Review Report for "Monetizing Serialized Content: How 'Wait for Free' Impacts Paid and Free Consumption"

Summary

The paper investigates the impact of a reduction in restriction time to access serialized content under the "wait for free" (WFF) scheme on the extent of episode consumption and consumer payments on a serial fiction novel platform. Using difference-in-differences estimations with propensity score matching, the authors demonstrate that the reduction in the wait time significantly increases consumer content consumption and payment.

Overall, the paper is very well-written, and the empirical analyses are comprehensively conducted. I appreciate the authors' rigor in the causal inference analyses. The WFF context is a unique and interesting marketing practice that broadly relates to freemium product versioning strategies, and the paper could potentially make important contributions. I have a few major comments.

Contribution

The major contribution of the paper is in a few directions: 1. assessing the impact of wait time reduction for WFF on content consumption and payment. 2. validating the mechanism (decreased complementarity between episodes over time) for increased payment despite the cannibalization effect under a reduced wait time threshold.

1. Overall, the academic contribution of the current manuscript is limited. The authors have done an excellent job conducting this event study, yet the generalizability and broader academic and managerial contributions are rather limited. On one side, whether a reduced wait time would be associated with higher consumer payment or platform revenue is an empirical question that highly depends on the context and the existing levels of wait time. I would expect a non-monotonic relationship between wait time and consumer payment. In the extreme case, if the wait time is reduced to be sufficiently low, such as 1 minute, then we would expect almost none of the consumers would pay to access the content. The authors indicate they do not intend to explore the "optimal" level of wait time design under the context and presume the effect would be monotonic (they assumed the policy impact would be homogeneous in the main analyses even though there are different degrees of wait time reduction). This approach limits the potential academic and practical contributions.

   Second, existing papers have documented the spillover effects between episodes for serial content captured in models for binge consumption (Zhao et al., 2022 etc.). Thus, the paper's proposal of complementarity for serial content is not new to the literature. Indeed, previous papers have more directly captured these effects with more granular data and detailed modeling. In addition, the paper by Choi et al. (2022) highlighted the positive effect of the introduction of WFF on paid episode views from an expanded user base. The current paper's contribution largely overlaps with Choi et al (2022). Third, the

authors provide evidence for the complementarity based on the aggregate level DID analysis, linking reduced wait time with increased overall payment. Yet, there is no direct test of the existence and magnitude of such complementary effect.

To elevate the contribution of the research, I recommend the authors consider the following potentials:

1. Extend the analyses to assess the non-linear or non-monotonic impacts of wait time reductions and draw more direct managerial implications.
2. Explicitly model the complementarity and other factors in consumers' reading and payment behavior. This might require a more structural approach on granular consumer-level data. The authors have a good starting point in the paper as they illustrate the impacts with the conceptual utility framework.

Context and Data Evidence

Overall, there is a lack of data evidence (before the formal DID tests) to support the potential mechanism. For example, while the authors have conceptually decomposed the overall impact into the extensive margin and intensive margins, some high-level statistics to show the changes would be helpful but are missing. In addition, for the proposed main mechanism, data evidence to support the complementarity between episodes would be critical. I do not see clearly how the patterns in Table 2 “lend empirical support to the notion of directed complementarity that diminishes over time.”

It would be very helpful to include in a visualization chart (similar to Figure 7) the percentage of paid consumptions for each episode sequence for both before and after the policy change.

Mechanism

The mechanism the authors propose is the existence of complementarity between episodes and the decreasing magnitude of this factor over time, such that once waited for too long, the consumers would have less value in reading the next episode. While I find the conceptual framework well-developed and the numerical illustration useful, I find the connection between the framework and the empirical DID analyses weak because DID only provides directionally consistent patterns with the framework but is not a direct test of the framework. In particular, the level of analyses are different, and the main model does not provide quantification to the parameters in the conceptual framework.

Conceptually, I am unclear how this “complementarity value” differs from a cost interpretation. A user can either pay a price to have immediate access or wait until it becomes free; both are costs. I believe the authors would also obtain qualitatively similar results under a cost specification where there is a fixed disutility to access the content, and a reduction in that component would help retain more readers and hopefully recoup the value from future payments.

Modeling Approach

Overall, I find the authors have done a comprehensive job estimating the causal effects using the DID approach. They have also performed a series of robustness checks. However, there are still some potential issues.

1. The majority of the policy change happens for series with a reduction of wait time from 3, 12, 34, and 48 hours to 1 hour. Given the significant differences in the wait times before the policy change, it is natural to expect substantial heterogeneity in the impact. While the authors have used the relative magnitude of wait time reduction as a robustness check specification, I believe it would be very helpful to create separate matched panels to assess the impact of changes at these different levels in the main analysis, instead of estimating a uniform average treatment effect. Indeed, Table 3 shows that the mean wait time between the treated and control, even after matching, is still significantly different. The authors should put more effort into creating a balanced panel.

2. Given that the treatment implicitly reduces the wait time for accessing content, I find the fixed time window for both the control and treatment groups a bit problematic. Specifically, consider a reader who would read all the episodes for free regardless. Under the reduced wait time scheme, she would naturally finish more chapters within the same 15-day window. In other words, the increase in consumption and payment for 15 days after the treatment would be partially attributed to free content being accessible more immediately. This raises questions about whether the authors have accurately captured the impact. The robustness checks using a 60-day window in the appendix showed a reduced magnitude of the effects, which is evidence of this potential issue.

3. Potential spillover effects. The authors argue that given the large number of series hosted on the platform, potential spillover effects from a small set of experimental series would be limited. They also try to mitigate the concerns for possible spillover effects into non-treated series by using only a subset of the sample with limited observed overlaps. I would suggest also looking at user-level data for more direct evidence. For example, do the readers of the same series before and after the treatment read, on average, the same number of series, consume the same number of episodes for the non-treated series, and pay the same amount for the non-treated series? This is indeed important, as the balance between exploration and exploitation would be managerially important for the platform.

4. Decomposition of the aggregate effect. One potential contribution of the paper is the decomposition of the treatment effect into multiple factors: the intensive margin for existing readers and the extensive margin for customer acquisition. The authors estimate three different DID models, and then decompose the aggregate impact based on the differences in the marginal effect at the mean of each model. I understand it is a rough decomposition, but I am not sure if this is the most rigorous way for the decomposition, given the potential heterogeneity in treatment effects. I suggest a more formal derivation of the equations.