Proceedings of the Nineteenth International AAAI Conference on Web and Social Media (ICWSM 2025)

Bootstrapping Social Networks: Lessons from Bluesky Starter Packs

Leonhard Balduf[1*], Saidu Sokoto[2*], Andrea Baronchelli[2], Ignacio Castro[3], Michał Król[2], Gareth Tyson[3,4], George Pavlou[5], Björn Scheuermann[1], Onur Ascigil[6]

[1]TU Darmstadt
[2]City St George's, University of London
[3]Queen Mary University of London
[4]Hong Kong University of Science and Technology (GZ)
[5]University College London
[6]Lancaster University
leonhard.balduf@tu-darmstadt.de, saidu.sokoto@city.ac.uk, andrea.baronchelli.1@city.ac.uk, i.castro@qmul.ac.uk, michal.krol@city.ac.uk, gtyson@ust.hk, g.pavlou@ucl.ac.uk, scheuermann@kom.tu-darmstadt.de, o.ascigil@lancaster.ac.uk

Abstract

Microblogging is a crucial mode of online communication. However, launching a new microblogging platform remains challenging, largely due to network effects. This has resulted in entrenched (and undesirable) dominance by established players, such as X/Twitter. To overcome these network effects, Bluesky, an emerging microblogging platform, introduced *starter packs* — curated lists of accounts that users can follow with a single click. We ask if starter packs have the potential to tackle the critical problem of social bootstrapping in new online social networks. We assess whether starter packs have indeed been helpful in supporting Bluesky growth. Our dataset includes 25.05×10^6 users and 335.42×10^3 starter packs with 1.73×10^6 members, covering the entire lifecycle of Bluesky. We study the usage of these starter packs, their ability to drive network and activity growth, and their potential downsides. We also quantify the benefits of starter packs for members and creators on user visibility and activity while identifying potential challenges. By evaluating starter packs' effectiveness and limitations, we contribute to the broader discourse on platform growth strategies and competitive innovation in the social media landscape.

1 Introduction

Microblogging platforms have become integral to modern communication. Over 4.5 billion people are active on social media, with microblogging platforms playing a critical role in this ecosystem (We Are Social 2024). Studies show that over 70% of users turn to social platforms to stay informed about breaking news and global events (Pew Research Center 2023). These platforms amplify diverse voices, enabling grassroots movements, social activism, and citizen journalism to thrive on a scale previously unattainable.

However, launching a new social platform is extremely difficult. Users leaving an established platform are forced to leave familiar content, interfaces, and, most importantly, their social network. Of course, users could try to convince their social network to follow their migration. However, such attempts are rarely fully successful due to the network effect (He et al. 2023). Worryingly, this exacerbates the dominance of established platforms, prevents innovation, and can constrain users' ability to migrate even if dissatisfaction with the platform is widespread (Mekacher, Falkenberg, and Baronchelli 2024).

Innovation attempts in microblogging are on the rise though. From blockchain-based microblogging such as Memo.cash (Zuo et al. 2024, 2023)) to decentralisation attempts like Mastodon (He et al. 2023). Even large social networks such as Facebook have tried to innovate in this space (Zhang et al. 2024). Bluesky is part of this wave of innovation in microblogging.

In 2022, Bluesky launched a new microblogging service. Bluesky resembles Twitter/X: users can follow each other and share short posts, including images and videos. A key innovation of Bluesky is decomposing and opening core platform functions into sub-components that can be provided by stakeholders other than Bluesky (Kleppmann et al. 2024). This approach has driven rapid adoption, increasing the user base by $\approx 10\times$ in a short time (Balduf et al. 2024). Over the past year, Bluesky's user base grew from 2.59×10^6 in January 2024 to 25.05×10^6 by the end of year. Bluesky is now the largest new social platform, with over 30 million users.

However, Bluesky still faces challenges persuading users to migrate from incumbent competitors (*e.g.* Twitter/X). To overcome this, Bluesky introduced *starter packs* in June 2024. Starter packs are curated lists of accounts users can follow in one click, enabling the rapid creation of a denser social network. Anyone can create starter packs, which aim to (re)create new or existing communities.

The rapid growth of Bluesky gives credence to the starter packs' ability to mitigate the challenge of network bootstrapping, enabling new users to quickly form social connections. We believe that understanding the efficacy of starter packs is critical, both to understand the success of Bluesky and to identify the potential for other platforms facing similar challenges. If proven effective, starter packs could help disrupt future social networks by becoming a standard tool for onboarding and fostering early engagement.

[*]These authors contributed equally.

Research Questions. To address this, we answer the following three research questions:

- **RQ1:** How are starter packs used in Bluesky, and to what extent are they employed?
- **RQ2:** How effective are starter packs in driving network and activity growth? Do members of starter packs experience tangible social benefits?
- **RQ3:** How do users perceive the starter packs, and do they speak positively of them? Are there any downsides to introducing starter packs into the network?

First (**RQ1**), we collect *all* starter packs, their changes, creators, members, and descriptions. We perform a temporal analysis, co-locate activity spikes with real-world events, and explore which communities use starter packs.

We find that starter packs have experienced considerable uptake, with 335,416 created over the 6 months since they were introduced. They are impactful, being responsible for up to 43% of daily follow operations at their peak. Yet, they include a relatively small number of users, with only 6.25 % users being members of at least one starter pack. At the same time, the starter packs played an important role during large user influx spikes caused by various political events. We find that they are popular among artists, journalists, and academic communities.

Second (**RQ2**), we perform a temporal analysis of the *follow* operations to estimate the number of new social graph edges created through starter packs. We then quantify the benefits of being included or creating a starter pack using Propensity Score Matching (PSM) and Difference-in-Differences (DiD). We focus on increased visibility in the network (*e.g.* a higher number of followers) and the activity of users (*e.g.* a higher number of posts). We then use graph analysis to asses the macro-level impact on the overall social graph.

Our analysis reveals that becoming a member or a creator of a starter pack yields substantial benefits. Starter pack members receive up to 85% more new followers and 70% more likes than similar users not included in starter packs. Starter pack members also generate 60% more posts and issue 71% more likes. This effect is even stronger for the starter pack creators, reaching 117% new followers and 100% created posts increase. On a macro-perspective, we notice a limited effect on the overall social graph. Starter packs strengthen links between already existing communities rather than creating new ones. Furthermore, we find evidence that starter packs contribute to the *rich get richer* effect, increasing popularity inequalities in the system.

Third (**RQ3**), we extract all Bluesky posts discussing starter packs, perform sentiment analysis, and categorize those posts into the most commonly discussed topics.

We show that starter packs were mostly perceived positively by the community, with more than 10× more positive than negative posts. At the same time, we notice multiple problems flagged by the users. For instance, starter packs enable their creators to add any new member without asking for their permission. This enables using popular and well-established accounts to promote malicious starter packs or use the feature as a tool for harassment. Furthermore, we discover traces of a market where users pay to be included in a given starter pack. We make our code for identifying starter packs available to support future research.

2 Background

Bluesky is a novel social network built on the Authenticated Transfer Protocol (ATProto). The system is decomposed into components that can be operated by the community. We now introduce the relevant components for this study and refer interested readers to previous work (Kleppmann et al. 2024; Balduf et al. 2024) for a deeper analysis of the architecture and its critical components.

User Data is stored in user-controlled repositories, each stored on a Personal Data Server (PDS). Repositories provide signed and ordered lists of a user's public records. Due to its open architecture, the repositories must be public and contain all the information required to operate the other components of the system. Repositories store, for example, a user's posts, likes and follows, as well as other information such as the list of blocked users. Repositories are updated via signed commits created by the user. These commits include the creation of new records, as well as deletions or updates of existing records. Commits are published via a publish/subscribe endpoint by the hosting PDS.

Firehose is an aggregated publish/subscribe endpoint, which is subscribed to all federated PDSes and re-publishes their commits as a single feed.

Feeds are Bluesky's algorithmically-driven content timeline creation mechanisms. Each custom feed is generated by a *feed generator*, which curates posts to be included in the feed. Feed generators can be operated by Bluesky or created by users. When a user subscribes to a feed, the curated posts become available in the user's timeline, in the order stipulated by the feed. There is no limit on the number of feeds that a user can subscribe to.

Starter Packs are an onboarding feature introduced to Bluesky on 2024-06-26. A starter pack can be created by any user without requiring any special privileges. Each starter pack consists of a list of up to 150 users and 3 feeds. We refer to the users and feeds included in a starter pack as *members*. Starter packs also include a name, a description, and the name of their creator. The contents of the starter pack are mutable and can be changed by its creator. Bluesky users can either *(i)* use the *follow-all* option that automatically follows all the members and subscribes to all the feeds in the starter pack; or *(ii)* follow starter pack members individually. Bluesky enables users without a Bluesky account to sign up via a starter pack. This triggers the regular sign-up process but also bootstraps the user's initial social network with a *follow-all* operation on the selected starter pack.

3 Data and Methodology

3.1 Bluesky Data Collection

We collect a complete snapshot of Bluesky by downloading data from all known PDSes on 2025-01-01. We obtain the list of all known PDSes from users' Decentralized Identifier

(DID) Documents. We then complete the data with Firehose updates, starting from 2024-06-10. Combined, this enables us to recreate Bluesky's complete state at any given time between 2024-06-10 and 2024-12-31. We gather public data only (*e.g.* no direct messages).

Our dataset contains the activity of all 25.05×10^6 Bluesky users at the end of 2024. It includes 1.55×10^9 follow relations in the social graph, 810.17×10^6 posts, and 3.87×10^9 likes. This includes all the 335,416 starter packs created before 2025-01-01, their metadata, and modifications (*e.g.* updates, deletions). Notably, around 20 % of the created starter packs were deleted before 2025-01-01. By the end of 2024, there were 265,595 starter packs.

3.2 Identifying Starter Pack Usage

Bluesky does not explicitly record when someone uses a starter pack. However, through manual analysis, we find that clicking on *follow-all* from a starter pack triggers a multi-follow operation, whereby multiple starter pack members are followed in rapid succession.[1] This manifests as a sequence of repository commits containing up to 50 follows each. We leverage this to identify candidate starter pack users by selecting all whose repositories contain such multi-follow operations. This represents a lower bound on the starter pack usage, as users can also manually follow specific starter pack members (instead of the whole starter pack) without triggering the multi-follow operations.

Mapping Multi-Follows to Starter Packs. We then attempt to assign the multi-follow operations to specific starter packs. First, for each day, we reconstruct members in each starter pack from the Firehose data. This is necessary because members can be added or deleted from starter packs over time. As such, we consider the state of the starter pack at the time the multi-follow operation took place.

We then match each multi-follow operation O_i to the starter packs S_j containing the most similar users. Intuitively, if a user follows a large number of starter pack members in a single commit, we can be confident they did so via using that starter pack. Thus, for every pair between multi-follow operation following members M_{O_i} and a starter pack containing members M_{S_j}, we calculate a weighted set overlap score $s_{ij} \in [0, 1]$:

$$s_{ij} = f_{ij} \frac{|M_{O_i} \cap M_{S_j}|}{\min(|M_{O_i}|, |M_{S_j}|)}$$

where f_{ij} denotes a factor to penalize large size differences:

$$f_{ij} = 1 - \frac{\mathrm{abs}\left(|M_{O_i} \cap M_{S_j}| - \max(|M_{O_i}|, |M_{S_j}|)\right)}{\max(|M_{O_i}|, |M_{S_j}|)}$$

We then assign each multi-follow operation O_i to the starter pack S_j with the highest s_{ij}. Note that $s_{ij} = 1$ if the multi-followed members match perfectly the starter pack members (*i.e.* $M_{O_i} = M_{S_j}$) and the user clicking on *follow-all* did not previously follow any starter pack members.

[1]The multi-follow operation includes only starter pack members who were not previously followed by the user clicking on *follow all*.

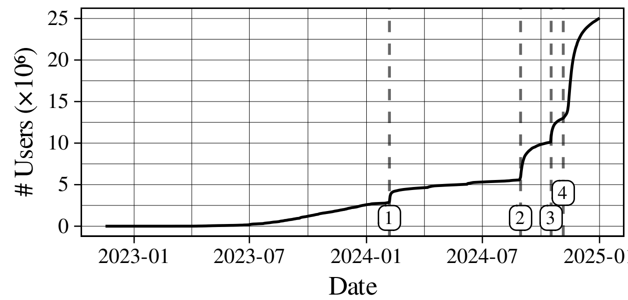


Figure 1: Number of registered Bluesky users.

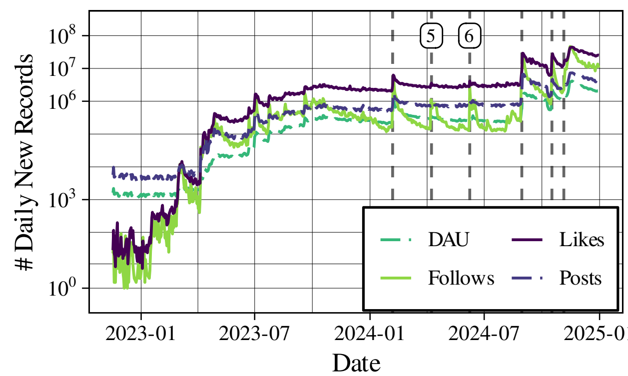


Figure 2: Number of Daily Active Users (DAU), likes, follows, and posts.

Out of the 5.69×10^6 multi-follow operations, we find matches (*i.e.* map the multi-follow to a specific starter pack) for 99.88 %, with a median best set overlap score of 0.75.

3.3 A Primer on Bluesky Growth

For context, we briefly present the current growth and trends of Bluesky. As a new social platform, this is intended to lay a foundation for the rest of the paper.

User Growth Figure 1 presents the number of registered Bluesky users over time with event annotations. We observe substantial growth, especially since mid-2024. A number of events seem to have fueled Bluesky adoption: 1 Bluesky opening registrations to the public (*i.e.* requiring no invites); 2 Twitter/X banning in Brazil; 3 Twitter/X's controversial change making content visible to blocked users, and 4 The 2024 US elections.

Activity Growth We now examine whether this user growth results in an increasing level of activity. Figure 2 plots the number of new follows, posts, and likes per day. We confirm notable growth aligned with the influx of users. 5 and 6 mark migrations and surges in activity from users from Brazil[2] and Indonesia,[3] respectively, both due to announced changes to Twitter/X. [4] That said, since early 2024,

[2]https://bsky.app/profile/bsky.app/post/3kpllkdgtqu2v

[3]https://bsky.app/profile/bsky.app/post/3kufde3xvol2j

[4]cf. https://www.bbc.com/portuguese/articles/cm2nkdkypk7o and https://www.reddit.com/r/BlueskySocial/comments/1dfiry8/

the number of posts has not grown linearly with the number of users. Indeed, in December of 2024, there were an average of 2.42×10^6 daily active users, representing just 10.09 % of daily registered users. This suggests a large number of more experimental users, who are yet to actively engage in the platform. The most common user activity is liking with 3.87×10^9 likes by the end of 2024, compared to 1.55×10^9 follow operations and 810.17×10^6 posts.

4 Measuring Starter Pack Use (RQ1)

Before assessing the impact of starter packs, we inspect the usage trends since Bluesky introduced them in June 2024.

Starter Pack Growth Figure 3 presents a time series of the number of starter packs released over time. We observe considerable uptake, with 265,595 starter packs as of 2025-01-01. Recall, a total of 335,416 starter packs have been created throughout the entire lifespan of Bluesky— this smaller number reveals that 69,821 have since been deleted. This uptake grew particularly after the first week of November 2024 (following the US elections), with a growth of 74% from then to January 2025. The number of creators is noticeably smaller than the number of starter packs, confirming that a subset of users create multiple ones. Indeed, 13.8% of creators have two or more starter packs, and 3.4% have more than three. We also find that starter packs are actively maintained. Recall that the creators can modify their starter packs — 99.8% of starter packs have at least one update, and 30.5% have over 50. This suggests considerable investment by their creators and an active and evolving community. By the end of 2024, 238.51×10^3 (0.95 % of all) users had created at least one starter pack, 1.56×10^6 (6.25 %) users were members of at least one starter pack, and 1.1×10^6 (4.37 %) users had employed the *follow-all* operation on a starter pack. This indicates that the new feature was used by a relatively small portion of users. We note again that this is a lower bound on starter pack-mediated follows, as we can only confidently detect bulk follow operations (rather than users who only follow one or two people from a given starter pack). We observe that only around 60,000 users have taken advantage of the "sign-up via starter pack" feature to join Bluesky, a figure significantly smaller than the number of starter pack users overall.

Starter Pack Followers We now estimate the number of follows created by starter packs using our matching of multi-follow operations to starter packs, plotted as a timeseries in Figure 4. The starter-pack-induced follow operations closely match the system-wide trend. Their impact on the social graph increases over time, surpassing 40 % of all the follow operations in December 2024. Starter packs created a total of 308.57×10^6 unique edges in the follower graph. This represents a remarkable 19.95 % of all follow edges of the network, indicating a large impact of starter packs on the overall social graph. Follows resulting from starter packs are also long-lasting. We observe that by the end of 2024, 93.82 % of them are still present.

daily_follows_have_surged_it_might_not_be_in_your/

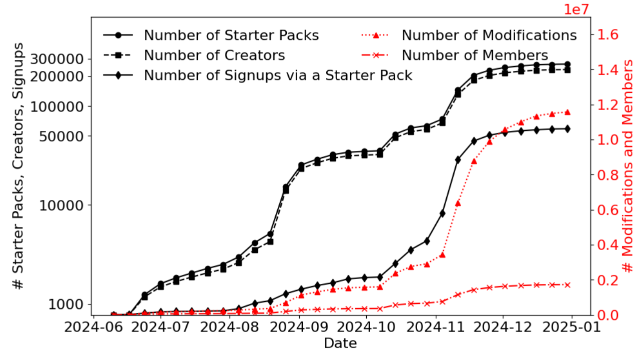


Figure 3: Evolution of the number of starter packs, their members, modifications, and creators over time.

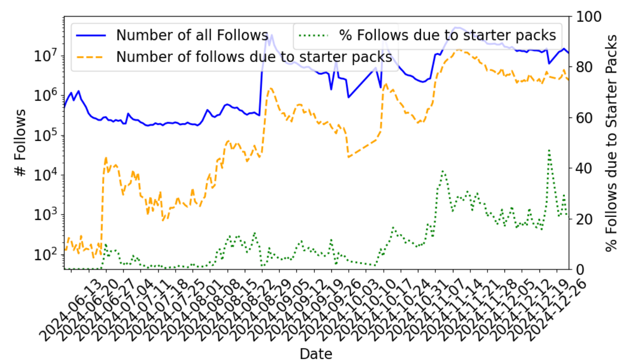


Figure 4: Daily count and percentages of all follow operations, and follow operations due to starter packs.

Figure 5 plots the number of followers created per starter pack. As expected, the distribution is highly skewed, with the top 20 % of starter packs creating 97.17 % of all starter-pack-induced follow edges. The most popular starter pack (by the number of follow edges created), with 7.01 M follow edges created, lists "pro-democracy accounts". The following top 10 show a similar focus, with politics, journalism, and media as the main themes.

Starter Pack Languages We observe substantial use of starter packs across languages. Figure 6 shows the distribution of the top 10 languages in the starter pack descriptions, out of the 47 that we detect using `langdetect`. We find that the prominence of the different language starter packs reflects underlying trends in new user arrivals (depicted previously in Figure 1). For instance, the number of starter packs with a description in Portuguese spikes during the banning of Twitter/X in Brazil (2), which corresponds to the period around the Twitter/X ban in Brazil (2024-08-30). This dominance soon gives way to English, with two spikes that coincide with the (3) Twitter/X change in the visibility of blocked users (2024-10-17), and the (4) the US elections on 2024-11-05.

Starter Pack Communities To briefly explore the topics that these starter packs cover, we inspect the starter pack de-

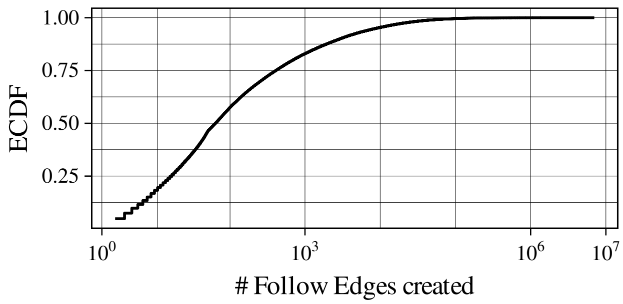


Figure 5: Distribution of follow edges created per starter pack.

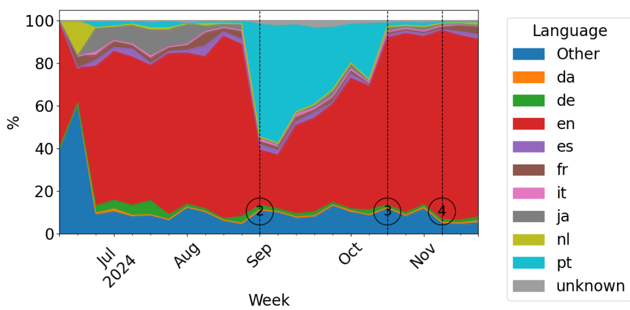


Figure 6: Language distribution in starter pack descriptions over time.

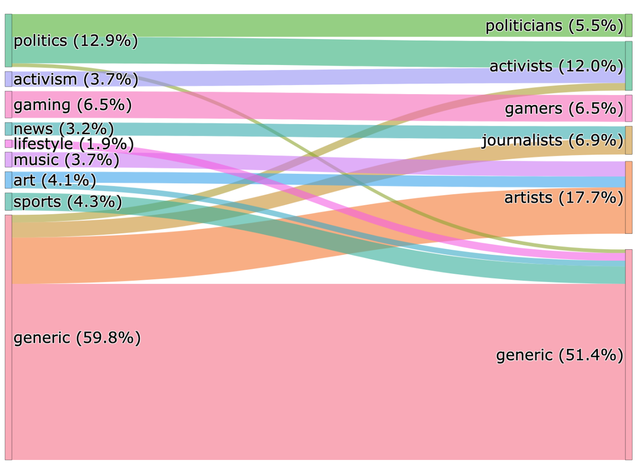


Figure 7: Sankey diagram showing the top 10 themes of starter packs (left) and the occupation or focus of their members (right).

scriptions. We employ the Large Language Model (LLM) Mistral (Jiang et al. 2023) (see prompts and details in the Appendix A) for this classification. We prompt the LLM to provide both the focus of the starter pack as well as the list of potential members, often listed in the pack description. For instance, a starter pack focused on climate might be explicit about potential members being journalists, politicians, activists, or scientists.

We identify a large number of "generic" starter packs. These are often the creator's personal selection (*e.g.* "My favorite Bluesky accounts") without a specific focus indicated in the description. Similar to related work (Balduf et al. 2024), we notice well-developed art and gaming communities. We also observe starter packs focused on emerging communities such as politicians, journalists, or activists, as shown in Figure 7. This suggests that certain professional communities have been migrating to Bluesky and are using starter packs to bootstrap their social graphs.

5 Measuring Starter Pack Impact (RQ2)

Starter packs have seemingly helped Bluesky to quickly bootstrap a large and active social network. We now assess this by quantifying the impact of being included in or creating a starter pack. This is non-trivial. Many factors may help a user gain followers, agnostic of their inclusion in a starter pack. We address this challenge with Propensity Score Matching (PSM). PSM has proven effective in related tasks where controlling for several confounding variables is necessary (Bhattacharjee and Mohanty 2022; Valenzuela, Arriagada, and Scherman 2014; Dos Reis and Culotta 2015). We then assess the macro-level impact of the starter packs using social graph analysis.

5.1 Methodology

Propensity Score Matching (PSM) is a statistical technique that estimates the effect of a treatment or policy by accounting for the factors that predict treatment (*i.e.* focusing on causation rather than correlation).

To employ PSM in our data, we first divide Bluesky accounts into treated and control groups. We define two independent treatment indicators specifying whether: *(i)* a user has been a member of at least one starter pack; and *(ii)* a user has created at least one starter pack. For each experiment, we calculate success indicators measuring the increase in the *(i)* number of followers; *(ii)* likes received; *(iii)* issued likes; and *(iv)* average daily posts.

Indicators *(i)* and *(ii)* (*i.e.* followers and likes received) measure whether starter packs increase the visibility (i.e., popularity) in the network, while indicators *(iii)* and *(iv)* (*i.e.* likes issued and average daily posts) focus on the increase in the activity of the users involved. We measure each success indicator at time t specifying the time of the first inclusion or creation of a starter pack (depending on the treatment group) and after 7-day intervals (*i.e.* $t+7$, $t+14$, $t+21$, $t+28$). To ensure that success indicators are calculated reliably at all intervals, we exclude accounts that were included in or created a starter pack after 2025-12-03 (*i.e.*, 28 days before the end of our dataset), this discards 99,386 ($0.4\,\%$) accounts.

To measure the effectiveness of starter packs, we match treated accounts with the most similar accounts from the control (*i.e.* untreated) group and compare their success indicators. This requires selecting t for the control group as well. We thus set t at a random time after their creation between 2024-01-01 and 2024-12-03, then calculate the same success indicators at 7-day intervals (up to 28 days before and after treatment). The robustness of the PSM is determined by the level of similarity between treated and non-treated accounts (*i.e.* the higher, the better). To maximize the matching quality, we choose three different t for each account in the control group, effectively tripling its size.

PSM Confounding Factors PSM requires a set of confounding factors that may impact the dependent variables. To control for confounding factors and ensure a robust analysis, we select a comprehensive set of covariates that capture key characteristics of the accounts and their activity. All the covariates are calculated at t. **Number of followers** reflects the existing popularity of an account, which could independently influence subsequent follower growth (Kwak et al. 2010). **Account age (in days)** captures the amount of time an account has had to accumulate followers and activity, mitigating the effects of longer-established accounts naturally having larger followings (Mislove et al. 2007). **Number of posts** and **Number of received likes** (on posts) measure engagement levels and content attractiveness, which are likely to affect follower acquisition (Gilbert and Karahalios 2009). **Number of issued likes** represents the degree of interaction initiated by the account, capturing a behavioral aspect of network participation (Golder, Wilkinson, and Huberman 2007). **Followers-to-following ratio** serves as an indicator of influence and credibility, as accounts with a higher ratio may be perceived as more authoritative or desirable to follow (Cha et al. 2010). Last, **network size** is included to account for temporal variations in the overall network's growth, ensuring that follower trends are not conflated with the increasing pool of users on the platform (Ugander et al. 2011). Together, these covariates provide a nuanced representation of the factors that may influence follower growth, allowing us to isolate the causal effect of inclusion in starter packs.

Graph Analysis To compare the Bluesky social graph, we use the directed graph G of all follows at the end of 2024. Within G, we mark edges created via starter pack multi-follows S. For the analysis, we investigate properties of G with and without S, *i.e.* G and $G \setminus S$.

To obtain G, we take all follows on 2024-12-31. We do not include accounts without any incoming or outgoing follow edges. Furthermore, we remove self-loops and duplicate edges from G, which do not occur naturally, but can be added manually by tech-savvy users. To measure the relevance of the starter packs, we create a second graph, $G \setminus S$, where we remove from G any edges that were created via starter pack multi-follow operations. We use NetworKit to analyze and contrast the resulting graphs (Angriman et al. 2022). Note that, for the analysis of the follower graph at different points in time, we omit from G any edges created after the selected date.

5.2 Results

We first quantify the impact of the starter packs for both their creators and members using PSM.

Popularity Gains for Starter Pack Members. We investigate differences between accounts that are included in starter packs vs. those that are not. We observe notable differences across all metrics. We find that the popularity of a member grows after its inclusion in a starter pack. In the first week after its inclusion, the members receive on average 39% more follow operations (Figure 8a). This trend increases over time, reaching 57%, 71% and 85% after two, three, and four weeks, respectively. We observe a similar trend for the number of likes received on published posts (Figure 8b). The accounts included receive 23%, 42%, 51%, and 70% more likes in each of the four consecutive weeks. This confirms that inclusion in a starter pack has a substantial positive impact on the visibility and popularity of the accounts included, and that the increase in popularity is not ephemeral.

Activity Growth of Starter Pack Members We then inspect the impact on the activity of the users who are included in a starter pack. We conjecture that inclusion in a starter pack may create a positive feedback loop that encourages greater activity (*e.g.* posting). Indeed, we observe a significant increase in the activity of the included accounts. The number of likes issued by members increases by 25%, 45%, 59%, and 71% in the 4 consecutive weeks (Figure 9b). At the same time, the number of posts created increases by 20%, 36%, 50%, and 60% (Figure 9a). We hypothesize that the increase in popularity (and the related notifications) makes users more likely to visit the social network, and engage more widely. This confirms that starter packs have a positive impact on encouraging engagement from both those who subscribe to them and those who are included in them.

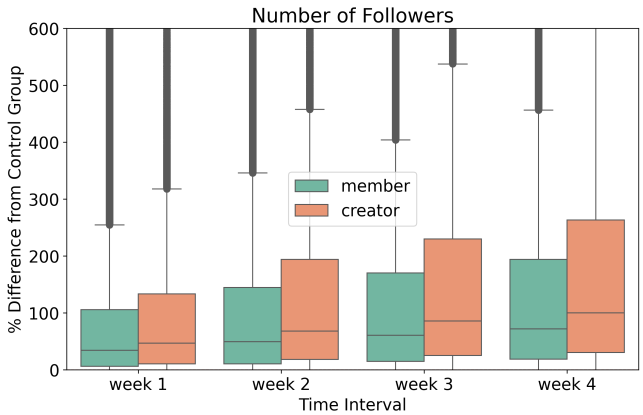
Benefits for Starter Pack Creators Next, we investigate the benefits that accrue to the *creator* of a starter pack. We conjecture that users who create starter packs may gain increased popularity and activity. As shown in Figure 8a and Figure 8b, creators show strong post-treatment increases in followers (51% to 117%) and received likes (46% to 115%). Similarly, their activity grows across all metrics, with increases in posts and issued likes reaching up to 100% (Figure 9a, Figure 9b). However, as we show below in the DiD analysis, the observed differences may reflect pre-existing behavioral patterns rather than the causal effect of creating a starter pack.

Notably, 99.91% of starter pack creators include themselves in the pack,[5] which may partly explain the observed gains. When comparing creators who include themselves to those who do not, we find that the latter see moderate growth by $t{+}28$—including a 22% increase in followers and 14% more posts; however, these effects are considerably smaller.

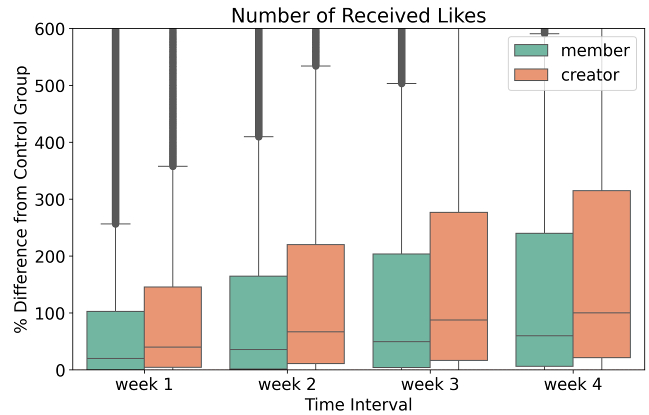
Difference-in-differences (DiD) analysis To evaluate the robustness of our conclusions to potential hidden biases and validate the assumptions underlying our PSM approach, we conduct additional robustness checks. Specifically, we first conduct a Difference-in-Differences (DiD) event study. Then, to validate the matching process in PSM, we report covariate balance analysis in Appendix B.

The primary goal of the DiD analysis is to assess the validity of the *parallel trends* assumption before treatment over a 4-week window, between the treated and control groups—an assumption required for causal inference in observational studies. The parallel trends assumption requires that the

[5] This is the default behavior, though it can be overridden.

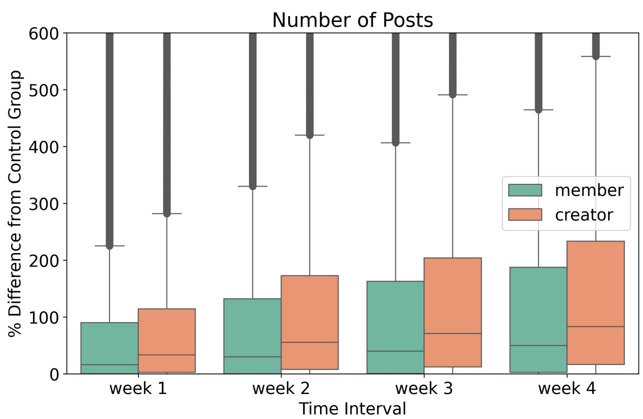


(a) Followers number increase.

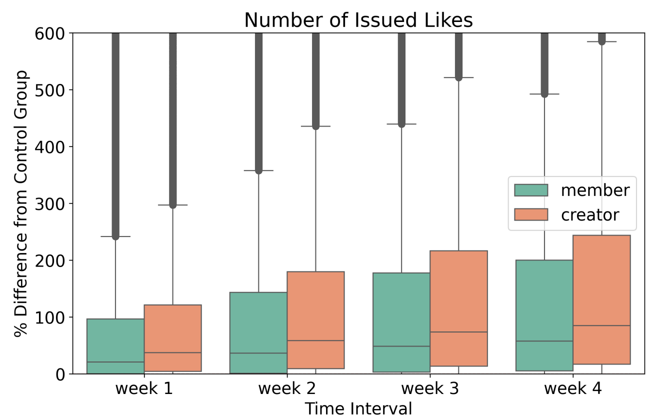


(b) Received likes increase.

Figure 8: Weekly visibility increase for starter pack members and creators w.r.t. accounts in the control group (*i.e.* neither creators nor members of a starter pack).



(a) Number of written posts increase.



(b) Number of issued likes increase.

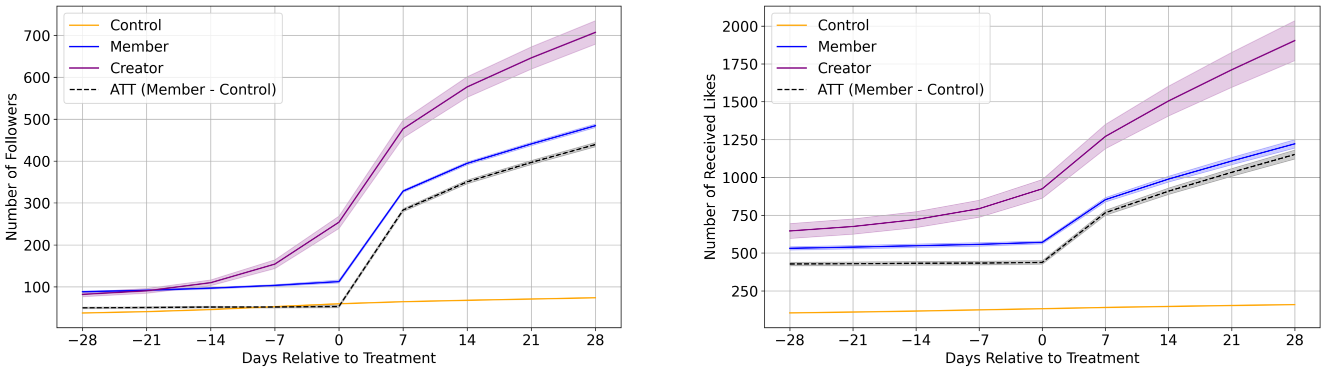
Figure 9: Weekly activity increase for starter pack members and creators w.r.t. accounts in the control group (*i.e.* neither creators nor members of a starter pack).

treated and control groups experience similar changes in the outcome before the treatment, *i.e.* prior to the treatment, their trajectories are parallel even if their levels differ.

In Figure 10 and Figure 11, we present the temporal evolution of each metric for the control group, as well as two subgroups within the treated population for four weeks before and after joining or creating a starter pack. The treatment group consists of *(i) members* of a starter pack (excluding the creator) and *(ii)* starter pack *creators*. We also measure the *Average Treatment Effect on the Treated (ATT)*; that is, the difference between the average outcome for the treated and the control group. We focus the ATT analysis on members only because, unlike creators, their pre-treatment trends closely align with those of the control group, as we discuss below. The shaded regions around each line represent 95% confidence intervals.

The results provide strong evidence that the pre-treatment parallel trends assumption holds for the members of starter packs. As shown in Figure 10a and Figure 10b, both the member and control groups follow identical trajectories in the pre-treatment period (*i.e.* $t-28$ through t). This is reflected in the ATT line, which remains constant before treatment, demonstrating the fixed difference between the members and the control group. Following the treatment, we observe a substantial increase in both followers and received likes for the member group, while the control group remains on its prior trajectory. This suggests a clear causal effect of the starter pack treatment on members. The post-treatment trends are aligned with our earlier results in Figure 8 comparing each member against the matching control account.

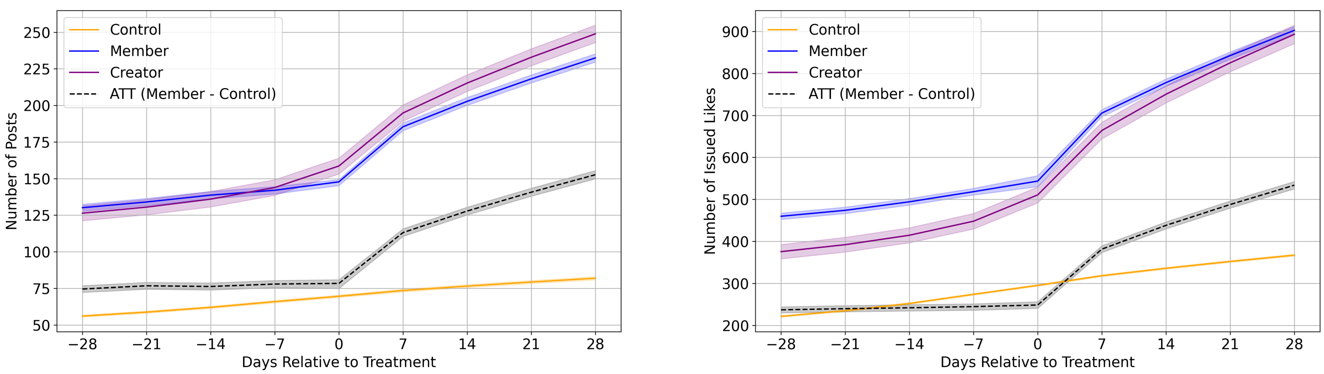
The causal effect for starter pack creators is less clear due to deviations from parallel trends. While creators and control users follow similar trajectories earlier in the pre-treatment period, their trends diverge, especially in the final week before treatment (*i.e.* $t-7$ to t). This appears driven by creators becoming more active and visible as they prepare



(a) Followers number increase.

(b) Received likes increase.

Figure 10: Difference-in-differences analysis of popularity metrics.



(a) Number of written posts increase.

(b) Number of issued likes increase.

Figure 11: Difference-in-differences analysis of activity metrics.

their starter packs—we observe that creators often announce plans to solicit interest from potential members. Such behavior generates discussion and visibility prior to the starter pack's release. As a result, post-treatment gains in activity and popularity may reflect pre-existing trends and latent factors rather than the causal impact of creating a starter pack.

Finally, we find that the parallel trends assumption also holds for activity-based metrics in the member group, as shown in Figure 11. The number of posts (Figure 11a) and issued likes (Figure 11b) follow similar pre-treatment trajectories for members and controls, with the ATT line again remaining flat before treatment. As with popularity metrics, the creator group again diverges from control users, especially in the week prior to treatment.

Graph Properties We next assess the starter packs' macro-level impact on the Bluesky social graph by comparing the follow relations graphs with (G) and without ($G \setminus S$) the starter pack edges. Surprisingly, the removal of the $\approx 300\,\mathrm{M}$ starter-pack-induced edges ($\approx 20\,\%$ of all the edges) has little impact on the social graph from a macro perspective. The number of strongly connected components increases from 8.25×10^6 to only 8.27×10^6 (*i.e.*, a $0.002\,\%$ increase).[6] The size of the largest strongly connected component decreases from 16.61×10^6 to 16.58×10^6 (*i.e.* $0.002\,\%$ decrease). The average in- and out-degrees unsurprisingly do change, from 62 to 50, a decrease of $24\,\%$. This suggests that starter packs provide tighter connections within existing communities rather than creating inter-community connections and promoting connections across the entire system. This raises the possibility of starter packs exacerbating potential echo chambers.

To better understand this phenomenon, we investigate the out-degree distribution (*i.e.* number of followed accounts) for G and $G \setminus S$ comparing starter pack members with the remaining users (Figure 12).[7] Interestingly, we find that removing starter pack induced edges almost exclusively affects starter pack members, while the other users remain mostly unaffected. This indicates that starter pack members

[6] Strongly connected components are maximal subsets where a path exists between any two members. Adding an edge can only reduce the number of components by merging smaller ones into a larger one.

[7] The in-degree changes affect only starter pack members, by definition, and are thus omitted here.

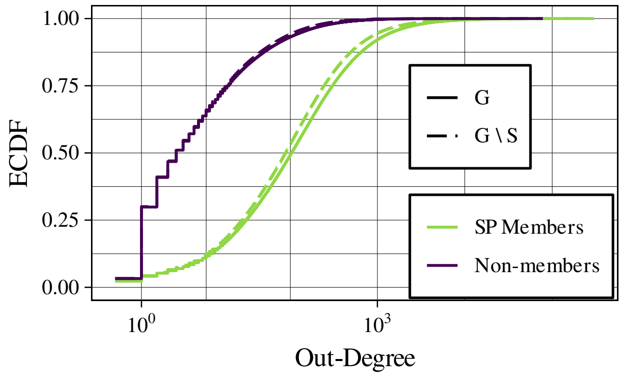


Figure 12: Out-degree distribution for the G complete graph, including all edges and the $G \setminus S$ graph without starter-pack induced edges, members and non-members of starter packs.

are also the ones using starter packs, while non-members rarely use the starter pack *follow-all* operations.

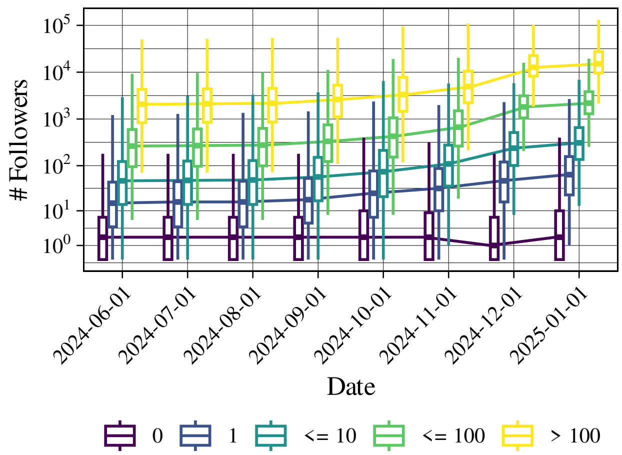


Figure 13: Evolution of the in-degree distribution of starter pack members depending on the number of starter packs they are included in at the end of 2024. Outliers omitted for visibility.

The above points towards the fact that starter pack inclusion may be limited to already popular accounts in the network. To study this, Figure 13 shows the number of followers over time for users who are included in different numbers of starter packs. For each user, we calculate the number of starter packs they are included in on 2024-12-31. Indeed, we observe that accounts included in many starter packs tend to be more popular at *any* point in time, even before the introduction of starter packs—note that the earliest data point in the figure predates the introduction of starter packs on 2024-06-26. The graph shows the Matthew effect, where popular accounts are more likely to be included in starter packs, increasing their popularity even further. We observe the gap between more and less popular accounts widening over time. We note that we can only detect the follow-all operations, and that users might instead cherry-pick the specific accounts within a starter pack that they wish to follow. These starter-pack-enabled follow operations might produce different results, although we posit that selecting specific individuals is more likely to increase the Matthew effect than bulk following all accounts in the starter pack.

6 Perceptions & Downsides of Starter Packs (RQ3)

Finally, we briefly explore the users' perception of starter packs to understand whether they are appreciated by users.

6.1 Methodology

We extract a total of $363{,}999$ Bluesky posts ($0.06\,\%$ of all posts since June 2024) containing the term "starter pack" and analyze the sentiment of each post using the TextBlob library (Steven Loria 2024) commonly used in previous social platform studies (Abiola et al. 2023; Diyasa et al. 2021). TextBlob categorizes sentiment into three categories based on a polarity score: Positive (values > 0.1), Neutral (values between -0.1 and 0.1), or Negative (values < -0.1).

To better understand the content of the posts, we manually annotate 4,000 (1% of all the posts mentioning starter packs) and identify the 10 most common themes expressed in positive, neutral, and negative posts. We then classify the remaining posts into those 10 themes using the Mistral LLM (see Appendix A.2 for further methodological details). Note, we include an "other" category for all remaining posts that do not fall into one of the 10 themes.

6.2 Results

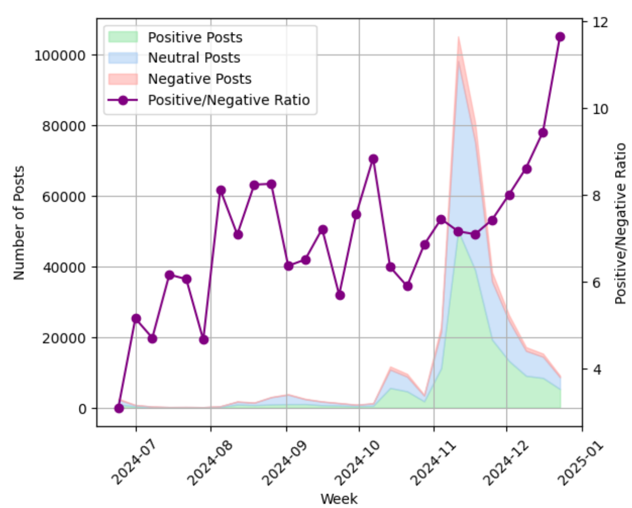


Figure 14: Distribution of the sentiment of posts discussing starter packs.

Sentiment Analysis We identify a total of $176{,}648$ ($49\,\%$) positive and $164{,}225$ ($45\,\%$) neutral posts, yet only $24{,}127$ ($6\,\%$) negative posts. Figure 14 shows the volume of posts mentioning starter packs over time, classified into their corresponding sentiment. Initially, starter packs are rarely discussed. Users start to discuss them more actively during the large influx of users to Bluesky (*i.e.* September–November, 2024). We note that the number of posts discussing starter packs amounts to just $\approx 0.06\,\%$ of all posts since June 2024.

For the users who do discuss them, starter packs are positively perceived, with $\approx 5\times$ more positive than negative comments in July (one month after the introduction of starter packs). The ratio increases over time, with ≈ 10 times more positive at the end of the year. As there were no major changes to how the starter packs work, our results suggest that the starter pack perception improves as the users start to use and learn more about them.

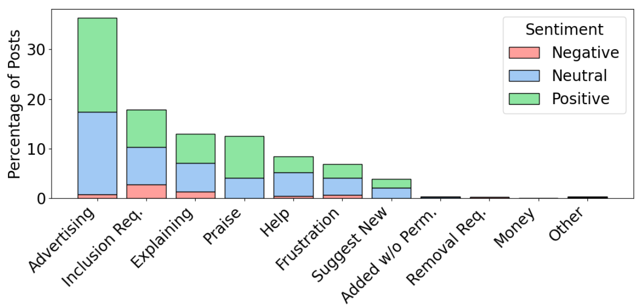


Figure 15: Distribution of posts mentioning starter packs by theme and sentiment.

Theme analysis To gain a deeper insight into the users' perceptions of starter packs, we also look into the themes discussed by the community. Figure 15 plots a histogram of the 10 manually identified themes. The most popular themes are advertising an existing starter pack or requesting to be included in an existing pack. These are followed by posts where users explain how the system works, share their experiences with starter packs, or express praise, typically highlighting the usefulness or appeal of a particular starter pack. Posts asking for help (*e.g.* instructions on how to create a new starter pack) are also common. Some users also express frustration with the current system and suggest features to be added. Those posts focus mostly on the lack of a search feature to discover new starter packs and users suggesting a notifications system to be informed when an account is added to a starter pack.[8] Interestingly, we identified multiple instances of starter pack creators asking for money to include new members. This suggests the existence of an emerging market where popular creators can sell membership in their starter packs. Multiple posts suggest that this exchange happens mostly over direct messages, and its scale may be larger than indicated by the public post analysis.

User concerns Finally, we manually analyze the posts classified under the "frustration" theme $1{,}699$ ($0.47\,\%$ of all posts containing starter packs) to understand the potential concerns of users. We find multiple users reporting that the starter packs are used in a negative way. This includes using starter packs as lists of accounts to block or even harass. Such usage is usually inspired by political and ideological differences. Many users also dislike being added to starter packs without their consent, finding it stressful or invasive, especially when the resulting follower spike disrupts their usual interactions. This includes popular and well-established accounts being added without their consent to malicious starter packs. It seems that creators do this to boost their starter pack credibility. We note that all these problems could be alleviated by requiring members to first agree to be added to a selected starter pack.

Multiple users also criticize the *follow-all* operation. These users perceive that it artificially inflates the social graph, leading to shallow interactions and timeline pollution. Finally, we observe an emotional strain reported by some starter pack members and creators. Members worry that being added to starter packs may lead to followers expecting content misaligned with their usual posts, causing misunderstandings or negative feedback. Creators report feeling pressured to include everyone—an impossible task given the cap on the maximum number of members of starter packs. They therefore fear backlash from excluded users.

7 Related Work

There are two core areas of related work: (*i*) social bootstrapping, often referred to as the cold start problem; and (*ii*) social network migration.

Social Bootstrapping There has been extensive work looking at the cold-start problem in social networks (aka "social bootstrapping"). Traditionally, this has been treated as a recommendation system problem (Yuan and Hernandez 2023), whereby "new" friends must be recommended to the incoming user. Early work focused on accelerating friend recommendations by quickly identifying other users with similar interests or friends-in-common (Sahebi and Cohen 2011). However, this body of work does not consider user migrations from other platforms, instead, working on the assumption that users come afresh to the platform. We argue that this is a missed opportunity as other studies have shown that prior social links have high predictive power in determining which newcomers will continue to engage with services (Burke, Marlow, and Lento 2009).

Consequently, there has also been work looking at how new social networks can "borrow" links from older ones. Gong et al. (Gong et al. 2021, 2018) show that a user's existing social network (*e.g.* on Facebook) can effectively be used to predict emergent links on a new social network (*e.g.* Pinterest). Zhong et al. (Zhong et al. 2014) also study the benefits of migrating links from prior social networking platforms. They proposed a mechanism to copy social links for existing social graphs, called Link Bootstrapping Sampling. They show that borrowing links from existing user social networks improves the robustness of the fledgling one. Since these studies, various social networks have introduced such bootstrapping techniques. For example, Zhang et al. (Zhang et al. 2024) explore how the creation of Threads (by Meta) benefits from the importation of links and users from Instagram. Others have studied alternative applications of link transfers across social networks. For example, Venkatadri et al. (Venkatadri et al. 2016) use link transfers to estab-

[8] We note that leveraging the open nature of Bluesky, the community has created ad hoc solutions for both of these cases.

lish trust in new social networks, which have not yet been bootstrapped. Bluesky is rather different in that it does not migrate links directly from prior social networks. Instead, key people create starter packs that contain well-known people from the community. We show that this novel approach brings similar benefits to network bootstrapping.

User Migrations A small set of recent studies have investigated migration patterns between social networks, primarily of users from Twitter/X (Bittermann, Lauer, and Peters 2023). Rather than borrowing links, these primarily study cases where users entirely abandon the previous social network. He et al. (He et al. 2023) study the recent migration of users from Twitter/X to Mastodon. They show that the social network plays a major role, with users becoming more likely to migrate once their friends have migrated. Cava et al. (Cava, Aiello, and Tagarelli 2023) found similar patterns, showing that a user's social network is a key factor in driving their migrations. Jeong et al. (Jeong et al. 2024) also investigate the behavior of users who perform this migration. We complement this work, by studying how such social networks are migrated on Bluesky. Importantly, Mastodon lacks any concept of starter packs, forcing users to manually rebuild their network. Importantly, our work differs in that we are not investigating *why* users migrate to Bluesky. Instead, we focus on how starter packs simplify this migration and accelerate social bootstrapping.

8 Conclusion & Future Work

This paper has studied the use of starter packs on Bluesky and how they can help bootstrap a robust social network. We began by collecting all 335.42×10^3 starter packs on Bluesky, tracking their changes, creators, members, and descriptions (**RQ1**). Our temporal analysis revealed activity spikes, particularly during real-world events, confirming the central role of starter packs in organic community development. We then examined temporal patterns on Bluesky (**RQ2**) and found that inclusion in a starter pack provides clear benefits: included users gain significantly more followers and share more posts than others. However, these effects are largely limited to starter pack members and may reinforce the *rich get richer* dynamic, potentially increasing inequalities. Next, we analyzed posts discussing starter packs (**RQ3**). While most users view them positively, our perception analysis also uncovered complaints about their impact on conversation quality and the dynamics of inclusion.

By addressing these research questions, we lay the groundwork for studying how starter packs influence social graph dynamics. We identify several future directions. First, we plan to extend our analysis over a longer period. Even in the few months studied, we observe clear temporal patterns, and we aim to explore how these evolve outside of the major user influxes in late 2024. Second, we argue that starter pack operators may take on increasingly central roles, influencing community formation. This could introduce power imbalances and lucrative opportunities (e.g., selling spots in starter packs). We aim to study these dynamics as they unfold, including whether they foster echo chambers and polarization. Finally, we make our code and datasets available to support further research into starter packs and Bluesky.[9]

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions, and Gabriele Piazza for his generous support and insightful discussions. This work was supported in part by the Guangzhou Science and Technology Bureau (2024A03J0684), Guangdong provincial project 2023QN10X048, the Guangzhou Municipal Key Laboratory on Future Networked Systems (2024A03J0623), the Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No.2023B1212010007), the Guangzhou Municipal Science and Technology Project (2023A03J0011), Guangdong provincial project (2023ZT10X009), and the 111 Center (No. D25008). This work was supported in part by AP4L: Adaptive PETs to Protect & emPower People during Life Transitions (EP/W032473/1). This work was co-funded by the LOEWE initiative (Hessen, Germany) within the *emergenCITY* center [LOEWE/1/12/519/03/05.001(0016)/72].

References

Abiola, O.; Abayomi-Alli, A.; Tale, O. A.; Misra, S.; and Abayomi-Alli, O. 2023. Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser. *Journal of Electrical Systems and Information Technology*, 10(1): 5.

Angriman, E.; van der Grinten, A.; Hamann, M.; Meyerhenke, H.; and Penschuck, M. 2022. *Algorithms for Large-Scale Network Analysis and the NetworKit Toolkit*, 3–20. Cham: Springer Nature Switzerland.

Balduf, L.; Sokoto, S.; Ascigil, O.; Tyson, G.; Scheuermann, B.; Korczyński, M.; Castro, I.; and Król, M. 2024. Looking AT the Blue Skies of Bluesky. In *Proceedings of the 2024 ACM on Internet Measurement Conference*, 76–91.

Bhattacharjee, D.; and Mohanty, P. C. 2022. Does social network impact rising out-of-pocket health care expenditure in India? A propensity score matching approach. *The International Journal of Health Planning and Management*, 37(5).

Bittermann, A.; Lauer, T.; and Peters, F. 2023. Academic# TwitterMigration to Mastodon: The role of influencers and the Open Science movement.

Burke, M.; Marlow, C.; and Lento, T. 2009. Feed me: motivating newcomer contribution in social network sites. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 945–954.

Cava, L. L.; Aiello, L. M.; and Tagarelli, A. 2023. Drivers of social influence in the Twitter migration to Mastodon. *Scientific Reports*, 13(1): 21626.

Cha, M.; Haddadi, H.; Benevenuto, F.; and Gummadi, K. 2010. Measuring user influence in twitter: The million follower fallacy. In *Proceedings of the international AAAI conference on web and social media*, volume 4, 10–17.

[9]https://bsky.leobalduf.com/.

Diyasa, I. G. S. M.; Mandenni, N. M. I. M.; Fachrurrozi, M. I.; Pradika, S. I.; Manab, K. R. N.; and Sasmita, N. R. 2021. Twitter sentiment analysis as an evaluation and service base on python textblob. In *IOP conference series: materials science and engineering*, volume 1125, 012034.

Dos Reis, V. L.; and Culotta, A. 2015. Using matched samples to estimate the effects of exercise on mental health via twitter. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29.

Gilbert, E.; and Karahalios, K. 2009. Predicting tie strength with social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 211–220. ACM.

Gligorić, K.; Zrnic, T.; Lee, C.; Candès, E. J.; and Jurafsky, D. 2024. Can Unconfident LLM Annotations Be Used for Confident Conclusions? *arXiv preprint arXiv:2408.15204*.

Golder, S. A.; Wilkinson, D.; and Huberman, B. A. 2007. The rhythms of social interaction: Messaging within a massive online network. *Communications Review*, 2007(2007).

Gong, Q.; Chen, Y.; He, X.; Xiao, Y.; Hui, P.; Wang, X.; and Fu, X. 2021. Cross-site prediction on social influence for cold-start users in online social networks. *ACM Transactions on the Web (TWEB)*, 15(2): 1–23.

Gong, Q.; Chen, Y.; Hu, J.; Cao, Q.; Hui, P.; and Wang, X. 2018. Understanding cross-site linking in online social networks. *ACM Transactions on the Web (TWEB)*, 12(4): 1–29.

He, J.; Zia, H. B.; Castro, I.; Raman, A.; Sastry, N.; and Tyson, G. 2023. Flocking to mastodon: Tracking the great twitter migration. In *Proceedings of the 2023 ACM on Internet Measurement Conference*, 111–123.

Jeong, U.; Sheth, P.; Tahir, A.; Alatawi, F.; Bernard, H. R.; and Liu, H. 2024. Exploring platform migration patterns between twitter and mastodon: A user behavior study. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, 738–750.

Jiang, A. Q.; Sablayrolles, A.; Mensch, A.; Bamford, C.; Chaplot, D. S.; Casas, D. d. l.; Bressand, F.; Lengyel, G.; Lample, G.; Saulnier, L.; et al. 2023. Mistral 7B. *arXiv preprint arXiv:2310.06825*.

Kleppmann, M.; Frazee, P.; Gold, J.; Graber, J.; Holmgren, D.; Ivy, D.; Johnson, J.; Newbold, B.; and Volpert, J. 2024. Bluesky and the AT Protocol: Usable Decentralized Social Media. In *Proceedings of the ACM Conext-2024 Workshop on the Decentralization of the Internet*, 1–7.

Kwak, H.; Lee, C.; Park, H.; and Moon, S. 2010. What is Twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web (WWW'10)*, 591–600. ACM.

Mekacher, A.; Falkenberg, M.; and Baronchelli, A. 2024. The Koo Dataset: An Indian Microblogging Platform With Global Ambitions. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18.

Mislove, A.; Marcon, M.; Gummadi, K. P.; Druschel, P.; and Bhattacharjee, B. 2007. Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, 29–42. ACM.

Pew Research Center. 2023. Social Media Use in 2023. https://www.pewresearch.org/. Accessed: 2025-01-08.

Sahebi, S.; and Cohen, W. W. 2011. Community-based recommendations: a solution to the cold start problem. In *Workshop on recommender systems and the social web, RSWEB*, volume 60.

Steven Loria. 2024. TextBlob: Simplified Text Processing. https://textblob.readthedocs.io/en/dev/. Accessed: 2025-01-14.

TeBlunthuis, N.; Hase, V.; and Chan, C.-H. 2024. Misclassification in Automated Content Analysis Causes Bias in Regression. Can We Fix It? Yes We Can! *Communication Methods and Measures*, 18(3): 278–299.

Ugander, J.; Karrer, B.; Backstrom, L.; and Marlow, C. 2011. The anatomy of the Facebook social graph. *arXiv preprint arXiv:1111.4503*.

Valenzuela, S.; Arriagada, A.; and Scherman, A. 2014. Facebook, Twitter, and youth engagement: A quasi-experimental study of social media use and protest behavior using propensity score matching. *International Journal of Communication*, 8: 25.

Venkatadri, G.; Goga, O.; Zhong, C.; Viswanath, B.; Gummadi, K. P.; and Sastry, N. 2016. Strengthening weak identities through inter-domain trust transfer. In *Proceedings of the 25th International Conference on World Wide Web*, 1249–1259.

We Are Social. 2024. Digital 2024: Global Overview Report. https://wearesocial.com/. Accessed: 2025-01-08.

Yuan, H.; and Hernandez, A. A. 2023. User Cold Start Problem in Recommendation Systems: A Systematic Review. *IEEE Access*, 11: 136958–136977.

Zhang, P.; He, Y.; Haq, E.-U.; He, J.; and Tyson, G. 2024. The Emergence of Threads: The Birth of a New Social Network. *IEEE/ACM ASONAM*.

Zhong, C.; Salehi, M.; Shah, S.; Cobzarenco, M.; Sastry, N.; and Cha, M. 2014. Social bootstrapping: how pinterest and last. fm social communities benefit by borrowing links from facebook. In *Proceedings of the 23rd international conference on World wide web*, 305–314.

Zhu, Y.; Zhang, P.; Haq, E.-U.; Hui, P.; and Tyson, G. 2024. Exploring the Capability of ChatGPT to Reproduce Human Labels for Social Computing Tasks. In *International Conference on Advances in Social Networks Analysis and Mining*.

Zuo, W.; Mondragon, R. J.; Raman, A.; and Tyson, G. 2024. Understanding and Improving Content Moderation in Web3 Platforms. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, 1859–1870.

Zuo, W.; Raman, A.; Mondragón, R. J.; and Tyson, G. 2023. Set in Stone: Analysis of an Immutable Web3 Social Media Platform. In *Proceedings of the ACM Web Conference 2023*, WWW '23, 1865–1874. New York, NY, USA: Association for Computing Machinery. ISBN 9781450394161.

Ethical Checklist

1. For most authors:
   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? Yes
   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes
   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes
   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes
   (e) Did you describe the limitations of your work? Yes
   (f) Did you discuss any potential negative societal impacts of your work? NA
   (g) Did you discuss any potential misuse of your work? NA
   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? NA. Our research relies on public data only.
   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes
2. Additionally, if your study involves hypotheses testing:
   (a) Did you clearly state the assumptions underlying all theoretical results? Yes
   (b) Have you provided justifications for all theoretical results? Yes
   (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? Yes
   (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? Yes
   (e) Did you address potential biases or limitations in your theoretical framework? Yes
   (f) Have you related your theoretical results to the existing literature in social science? Yes
   (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? Yes
3. Additionally, if you ran machine learning experiments:
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? Yes. We provide all the prompts in the appendix. We will release the code after the review process to avoid violating double-blind rules.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? Yes, see the appendix
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? NA
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? Yes
   (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? Yes
   (f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? NA

A Appendix: Methodology for Large Language Model Usage

We classify started pack descriptions and posts using a prompt-based approach with the Mistral Instruct language model. The analysis was conducted on an HPC cluster with over 3,500 CPU cores and 30 TB of RAM. Each node has two Intel® Xeon® Gold 6248R processors and 384 GB RAM. GPU-accelerated tasks were run on nodes with NVIDIA RTX 8000 GPUs (40 GB memory).

A.1 Starter Packs Classification

LLMs have demonstrated robust performance at various text classification tasks (Zhu et al. 2024). In our starter pack classification (§4), we prompt the LLM to classify starter packs based on their descriptions. To minimize misclassification, we explicitly instruct the model to classify an entry as "unknown" when the available information is insufficient.

This results in 3.63% of entries being marked as "unknown" in both the LLM classifications, ensuring that uncertainty was appropriately reflected in cases where classification was genuinely indeterminate. Our approach aligns with prior research on black-box-based content classification, which has identified challenges in contextual interpretation and the risks of misclassifications due to incomplete or ambiguous input (Gligorić et al. 2024; TeBlunthuis, Hase, and Chan 2024). By combining LLM-based classification with targeted validation and bias mitigation strategies, we ensure that our method remains both rigorous and scalable for large-scale social data analysis.

Below, we provide the prompt for the classification of starter packs and their participants.

```
Based solely on the provided name and
description of the starter pack,
classify it into a community/category.
Do not provide any other text!
Name: {name}
Description: {description}

Follow these instructions for the
response:

1. Provide two classifications:
    - The first classification should
    represent the starter pack itself.
    If there is insufficient information
     or if the description is unclear,
    classify as "unknown".
```

```
    - The second classification should represent the participants (e.g., "artists," "musicians," "politicians," "activists," "scientists," "unknown"), as you deem appropriate.

    If there is insufficient information or if the description is unclear, classify as "unknown".

2. Each classification should be a category that represents the core idea of the community or its members.
    - Do not use two or more words (e.g., "sports or water sports").
    - Avoid ambiguity or overlapping terms. Select only the most appropriate classification based on the description and do not add any other details, just the classification.
    - If the description is unclear or if there is insufficient information, classify as "unknown".
    - Do not provide anything other details alongside the classifications. Not even in brackets ().

Guidelines:
Provide your response in the following format and do not output any other text:
- Starter Pack Classification: [only a suitable classification for the starter pack or "unknown" and nothing else (no explanations, no justifications, nothing)]
- Participants Classification: [only a suitable classification for the participants (plural) or "unknown" and nothing else (no explanations, no justifications, nothing)]
```

LLM Validation To validate the LLM's efficacy, we conduct a manual validation of 500 randomly sampled starter pack descriptions using a structured annotation methodology. We find that 87.6% (438/500) of classifications fully aligned with the LLM's predictions, confirming strong overall agreement. The remaining 12.4% (62/500) of cases are ambiguous, where the starter pack description lacks clarity or suggests multiple plausible classifications. We did not observe any erroneous classifications. Importantly, ambiguous cases are not automatically labeled as "unknown"; rather, they represent instances where a classification *could* be justified but lacks definitive clarity. This distinction ensures that our results accurately capture both model performance and the inherent complexities of the data.

A.2 Posts Classification Details

For posts classification, we use an approach similar to the one described above (Appendix A). However, we provide a manually created list of categories to be used for classification. In cases where the LLM returns text outside the pre-defined categories and not "Other," this typically occurs because it decides to provide its own classification as the post does not fit into any one category. These instances are considered as "Other" in Figure 15. The classification of posts is performed using the following prompt:

```
Classify the following post into one of these categories only.
Provide no additional text, explanation, or reasoning, just the category.

Categories:
    1: "Praising a Starter Pack or Starter Packs in General",
    2: "Explaining How the System Works or Reporting Starter Pack Experience",
    3: "Desire to Be Added to a Starter Pack",
    4: "Advertising a Starter Pack (including asking for members or inviting others to join)",
    5: "Expressing Frustration with the Current System (e.g., mass follow but zero engagement)",
    6: "Added Without Permission",
    7: "Suggesting Someone Create a New Starter Pack",
    8: "Asking for help (e.g., understanding how the system works or looking for a specific Starter Pack)",
    9: "Asking for money to include someone in a starter pack",
    10:"Asking to be removed from a starter pack",
    11: "Other"

Post: {post}

Your response should follow this exact format:
Category: [chosen category from the 9 options]

Guidelines:
- Do not add any text, parentheses, explanations, or reasoning.
- If unsure, select: "Other"
- Output only the category in the specified format.
```

LLM Validation For validation, we manually reviewed 500 randomly sampled classifications, comparing the LLM-assigned category with human judgment. Our findings indicate that 81.4% of classifications were accurate, 7% exhibited some ambiguity, and 11.6% were incorrect. The ambiguous cases often resulted from overlapping themes, leading to plausible yet imprecise classifications. For instance, ≈80% of posts incorrectly classified as "advertisements" contain a hyperlink that the model could not reach. Furthermore, ≈70% of wrongly classified posts were written in languages other than English, suggesting a lower model performance for those languages.

B Appendix: Supplementary Analysis and Validation of PSM Results

Balance Table. We assess the covariate balance before and after matching using *Standardized Mean Differences* (SMDs). This is a common diagnostic step to ensure that the matched treatment and control groups are statistically comparable across the key confounding factors discussed in Section 5.1. SMD quantifies the difference in means between two groups in units of pooled standard deviation and is widely used to evaluate the quality of matching in observational studies. An SMD below 0.1 is typically considered indicative of good balance.

Figure 16 shows the SMDs for each covariate before and after matching. The analysis includes variables measured at the time of treatment (t_0), such as number of followers, account age, number of posts, number of received likes, number of issued likes, and number of followed accounts. After matching, all covariates achieve an SMD below 0.1, indicating strong covariate balance and therefore fulfilling the criteria. Notably, the SMD for account age drops from 1.1 to 0.03, showing substantial improvement. All covariates also show reduced SMDs post-matching compared to pre-matching. The only exception is the followers-to-following ratio, which nevertheless remains always below the 0.1 threshold. These results confirm the quality of the PSM-based matching procedure to pair the treatment and control groups.

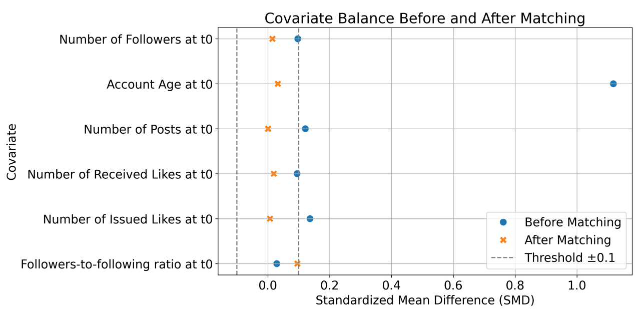


Figure 16: Balance table with covariate balance before and after matching. The dashed vertical lines at ±0.1 indicate the commonly used threshold for acceptable balance.