted the basis of the higher-intensity contract. We then aggregate these data at the writer-day level, summing views across articles whenever writers published multiple articles on the same day. For context, writers published more than one article on 40% of their active days. The second outcome is writer earnings (in KES), which we also aggregate at the writer-day level, and assign to the article's publication date.

Article Performance. To separate article volume from article engagement, we also report the number of pageviews that writers receive per article. We report this measure alongside writer earnings per article.

3.4.2 Effort and Engagement Predictions

This section describes two pre-registered outcomes: writer effort and prediction errors of an article's pageviews.

Submission Surveys We began collecting submission-level survey data on July 27, 2017 (five weeks before the intervention start). On the submission page, each time writers submitted an article, they also reported (i) the approximate number of minutes they spent preparing the article, and (ii) their expectations about how many pageviews the article would get in its first week. Writers were assured in writing that editors could not see these answers and that their responses would not affect their article's publication chances. Rather, they were encouraged to report truthfully so that "the firm can learn to better serve them." The pre-submission survey was compulsory and more than 99% of responses were non-missing. Since writers could view page-level traffic for past publications, we interpret their predictions as measuring informed expectations about article popularity.[9]

Effort and Prediction Errors We aggregate the survey-based effort reports at the writer-day level. If a writer submitted multiple articles on the same day, we average their reported effort across submissions on that day. We also compute prediction errors by comparing reported and actual pageviews after seven days. We report both the absolute prediction error (the absolute value of the difference between predicted and actual pageviews after 7 days) and the relative prediction error (as a percentage difference).

Predictors of Engagement Finally, we regress both predicted and actual pageviews on article observable characteristics to analyze the factors associated with predicted and actual engagement. Covariates include whether the article covers a national story or politics, the number of positive and negative words in the article headline, article tone, and article toxicity (see Section 3.4.3 for variable definitions). This analysis complements the causal estimates by examining observable correlates of engagement.

3.4.3 Article Characteristics

We next examine how contract assignment affects article content.

Topic Choices. The topical content of articles, a pre-registered outcome, directly contributes to the quality of journalism that the firm offers. High-quality journalism balances national and

[9]We did not include financial incentives to encourage accurate time and belief responses. Effort reporting could not be incentivized, as we would be unable to verify reports. The evidence on how much incentives improve the accuracy of reported beliefs is mixed. For example, Hoffman and Burks (2017) compare predicted productivity for truck drivers (miles per week) using both incentivized and un-incentivized elicitation. They find no evidence that reported beliefs were different when drivers are rewarded for accuracy.

local coverage across a broad set of issues. These features were central to the firm's mission and may shape local political discourse. Higher-intensity contracts risk shifting coverage towards highly engaging but narrower and/or lower-quality content. To measure topical shifts, we record whether each article covers a locally- or nationally-relevant story. We also record whether the article concerns politics or a non-political topic. We aggregate both outcomes at the writer-day level, computing the share of articles in each category that a writer published on a given day.

Editor's Rating. Our pre-registration proposed social media sharing as a proxy for quality. For technical reasons, we were unable to reliably capture this outcome. Instead, we use editor-provided quality ratings as an alternative measure. Editors rated each article on a 1-5 scale, with 5 denoting the highest quality, before deciding whether or not the article should be published. Editors were instructed to base their quality assessment on the accuracy, depth, impartiality, and style of each article—not on expected popularity. We regularly reminded editors that their quality assessment should be independent of the number of views they expected articles to receive. Because we started collecting editors' ratings late in the baseline period, the panel is shorter for this outcome. We aggregate ratings at the writer-day level, averaging across articles whenever a writer published multiple articles on the same day.

Sentiment and Toxicity. We also analyze article tone and toxicity of each submitted article's title and body.[10] Additionally, given strong evidence in the literature that negativity and toxicity generate user engagement (Baumeister et al., 2001; Beknazar-Yuzbashev et al., 2025; Robertson et al., 2023), we investigate whether writers with the PPC contract adjust their writing to leverage these phenomena. We report several relevant measures based on the textual analysis.

First, we analyze headline sentiment, using the Lexicoder Sentiment Dictionary (Young and Soroka, 2012), which includes a pre-defined sentiment dictionary of 2,858 negative and 1,709 positive sentiment words. For each article title, we count the number of negative and positive words from the dictionary.

Second, we compute article body tone as the number of positive words less the number

[10] These outcomes replace our pre-registered plan to employ MTurk auditors to categorize articles according to characteristics such as negativity towards ethnic groups or political parties. Given the high volume of articles and the limited availability of Kenyan MTurkers—there were only 9 Kenyans registered on the platform at the time—we instead rely on dictionary- and model-based measures that offer consistency and broader coverage.

of negative words, scaled by total word count. Before analysis, we pre-process the text using standard methods: we remove stop words and reduce the complexity of the data using a stemming algorithm in Natural Language Toolkit (NLTK).[11] This pre-processing reduces vocabulary size and allows us to aggregate text with similar sentiment more easily.

Separately, we look at article toxicity using Unitary's Detoxify library (Hanu and Unitary, 2020), a Python-based toxicity detection algorithm with multiple research applications (e.g., Rizzi, 2024; Beknazar-Yuzbashev et al., 2025). The concept of online toxicity spans several types of harmful content, including hate speech, harassment, and profanity. Many prominent state-of-the-art machine learning algorithms, such as Unitary's Detoxify or Perspective API, rely on the following definition of toxicity: "a rude, disrespectful, or unreasonable comment that is somewhat likely to make you leave a discussion or give up on sharing your perspective." The definition involves two distinct components: toxicity either involves a violation of norms of interpersonal communication or excludes individuals from a conversation. The classifier assigns each text a toxicity score from 0 to 1 predicting the share of individuals who would classify it as toxic according to the above definition.

For each article, we compute toxicity scores based on the text of its body. We aggregate the data at the writer-day level, averaging toxicity scores if a writer submitted multiple articles on the same day. As an additional measure, we report the share of articles with a toxicity score exceeding 0.1. Although a score of 0.1 does not strongly indicate toxicity for short statements, it serves as a good indicator, given that we analyze entire articles. The average article in our sample has a toxicity score of just 0.01 during the baseline period.[12]

3.4.4 Heterogeneity Analysis

We examine treatment heterogeneity for two outcomes: the number of publications and per-publication pageviews. Our primary focus is heterogeneity with respect to writers' risk preferences. As discussed in Section 3.5, theory predicts that risk aversion should influence how the PPC contract affects the number of articles writers supply, but not how engaging their articles are.

We measure risk preferences using a firm-administered survey prior to the intervention,

[11] A stemming algorithm simplifies words to their root form, e.g., by removing suffixes. This reduces the vocabulary size—the number of unique words—to consider. NLTK is a Python library used for processing human language data.

[12] The average article in our sample is 213 words long. The toxicity detection algorithm was trained on text averaging 67 words (Feng, 2023).

which 58.5% of writers completed. The survey asked respondents, on a scale from 0 to 10, "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" (where 10 means "fully prepared to take risks"). This question has been widely used in the literature as a proxy for risk preferences and has been validated through comparisons with behavior in incentivized lottery tasks (Dohmen et al., 2011).

In addition, we explore heterogeneity by other writer characteristics: the number of published articles, the share of political articles, the share of articles covering national stories, per-article pageviews, and tenure with the firm. These dimensions are relevant in their own right and help benchmark the role of risk preferences in interpreting shaping contract responses.

3.5 Theoretical Predictions

We discuss theoretical predictions for a subset of our outcomes. Appendix A presents a variation of the classic linear-contract moral-hazard model with normally distributed noise and CARA preferences (e.g. Holmstrom and Milgrom, 1987, 1991). An important difference is that our model allows agents to endogenously choose how many tasks to perform. The model generates two clear predictions. First, writers will expend more effort under the PPC contract than under the piece-rate. Second, this effort should translate into a higher average number of pageviews for PPC writers. Our experimental design lets us test both predictions.

The model's implications for how the PPC contract affects the optimal number of articles that writers choose to produce is more subtle. It predicts that highly risk averse writers will submit fewer articles under the PPC contract, while less risk averse writers may not reduce their output. As a result, the expected sign of the treatment effect on the number of submissions was ambiguous ex ante, with the result depending on the distribution of risk preferences. The model also predicts that more risk averse writers in the Choice treatment are more likely to choose the pay-per-article contract over the pay-per-click contract.

Finally, we consider how PPC contracts might affect article characteristics. Although the model predicts higher effort under the PPC contract, this does not necessarily imply better quality by objective metrics such as accuracy, reporting depth, impartiality, style, or topical variety. On the contrary, since the contract rewards engagement, writers may instead trade quality for user engagement (pageviews) by narrowing their topics, using click-bait, or resorting to negativity, toxicity, and divisiveness. A theoretical framework by Beknazar-Yuzbashev et al. (2024) shows that exposure to harmful or toxic content may increase user engagement even if they have a negative marginal utility of consuming such content. Empirical evidence supports this: toxicity (Beknazar-Yuzbashev et al., 2025) and negativity (Robertson et al., 2023) both raise engagement. If writers understand these dynamics and act on them, we would expect increased reliance on negativity or toxicity in the PPC contract.

3.6 Empirical Strategy

Our empirical strategy leverages three features of our setting: (i) a five-month intervention, during which we observe writers' choices at a daily level, (ii) a long pre-intervention baseline period, and (iii) simultaneous treatment rollout for all writers. Drawing on insights from McKenzie (2012), we use both between- and within-subject variation to maximize statistical power. Given these features, we adopt a two-way fixed effects model (TWFE) as our main specification.

Specifically, we estimate the following model at the writer-day level:

$$Y_{it} = \alpha_i + \delta_t + \beta PPC_{i,t} + \gamma Choice_{i,t} + \epsilon_{i,t}, \tag{1}$$

where $Y_{i,t}$ is the outcome for writer i on day t, α_i is the writer fixed effect, δ_t is the day fixed effect, $PPC_{i,t}$ is a dummy variable indicating whether writer i on day t experienced a pay-per-click contract, and $Choice_{i,t}$ is a dummy variable indicating whether writer i on day t experienced a choice contract. The two coefficients of interest are β, corresponding to the treatment effect of the PPC contract, and γ, corresponding to the treatment effect of the choice contract.

We use Driscoll and Kraay as our default standard errors (Driscoll and Kraay, 1998). This nonparametric estimator is robust to heteroscedasticity and very general forms of serial and cross-sectional dependence. We leverage this approach due to the large number of day-periods for writers in our dataset, which is the main requirement of this method (Cameron and Miller, 2015).[13] For outcomes like article publications, writer pageviews, and earnings, we observe more than 150 writer-day periods during the intervention alone.[14]

We test the robustness of our results in several ways. Section 4.4 shows that alternative ways of computing standard errors and clustering at the writer level leave the main conclusions of the paper unchanged. We also demonstrate that replacing our difference-in-differences (TWFE) approach with ANCOVA does not alter our conclusions. In Section 4, we present

[13] See Alvarez and Argente (2022) for an interesting use case and Hoechle (2007) for performance analysis.

[14] Even for outcomes conditional on publication, coverage is good: the number of writer-day time periods is considerable, with writers publishing an average of 317 articles across 83 different publication days (the medians are 160 and 58, respectively). See Table C2 in the appendix for details.

week-by-week visualizations of raw data by treatment for all our main outcomes, allowing for a visual inspection of pre-trends and the treatment effects.

We consider treatment effect heterogeneity, with respect to the variables described in Section 3.4.4, using the following specification:

$$Y_{it} = \alpha_i + \delta_t + \beta PPC_{i,t} + \gamma Choice_{i,t} + \xi(PPC_{i,t} \times High_i) + \psi(Choice_{i,t} \times High_i) + \epsilon_{i,t}, \quad (2)$$

where $High_i$ is a binary indicator equal to 1 if writer i has an above-median value of the characteristic in question. We cluster standard errors at the writer level for all heterogeneity regressions.

4 Results

4.1 Article Supply and Writer Performance

4.1.1 Overall Effects

We begin by examining how contract assignment affected article supply and pageviews. Table 2 presents the difference-in-differences estimates. Writers on the PPC contract published 0.46 fewer articles per day, a 74% drop relative to the control group. Moreover, PPC writers' likelihood of submitting an article on a given day fell by 8 pp. Both results are significant at the 1% level. These large declines are consistent with relatively high risk aversion levels among the writers.

As predicted by the theoretical model, the performance contract increased writers' total pageviews. The PPC group received 771 more pageviews per day, a 107% increase. At the same time, their overall earnings fell by 30 KES (49%) per day, mostly due to the drop in article supply. Since the PPC contract was calibrated to have ex ante identical expected payoffs (see Section 3.3.1), this result is not explained by the PPC writers facing worse contract terms.

Combined, these results show that the PPC contract improved the firm's profitability. It generated more pageviews—hence more advertising impressions—while lowering costs through reduced writer payments and fewer article submissions.

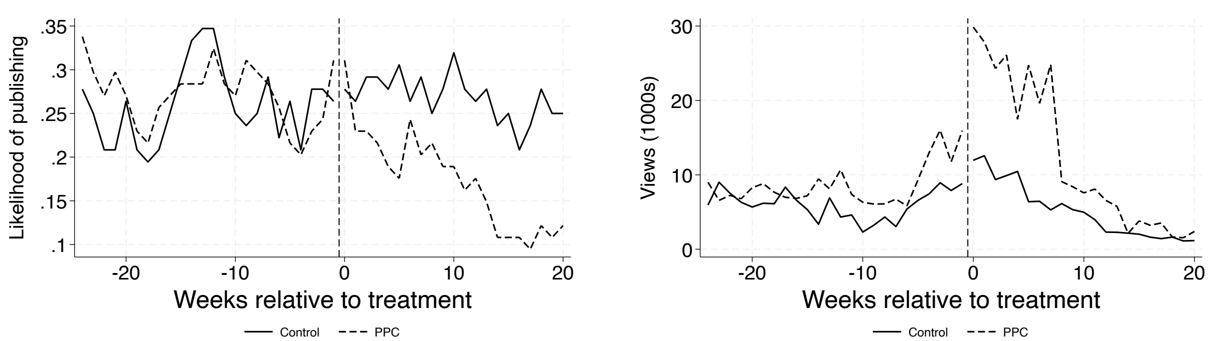
Figure 4 shows the raw data for the likelihood that a writer publishes at least one article and week-level pageviews, with solid lines representing the control group and dashed lines showing the PPC writers. Panel A shows that the introduction of the PPC contract led to a marked reduction in the likelihood of publications throughout the intervention period, with

Table 2: Treatment Effects on Supply, Views, and Earnings

	Published>0	Published	Views	Payment
Pay per click	-0.080***	-0.46***	771.3***	-30.2***
	(0.0082)	(0.041)	(200.1)	(6.33)
Choice	-0.054***	-0.35***	-29.9	-33.7***
	(0.0089)	(0.045)	(147.6)	(4.55)
Observations	60487	60487	60487	60487
Mean	0.16	0.62	723.6	61.7

Note: This table reports estimates from Equation (1) for our main experimental sample. The dependent variables are a dummy for whether a writer submitted at least one article (column 1), the number of articles writers published (column 2), the number of pageviews a writer's article receives in the 7 days after publication (column 3), and writer earnings (in KES) (column 4). We report the control group mean for each dependent variable during the intervention period. The unit of observation is the writer-day, where day refers to a specific calendar date. Driscoll-Kraay standard errors in parentheses. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

the gap becoming more pronounced over time. As expected, both groups follow similar trends during the baseline period. Panel B illustrates the effect on the number of pageviews, with little variation between the groups in the baseline. During the intervention period, we see a large gap between the PPC (high views) and the control (low views), especially pronounced in the first three months. This period coincides with the time around the Kenyan elections, which may have offered more opportunities for attention-grabbing headlines. Taken together, these raw data align with the conclusions derived from the regression analysis.



(A) Publishing Any Articles (B) Views

Note: Panel A depicts the likelihood of publishing an article in any given week by treatment group (PPC vs. Control). Panel B shows the total number of pageviews that writers' articles received in any given week. For each article, we include views during the 7 days following publication. The figure is based on the main experimental sample.

Figure 4: Supply and Views by Week

We also discuss treatment effects for the Choice condition relative to the control. At the start of the intervention, 31% of writers in the Choice group selected the PPC contract. The Choice contract reduced the number of publications per week, though less sharply than the

PPC contract. For comparison, the drop in the likelihood of submitting an article on a given day is 5.4 pp for Choice, compared to 8 pp for the PPC group. Unlike the PPC contract, the Choice contract had no impact on pageviews. With fewer publications and pageview gains, writers earnings declined. Allowing writers to self-select therefore appears to have improved profitability by lowering cost for a given number of pageviews, though the gains were smaller than under the mandatory PPC contract.

Table 3: Treatment Effects on View and Payments per Article

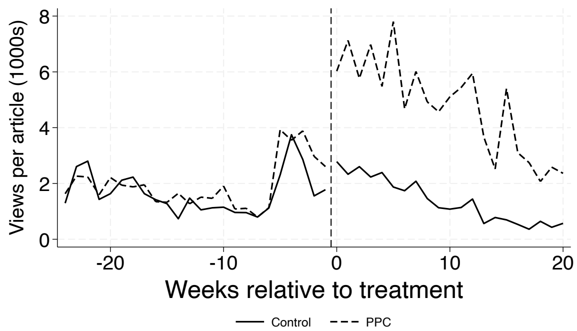
	Views/Pub	Payment/Pub
Pay per click	3924.6***	53.5***
	(278.4)	(5.70)
Choice	481.4***	2.37*
	(181.7)	(1.42)
Observations	8474	8474
Mean	1320.3	100

Note: This table reports estimates from Equation (1) for our main experimental sample. The dependent variables are per-publication pageviews, measured over the 7 days after publication, and writer earnings per article (in KES). We report the control group mean for each dependent variable during the intervention period. The unit of observation is the writer-day, where day refers to a specific calendar date. Driscoll-Kraay standard errors are in parentheses. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

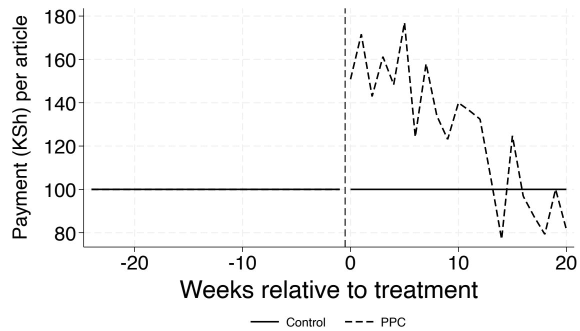
4.1.2 Per-Article Effects

Table 3 investigates the treatment effects on a per-article level. The PPC contract raised both per-article pageviews and the payments that writers received per article. The effect sizes are large and significant at the 1% level: per-article pageviews tripled on average, and earnings per article increased by more than 50% against the control. This indicates that the writers can produce more engaging content when given the financial incentives to do so. The writers who continue to publish under the PPC contract earned more per article, suggesting a potential financial benefit. However, as previously discussed, the intervention reduced writers' overall earnings (see Table 2), because writers published fewer articles—likely due to some writers struggling to consistently produce content that draws engagement.

Panel A of Figure 5 shows that per-article pageviews in the PPC group and the control evolved in an almost identical way during the baseline period. After the intervention began, views per publication rose sharply and immediately in the PPC group, creating a large gap relative to control. The gap persisted throughout the intervention period, though it narrowed slightly over time. Panel B of Figure 5 presents earnings per publication. Consistent with the effect on pageviews, the PPC contract led to higher earnings per publication. The overall



(A) Views per Article



(B) Payment per Article

Note: Panel A of the figure depicts the number of per-publication pageviews that writers' articles received in any given week by treatment group (PPC vs. Control). For each article, we include all views over the 7 days after publication. Panel B shows writers' earnings per article in any given week. The figure is based on the main experimental sample.

Figure 5: Views and Payments per Publication by Week

effect is driven by the first three months of the intervention period, which were closer to the elections, offering more opportunities for writing articles with engaging headlines.

Table 3 shows that the Choice contract increased per-article pageviews by 36%—a statistically significant result at the 1% level, but a much smaller effect than in the PPC group, where pageviews tripled. Writers' earnings per article did not differ from the control under the Choice contract. These patterns echo earlier results: while the Choice contract raised the firm's profit over the status quo, it is less effective than simply imposing the PPC contract.

The lack of a detectable per-publication earnings effect under the Choice contract is worth noting. We have seen that the PPC successfully incentivizes writers to supply fewer, but more engaging, articles. Our results show that the PPC writers who continued publishing are successful and raise the firm's profits. Some writers likely faced higher costs of generating engaging content, which discouraged them from submitting and reduced costly publications that received low traffic. The Choice contract, on the other hand, was designed to allow writers to select into whichever arrangement best matched their ability to generate pageviews. Those who expected to struggle to generate engaging content should have chosen the status quo while those with lower cost of engagement generation should have selected into the PPC contract, yielding high views and high earnings for each article. However, this sorting failed. Despite access to detailed data on prior article pageviews, writers did not necessarily choose optimally. This highlights the difficulty in making optimal contract commitments despite ample available information. Our results showcase a limitation of flexible contracting arrangements: even with good information, self-selection may not produce efficient matches, and therefore may not bring the intended engagement.

4.2 Effort and Engagement Predictions

Table 4 presents the difference-in-differences estimates of the treatment effects on writers' effort and prediction errors regarding pageviews.

In line with the theoretical predictions, the PPC contract increased writers' effort. The intervention raised average self-reported preparation time by 7.9 minutes, or 25%, an effect significant at the 1% level. This extends the results presented in Section 4.1 by showing that writers can attract more engagement by exerting more effort. However, this effort requires more time, which likely contributes to the decline in publication rates.

We find no evidence that the PPC contract attenuates writers' prediction error about article pageviews. If anything, the absolute prediction error increased, though the result is not very robust (the effect on the relative prediction error is statistically insignificant). Since earnings in the PPC group depend on understanding how to generate engagement, one would expect that writers might try to learn how to accurately predict article impact. Our point estimates suggest that such learning might be difficult.

Table 4: Treatment Effects on Effort and Prediction Error

	Effort (mins)	Error	Error (P)
Pay per click	7.92***	3724.3***	-0.71
	(1.49)	(468.8)	(0.59)
Choice	10.9***	280.6	-0.89**
	(1.60)	(723.3)	(0.39)
Observations	4897	4892	4892
Mean	31.4	6248.7	2.40

Note: This table reports estimates from Equation (1) for our main experimental sample. The dependent variables are the self-reported number of minutes spent by writers on preparing a submission (column 1), the absolute prediction error when writers predict the number of views generated by their articles (column 2), and the relative prediction error as a percentage change (column 3). The dependent variables were winsorized at the 99th percentile. We report the control group mean for each dependent variable during the intervention period. The unit of observation is the writer-day, where day refers to a specific calendar date. Driscoll-Kraay standard errors are in parentheses. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

Lastly, Table C3 in the appendix presents article-level analyses of factors associated with both predicted and actual engagement. Covering a national story, addressing political topics, and toxicity of the article's text are positively correlated with actual views. In contrast, the presence of positive words in the headline and a more positive tone in the article's body are negatively associated with actual views. These correlations align with the treatment effects of the PPC contract discussed throughout the paper—writers appear to rely on all of the engagement-enhancing factors identified above.

In line with the treatment effects on prediction errors, these factors are not strongly correlated with predicted views. This suggests that other influences, such as overconfidence or self-image concerns, may have shaped the writers' expectations.

4.3 Article Characteristics

4.3.1 Topical Composition

Columns 1-2 of Table 5 report the treatment effects on topical composition. These outcomes matter because part of journalism's, and our partner firm's, mission is to cover a broad set of issues across a variety of geographical locations.

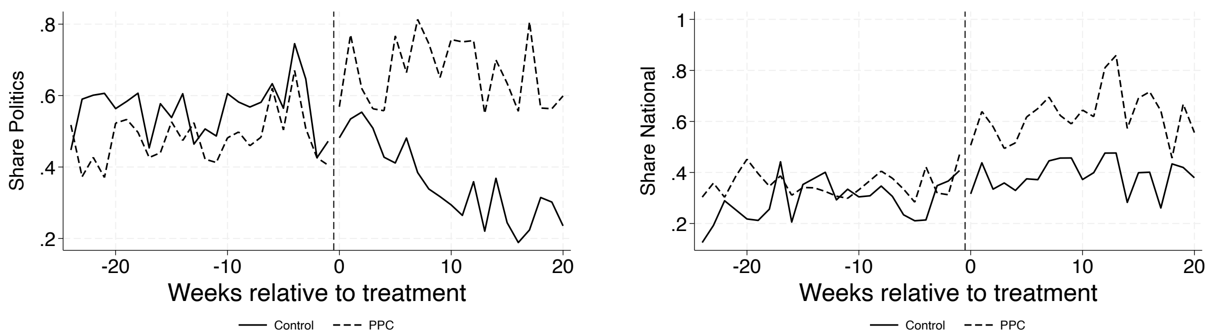
The PPC contract increased the share of articles about politics by 33 pp. The effect is large—the share is almost double the rate in the control group—and significant at the 1% level. The shift suggests that writers realigned their portfolio and that political coverage crowded out other topics. The treatment also raised the share of articles covering national stories by 19 pp, a 50% increase. This may reduce coverage of local news, which in our context may negatively affect exposure to important local information and thereby limit the platform's value to local communities. Our results offer a cautionary tale: if media outlets adopt engagement-based incentives at scale, it could lead to a narrowing of coverage and the omission of issues important to minority audiences.

Table 5: Treatment Effects on Topical Composition and Editor Rating

	Politics	National	Editor Rating
Pay per click	0.33***	0.19***	-0.040
	(0.029)	(0.022)	(0.081)
Choice	0.033	0.084***	-0.064
	(0.023)	(0.019)	(0.081)
Observations	8474	8474	3661
Mean	0.35	0.38	2.69

Note: This table reports estimates from Equation (1) for our main experimental sample. The dependent variables are the share of articles on political issues, the share of articles covering national (as opposed to local) stories, and editor's quality rating of the articles (on a scale from 1 to 5). We report the control group mean for each dependent variable during the intervention period. The unit of observation is the writer-day, where day refers to a specific calendar date. Driscoll-Kraay standard errors are in parentheses. ∗, ∗∗, and ∗∗∗ denote significance at the 10%, 5%, and 1% levels, respectively.

Figure 6 summarizes the raw data on the two outcomes, supporting the regression analysis. Panel A shows that, after minimal differences between the groups in the baseline, a large gap opened up between the PPC condition (high focus on politics) and the control (low focus on politics) following the introduction of the PPC contract. The effect is driven by both an increase in political articles in the PPC group and a decrease in such articles in the control group. Competition is a likely contributing factor—if PPC writers focused on submitting highly engaging political topics quickly, control writers who did not face incentives to maximize views may have shifted attention to other topics. This is plausible as editors often used "duplication" as a rejection criterion. Panel B shows that, following little cross-group variation in the share of national-focused articles during the baseline, the PPC contract led to a higher prevalence of national stories, creating a persistent gap between the treatment groups.



(A) Politics (B) National

Note: Panel A of the figure depicts the mean share of articles on political issues published by writers in any given week by treatment group (PPC vs. control). Panel B presents the mean share of articles covering national (as opposed to local) stories in any given week. The figure is based on the main experimental sample.

Figure 6: Article Topical Composition

The Choice contract had a less pronounced effect on topical composition than the PPC contract. It raised the share of national stories by 8.4 pp—less than half the 19 pp effect in the PPC group), an effect significant at the 1% level. We find no change in the share of political articles. In sum, the Choice contract poses fewer risks to coverage breadth, but it also offers lower returns in terms of engagement.

4.3.2 Editor Ratings

We also investigate how the treatments affected editors' assessments of journalistic quality. Recall that editors received repeated instructions to not consider engagement or popularity in their ratings. Column 3 of Table 5 presents the regression results. Neither the PPC nor the Choice contract had a significant effect on editors' ratings. Despite higher reported effort (see Section 4.2), we find no improvement in assessments based on accuracy, writing style, or the other basic features editors were asked to rate. This suggests that writers used the extra time spent preparing submissions on increasing engagement—likely by adjusting an article's framing or optimizing its tone. Our results on topical composition (Section 4.3.1) support this conclusion. We now turn to toxicity and sentiment (Section 4.3.3) to examine whether these are relevant margins of adjustment.

4.3.3 Sentiment and Toxicity

We examine both article headlines and the article body in our analysis of sentiment and toxicity. Attention-grabbing headlines are central to attracting readers, but we expect the treatments to also affect article content. To pass editorial review, headlines have to match the article's content. As a result, an overly provocative or negative headline might require the article to address a sensitive or controversial topic. In addition, since the author's name appears prominently and is salient to the readers, misleading or click-bait headlines without matching content could reduce readers' future engagement with that writer's work.

Table 6 presents results that support our predictions. First, the PPC contract reduced the number of positive words in headlines by 17%, an effect significant at the 5% level. We find no evidence of an increase in the number of negative words in article headlines. Second, the PPC contract shifted the tone of article body. Articles became less positive, with tone defined as the number of positive words minus the number of negative words, scaled by article length. This effect is large—a 58% more negative tone—and statistically significant at the 1% level.

Table 6: Treatment Effects on Toxicity and Article Sentiment

	Negative	Positive	Tone Body	Tox Body	Tox(1) Body
Pay per click	-0.013	-0.052**	-0.45***	0.0035***	0.010***
	(0.034)	(0.022)	(0.16)	(0.0012)	(0.0038)
Choice	-0.035	0.041*	0.14	0.0040***	0.011***
	(0.032)	(0.023)	(0.14)	(0.0012)	(0.0039)
Observations	8474	8474	8474	8474	8474
Mean	0.66	0.30	-0.77	0.0042	0.0069

Note: This table reports estimates from Equation (1) for our main experimental sample. The dependent variables are the number of negative words in the title, the number of positive words in the title, the tone of the body of the article, defined as the number of positive words less the number of negative words scaled by the overall length of the article, the toxicity score of the body of the article measured by Unitary's Detoxify library, the proportion of articles where the toxicity score of the body exceeds 0.1. We report the control group mean for each dependent variable during the intervention period. The unit of observation is the writer-day, where day refers to a specific calendar date. Driscoll-Kraay standard errors are in parentheses. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

The higher-intensity contract shaped more than article sentiment. The PPC contract significantly increased article toxicity by 83%. It also increased the proportion of articles with a body toxicity score over 0.1 by 1 percentage point—nearly doubling the rate in the control group. These findings suggest that the financial incentives associated with the PPC contract both negatively affected the sentiment of article headlines and body and increased readers' exposure to toxic content. These results complement evidence from prior literature that negativity and toxicity increase user engagement and shows that writers, when incentivized to do so, can strategically adjust content to exploit the demand side.

The effects of the Choice contract are more mixed. Articles in this treatment group became more toxic, but we find no shift in the tone of headlines or article bodies. We interpret these results with caution as it is hard to model how selection into contracts in the Choice condition was correlated with preferences. Selection into the PPC contract in the Choice group may reflect unobserved traits, such as lower risk aversion, some of which may correlate with a greater willingness to write provocatively or with more toxicity.

4.3.4 Heterogeneity Results

Following Section 3.4.4, we present results on heterogeneity in the PPC treatment effect relative to the Control group. Figure B2 in the appendix summarizes the findings. Panel A shows the impact on the number of published articles. Consistent with our theoretical predictions, the PPC contract caused a sharp drop in output among risk averse writers but had little effect on the publication behavior of risk-loving writers.

We also detect treatment effect heterogeneity along other dimensions. Writers who published more at baseline showed a more limited decline in article supply in response to the PPC treatment. Similarly, those who previously focused on political or national topics—both associated with higher engagement—were less affected by the treatment, likely because their existing content strategies already aligned with the PPC contract's incentives. Nonetheless, these patterns are less pronounced than the heterogeneity observed by risk preferences.

Panel B of Figure B2 repeats the heterogeneity analysis for views per publication. Here, we find no differential treatment effect by risk aversion, in line with the theoretical prediction that the PPC contract should increase views per publication regardless of risk preferences. However, we do observe heterogeneity across other dimensions. Writers with a higher baseline share of political content experience larger gains in engagement under the PPC contract, suggesting that topic expertise facilitates adaptation. In addition, those with higher baseline engagement per publication and longer tenure at the firm also generate relatively more engagement under the new contract. These results underscore the role of prior experience in shaping treatment responsiveness.

4.4 Robustness

4.4.1 Robustness Check Description

We conduct three types of robustness checks. First, we address concerns about the influence of highly active writers who may act as outliers. To this end, we repeat the regression analysis after excluding writers whose total number of articles published in the pre-period falls above the 99th percentile. Table C1 in the appendix confirms that the resulting sample remains well balanced across treatment groups.

Second, we assess the robustness of our regression results to an alternative clustering of standard errors—specifically, clustering at the writer level. Because this approach does not fully leverage the richness of our panel data across pre- and post-intervention periods, it substantially reduces statistical efficiency. As a result, it serves as a stringent test of the robustness of our findings.

Third, we discuss robustness to using an ANCOVA specification as opposed to TWFE. Following McKenzie (2012), we estimate ANCOVA models, controlling for mean baseline outcomes $\bar{Y}_{i,pre}$, individual characteristics X_i, strata fixed effects ν_s, and day fixed effects δ_t for post-treatment days. In particular, we estimate:

$$Y_{it} = \gamma_1 PPC_{it} + \gamma_2 Choice_{it} + \phi \bar{Y}_{i,pre} + X_i + \delta_t + \nu_s + \epsilon_{it}, \tag{3}$$

where Y_{it} is computed at the writer-day level. Consistent with our approach to TWFE, when we estimate impacts on total publications and total views with ANCOVA, we set values to zero for days during which a writer did not publish anything. When we look at per-article measures, we drop writer-days in which the writer did not publish articles. Similarly to our second robustness check, we cluster our standard errors at the writer level.

4.4.2 Robustness Check Results

We report robustness results grouped by outcome category. Table C4 in the appendix shows that our findings that the PPC contract (1) reduces the likelihood of publishing articles, (2) reduces the total number of publications, (3) increases overall views generated by writers, and (4) reduces writers' overall earnings are all robust to excluding the most active writers based on pre-period publishing activity. Moreover, all four results are robust to clustering the standard errors at the writer level. Lastly, results (1), (2), and (4) are robust to using the ANCOVA specification instead of TWFE, with the effect on overall views being marginally insignificant.

We conclude that the implication that the PPC treatment increased the firm's profit is robust to the alternative specifications.

Table C5 in the appendix demonstrates that the positive impacts of the PPC contract on views per publication and earnings per publication remain significant in all of our robustness specifications. Moreover, Table C6 in the appendix corroborates our results on effort, with the PPC treatment effects robust to all of the alternative specifications.

Table C7 in the appendix shows that the findings on the topical composition of articles—both the share of political articles and the share of national vs. local articles—are robust to all three robustness specifications.

Lastly, Table C8 in the appendix shows that the finding that the PPC contract reduces the positivity of article headlines is robust across all alternative specifications. The results for article bodies are less robust: the treatment effects on toxicity are robust to the exclusion of the top 1% of writers by pre-period publishing activity but not to other specifications. This asymmetry is unsurprising, as headlines play a central role in attracting pageviews, making changes in their style more detectable. In contrast, article bodies are longer and more variable, which increases noise in textual analysis. Nevertheless, our finding that the PPC contract negatively affects the tone of article bodies—defined as the number of positive words minus the number of negative words, scaled by article length—is robust to both the exclusion of outlier writers and the use of the ANCOVA specification. This strengthens our overall conclusion about the treatment's effect on sentiment and its role in driving engagement.

5 Conclusion

We experimentally evaluate the effects of high-intensity performance pay contracts for journalists at an impactful digital media outlet in Kenya. Journalists assigned to pay-per-click (PPC) contracts generate more pageviews, submit fewer articles, and receive lower overall pay—considerably improving the firm's profits. These efficiency gains came, at least in part, from higher writer effort levels, as they spent more time preparing their articles. However, writers did not direct this extra effort towards improving basic journalistic quality, such as accuracy, depth, or writing style. Instead, writers focused on generating engagement through other means, shifting towards topics that attract attention, reducing coverage of local news, and increasing their articles' negativity and toxicity.

These findings highlight tradeoffs from applying higher-intensity contracts for journalists at scale. While PPC contracts boost engagement and profits, they risk weakening the breadth and quality of coverage. Insufficient reporting on local issues can harm social cohesion and reduce political participation. Increased use of negative and inflammatory content may encourage violence and hate crime (Müller and Schwarz, 2023). These considerations make the design of incentives for creators to supply engaging content an important topic of study.

Future research could explore contract structures that better balance engagement with quality. One option we examined—letting writers self-select into pay-per-click or pay-per-article contracts—reduced profit gains compared to the mandatory PPC contracts, and did not eliminate the negative side effects. Better alternatives are needed. Further work could also explore how the impact of higher-intensity contracts generalize to other media settings, including social media platforms. Finally, the rise of AI tools raises new questions about how journalists and content creators, who may now rely on the new tools, respond to incentives.

References

Allcott, H., L. Braghieri, S. Eichmeyer, and M. Gentzkow (2020): "The welfare effects of social media," *American Economic Review*, 110, 629–676.

Alvarez, F. and D. Argente (2022): "On the effects of the availability of means of payments: The case of Uber," *The Quarterly Journal of Economics*, 137, 1737–1789.

Angelucci, C. and J. Cagé (2019): "Newspapers in times of low advertising revenues," *American Economic Journal: Microeconomics*, 11, 319–364.

Angelucci, C., J. Cagé, and M. Sinkinson (2024): "Media competition and news diets," *American Economic Journal: Microeconomics*, 16, 62–102.

Aridor, G., R. Jiménez-Durán, R. Levy, and L. Song (2024): "The economics of social media," *Journal of Economic Literature*, 62, 1422–1474.

Bandiera, O., I. Barankay, and I. Rasul (2005): "Social preferences and the response to incentives: Evidence from personnel data," *The Quarterly Journal of Economics*, 120, 917–962.

——— (2007): "Incentives for managers and inequality among workers: Evidence from a firm-level experiment," *The Quarterly Journal of Economics*, 122, 729–773.

——— (2011): "Field experiments with firms," *Journal of Economic Perspectives*, 25, 63–82.

Baron, D. P. (2006): "Persistent media bias," *Journal of Public Economics*, 90, 1–36.

Bartlett, R. (2013): "The Forbes contributor model: Technology, feedback and incentives," Accessed June 29, 2025.

Baumeister, R. F., E. Bratslavsky, C. Finkenauer, and K. D. Vohs (2001): "Bad is stronger than good," *Review of General Psychology*, 5, 323–370.

Beattie, G. (2020): "Advertising and media capture: The case of climate change," *Journal of Public Economics*, 188, 104219, 1–17.

Beknazar-Yuzbashev, G., R. Jiménez-Durán, J. McCrosky, and M. Stalinski (2025): "Toxic content and user engagement on social media: Evidence from a field experiment," *CESifo Working Paper*.

Beknazar-Yuzbashev, G., R. Jiménez-Durán, and M. Stalinski (2024): "A model of harmful yet engaging content on social media," *AEA Papers and Proceedings*, 114, 678–683.

Bénabou, R. and J. Tirole (2006): "Incentives and prosocial behavior," *American Economic Review*, 96, 1652–1678.

Benton, J. (2009): "With ad revenue up 35%, Gawker Media returns to pageview bonuses—and plans checkbook journalism," Accessed June 29, 2025.

Bolton, P., J. Scheinkman, and W. Xiong (2006): "Executive compensation and short-termist behaviour in speculative markets," *The Review of Economic Studies*, 73, 577–610.

Bursztyn, L., G. Egorov, R. Enikolopov, and M. Petrova (2024): "Social media and xenophobia: Evidence from Russia," Tech. Rep. 26567, National Bureau of Economic Research.

Butschek, S., R. González Amor, P. Kampkötter, and D. Sliwka (2022): "Motivating gig workers – evidence from a field experiment," *Labour Economics*, 75, 102105, 1–19.

Butschek, S., P. Kampkötter, and D. Sliwka (2017): "Paying gig workers," .

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, 1–24.

Cagé, J., M. Hengel, N. Hervé, and C. Urvoy (2025): "Political bias in the media: Evidence from the universe of French broadcasts, 2002-2020," Tech. Rep. 4036211, SSRN.

Cameron, A. C. and D. L. Miller (2015): "A practitioner's guide to cluster-robust inference," *Journal of Human Resources*, 50, 317–372.

Campante, F., R. Durante, and A. Tesei (2022): "Media and social capital," *Annual review of economics*, 14, 69–91.

Chandler, D. and A. Kapelner (2013): "Breaking monotony with meaning: Motivation in crowdsourcing markets," *Journal of Economic Behavior & Organization*, 90, 123–133.

De Quidt, J. (2018): "Your loss is my gain: A recruitment experiment with framed incentives," *Journal of the European Economic Association*, 16, 522–559.

DellaVigna, S. and D. Pope (2018): "What motivates effort? Evidence and expert forecasts," *The Review of Economic Studies*, 85, 1029–1069.

Di Tella, R. and I. Franceschelli (2011): "Government advertising and media coverage of corruption scandals," *American Economic Journal: Applied Economics*, 3, 119–151.

Dohmen, T. and A. Falk (2011): "Performance pay and multidimensional sorting: Productivity, preferences, and gender," *American Economic Review*, 101, 556–590.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the european economic association*, 9, 522–550.

Driscoll, J. C. and A. C. Kraay (1998): "Consistent covariance matrix estimation with spatially dependent panel data," *The Review of Economics and Statistics*, 80, 549–560.

Feng, J. (2023): "Classifying Toxic Comments with NLP," https://johnfengphd.medium.com/classifying-toxic-comments-with-nlp-3418ee9f399, accessed: 2025-06-07.

Finan, F., B. A. Olken, and R. Pande (2017): "The personnel economics of the developing state," in *Handbook of Field Experiments*, ed. by A. V. Banerjee and E. Duflo, North-Holland, vol. 2, chap. 6, 467–514.

Freedom House (2017): "Freedom of the press 2017," Tech. rep., Freedom House.

Gentzkow, M. and J. M. Shapiro (2010): "What drives media slant? Evidence from US daily newspapers," *Econometrica*, 78, 35–71.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2011): "The effect of newspaper entry and exit on electoral politics," *American Economic Review*, 101, 2980–3018.

Gicheru, C. and G. Nyabuga (2024): "Reuters Institute Digital News Report 2024: Kenya," Tech. rep., Reuters Institute for the Study of Journalism, University of Oxford, accessed: 2025-05-31.

Groseclose, T. and J. Milyo (2005): "A measure of media bias," *The Quarterly Journal of Economics*, 120, 1191–1237.

Hanu, L. and Unitary (2020): "Detoxify," https://github.com/unitaryai/detoxify.

Hoechle, D. (2007): "Robust standard errors for panel regressions with cross-sectional dependence," *The stata journal*, 7, 281–312.

Hoffman, M. and S. V. Burks (2017): "Worker overconfidence: Field evidence and implications for employee turnover and returns from training," Tech. Rep. 23240, National Bureau of Economic Research.

Holmstrom, B. and P. Milgrom (1987): "Aggregation and linearity in the provision of intertemporal incentives," *Econometrica*, 55, 303–328.

——— (1991): "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design," *Journal of Law, Economics, & Organization*, 7, 24–52.

Hong, B., L. Kueng, and M.-J. Yang (2019): "Complementarity of performance pay and task allocation," *Management Science*, 65, 5152–5170.

House, F. (2023): "Freedom on the Net 2023: Kenya Country Report," Tech. rep., Freedom House.

Kerkhof, A. (2024): "Advertising and content differentiation: Evidence from YouTube," *The Economic Journal*, 134, 2912–2950.

Konopliov, A. (2024): "US Newspaper Industry Statistics & Facts (2024)," Accessed: 2025-05-31.

Larcinese, V., R. Puglisi, and J. M. Snyder, Jr (2011): "Partisan bias in economic news: Evidence on the agenda-setting behavior of US newspapers," *Journal of Public Economics*, 95, 1178–1189.

Lazear, E. P. (2000): "Performance pay and productivity," *American Economic Review*, 90, 1346–1361.

List, J. A. and F. Momeni (2021): "When corporate social responsibility backfires: Evidence from a natural field experiment," *Management Science*, 67, 8–21.

Lohner, J., S. Banjac, and I. Neverla (2016): "Mapping structural conditions of journalism in Kenya," Tech. rep., MeCoDEM.

McKenzie, D. (2012): "Beyond baseline and follow-up: The case for more T in experiments," *Journal of Development Economics*, 99, 210–221.

Metzger, Z. (2024): "The State of Local News 2024," Tech. rep., Medill Local News Initiative, Northwestern University, accessed: 2025-05-31.

Miller, G. and K. S. Babiarz (2013): "Pay-for-performance incentives in low- and middle-income country health programs," Tech. Rep. 18932, National Bureau of Economic Research.

Mullainathan, S. and A. Shleifer (2005): "The market for news," *American Economic Review*, 95, 1031–1053.

Müller, K. and C. Schwarz (2023): "From hashtag to hate crime: Twitter and antiminority sentiment," *American Economic Journal: Applied Economics*, 15, 270–312.

Namwaya, O. (2018): "Kenya: Government crackdown threatens rights," *allAfrica*.

Nyabuga, G. (2023): "Policy paper media under pressure: The trouble with press freedom in Kenya," Policy brief, Friedrich Naumann Foundation for Freedom.

Oberholzer-Gee, F. and J. Waldfogel (2009): "Media markets and localism: Does local news *en Español* boost Hispanic voter turnout?" *American Economic Review*, 99, 2120–2128.

Pew Research Center (2023): "Newspapers Fact Sheet," Fact sheet, Pew Research Center, Washington, DC, accessed 29 June 2025.

Prat, A. and D. Strömberg (2013): "The political economy of mass media," in *Advances in Economics and Econometrics: Tenth World Congress, Volume II, Applied Economics*, ed. by D. Acemoglu, M. Arellano, and E. Dekel, Cambridge University Press, Econometric Society Monographs, chap. 3, 135–187.

Prendergast, C. (2002): "The tenuous trade-off between risk and incentives," *Journal of Political Economy*, 110, 1071–1102.

——— (2015): "The empirical content of pay-for-performance," *Journal of Law, Economics, & Organization*, 31, 242–261.

Raith, M. (2008): "Specific knowledge and performance measurement," *The RAND Journal of Economics*, 39, 1059–1079.

Reporters Without Borders (2024): "Kenya," Accessed: 2025-05-31.

Rizzi, M. (2024): "Self-regulation of social media and the evolution of content: A cross-platform analysis," Tech. Rep. 5018309, SSRN.

Robertson, C. E., N. Pröllochs, K. Schwarzenegger, P. Pärnamets, J. J. Van Bavel, and S. Feuerriegel (2023): "Negativity drives online news consumption," *Nature Human Behaviour*, 7, 812–822.

Schulhofer-Wohl, S. and M. Garrido (2013): "Do newspapers matter? Short-run and long-run evidence from the closure of *The Cincinnati Post*," *Journal of Media Economics*, 26, 60–81.

Shearer, B. (2004): "Piece rates, fixed wages and incentives: Evidence from a field experiment," *The Review of Economic Studies*, 71, 513–534.

Simiyu, T. F. (2014): "Media ownership and the coverage of the 2013 general election in Kenya: Democracy at the crossroads," *Global Media Journal - African Edition*, 8, 114–145.

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

Starkman, D. (2013): "The mission sounds simple: Pay for clicks," Accessed June 29, 2025.

Szeidl, A. and F. Szucs (2021): "Media capture through favor exchange," *Econometrica*, 89, 281–310.

Wu, Y. (2017): "Authority, incentives, and performance: Evidence from a Chinese newspaper," *The Review of Economics and Statistics*, 99, 16–31.

Wu, Y. and F. Zhu (2022): "Competition, contracts, and creativity: Evidence from novel writing in a platform market," *Management Science*, 68, 8613–8634.

Young, L. and S. Soroka (2012): "Affective news: The automated coding of sentiment in political texts," *Political Communication*, 29, 205–231.

Zhuravskaya, E., M. Petrova, and R. Enikolopov (2020): "Political effects of the internet and social media," *Annual Review of Economics*, 12, 415–438.

A Theoretical Framework

The model presented in this section is based on the classic linear-contract moral-hazard model with normal noise and CARA preferences (e.g., Holmstrom and Milgrom, 1987, 1991). One key difference in our setting is that we allow agents to choose the number of tasks to perform. It is very similar in the general question and approach to Butschek et al. (2017). Our model structure is similar to theirs, but their focus, questions, and results differ significantly from ours.

Assume that the decision maker chooses n, the number of articles to write and submit, and e, how much effort to expend on each article on average. To avoid uninteresting discreteness issues, we assume that n can be any non-negative real number: i.e., $n \geq 0$. Likewise, let $e \geq 0$, where we normalize 0 to be the lowest level of effort that guarantees meeting the editorial standards and subsequent publication.

Conditional on a chosen effort level e, the number of pageviews that an article i attracts is $e + \eta_i$, where each η_i is an iid normal variable with mean $m > 0$ and variance σ^2. Thus, for a given production plan (n, e), the total pageviews of all articles is $ne + \varepsilon$, where ε is a normal variable with mean nm and variance $n\sigma^2$.

A payment contract is a pair of non-negative parameters (α, β), where α is the fee per article and β is a scaling parameter that determines the pageviews-contingent payment. Thus, if an article attracts p pageviews, the payment for that article is $\alpha + \beta p$. The agent has CARA utility with risk-aversion parameter $\rho > 0$ over her monetary payment minus the cost of effort given by the function $(n, e) \mapsto c_1 ne^2 + c_2 n^2$ for some strictly positive parameters c_1 and c_2:

$$U((\alpha, \beta), (n, e)) = -\exp\left(-\rho\left(\alpha n + \beta(ne + \varepsilon) - c_1 ne^2 - c_2 n^2\right)\right).$$

We compare a flat payment contract (where $\alpha_{FC} > 0$ and $\beta_{FC} = 0$) to a linear (fee-per-pageview) one (where $\alpha_{LC} = 0$ and $\beta_{LC} > 0$).

Proposition 1 *The utility-maximizing production plan under the flat contract* $(\alpha, 0)$ *is*

$$(n^*_{FC}, e^*_{FC}) = \left(\frac{\alpha}{2c_2}, 0\right). \tag{4}$$

The utility-maximizing production plan under the linear contract $(0, \beta)$ *is*

$$(n^*_{LC}, e^*_{LC}) = \left(\max\left\{0, \frac{\beta}{2c_2}\left(\frac{\beta}{4c_1} + m - \frac{\rho}{2}\beta\sigma^2\right)\right\}, \frac{\beta}{2c_1}\right). \tag{5}$$

As expected, writers exert more effort under the linear than under the flat contract. The comparison for the optimal number of articles under the two contracts depends on the parameters of the model. If risk aversion and/or the marginal cost of effort are relatively high, it is optimal for the agent to choose small n^*_{LC} or even $n^*_{LC} = 0$ under the linear contract. As we observe below, it is possible that a writer may choose to produce *more* articles under the linear than under the flat contract.

In what follows, assume the following relationship between the parameters of the two contracts:

$$1 = \alpha_{FC} = \beta_{LC} m \Leftrightarrow \beta_{LC} = \frac{1}{m}, \tag{6}$$

where we have normalized $\alpha_{FC} = 1$. These parameters are based on the way our partner firm chose its linear contract. Namely, the per-pageview payment (β_{LC} in our model) was chosen so that the payout under the linear contract, *when calculated using the average per-article pageviews* for the flat contract (m as $e^*_{FC} = 0$), would equal the flat-contract per-article payment (α_{FC}). Note that with the parametrization from (6), (4) and (5) imply that the agent would write strictly more articles under the linear than under the flat contract whenever ρ is sufficiently small.

Proposition 2 *There exist* $\bar{\rho}_u > 0$ *and* $\bar{\rho}_{pv} > 0$*, such that* $\bar{\rho}_u < \bar{\rho}_{pv}$ *and*

- *for all* $\rho < \bar{\rho}_u$*, the agent strictly prefers the linear contract* $\left(0, \frac{1}{m}\right)$ *over the flat contract* $(1, 0)$ *and conversely for* $\rho > \bar{\rho}_u$*; and*
- *for all* $\rho < \bar{\rho}_{pv}$*, the linear contract* $\left(0, \frac{1}{m}\right)$ *induces more pageviews than the flat contract* $(1, 0)$ *at the agent's optimum and conversely for* $\rho > \bar{\rho}_{pv}$.

Proposition 2 is consistent with our finding that the average number of pageviews per writer increases under the main linear-contract treatment while, at the same time, relatively few writers opt into the linear contract when given the choice. This implies that most writers have coefficient of absolute risk aversion in the interval $(\bar{\rho}_u, \bar{\rho}_{pv})$. Consequently, most writers are made worse off by the linear contract relative to the flat one.

Several testable implications emerge from the model: first, we expect average per-article effort to increase under the linear contract. This should translate into higher average pageviews for treatment writers. Second, the number of submissions under the PPC contract should diminish in writer risk aversion, but risk aversion should not affect article quantity in the control group. Third, writers should sort based on risk aversion, with only sufficiently risk tolerant writers choosing the performance contract in the Choice treatment.

Proofs

Proof of Proposition 1. Given a payment contract (α, β), an agent with a coefficient of absolute risk aversion ρ solves

$$\max_{(n,e)} \left\{ \mathbb{E}\left[-\exp\left(-\rho \left(\alpha n + \beta n e + \beta \varepsilon - c_1 n e^2 - c_2 n^2 \right) \right) \right] \right\}.$$

As $\mathbb{E}[-\exp(-\rho x)] = -\exp(-\rho(\mu - (\rho/2)\sigma^2))$ for any $x \sim N(\mu, \sigma^2)$, the optimization problem is maximizing the certainty equivalent:

$$\max_{(n,e)} \left\{ \underbrace{\alpha n + \beta e n + \beta n m - c_1 n e^2 - c_2 n^2 - \frac{\rho}{2} \beta^2 n \sigma^2}_{=: F((\alpha,\beta),(n,e))} \right\}.$$

Under the flat contract $(\alpha, 0)$, the agent does not benefit from choosing any $e > 0$ and her optimization problem can be simplified to:

$$\max_{n} \left\{ \alpha n - c_2 n^2 \right\}. \tag{7}$$

This objective function is strictly concave and $n^*_{FC} = \frac{\alpha}{2c_2}$.

Under the linear contract $(0, \beta)$, the agent solves

$$\max_{(n,e)} \left\{ \beta e n + \beta n m - c_1 n e^2 - c_2 n^2 - \frac{\rho}{2} \beta^2 n \sigma^2 \right\}. \tag{8}$$

If $n = 0$, the value of e does not change the value of the objective function and e^*_{LC} is indeterminate. If $n > 0$ instead, $F((0, \beta), (n, e))$ is strictly concave in e and the optimal value $e^*_{LC} = \frac{\beta}{2c_1}$ does not depend on n. Plugging $\frac{\beta}{2c_1}$ for e in (8) reduces the maximization problem to:

$$\max_{n \geq 0} \left\{ n \left(\frac{\beta^2}{4c_1} + \beta m - c_2 n - \frac{\rho}{2} \beta^2 \sigma^2 \right) \right\}.$$

This objective function is strictly concave in n and is positive for some $n > 0$ if and only if $\frac{\beta}{4c_1} + m - \frac{\rho}{2} \beta \sigma^2 > 0$. This, together with the first-order condition for n, gives us $n^*_{LC} =$

$\max\left\{0, \frac{\beta}{2c_2}\left(\frac{\beta}{4c_1} + m - \frac{\rho}{2}\beta\sigma^2\right)\right\}$. ■

Proof of Proposition 2. If $n^*_{LC} = 0$, both the optimal expected utility and the corresponding expected number of pageviews are lower under the linear contract than under the flat contract. In what follows, assume instead that $n^*_{LC} > 0$ for $(\alpha_{LC}, \beta_{LC}) = \left(0, \frac{1}{m}\right)$ or

$$\begin{aligned} &\frac{\beta_{LC}}{2c_2}\left(\frac{\beta_{LC}}{4c_1} + m - \frac{\rho}{2}\beta_{LC}\sigma^2\right) > 0 \\ \Leftrightarrow &\frac{1}{4c_1 m} + m - \frac{\rho\sigma^2}{2m} > 0 \\ \Leftrightarrow &\rho < \bar{\rho}_n := \frac{1 + 4m^2 c_1}{2\sigma^2 c_1}. \end{aligned} \tag{9}$$

The expected number of total pageviews under the linear contract $(\alpha_{LC}, \beta_{LC}) = \left(0, \frac{1}{m}\right)$ is

$$\begin{aligned} n^*_{LC}(e^*_{LC} + m) &= \frac{\beta_{LC}}{2c_2}\left(\frac{\beta_{LC}}{4c_1} + m - \frac{\rho}{2}\beta_{LC}\sigma^2\right)\left(\frac{\beta_{LC}}{2c_1} + m\right) \\ &= \frac{1}{2c_2 m}\left(\frac{1}{8(c_1)^2 m^2} + \frac{3}{4c_1} + m^2 - \frac{\rho\sigma^2}{4c_1 m^2} - \frac{\rho\sigma^2}{2}\right). \end{aligned} \tag{10}$$

.

The expected number of total pageviews under the flat contract $(\alpha_{FC}, \beta_{FC}) = (1, 0)$ is

$$n^*_{FC}(e^*_{FC} + m) = \frac{\alpha_{FC}}{2c_2}(0 + m) = \frac{m}{2c_2}. \tag{11}$$

Subtract (11) from (10) to find that the expected number of total pageviews is greater under the linear contract if and only if

$$\begin{aligned} &\frac{1}{2c_2 m}\left(\frac{1}{8(c_1)^2 m^2} + \frac{3}{4c_1} + m^2 - \frac{\rho\sigma^2}{4c_1 m^2} - \frac{\rho\sigma^2}{2}\right) - \frac{m}{2c_2} > 0 \\ \Leftrightarrow &\frac{1}{2c_2 m}\left(\frac{1}{8(c_1)^2 m^2} + \frac{3}{4c_1} - \frac{\rho\sigma^2}{4c_1 m^2} - \frac{\rho\sigma^2}{2}\right) > 0 \\ \Leftrightarrow &\rho < \bar{\rho}_{pv} := \frac{1 + 6c_1 m^2}{2\sigma^2 c_1 + 4\sigma^2 (c_1)^2 m^2}. \end{aligned}$$

The certainty equivalent of the lottery induced by (n^*_{LC}, e^*_{LC}) under $(\alpha_{LC}, \beta_{LC}) = \left(0, \frac{1}{m}\right)$ is

$$
\begin{aligned}
&F((\alpha_{LC}, \beta_{LC}), (n^*_{LC}, e^*_{LC})) \\
=&n^*_{LC}\left(\alpha_{LC} + \beta_{LC}e^*_{LC} + \beta_{LC}m - c_1(e^*_{LC})^2 - c_2 n^*_{LC} - \frac{\rho}{2}(\beta_{LC})^2\sigma^2\right) \\
=&\frac{\beta_{LC}}{2c_2}\left(\frac{\beta_{LC}}{4c_1} + m - \frac{\rho}{2}\beta_{LC}\sigma^2\right) \times \\
&\times\left(\alpha_{LC} + \beta_{LC}\frac{\beta_{LC}}{2c_1} + \beta_{LC}m - c_1\left(\frac{\beta_{LC}}{2c_1}\right)^2 - c_2\frac{\beta_{LC}}{2c_2}\left(\frac{\beta_{LC}}{4c_1} + m - \frac{\rho}{2}\beta_{LC}\sigma^2\right) - \frac{\rho}{2}(\beta_{LC})^2\sigma^2\right) \\
=&\frac{1}{2c_2 m}\left(\frac{1}{4c_1 m} + m - \frac{\rho\sigma^2}{2m}\right)\left(\frac{1}{8c_1 m^2} + \frac{1}{2} - \frac{\rho\sigma^2}{4m^2}\right) \\
=&\frac{1}{4c_2}\left(\frac{1}{4c_1 m^2} + 1 - \frac{\rho\sigma^2}{2m^2}\right)^2.
\end{aligned} \tag{12}
$$

The corresponding certainty equivalent for (n^*_{FC}, e^*_{FC}) and $(\alpha_{FC}, \beta_{FC}) = (1, 0)$ is

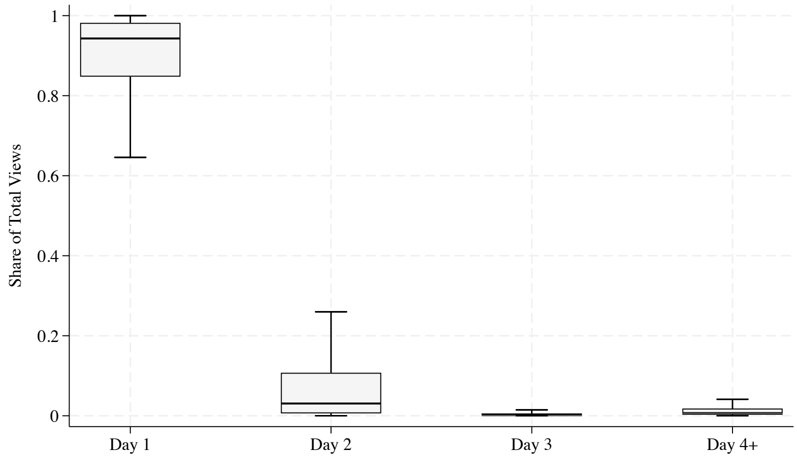
$$
\begin{aligned}
&F((\alpha_{FC}, \beta_{FC}), (n^*_{FC}, e^*_{FC})) \\
=&\alpha_{FC}n^*_{FC} + \beta_{FC}e^*_{FC}n^*_{FC} + \beta_{FC}n^*_{FC}m - c_1 n^*_{FC}(e^*_{FC})^2 - c_2(n^*_{FC})^2 - \frac{\rho}{2}(\beta_{FC})^2 n^*_{FC}\sigma^2 \\
=&\frac{1}{2c_2} - c_2\left(\frac{1}{2c_2}\right)^2 \\
=&\frac{1}{4c_2}.
\end{aligned} \tag{13}
$$

Subtracting (13) from (12), we find that the agent's expected utility is greater under the linear contract if and only if

$$
\begin{aligned}
&\frac{1}{4c_2}\left(\frac{1}{4c_1 m^2} + 1 - \frac{\rho\sigma^2}{2m^2}\right)^2 - \frac{1}{4c_2} > 0 \\
\Leftrightarrow&\frac{1}{4c_2}\left(\frac{1}{4c_1 m^2} - \frac{\rho\sigma^2}{2m^2}\right)\left(\frac{1}{4c_1 m^2} + 2 - \frac{\rho\sigma^2}{2m^2}\right) > 0 \\
\Leftrightarrow&\frac{1}{4c_1 m^2} - \frac{\rho\sigma^2}{2m^2} > 0 \\
\Leftrightarrow&\rho < \bar{\rho}_u := \frac{1}{2\sigma^2 c_1},
\end{aligned}
$$

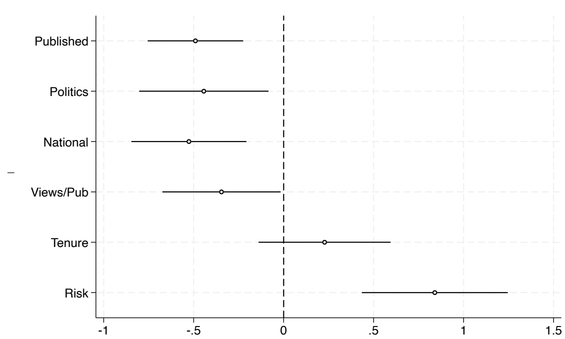
where the third line uses the fact that (9) implies $\frac{1}{4c_1 m^2} + 2 - \frac{\rho\sigma^2}{2m^2} > 0$. Finally, simple algebra can be used to verify that $\bar{\rho}_u < \bar{\rho}_{pv} < \bar{\rho}_n$. ■

B Additional Figures

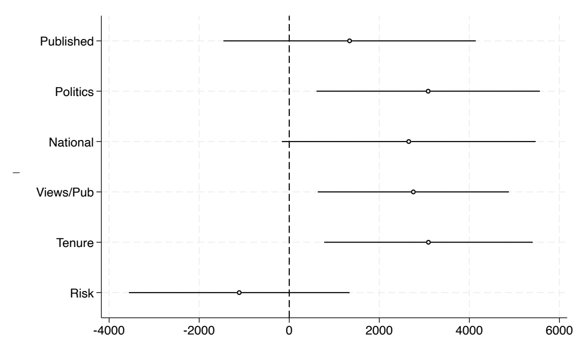


Note: The figure shows the distribution of the share of views each article received on the day of publication (first box plot), the second day, the third day, and the fourth day or later (last box plot). Each box plot displays the first quartile (bottom edge of the box), the median (line inside the box), and the third quartile (top edge of the box). The whiskers extend to the smallest and largest data points within 1.5 times the interquartile range (IQR) from the box. The plot is based on a sample of N=31,556 articles, which includes all articles published by writers in the main sample over the course of the study.

Figure B1: Distribution of Share of Views by Day since Publication



(a) Number of published articles



(b) Views per publication

Note: This figure reports estimates from Equation (2). The dependent variables are the number of published articles (Panel A) and the number of views per publication (Panel B). The reported coefficients correspond to the interaction between the treatment dummy for the PPC contract and an indicator for whether a writer's value of the characteristic under consideration is above the median (as part of the heterogeneity analysis). The y-axis of both panels lists the characteristics, in order: (1) number of published articles at baseline, (2) share of political articles at baseline, (3) share of articles covering national stories at baseline, (4) views per publication at baseline, (5) pre-intervention tenure with the company, and (6) risk preferences from the baseline survey (with higher values indicating greater risk tolerance). All categories except risk are based on the main experimental sample of 217 writers. The risk preference analysis uses a subsample of 127 writers who completed the relevant baseline survey question. We report point estimates with 95% confidence intervals. In all regressions, the unit of observation is the writer-day, and standard errors are clustered at the writer level.

Figure B2: Heterogeneity Analysis (PPC vs. Control)

C Additional Tables

Table C1: Balance on Writer Characteristics: Top 1% by Publishing Excluded

		Treatment Arm			
	All (1)	Pay per article (2)	Pay per click (3)	Choice (4)	p-value (5)
Tenure (months)	7.82	7.92	8.08	7.43	0.77
	(0.42)	(0.76)	(0.69)	(0.74)	
Articles published	56.83	55.07	60.53	54.75	0.87
	(7.00)	(10.97)	(12.73)	(12.77)	
Total views	98386.15	90695.97	107132.65	97157.13	0.73
	(13686.52)	(21020.36)	(24744.22)	(25502.72)	
Views/Published	1477.48	1524.54	1365.13	1547.24	0.72
	(115.74)	(202.40)	(167.25)	(232.11)	
Share national	0.36	0.36	0.36	0.35	0.98
	(0.03)	(0.05)	(0.04)	(0.05)	
Share politics	0.37	0.36	0.30	0.44	0.02
	(0.02)	(0.04)	(0.04)	(0.05)	
Negative (Title)	0.63	0.68	0.59	0.63	0.36
	(0.03)	(0.06)	(0.05)	(0.05)	
Positive (Title)	0.33	0.29	0.31	0.38	0.27
	(0.02)	(0.03)	(0.03)	(0.04)	
Tone (Body)	-0.33	-0.44	-0.35	-0.18	0.72
	(0.17)	(0.31)	(0.31)	(0.27)	
Toxicity	0.01	0.01	0.01	0.01	0.82
	(0.00)	(0.00)	(0.00)	(0.00)	
Toxicity (Binary)	0.02	0.03	0.01	0.02	0.28
	(0.01)	(0.02)	(0.01)	(0.01)	
Observations	214	72	73	69	

Note: This table reports means and standard deviations (in parentheses) for 11 observables, overall (column 1) and by treatment condition (columns 2-4). Column 5 summarizes balance tests based on regressions of each variable on the treatment indicators (PPC dummy and Choice dummy) with strata fixed effects. Specifically, it reports the p-values for a joint test that the two treatment dummies equal zero. The table is based on the main experimental sample, excluding writers whose number of publications in the pre-period falls above the 99th percentile. The resulting sample consists of N=214 writers. Summary statistics cover the full baseline period, except for tenure with the company, which measures months from a writer's registration with the firm and the intervention start date. We report the following covariates: (1) tenure at the firm at the time of randomization (in months), (2) the number of articles published (winsorized at the 95th percentile), (3) total pageviews (winsorized at the 95th percentile), (4) per-article pageviews, (5) the share of articles covering national issues, (6) the share of articles covering politics, (7) the number of negative words in the title, (8) the number of positive words in the title, (9) article body tone, defined as the number of positive words less the number of negative words scaled by the overall length of the article, (10) the toxicity score of the article body, measured by Unitary's Detoxify library, (11) the proportion of articles with a toxicity score over 0.1.

Table C2: Patterns of Article Publications

	Number of Articles		Number of Days	
Period	Mean	Median	Mean	Median
Baseline	159.65	49.50	40.34	23.00
Experiment	156.91	71.00	42.36	30.50
Total	316.56	159.50	82.70	57.50

Note: The table presents descriptive statistics on article publication patterns among writers in the main experimental sample (N=217). It reports the average and median number of unique articles published (columns 1–2), as well as the average and median number of distinct days on which writers published articles (columns 3-4).

Table C3: Correlates of Views and Predicted Views

	Baseline			Intervention		
	L(Views)	L(Prediction)	L(Prediction)	L(Views)	L(Prediction)	L(Prediction)
National	0.45***	-0.0090	0.011	0.25***	-0.074**	0.35***
	(0.070)	(0.059)	(0.053)	(0.034)	(0.031)	(0.055)
Politics	0.47***	0.070	0.090*	0.73***	0.083***	-0.056
	(0.067)	(0.052)	(0.053)	(0.031)	(0.027)	(0.055)
Negative	0.0073	-0.025	-0.018	-0.052***	-0.010	-0.066**
	(0.042)	(0.031)	(0.035)	(0.018)	(0.016)	(0.031)
Positive	-0.17***	-0.0026	-0.018	-0.14***	-0.043*	0.010
	(0.049)	(0.035)	(0.041)	(0.026)	(0.022)	(0.046)
Tone	-0.050***	-0.0045	-0.0059	-0.050***	0.00014	-0.0041
	(0.0087)	(0.0062)	(0.0075)	(0.0040)	(0.0035)	(0.0073)
Toxic	0.79***	0.11	0.67	0.82***	0.16	0.18
	(0.25)	(0.13)	(0.56)	(0.16)	(0.14)	(0.31)
Observations	2366	1095	1271	9581	6523	3059
Mean	6.90	7.66	7.46	6.34	7.61	8.12
Sample	Combined	Control	PPC	Combined	Control	PPC

Note: This table reports OLS regressions of the number of article views and the predicted number of article views (by its author) on six observables: (1) a dummy equal to 1 if the article covers a national story, (2) a dummy equal to 1 if the article covers politics, (3) the number of positive words in the article's headline, (4) the number of negative words in the headline, (5) the tone of the body of the article, defined as the number of positive words less the number of negative words scaled by the overall length of the article, and (6) a dummy equal to 1 if the toxicity score of the body of the article measured by Unitary's Detoxify library exceeds 0.1. The dependent variables are expressed in logarithmic terms. Different columns pertain to different samples of articles (based on the Control group writers, PPC group writers, or both) and collection periods (baseline period or the intervention period). The regressions were estimated at the article-level. All regressions cover only articles for which both the predicted views are actual views are available. Standard errors clustered at the writer level are parenthesized. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

Table C4: Robustness: Supply, Views, and Earnings

	Published>0	Published>0	Published>0	Published	Published	Published	Views	Views	Views	Payment	Payment	Payment
Pay per click	-0.079***	-0.080***	-0.077***	-0.46***	-0.46***	-0.38***	526.1***	771.3**	546.6	-35.8***	-30.2*	-25.3*
	(0.0081)	(0.030)	(0.026)	(0.036)	(0.16)	(0.13)	(121.0)	(384.1)	(354.5)	(4.77)	(16.1)	(13.9)
Choice	-0.052***	-0.054*	-0.035	-0.26***	-0.35*	-0.20	-173.0*	-29.9	-9.50	-25.6***	-33.7*	-20.5
	(0.0088)	(0.031)	(0.027)	(0.038)	(0.18)	(0.14)	(90.7)	(257.8)	(283.8)	(3.88)	(17.7)	(14.0)
Observations	59530	60487	32767	59530	60487	32767	59530	60487	32767	59530	60487	32767
Mean	0.16	0.16	0.16	0.62	0.62	0.62	723.6	723.6	723.6	61.7	61.7	61.7
Specification	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA
S.E.	DK	Writer	Writer	DK	Writer	Writer	DK	Writer	Writer	DK	Writer	Writer
Sample	p99	Main	Main	p99	Main	Main	p99	Main	Main	p99	Main	Main

Note: This table presents robustness analyses for four dependent variables: (i) a binary indicator for whether a writer published at least one article (columns 1–3); (ii) the number of articles published (columns 4–6); (iii) the number of pageviews generated by a writer's articles within seven days of publication (columns 7–9); and (iv) writer earnings, measured in KES (columns 10–12). For each outcome, the first column reports estimates from Equation (1) using our main experimental excluding writers whose pre-period publication count exceeded the 99th percentile. Driscoll-Kraay standard errors are parenthesized. The second column reports estimates from Equation (1) using the main sample of writers with standard errors clustered at the writer level. The third column reports estimates from Equation (3), also using the main sample with writer-level clustered standard errors. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

Table C5: Robustness: Views and Payments per Article

	Views/Pub	Views/Pub	Views/Pub	Payment/Pub	Payment/Pub	Payment/Pub
Pay per click	4013.9***	3924.6***	3571.3***	49.8***	53.5***	41.9***
	(269.8)	(913.6)	(828.6)	(6.02)	(15.6)	(14.4)
Choice	196.4	481.4	665.3	3.06*	2.37	5.05
	(173.1)	(403.5)	(405.3)	(1.78)	(1.64)	(4.52)
Observations	7809	8474	3982	7809	8474	3982
Mean	1320.3	1320.3	1320.3	100	100	100
Specification	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA
S.E.	DK	Writer	Writer	DK	Writer	Writer
Sample	p99	Main	Main	p99	Main	Main

Note: This table presents robustness analyses for two dependent variables: (i) the number of pageviews per publication generated by writers' articles over the period of 7 days since the publication date (columns 1–3); and (ii) writer earnings per article, in KES (columns 4–6). For each outcome, the first column reports estimates from Equation (1) using our main experimental excluding writers whose pre-period publication count exceeded the 99th percentile. Driscoll-Kraay standard errors are parenthesized. The second column reports estimates from Equation (1) using the main sample of writers with standard errors clustered at the writer level. The third column reports estimates from Equation (3), also using the main sample with writer-level clustered standard errors. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

Table C6: Robustness: Effort and Prediction Error

	Effort (mins)	Effort (mins)	Effort (mins)	Error	Error	Error	Error (P)	Error (P)	Error (P)
Pay per click	5.18***	7.92*	5.88*	3602.3***	3724.3***	2018.3	-1.17*	-0.71	0.98
	(1.36)	(4.46)	(3.21)	(490.1)	(716.7)	(1631.8)	(0.63)	(2.08)	(1.21)
Choice	11.7***	10.9**	10.5***	540.5	280.6	-555.8	-1.10***	-0.89	-0.40
	(1.80)	(4.57)	(3.07)	(750.6)	(852.1)	(1465.3)	(0.41)	(1.43)	(1.03)
Observations	4548	4897	3953	4543	4892	3948	4543	4892	3948
Mean	31.4	31.4	31.4	6248.7	6248.7	6238.5	2.40	2.40	2.39
Specification	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA
S.E.	DK	Writer	Writer	DK	Writer	Writer	DK	Writer	Writer
Sample	p99	Main	Main	p99	Main	Main	p99	Main	Main

Note: This table presents robustness analyses for three dependent variables: (i) the self-reported number of minutes spent by writers on preparing a submission (columns 1–3); (ii) the absolute prediction error when writers predict the number of views generated by their articles (columns 4–6); and (iii) the relative prediction error, percentage change (columns 7-9). For each outcome, the first column reports estimates from Equation (1) using our main experimental excluding writers whose pre-period publication count exceeded the 99th percentile. Driscoll-Kraay standard errors are parenthesized. The second column reports estimates from Equation (1) using the main sample of writers with standard errors clustered at the writer level. The third column reports estimates from Equation (3), also using the main sample with writer-level clustered standard errors. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

Table C7: Robustness: Topical Composition and Editor Rating

	Politics	Politics	Politics	National	National	National	Editor Rating	Editor Rating	Editor Rating
Pay per click	0.33***	0.33***	0.33***	0.14***	0.19***	0.18***	-0.093	-0.040	0.12
	(0.028)	(0.051)	(0.045)	(0.022)	(0.070)	(0.064)	(0.097)	(0.18)	(0.15)
Choice	0.0099	0.033	0.0097	0.046**	0.084	0.033	-0.069	-0.064	0.083
	(0.025)	(0.046)	(0.043)	(0.021)	(0.062)	(0.048)	(0.090)	(0.20)	(0.15)
Observations	7809	8474	3982	7809	8474	3982	3388	3661	3238
Mean	0.35	0.35	0.35	0.38	0.38	0.38	2.69	2.69	2.69
Specification	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA
S.E.	DK	Writer	Writer	DK	Writer	Writer	DK	Writer	Writer
Sample	p99	Main	Main	p99	Main	Main	p99	Main	Main

Note: This table presents robustness analyses for three dependent variables: (i) the share of articles on political issues (columns 1–3); (ii) the share of articles covering national, as opposed to local, stories (columns 4–6); and (iii) editor's quality rating of the articles on a scale from 1 to 5 (columns 7-9). For each outcome, the first column reports estimates from Equation (1) using our main experimental excluding writers whose pre-period publication count exceeded the 99th percentile. Driscoll-Kraay standard errors are parenthesized. The second column reports estimates from Equation (1) using the main sample of writers with standard errors clustered at the writer level. The third column reports estimates from Equation (3), also using the main sample with writer-level clustered standard errors. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

Table C8: Robustness: Toxicity and Sentiment of Articles

	Negative	Negative	Negative	Positive	Positive	Positive	Tone Body	Tone Body	Tone Body	Tox Body	Tox Body	Tox Body	Tox(1) Body	Tox(1) Body	Tox(1) Body
Pay per click	-0.017	-0.013	-0.0027	-0.055**	-0.052*	-0.063***	-0.44**	-0.45	-0.68***	0.0039***	0.0035	0.000023	0.011***	0.010	0.0011
	(0.035)	(0.042)	(0.044)	(0.023)	(0.027)	(0.018)	(0.17)	(0.29)	(0.23)	(0.0013)	(0.0032)	(0.0015)	(0.0041)	(0.0083)	(0.0036)
Choice	-0.051	-0.035	-0.054	0.043*	0.041	0.0092	0.16	0.14	0.088	0.0046***	0.0040	0.00066	0.013***	0.011	0.0022
	(0.033)	(0.048)	(0.045)	(0.025)	(0.030)	(0.022)	(0.15)	(0.30)	(0.27)	(0.0012)	(0.0033)	(0.0018)	(0.0042)	(0.0090)	(0.0050)
Observations	7809	8474	3982	7809	8474	3982	7809	8474	3982	7809	8474	3982	7809	8474	3982
Mean	0.66	0.66	0.66	0.30	0.30	0.30	-0.77	-0.77	-0.77	0.0042	0.0042	0.0042	0.0069	0.0069	0.0069
Specification	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA	TWFE	TWFE	ANCOVA
S.E.	DK	Writer	Writer	DK	Writer	Writer	DK	Writer	Writer	DK	Writer	Writer	DK	Writer	Writer
Sample	p99	Main	Main	p99	Main	Main	p99	Main	Main	p99	Main	Main	p99	Main	Main

Note: This table presents robustness analyses for five dependent variables: (i) the number of negative words in the title (columns 1–3); (ii) the number of positive words in the title (columns 4–6); (iii) the tone of the body of the article, defined as the number of positive words less the number of negative words scaled by the overall length of the article (columns 7–9); (iv) the toxicity score of the body of the article measured by Unitary's Detoxify library (columns 10–12); and (v) the proportion of articles where the toxicity score of the body exceeds 0.1 (columns 13-15). For each outcome, the first column reports estimates from Equation (1) using our main experimental excluding writers whose pre-period publication count exceeded the 99th percentile. Driscoll-Kraay standard errors are parenthesized. The second column reports estimates from Equation (1) using the main sample of writers with standard errors clustered at the writer level. The third column reports estimates from Equation (3), also using the main sample with writer-level clustered standard errors. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.