HARVARD | BUSINESS | SCHOOL

Sampling Bias in Entrepreneurial Experiments

Ruiqing Cao
Rembrand Koning
Ramana Nanda

Working Paper 21-059

Sampling Bias in Entrepreneurial Experiments

Ruiqing Cao
Stockholm School of Economics

Rembrand Koning
Harvard Business School

Ramana Nanda
Imperial College London

Working Paper 21-059

Funding for this research was provided in part by Harvard Business School and the Kauffman foundation. During the academic years 2019-2021, Nanda is a Visiting Professor at Imperial College, London.

Sampling Bias in Entrepreneurial Experiments*

Ruiqing Cao
Stockholm School of Economics

Rembrand Koning
Harvard Business School

Ramana Nanda
Imperial College London

July 29, 2022

Abstract

Using data from a prominent online platform for launching new digital products, we document that 'sampling bias'—defined as the difference between a startup's target customer base and the actual sample on which early 'beta tests' are conducted—has a systematic and persistent impact on the venture's success. Specifically, we show that products with a female-focused target market launching on a typical day, when nine in ten users on this platform are men, experience 45% less growth a year after launch than those for whom the target market is more male-focused. By isolating exogenous variation in the composition of beta testers unrelated to the characteristics of launched products on that day, we find that on days when there are unexpectedly more women beta testers on the platform—reducing the amount of sampling bias for female-focused products—the gender-performance gap shrinks towards zero. Our results highlight how sampling bias can lead to fewer successfully commercialized innovations for consumers who are underrepresented among early users.

*Cao: rcao@hbs.edu; Koning: rem@hbs.edu; Nanda: ramana.nanda@imperial.ac.uk. Author order is alphabetical and all authors contributed equally. We are extremely grateful to Kevin Laws and Ryan Hoover for providing us access to data, to Daniel Yue and Wes Cohen for feedback on the learning mechanism, and to participants at Berkeley Haas, Columbia Business School, FDIC-Duke Fintech Conference, HBS, SIE Workshop, Chalmers University, Strategy Science Conference, London Business School, Duke Fuqua, NBER Productivity Lunch, Max Plank Institute for Innovation and Competition, Rice University, Rotman's Workshop on Gender, Race and Entrepreneurship, and Wharton School of the University of Pennsylvania for helpful comments. Koning and Nanda thank the Division of Research and Faculty Development at HBS for financial support. Koning thanks the Kauffman foundation for financial support. All errors are our own.

1 Introduction

Given the irreducible uncertainty associated with new ventures (Hayek 1948), entrepreneurship is increasingly studied and practiced as a process of experimentation. In this approach, entrepreneurs test initial versions of a new product with early users and learn from these tests to improve subsequent decision-making and future strategic investments (Ries 2011; Kerr, Nanda, and Rhodes-Kropf 2014; Gans, Stern, and Wu 2019). Consistent with this view, a growing body of research highlights the value of a scientific approach and tools like A/B testing that help startups run more effective experiments with early users (Camuffo et al. 2020; Koning, Hasan, and Chatterji 2022).

An implicit assumption behind the efficacy of such early experiments is that the preferences of early users reflect those of the startup's target customer base. In such instances, 'traction' with early users provides an unbiased estimate of potential traction with the target user base. However, if there is sampling bias, that is—*if early users of a new product are not representative of the startup's target market*—then the information gleaned from such early experiments might lead to weak initial traction, biased measures of the startup's potential, and lead to inefficient termination or pivoting by the venture. Beyond the impact of sampling bias on an individual entrepreneur's success, a *systematic* under-representation of key consumers amongst early users of a product has the potential to have aggregate effects, potentially leading to gaps in terms of who benefits from startup innovation. While practitioner-oriented work has identified this challenge as one to be aware of in the context of the mismatch between the needs of "early adopters" and the startup's larger target market (Moore 1991; Eisenmann 2021), there remains scant large sample evidence of either the presence and consequences of such sampling bias.

One reason for this dearth of evidence on sampling bias is data constraints. Identifying sampling bias across a large number of startups requires a scalable way for researchers to identify a wedge between a venture's *target* customer base and its *actual* early users. Moreover, studying the impact of sampling bias on firm outcomes requires exogenous variation in the sample of early users for these startups unrelated to the startup's target market.

We overcome these challenges through a combination of machine learning tools and unique data on nearly 6,000 startup launches on the prominent online product discovery platform Product Hunt. This platform, similar to the crowdfunding platform Kickstarter or Y Combinator's tech news platform Hacker News, lets users discover, vote on, and share products from nascent and emerging startups (Mollick 2014). Users on the platform come from the technology startup ecosystem—they are product managers, VCs, engineers, and other startup employees. Startups use these launches as "experiments" to gain traction, validate ideas, and even as a signal of quality when raising VC funding (Cao 2019).

Our analysis of sampling bias on Product Hunt progresses in three distinct steps. First, we use contemporary word embedding machine learning methods to characterize each product's pre-launch description in terms of its predicted appeal to female vs. male target customers. We show that our measure has strong face validity: startups such as Thinx (direct-to-consumer period-proof underwear) and Clue (menstruation tracking app) are above the 95th percentile in terms of this measure of predicted appeal to female users while a startup like Hims (direct-to-consumer hair loss and erectile dysfunction treatments) are below the 5th percentile. Moreover, leveraging proprietary data from Product Hunt on the gender of early users on the platform, we demonstrate a strong correlation between our measure of the product's predicted appeal to that gender and the actual gender-based preferences for that product as measured by upvotes. For example, even after accounting for product and user fixed effects, we find that men providing feedback on the platform are 25% less likely to vote for products that are in the top quartile of our measure of predicted appeal to female consumers.[1]

In our second step, we use show that on average products that are in the top quartile of our measure of predicted appeal to female consumers (what we refer to as 'female-focused products') have similar growth trajectories as products in the bottom quartile ('male-focused products') in the 6 months before launching on Product Hunt, but experience 45% less growth in the 12 months after launch. In other words, we find that products predicted to be more likely to appeal to female customers see less benefit from launching on Product Hunt and this holds no matter the gender of the entrepreneurs. Since such platforms are, on average, dominated by male early users—75% of visitors to Kickstarter are men, 79% for Hacker News, and 90% for Product Hunt[2]—that female-focused ventures benefit less from launching on Product Hunt is consistent with the presence of

[1]This variation is not reducible to a founder's gender, to differences in startup quality, or the "harshness" of male versus female users. That female-focused ideas launch with Product Hunt suggests that there may be limited alternatives, a point to which we return to in our conclusion.

[2]Estimates produced by the authors using data from the analytics platform SimiliarWeb as of Q4 of 2019.

sampling bias.

In our third step, we go beyond the correlation noted in our second step to provide causal evidence of sampling bias impacting the female-product growth gap we documented above. To do so, we isolate plausibly exogenous variation in the composition of female early users on the platform and show that when female-focused products launch on such days, the growth penalty we document relative to male-focused startups declines and may even go to zero. In other words, for two startups with the same predicted appeal to female target customers but launching on different days, differences in the gender composition of early users on the day the two startups launch has a systematic and measurable impact on the growth differentials of those startups at least one year after their respective launches. We see this as compelling evidence that sampling bias among early users has a real and potentially long-term effect on startup outcomes.

Having established the presence of sampling bias among startups, we then turn to the mechanisms by which sampling bias might impact systematic differences in startup growth after launch. Most directly, underrepresented early users make customer discovery for startups targeting these users harder, impacting future referrals and worth-of-mouth user growth; each of which can drive negative social learning dynamics that stunt consumer demand (Moretti 2011; Boudreau 2021). In addition, recent work on entrepreneurship highlights how entrepreneurial failure is often rooted in an entrepreneur failing to learn (Camuffo et al. 2021). Building on this work, if sampling bias yields signals of consumer demand that, because few women engage with the product, are both biased downward and *less precise* due to the small number of female users then female-focused entrepreneurs may incorrectly conclude their ideas have limited promise. As a result, entrepreneurs and investors—even those who are aware of and adjust for the sampling bias—may abandon promising female-focused ideas because they make incorrect or noisy inferences about the idea's potential. Indeed, using data on technology investments by entrepreneurs and capital investments by VCs for up to four years after launching on Product Hunt, we show that entrepreneurs and VCs are less likely to invest in female-focused startups – in a manner related to the composition of early users on the day the startup launches. This pattern is consistent with the idea that sampling bias impacts the ability to effectively learn from entrepreneurial experimentation.

Our results are relevant to scholars and practitioners interested in entrepreneurship, innovation, and gender. First, our findings enrich our understanding of the benefits and costs of experimental strategies (Levinthal 2017; Gans, Stern, and Wu 2019; Koning, Hasan, and Chatterji 2022; Camuffo et al. 2020). While prior work has largely focused on the benefits of business experimentation, our findings shed light on enabling conditions for these benefits to be fully realized. Second, our findings contribute to work on the rate and direction of innovation. While most work on the gender gap in innovation and entrepreneurship is focused on the entrepreneurs themselves (Gompers and Wang 2017; Scott and Shu 2017; Howell and Nanda 2019; Guzman and Kacperczyk 2019; Ewens and Townsend 2020), an emerging body of work has begun to show that product innovations and even ideas appear to be oriented towards the needs of men over women (Feng and Jaravel 2019; Koning, Samila, and Ferguson 2021; Truffa and Wong 2022). Our finding that female-focused products experience 45% less growth and are five percentage points more likely to be inactive provide further evidence that this appears to be the case. Moreover, we show how demographic biases need not only operate at the level of the founder or worker but appear to also shape what types of products succeed and who benefits from these innovations.

Finally, our paper's limitations point to promising puzzles for future research. First, why do our results persist for years after launch? While we are fortunate to have measures of performance from outside the Product Hunt platform, we can only partially observe what the founders' beliefs and strategic decisions were over this period. What frictions prevent them from making up the gap by launching on other platforms or testing the market through other means? Is it that entrepreneurs are "unscientific" and fail to correct for Product Hunt's biased signal or is it that the signals are simply too noisy to learn from given the limited sample size of female early users (Camuffo et al. 2020, 2021)? Is it that alternative means of discovering demand are harder to find or more costly to use for female-focused startups? If so, given competition between platforms, why do we observe so few platforms with a majority of women? These questions are especially important given the growing use of entrepreneurial testing and the dominance of a few platforms that de facto serve as gatekeepers to the types of ideas that are ultimately successful.

2 Empirical Context and Data: Product Hunt

2.1 Empirical Context

Our empirical context is the online platform Product Hunt, founded in 2013 and acquired by AngelList in 2016. Product Hunt serves as a community for technology enthusiasts and early adopters, who share new and emerging products on a daily basis. It has evolved over time into a platform for product launches, with early-stage startups using it to gain traction, get feedback, and build interest with investors. Successful products launched on the platform include RobinHood (day trading app which has raised more than $300 million in VC funding), Eero (interconnected wifi routers, acquired by Amazon), and Front (shared inbox for teams that has raised more than $60 million in funding). While most products posted are from small entrepreneurial teams, new products from companies like Stripe and Amazon are also listed on the platform.

The daily mechanics of Product Hunt are relatively straightforward. Each day around 20 newly launched technology products are featured on the platform and displayed on the homepage. Though products are submitted to the platform throughout the day, the vast majority of product launches occur in the early morning (Pacific Time) to maximize exposure and engagement over the course of the day.[3] Typically, within an hour of posting, products are screened by the platform's curators as both appropriate (i.e., not explicit) and of a minimum quality threshold (i.e., not spam apps) to be featured on the homepage.[4] Our study focuses on these featured products.

A product submission includes photos, sometimes videos, a detailed text description of the product, and links to the product's own website. It also includes a profile picture and the name of each of the makers—overwhelmingly the entrepreneurs behind the product. The products are then voted on by Product Hunt users and sorted on the homepage so that products that receive more votes are displayed prominently at the top of the page. Again, this incentivizes entrepreneurs to post as early as possible to gain the most votes and so the most visibility. The top five products of the day get badges for their rankings, products that perform well are often featured on the website

[3]Products are submitted to the platform by power users called "hunters.", 40% of the time, the hunter is the same person who builds the product (makers). Otherwise, makers reach out to hunters to post their products onto the platform. Either way, given the prominence of Product Hunt in the technology community, launches are overwhelming planned in collaboration with the firm.

[4]Occasionally, Product Hunt will ask a posted product to make a few changes to the descriptive texts, images, or videos, and the featuring will be delayed to the next day.

at later dates, and success on the platform often piques the interest of journalists and investors (Cao 2019).

2.2 Data

Product Hunt We use detailed data on products launched on the Product Hunt platform between October 2016 and October 2018. As discussed, this data includes the "makers" who built the product, a description of the product, and information on who visits and votes for products on the platform. Before we move to describe our additional sources of data, we first outline several restrictions we impose to ensure that our sample covers *new* products by new companies as against products launched by technology incumbents or posts that are not new technology products.

First, we restrict our sample to featured technology products launched on Product Hunt.[5] We make sure to discard listings that are blog posts, news articles, infographics, surveys, events, newsletters, email lists, political organizations, books, podcasts, and governmental agencies. In February 2017, Product Hunt made a concerted effort to exclude these sorts of postings from its featured list, so this sample restriction also maintains consistency in the types of products included across our sample period.

Second, as noted above, Product Hunt is also used by large companies such as Stripe and Microsoft to build buzz around new technology product releases. However, since we are interested in studying startups, we restrict our sample only to include the set of products that early-stage entrepreneurial firms launch. We define entrepreneurial firms as private companies that have raised at most a single round of Series A or Seed Financing, which we measure using data from Crunchbase and Preqin.[6] Since companies sometimes post to the platform multiple times, we restrict our analysis to the first post from a given website domain to look at new ventures and not subsequent iterations from already established companies. To further ensure our sample consists of early-stage

[5]As mentioned above, while roughly 20 technology products are featured on the homepage each day, there are about 60 additional non-featured postings too. Featuring is done by Product Hunt staff, who weed out dozens of junk and spam submissions that they feel will not be of general interest to the larger Product Hunt community. Non-featured products do not appear on the home page and receive little organic attention from the community. It is, of course, possible that this curating is biased against products that appeal to women; however, such a bias would only bias our estimates downwards. We leave it to future work to explore how the curation process impacts which products succeed and fail.

[6]This means that we exclude all firms that have already raised VC series B or beyond, that have raised multiple rounds of financing, and firms that have already gone public. Firms included in our sample include those without any external funding or those who have only raised Pre-Seed, Seed, Accelerator, Convertible Notes, Angel or Series A financing.

startups, we remove any companies where Crunchbase or Preqin have a listed founding year of 2013 or earlier.

Third, we restrict our sample to only include featured products launched on weekdays before 7AM Pacific Time. Many fewer users visit Product Hunt on weekends, and products launched on Saturday and Sunday are of noticeably lower quality. Nearly all of the most promising products launch early in the day to accrue as many votes as possible. Further, by only retaining products submitted before 7AM Pacific Time we ensure the product was submitted before the daily newsletter is sent out. As described below, this is relevant to how our empirical design isolates exogenous variation on the composition of users on the platform on a given day.

Finally, we restrict our analysis to the days when a "product update" newsletter is sent out to the Product Hunt community. As described in more detail in Section 4.5, Product Hunt sends out a daily newsletter just after 7AM Pacific Time. The majority of these newsletters provide updates on a handful of products previously launched on Product Hunt. We use the fact that newsletters with updates on a very female-focused product increase the number of female users on the platform, but that these updates are unrelated to the types of products launched that day.[7]

Our final sample is a balanced panel at the startup-month level comprising 5,742 nascent startups that have raised at most a single round of financing and launched on a "product update" newsletter day for the first time on Product Hunt between October 2016 and October 2018. We then supplement this sample of products with several additional data sources, each described below.

Genderize.io We use the first name of the makers and users on Product Hunt to estimate the gender composition of each startup team and the larger Product Hunt user base. We do so by taking the first names and feeding them through Genderize.io's public API, which returns a predicted probability that a person with the given first name is male or female. As described in detail in the appendix, we then use these probabilities to assign users as male, female, or unknown.

SimilarWeb To measure venture growth, we merge longitudinal data on website visits from SimilarWeb using each product's website URL. SimilarWeb provides web analytics services to businesses that allow them to benchmark competitor growth, keyword effectiveness, and a host of other digital trends. Using data from ISPs and a large panel of web browsers, SimilarWeb generates estimates of the number of users who visit a website each month. Crucially, web traffic is a key measure

[7] Our non-newsletter results are unchanged when run on all days, not just days with "update" newsletters.

of digital startups' initial traction and predictor of future investment and revenue (Koning, Hasan, and Chatterji 2022). Further, web traffic allows us to measure overall venture growth, not just success on the ProductHunt platform. For example, female-focused products may well get fewer votes on the Product Hunt, but could well find other avenues to build demand. Specifically, for each product, we measure monthly URL traffic for the 6 months before its launch on Product Hunt and 12 months after the launch. Building on Koning, Hasan, and Chatterji (2022) in the appendix, we show in Figure A1 that financing and page visits are strongly correlated, with startups in the bottom decile of visits having less than a 1% chance of raising venture funding and those in top decile a 12% chance.

Crunchbase and Preqin As briefly mentioned above, we linked products on Product Hunt to data from Crunchbase and Preqin venture capital databases. We linked the datasets using the product's name along with the listed URL. These datasets allow us to track which startups had raised funding, when they raised funding, and to measure the date of founding for more established firms. Using this data, we have up-to-date funding information as of October 2020.

BuiltWith Finally, we use data from `builtwith.com`. As described in more detail in Koning, Hasan, and Chatterji (2022) and in Roche, Oettl, and Catalini (2020) BuiltWith tracks the technologies startups use to run their websites. Since many of these technologies need to be 'client' facing, BuiltWith scrapes the website to see if it uses Google Analytics, Optimizlely A/B testing, Facebook tracking pixels, Shopify payment tools, and a myriad of other technologies. Using BuiltWith's free API, we measured the size of the technology stack as of October 2020 for 5,312 startups. We use the size of a website's technology stack as a proxy for the amount of product development. Indeed, given that the vast majority of products launched are digital, using the technology stack allows us to see if the entrepreneurs have continued developing the idea or halted their efforts. In the appendix, we show that products in the top decile of technology stack size relative to their cohort have a one-in-five chance of raising venture funding, whereas those in the bottom decile have essentially no chance of raising venture financing. Product development efforts also appear to correlate with more traditional measures of venture growth and success. See Appendix Section A for further details on these data sources.

3 Identifying the Target Market of Startups

Testing our sampling bias argument requires a measure of each startup's intended target market when they launch on the platform. However, generating such a measure is non-trivial. For example, a successful startup may have initially catered to women, but after launching on a male-dominated sample of early users, may have pivoted towards the needs of men. This example illustrates why using *realized* gender shares–even if available to the researcher—is not sufficient to measure the intended target market at launch. Instead, we need a measure of the startup's target market just before the time of launch.

3.1 A Measure of the Gender-Focus of Each Startup's Product

We create such a measure by analyzing the text associated with a product's description on the day it is launched. Our approach is similar to Koning, Samila, and Ferguson (2021), who use machine learning tools and the text of biomedical patents to classify inventions as more or less likely to benefit women. Our approach creates a continuous measure of a products' predicted appeal by gender, on a spectrum from catering primarily to women to catering primarily to men. Conceptually, our approach involves an algorithmic mapping from text-based product descriptions to a uni-dimensional measure of a product's gender focus. We describe this process below.

We begin by concatenating the text describing the product. This includes the new product's name, 'tagline' (a catchy one-liner attached to the product), a brief product description, and the initial comment by the makers describing the product in more detail.[8] In many regards, this text serves a similar function as the product descriptions in the 10Ks for public companies (Hoberg and Phillips 2016), but are for new products made by individual makers and non-public firms. Appendix Figure]A4 shows an example of each piece of text used to construct the measure of gender focus.

Using this text, we first remove common stopwords[9] and we keep only nouns, verbs, and ad-

[8]Note that the comment is also an ex-ante measure of the product's characteristics because of the particular way the Product Hunt platform works. To promote engagement with the Product Hunt community, makers often introduce products at the top of a public comment thread so that users are more likely to give feedback and test out the product.

[9]The stopwords are a union of the following lists: http://www.ranks.nl/stopwords; https://pypi.python.org/pypi/stop-words; https://msdn.microsoft.com/zh-cn/library/bb164590; http://snowball.tartarus.org/algorithms/english/stop.txt; https://bitbucket.org/kganes2/text-mining-resources/downloads/minimal-stop.txt; and Porter Stemmer stop words in NLTK.

jectives.[10] For the remaining words, we then use a pre-trained word embedding model[11] to map each word to a 300-dimensional numeric vector. This approach to text analysis—treating words as points in a high-dimensional vector space—is increasingly used by management scholars to capture hard-to-measure concepts including job-relatedness, firm competition, and organizational culture (Hasan, Ferguson, and Koning 2015; Hoberg and Phillips 2016; Srivastava et al. 2018).

That said, our approach extends this past work by relying on the fact that word embedding models produce vector spaces that preserve semantic meaning and context. Crucially, these embeddings appear to capture gender roles, stereotypes, preferences, and biases. For example, word embeddings are known to capture analogical reasoning. Taking the vector for the word "King," subtracting "man," and adding "woman" results in "Queen." These examples suggest that we can use the distance from clearly gendered words—male versus female, he versus she, man versus women—to measure whether a product is more or less likely to appeal to men or women.

Specifically, we calculate the extent to which the word is nearer in semantic space to words associated with women as against words associated with men. To do so we look up the a normalized word vector $\boldsymbol{v}_f$ that represents a word associated with women (e.g. she) and $\boldsymbol{v}_m$ a word associated with men (e.g. he). Then for any other word $\boldsymbol{w}$ (e.g. pregnancy) in a product's description we estimate its relative distance between the $\boldsymbol{v}_f$ and $\boldsymbol{v}_m$:

$$F_{\{f,m\}}(\boldsymbol{w}) = \frac{Cos\,(\boldsymbol{w}-\boldsymbol{v}_m, \boldsymbol{v}_f-\boldsymbol{v}_m)\cdot|\boldsymbol{w}-\boldsymbol{v}_m|}{|\boldsymbol{v}_f-\boldsymbol{v}_m|} - \frac{1}{2} = \frac{\langle \boldsymbol{v}_f-\boldsymbol{v}_m, w-\boldsymbol{v}_m\rangle}{|\boldsymbol{v}_f-\boldsymbol{v}_m|^2} - \frac{1}{2} \tag{1}$$

Note that geometrically, this is equal to the ratio of the length of vector $\widehat{\boldsymbol{w}-\boldsymbol{v}_m}$ – which is the projection of $(\boldsymbol{w}-\boldsymbol{v}_m)$ onto the vector defined by $(\boldsymbol{v}_f-\boldsymbol{v}_m)$ – to the length of the vector $(\boldsymbol{v}_f-\boldsymbol{v}_m)$, minus 0.5.

More generally, for any pair of keywords $\{f, m\}$ where f represents female and m represents male, we can define the relative appeal of a word represented by $\boldsymbol{w}$ to the female keyword using Equation 1. $F_{\{f,m\}}(\boldsymbol{w})$ increases in relative closeness to f, and a value close to 0 indicates that the

[10] Hoberg and Phillips (2016) keeps only nouns in product descriptions in 10Ks to construct word vectors. We also include verbs and adjectives because compared to the formal document such as 10Ks, our texts contain more vivid language, as product makers write these texts to encourage feedback from Product Hunt community, and use many verbs and adjectives that turn out to be very informative.

[11] We use the fastText package developed by Facebook Research and estimate the skip-gram model on the Wikipedia corpus as training texts. The vector space in 300 dimensions. For more details, see https://fasttext.cc/ and Bojanowski et al. (2017)

word is likely to be gender-neutral.

To measure gender focus at the product level, we calculate each word's closeness to 3 keyword pairs—{male,female}, {woman,man}, and {she,he}—and aggregate over all the words used to describe the product. Not all words appearing in product texts are counted equally. Following standard practice, for each word, we compute its TF-IDF (term-frequency inverse-document-frequency) weight, using texts of all products launched on Product Hunt as the corpus.[12] For each of the 3 keyword pairs, we calculate a measure at the product level that is the TF-IDF weighted sum of words' closeness to the the female vs. male keywords. Since each of these three keyword pairs likely capture slightly different and idiosyncratic meanings, we take the standardized first principal component as our measure of a product's female focus. The distribution of our final female-focus measure is presented in Figure A5. The histogram is bell-shaped and symmetric around the mean, but the tails are broader than those of a standard normal distribution.

3.2 Validating the Measure

We build our measure of a product's gender focus for all products ever launched on the platform. We then use this data to validate our measure in three ways. First, we examine the face validity of the measure by documenting examples from different points in the distribution. Table 1 presents examples from the 1st, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 99th percentiles of the female focus distribution, where higher percentiles correspond to being more focused on women.

A quick review of the table suggests that our measure captures differences in potential appeal. One of the most female-focused products is Babee on Board "Pregnant? Request a seat on public transport." One of the most male-focused products is Beard Bib 2.0 "Hair clippings catcher from Beard King." At the 90th percentile, you see products like Ropazi "Personal shopper for busy parents," which, while gender-neutral, seems reasonable to classify as more female-focused given the persistent fact that women do more housework and parenting than men (e.g. Fitzpatrick and Delecourt (2020)). At the 10th percentile, one sees products like Nikola "See your Tesla's battery percentage from your menubar," which again, while gender-neutral, seems reasonable to classify as more male-focused. The 50th percentile products such as Yomu "One place to read your favorite

[12]The IDF (inverse-document-frequency) down-weighs words that are common across all products (e.g. the word "product" itself), whereas the TF (term-frequency) weighs each word proportionally to its frequency of occurrence in the given product.

content from around the web" are also consistent with being completely gender-neutral. Finally, the 75th percentile features the example Joonko, which is a "Personal diversity and inclusion AI-coach for managers" which presumably would be something that would appeal more to female workers.

Second, we build on recent work showing that female-focused products are much more likely to be invented by women inventors and entrepreneurs (Feng and Jaravel 2019; Koning, Samila, and Ferguson 2021). Appendix Figure A7 is a binned scatterplot that documents a strong positive correlation between our measure of a product's female focus and the share of the founding teams that have at least one female maker (Starr and Goldfarb 2020). As can be seen from Figure A7, a one standard deviation increase in a product's estimated female focus – equivalent of moving from the 50th to 80th percentile – is associated with a 20% increase in the likelihood of at least one team member in the startup being a woman. Here we show results for all 19,388 products, including posts by incumbents, firms that have raised more than a round of financing, that were not featured, and posted on all days of the week. We find the same pattern holds in our sample of 5,742 new product-startup launches.

Third, we directly validate our measure using data from ProductHunt where we look at the correlation between our measure (which *predicts* the relative appeal of a product to women) and the *actual* appeal of that product as measured by the difference in upvotes from men and women. As described above, users can vote on the products launched each day. Crucially, we have an estimate of each user's gender. If our measure is valid, then we should see more women vote for products we estimate as female-focused and potentially fewer male users doing so too.

The binned scatter plots in Figure 1 show that more women and fewer men vote for female-focused products. In the figure, the y-axis in the left panel is the logged number of male votes and in the right panel the logged number of female votes for a product. The x-axis in both panels is the product's estimated female focus. To better isolate differences in female versus male preferences, as against overall differences in product quality that might be correlated with a product's female focus, both figures control for the logged number of votes of the opposite gender. Panel (a) on the left is strongly negative sloped and reveals that going from the most male-focused to most female-focused products corresponds to a nearly 50% decrease in the number of male votes. Panel (b) on the right shows a strong positive relationship and reveals that going from the most male-focused to female-focused products is associated with a a roughly 50% increase in the number of female votes.

To further validate that our female-focus measure captures differences in the gender appeal of the product and not merely differences product quality or gendered differences in voting "harshenss," in Figure 2 we test if female users are more likely to vote for female-focused products even after controlling for user- and product-level fixed effects. To do so, we leverage proprietary user-level data on who was active on the website on a given day, and for that day, which products they voted for. Using this data, we create a user-product level data set where each row represents a product that was launched on the day the user was active. Our assumption here is that if the user is active on the day a product is launched, they are at risk of voting for it. We then create a variable "voted for product" that is 1 if the user voted for the product and is 0 otherwise. Appendix Section E describes how we construct this proprietary data in detail.

This propriety browsing data lets us include individual fixed effects that account for differences in how "harsh" each voter is, which address concerns associated with potential differences in overall voting rates by gender. It also lets us include product-level fixed effects to account for observed and unobserved quality differences across products.[13] Figure 2 presents a binned scatter plot of the female-user-minus-male-user residuals plotted against the product's gender focus. The strong upward slope indicates that, even after accounting for potential differences in how women vote or the quality of female-focused products, women are more likely to vote for female-focused products than men. Given that the median browser upvotes roughly 1.47 of the 100 products they view, the graph suggests that going from the most-male to the most-female-focused product increases the absolute difference between female and male voters by 0.3 percentage points. Relative to the baseline, this represents an effect of about 20-25%, similar in size to the effects we find in Figure 1. Even on male-dominated product hunt, there are significant differences in each product's female-focus and whether women or men are more likely to vote for these products.

[13]We can estimate product fixed effects since we estimate the difference between male-female voters and not the overall appeal of the product. Put differently, within a product; we have variation in user gender that allows us to estimate an effect even after accounting for product-level fixed effects.

4 Results

4.1 Descriptive Statistics

Having established that product's meaningfully vary in the gender-focus of their target market, we next turn to understanding whether female-focused products benefit less from launching on the Product Hunt platform.

Table 2 first shows basic descriptive statistics for the 5,742 products used in our analysis and compares products in the top quartile of the female-focus distribution to those in the bottom quartile (i.e., the most male-focused products). Appendix Table A1 presents additional descriptive statistics for the full sample.

Product Hunt's topic categories reveal that the startups span many topics related to digital products. Comparing the distribution of products across those that are more female-focused vs. more male-focused reveals category differences: for example, products related to the topic "Developers" constitute a much larger share of the most male-focused products (21%) compared to female-focused products (8%). The average team has 2 makers (founders) and 19% of teams have at least one woman.

Unsurprisingly, products that we predict cater more to women compared to men are different on several of these dimensions. This suggests it is important to control for these covariates when studying the post-launch performance of startups. As we show below, our research design can control for these differences and any other time-invariant unobserved differences using product fixed effects. That said, we see little evidence that female-focused products are smaller or are less likely to have raised venture financing before launch. If anything, female-focused products have about 18% more monthly visits than male-focused products pre-launch.

4.2 Post-launch Performance Results

We begin our regression analysis by estimating the following simple fixed effect model:

$$y_{it} = \beta_1 POST_t + \beta_2 GenderFocus_i \times POST_t + \gamma_i + \delta_t + \epsilon_{it} \tag{2}$$

where i indexes products and t indexes time in months.

Y_t is the log of pageviews for a product in a month, and $POST_t$ is an indicator that takes the value of 1 after the product has launched and zero otherwise. The regression includes product fixed effects, γ_i, and year-month fixed effects, δ_t. Product fixed effects control for time-invariant unobserved differences in products such as their quality. Note that year-month fixed effects control for changes in traffic that happen over calendar time, for example, capturing points in the calendar when traffic might be especially high or low.The coefficient $\beta_1 POST_t$, which measures the average increase in visitor growth for products after they launch on the platform, continues to be separately identified due to the fact that products launch at different points in calendar time. Our coefficient of interest is $\beta_2 GenderFocus_i \times POST_t$ which captures the degree to which the post-launch visitor growth is related to the gender focus of products.

We report this regression in Panel A of Table 3. Looking first at the coefficient on the post-launch indicator in Column 1, we can see that on average, web traffic for a startup that launches on product hunt jumps by just over 300% in the post-period, relative to the average web traffic prior to launch. The average product startup goes from having hundreds of monthly visits pre-launch to thousands post-launch. Turning to Column 2, the β_2 coefficient shows that female-focused products appear to have systematically lower web traffic in the year following launch. The coefficient of -0.208 in column (2) implies that a one standard deviation increase in a product's female focus is associated with roughly 21% fewer visitors post-launch.

In Panel B, we look at the same relationship, but instead of imposing a linear functional form, we shift to a non-parametric approach where we compare the top and bottom quartiles of female-focused products to the middle two quartiles of our female-focus measure. Doing so is important because standard models from industrial organization would predict that male- and female-focused products would both see less growth. After all, by catering to only one gender, a product has cut the potential market size in half. That said, as Panel B shows, the most male-focused products do not experience any difference in web traffic from the more gender-neutral products. However, the most female-focused products experience nearly 45% lower web traffic post-launch.

Figures 3 provides a graphical illustration of the estimates produced by Table 3, but instead of reporting average post-launch growth, it reports the estimated visits in each month before and after launching, with the month before launch serving as the excluded baseline. The dashed line is the estimated number of monthly visits for products in the middle two quartiles of the female-focus distribution. The solid line is the estimated number of monthly visits for products in the top quartile. The bars reflect 95% confidence intervals.

Before launch, the growth trajectories look the same. Post-launch, female-focused products see less growth. It is important to note the estimated effects are on a log scale. So while the difference between male and female products is visually small, it represents substantive differences in relative growth. Figures 4 reports the difference between gender-neutral products and female-focused products. The growth-gap is immediate and persistent with female-focused products experiencing roughly 50% less growth post-launch.

Figure 5 shows that the estimated growth trajectory for male-focused products overlaps with gender-neutral products. Again, it does not appear that merely being "gendered" reduces growth. Figure 6 reports the difference between gender-neutral products and male-focused products. We again find no evidence that the most male-focused products experience less growth post-launch. These findings do not reflect limited growth opportunities for strongly gendered products. Rather they are distinctly related to a product being female-focused.

4.3 Distinguishing Between Target Market and Entrepreneur Gender

Figure A7 shows that female-focused products are much more likely to be built by teams with women. This suggests the product-gender gap might be the result of women founders benefiting less from launching on Product Hunt as against female-focused products receiving less of a boost. Indeed, in Column 3 of Table 3A, we regress logged monthly page visits on our post-launch dummy and this dummy interacted with whether the team includes a female maker. Consistent with the voluminous literature showing gender gaps in entrepreneurship and venture growth (Gompers and Wang 2017; Scott and Shu 2017; Howell and Nanda 2019; Guzman and Kacperczyk 2019; Ewens and Townsend 2020), we find that teams with female makers experience 38% less post-launch growth.

In Column 4, we include both our female-team and female-product measures and find that both remain negative and significant. Neither measure is reducible to the other. Furthermore, Panel B Column 4 shows that the effect sizes are comparable, with products in the top quartile and female-focused teams both seeing 40% less growth after launching. Finally, to completely rule out founder gender effects, in Column 5, we restrict our data to all-male teams. We again find a post-launch product-gender gap with female-focused products experiencing 50% less growth. Finally, in contrast to recent work documenting the importance of congruence between entrepreneurs and the gender focus of their products we find no evidence for an interaction effect between the entrepreneur's gender and the product's gender focus (Hebert 2018; Tak, Correll, and Soule 2019).

4.4 Startup Failure

We next examine the degree to which the gender gap we measure at the intensive margin in terms of website hits is also measured at the extensive margin – in terms of firm failure.

Measuring startup failure is challenging for very early-stage firms (Bennett and Chatterji 2019). The software firms in our sample may not be formally registered, may not have an independent physical office, and are often operated by founders who also hold full-time jobs at other companies. As a result, a startup launched on Product Hunt is unlikely to declare bankruptcy or be formally shut down; instead, the product and website are most likely to be abandoned by the founders if the venture fails. Fortunately, we can proxy whether a startup is still active by whether it has an active user base or not. Startups with no visitors in a month should be much more likely to have shut down or have failed. Following Koning, Hasan, and Chatterji (2022) we define a startup as active if it has more than zero visitors in a month according to SimilarWeb and inactive if the estimated number of visitors is zero.

Table 4 replicates Table 3 but swaps page visits with whether the startup has an active user base. Column 2 again shows that female-focused startups are 5 percentage points less likely to be active post-launch. As before, these results are present in both the parametric (Panel A) and non-parametric (Panel B) specifications, as well as with all-male teams (Column 5). The product-growth gap results in a survival gap where products that benefit women are more likely to exit than those that benefit men. This finding is consistent with our sampling bias mechanism and suggests that the direction of startup innovation skews more towards the needs of men than women.

4.5 Isolating the Causal Impact of Sampling Bias

While the emergence of a product-gender gap is consistent with the idea of sampling bias, it does not directly test if a shift in Product Hunt's gender composition would lead to different products succeeding. In fact, while the analysis thus far builds on the fact that on average 10% of Product Hunt users are female it does not directly test if variation in female engagement matters. Perhaps other features of the Product Hunt platform—like the daily tournament-style voting mechanism—drive the growth gap we observe. Moreover, even if we could establish a correlation between female platform engagement and female-focused product success such an estimate would be rife with potential selection bias. For example, it is easy to imagine that a talented female-entrepreneur who is part of her university's "women in tech" club is likely to both launch a more successful product *and* bring more of her female friends onto the platform the day she launches.

How then can we test our proposed sampling bias mechanism? An ideal experiment to tell these mechanisms apart would be to randomly shift the gender composition of the Product Hunt platform on any given day. If sampling bias is at play, female-focused startups should perform better when launched on days when more women were randomly assigned to visit the platform. Building on this logic, our sampling bias mechanism implies that any shock that tends to increase female engagement will dampen the product-gender gap.

Here we exploit the fact that the content of Product Hunt's daily email newsletter—which is designed to increase engagement, interest, and traffic to the website—sometimes appeals more to women and sometimes more to men, but in a manner unrelated to the quality of startups launching on a given day. This is because the daily newsletter mostly features prior products launched on the platform.[14] Appendix Figure A8 shows two example newsletters, one that is a sponsored newsletter by the birth control startup Nurx and the other a newsletter covering Lululemon's acquisition of Mirror, both products that had been featured on Product Hunt in the past. We argue that newsletters that feature particularly female-focused products—like Mirror and Nurx—bring a greater number of female users onto the platform that day. As a result, days with more female-focused newsletters should see a smaller growth gap.

Furthermore, there are several reasons to believe the newsletter's content is unrelated to the types of products launched on any given day. First, the newsletter comes out after nearly all products have been posted for the day. To maximize engagement, exposure, and growth outcomes, teams tend to launch just after midnight PST. This ensures that the product can rack up the most votes during the day, capture more media attention, and increase social media exposure. As the newsletter comes out between 7AM and 11AM PST teams that have already launched

[14] A small number of same-day launches are highlighted by the newsletter, which we exclude because it will directly influence traffic to the launch page of this product. We are primarily interested in the "spillover" effect of newsletter-induced female visits on other female-oriented products that are not themselves suggested by the newsletter.

are unable to strategically respond to the newsletter content.[15] Second, the newsletter's content is largely determined by exogenous events that impact previously launched products. The two examples in Appendix Figure A8 are illustrative. The newsletter featuring Nurx was a sponsored ad, and the timing likely the result of negotiations between Nurx and Product Hunt, something new startups are unlikely to know. The newsletter featuring Mirror resulted from Lululemon's acquisition announcement that day, information that was kept confidential by both parties before the deal closed.

Given these arguments, we create a proxy variable for female engagement "Female Newsletter" that is the gender score of the most female-focused product listed in the newsletter that day. To aid in the interpretation of the triple interactions we run, we rescale this variable to have a minimum of 0 and a maximum of 1.[16] A newsletter score of 0 corresponds to the newsletter day when the most female-focused product in the newsletter was the most male focused; a score of 1 to the day where the most female-focused newsletter product was the most female focused. Table 5 shows summary statistics for each of the 419 newsletters in our sample.

The institutional details outlined above suggest that the "Female Newsletter" variable can help us estimate sampling bias effects. However, as illustrated in Appendix Figure A10 for this variable to identify the impact of "sampling bias" a number of conditions must be met. First, the newsletter measure must actually drive female engagement. While the arguments above suggest the newsletter should shift who visits Product Hunt, if readership is small, the newsletter may well have no effect. Second, while we expect the newsletter to increase female user engagement with female-focused products across multiple channels—voting, comments, visits to the product's homepage, direct feedback to the founders, social media mentions, investment reach outs,...—it should not impact alternative mechanisms that are not rooted in sampling bias. For example, the newsletter should not cause male user to prefer female-focused products more nor should it increase women's interest in non-female-focused products. Finally, if the shock is exogenous, then we should see no correlation between our "Female Newsletter" measure and the characteristics of products launched that day.

[15] Furthermore, because of Product Hunt's launch policies, teams cannot strategically withdraw and relaunch the next day. New products get a single shot on the platform.

[16] In Appendix Section F.4 we explore alternative methods of constructing the variable. We find that the maximum, as compared to the mean or median of the newsletter product scores, best predicts female user engagement and voting behavior. It appears that especially female-focused newsletter products are what drive female user engagement with Product Hunt.

Do these three conditions hold? Fortunately, Product Hunt's detailed data lets us test each of these conditions. In regards to the first two conditions, Appendix F shows the female newsletter shock increases female engagement with Product Hunt overall, increase female votes for female-focused products launched on the day of the newsletter, and that the newsletter does not operate through alternative mechanisms. Appendix Table A2 shows that on female-focused newsletter days more women visit both the Product Hunt homepage and the pages of products launching that day.[17] Further, in Appendix Figures A11 and Table A3 we show the newsletter leads to more votes from female users for female-focused products. Appendix Figure A12 and A3 document that there is no increase in the number of votes from male users. Consistent with this latter result, Appendix Section F.3 shows the newsletter shock does not shift male's preferences for female versus male products.

In regards to the third condition that the newsletter is exogenous, Table 5 shows that the type of products launched on more or less female-focused newsletter days exhibit no observable differences. Appendix Table A7 presents a formal balance table showing that the female-focus of the newsletter is not significantly correlated with whether the products have female makers, are female-focused, or the pre-launch growth trajectories of the product. We find similar results in Appendix F.6 that the products in the newsletter are also balanced, varying only in their female appeal and not in terms of other characteristics. Especially when considered with the institutional details described above, the balance tests suggest that the newsletter content can be seen as random with respect to the products launched on that day.

To explore the impact of the newsletter shock on the product gender gap, we estimate the same differences-in-differences estimation as before, but now also interact our female-focused measure with how "female-focused" the daily newsletter is. Specifically, we run the regression:

[17] Again, the newsletter features previously launched products and we exclude the handful of days where the newsletter features a product launched the same day. As a result, our results are not merely driven by women viewing the newsletter products on Product Hunt. Instead, these findings show the newsletter drives engagement with the home page and so the products launched the day of the newsletter.

$$
\begin{aligned}
y_{it} = \beta_1 POST_t + \\
\beta_2 GenderFocus_i \times POST_t + \beta_3 NewsletterShock_i \times POST_t + \\
\beta_4 NewsletterShock_i \times GenderFocus_i \times POST_t \\
+\gamma_i + \delta_t + \epsilon_{it}
\end{aligned}
\tag{3}
$$

where $NewsletterShock_i$ is measured as the rescaled maximum female focus of all the suggested products mentioned in the daily newsletter. Table 6 shows results from this triple-differences model. Column 1 includes the triple interaction to test if the changes in the newsletter impact post-launch growth for female-focused products. We find a positive and statistically significant triple interaction term. The magnitude of the estimate, $1.28(SE = 037)$, suggests that the product gender gap, estimated to be $-0.88(SE = 0.20)$, is wiped out when moving from the most male-focused to the most female-focused newsletter. Column 2 focuses only on all-male maker teams. We find the same patterns. Table 7 replicates 6 but focuses on whether the startup still has an active user base. We find a similar pattern of results.

Figure 7 shows the estimated difference between a female-focused product (top quartile) and a gender-neutral product (middle quartiles) at different quartiles of the newsletter shock distribution. There is a clear pattern. Female Focused products that happen to launch on days when the newsletters were more female-focused perform as well as gender-neutral products. Female-focused products launched on days where more men are brought onto the platform suffer even more considerable growth penalties than gender-neutral products.

Crucially, this pattern is not apparent when we compare male-focused products to gender-neutral products. Figure 8 shows the estimated differences by newsletter quartile for male-focused vs. gender-neutral products. If anything, male-focused products do slightly worse when the newsletter is particularly female-focused, though the estimated drop is not statistically significant.

Additional robustness checks point to this sampling bias mechanism. When we restrict our analysis only to teams with a woman entrepreneur in Table A9 we find teams with a female-maker only benefit when the product is female-focused, not when the product is male-focused or gender neutral. Nor is the effect simply explained by shifts in the total number of votes which would be suggestive that the newsletter shock is simply serving as gender agnostic advertising, and not as shifting the gender composition of the sample. When we control for the total number of votes in Table A10 we find essentially no change in our estimates. Taken together, these findings provide strong evidence that the gender-composition of who engages with the platform on a day has a *systematic* and *persistent* impact on venture outcomes.

5 Sampling Bias Mechanisms

Our analysis thus far has established three facts. First, startups launched on Product Hunt meaningfully vary in whether their target market is primarily male or female. Second, female-focused startups benefit less from launching on Product Hunt than gender-neutral or male-focused firms. Third, on days when the newsletter is likely to bring more women onto the platform, this growth gap shrinks towards zero. Taken together, these findings suggest that sampling bias has a real impact on startup growth.

5.1 Customer-based Social Learning

In terms of the underlying mechanism, the most direct impact of launching on a sample not representative of a startup's target market is that such a launch makes it harder and more costly to reach the right users. Given the reality that most startups face some customer acquisition costs, with mismatched early users, the startup is less likely to benefit from the relatively cheap customer discovery, future referrals, and worth-of-mouth user growth that comes from launching on Product Hunt. Furthermore, initially tepid consumer interest can drive negative social learning dynamics that stunt consumer demand and startup growth, especially for the many socially enabled and networked products that are launched on Product Hunt (Boudreau 2021; Moretti 2011; Salganik, Dodds, and Watts 2006; Koning and Model 2013). These demand-side factors outline how, even if the entrepreneur still believes the idea has the same amount of potential after launching on Product Hunt, sampling bias may well stunt growth and lead to startup failure.

5.2 Investor and Entrepreneur Learning

Furthermore, recent work on entrepreneurship as experimentation suggests an equally important complementary "supply-side" mechanism: changes in entrepreneurial and investor beliefs. In this view, the launch serves as a signal of the startup's potential (Kerr, Nanda, and Rhodes-Kropf 2014; Camuffo et al. 2021). The votes and feedback from early users, and the rate of post-launch growth, are information. Entrepreneurs, along with investors, then use this information to update their beliefs about the startup's potential. If the signals are negative, they update downward, either abandoning the project or pivoting to something more promising. If signals are positive, entrepreneurs and investors double-down on their investments.

Moreover, models of entrepreneurial learning can help rationalize why founders of female-focused startups may abandon promising ideas even if everyone—including the entrepreneur—is aware of the fact that 90% of Product Hunt users are male. In short, learning depends both on the level of the signal and the signal's precision, both of which can limit firm learning (e.g. Cohen and Levinthal 1994). For example, a "sophisticated" female-focused entrepreneur who tests on 100 users may well discard the signal coming from the 90 men in the sample to avoid bias. However the trade-off between bias and variance holds not just for statisticians but also for entrepreneurs. With data only coming from 10 female early users, the unbiased signal is simply less informative, less convincing, and so less likely to lead entrepreneurs and VCs to continue to invest and scale the idea. Indeed, in Appendix H we formally illustrate this logic with a simple numerical Bayesian model to show that both "naive" entrepreneurs who learn from a biased sample, and "sophisticated" entrepreneurs who try and remove the sampling bias, are more likely to abandon an idea than if they had access to a larger representative sample of users from their target market.

To test if sampling bias impacts investor and entrepreneur decision-making, we turn to measures of startup funding and product development. Our measure of investor funding is straightforward: did the startup raise a round of venture financing after launching on Product Hunt? Our measure of product development effort comes from BuiltWith's technology stack database. If entrepreneurs are actively developing their products, they should be adding new technologies to their websites, which in turn should increase the size of the websites' technology stacks. Instead, if the entrepreneur is putting less effort into her idea, we should see less development and so fewer technologies being used on the website. Both measures let us test if decision-makers—be it entrepreneurs or investors—are learning from and responding to differences in product success that are rooted in the amount of sampling bias present on the day the product was launched.

Unlike our monthly web traffic measure, our funding and technology stack measures are observed much less frequently. For funding, we know whether the startup had raised venture funding before launch and in the period between launch and October 2020. For the size of the technology stack, we only have data from October 2020. As a result, we analyze both outcomes using basic cross-sectional regressions,. While this rules out the use of product fixed effects, we control for the number of page visits in the month before the product launched and for whether the startup had raised venture funding before launching. Though imperfect, these controls allow us to account for differences in quality between more and less female-focused products. The models do include fixed effects for the year-month of launch to account for differences in the amount of time startups have to raise venture funding and develop their product post-launch.

In Table 8 we test if sampling bias shapes funding decisions up to four years after the product launches on Product Hunt. Column 1 shows results from a linear probability model. Startups with more pre-launch visits and that have already raised funding are much more likely to raise a round after launching. In column 1, we find that a standard deviation increase in a product's female focus is associated with a 4.6 percentage point drop in the likelihood of raising funding after launching on the platform. In Appendix Table A12 we show that this pattern holds when we look at product quartiles, with products in the top female-focus quartile being 4.8 percentage points less likely to raise funding compared to gender-neutral products. The positive and significant interaction term suggests that as the newsletter shifts from pulling more men to more women onto the platform, the effect shrinks towards zero. While the coefficient on the interaction term is larger than the female-focus estimate, the difference in magnitudes is not statistically significant. Column 2 restricts the sample to all-male startup teams and finds a similar pattern of results, though the interaction term is only significant at the 10% level. Given that 3.4 percent of products raise funding post-launch, the magnitudes of these effects are economically significant.

Table 9 tests if sampling bias shapes an entrepreneur's product development effort as measured by the size of the technology stack, again up to four years after the product launched on Product

Hunt.[18] We log the dependent variable to account for the skewed distribution and to aid in interpretation.[19] In Column 1 we find that an entrepreneur who launched a one standard deviation more female-focused product ends up with a 30% smaller technology stack. In Appendix A13 we find that the top quartile of female-focused products has technology stacks that are 50% smaller. Both estimates suggest that entrepreneurs put less effort into developing female-focused products. Again, when the newsletter shifts towards pulling more women onto the platform, this effect, like with funding, shrinks to zero. In Column 2, we find these results hold when we look at products launched by all-male teams. Overall, it appears that entrepreneurs put less effort into post-launch product development when launching female-focused products on days when male users dominate the platform.

5.3 Exploring Heterogeneity in Entrepreneurial Learning

Does sampling bias impact all entrepreneurs equally? While there are myriad sources of heterogeneity, our sampling bias model suggests that entrepreneurs who depend more on Product Hunt's signals should see the most robust "sampling bias" effects. With a greater reliance on the platform, such teams should put less weight on outside signals and so be more swayed by the outcomes generated by launching on the Product Hunt platform, be they good or bad.

Conversely, entrepreneurs who engage and rely more on the Product Hunt platform may be more aware of the fact that the Product Hunt early user sample skews 90% male. Under this view, entrepreneurs who have engaged more with the platform and are now launching a female-focused product may be better at adjusting the biased growth signal the platform generates, in contrast to the theoretical arguments outlined in Appendix Section H. These divergent predictions suggest that checking for heterogeneity by a team's engagement and familiarity with Product Hunt is especially useful.

To test between these opposing predictions, we turn to the maker's prior Product Hunt records to generate a measure of prior engagement with the Product Hunt platform. Specifically, for each

[18] Our technology stack models only include 5,312 products for which BuiltWith had current technology stack information for. In Appendix Table A11 we show that while startups with more users and that had raised funding before launching are more likely to be tracked by BuiltWith our core variables—the gender focus of the product and the daily newsletter—do not predict whether we have technology stack data. This suggests that selection bias is unlikely to drive our technology stack findings.

[19] Our results are unchanged if we use the raw count instead.

team, we sum the number of prior products the makers have voted on. We then use this variable to split our sample into teams that have been more active on the platform (more than 20 prior votes) versus those that have not.

We then test for heterogeneity by running the same regression we estimate in Column 1 of Table 6 on each of these split samples. Columns 1 and 2 in Appendix Table A15 reveals that teams that have been more active on Product Hunt see equal, and if anything, slightly greater, sampling bias effects.[20] In fact, in each sample we find statistically significant effects (5% level) for the "Post-Launch × Female-focus" coefficient and for the "Post-Launch × Newsletter Shock × Female-focus" coefficient. Similarly, in Appendix Section K we also show that entrepreneurial experience does not mitigate nor amplify the impact of sampling bias on startup outcomes and entrepreneurial learning. Overall, we find little evidence for heterogeneity, instead, we find evidence for homogeneity in sampling bias effects. This homogeneity suggests that there are few entrepreneurs who can overcome the impact of sampling bias on learning and so points to the idea that the systematic under-representation of women as early users may well shape who benefits from startup innovation.

6 Conclusion

In this paper, we have argued that gender imbalance among early users systematically impacts the growth and survival of new ventures that target female customers. We find that after launching on Product Hunt, female-focused products experience 45% less growth and are five percentage points less likely to have any active users one year later compared to gender-neutral or male-focused products. Using the content of newsletters to isolate shifts in the composition of users that are unrelated to the products launched on a given day, we find that this gender gap shrinks towards zero on days when more women are active on the platform. The composition of users on the platform on the day of launch also impacts the likelihood of future VC funding and the entrepreneurs' product development effort years after they launch on Product Hunt.

While these findings highlight the role of sampling bias in driving real long-term outcomes,

[20] A potential concern with this analysis is that teams that are more active on Product Hunt simply don't launch a meaningful number of female-focused products. This could be because the makers are much more male or because being active on the platform causes makers to be less interested in female-focused ideas. No matter, strong imbalance would make it effectively impossible to compare the impact of the newsletter on female-focused newsletters across these groups. However, as shown in Appendix Figure A15, both types of teams are also equally likely to launch female-focused products.

our approach is not without limitations. Although the newsletter shock provides us with a source of exogenous variation and we cab show the shock impacts female user votes for female-focused products, we do not observe the multitude of channels—from simply logging in and using the product to giving feedback on social media—by which increased female engagement impacts outcomes. Second, while we have proxies for performance, we do not observe changes in the gender focus of the product itself. Do entrepreneurs reduce the female appeal of their products in response to the biased signals they receive? Finally, there is the question of generalizability. While Product Hunt is the most prominent platform for launching new digital products, it is only one platform. Thus, we do not know if our results generalize to other platforms like Kickstarter or other products like biomedical inventions (Greenberg and Mollick 2017; Koning, Samila, and Ferguson 2021). That said, given the over-representation of men throughout the process of invention and entrepreneurship, we think it is likely that our sampling bias mechanism is not merely an artifact of our "biased" sample.

Overall, our findings contribute to the growing body of work exploring entrepreneurial strategy and gender biases. On the startup strategy front, our findings imply that while user-focused, lean, and experimental strategies help founders quickly measure potential demand and pivot to the most promising ideas (Von Hippel 1986; Gans, Stern, and Wu 2019; Koning, Hasan, and Chatterji 2022; Camuffo et al. 2020), such methods have the potential to introduce bias into the venture growth process. If who an entrepreneur tests their ideas with is not representative of the larger market, then the signals they learn may be misleading of the idea's potential. Beyond gender, future work should explore what other dimensions early adopters are non-representative on and when such biases impact the direction of startup strategy and invention. For example, work on cultural markets shows that there likely exists unmet demand for racially diverse casts in movies and television shows (Kuppuswamy and Younkin 2019). Perhaps the under-representation of racial minorities in giving early feedback or in green lighting new movie ideas explains this racial-diversity gap.

Beyond illustrating a potential bias in experimental strategies, our findings also contribute to a related body of evidence that entrepreneurs are very much "boundedly rational." This work shows that much of the value of programs like accelerators is in overcoming both bias and noise in founder decision making (Cohen, Bingham, and Hallen 2019). Indeed, teaching entrepreneurs to be more "scientific" dramatically improves learning and startup performance, suggesting most entrepreneurs and innovators are far from the scientific frontier (Camuffo et al. 2020, 2021). However, our arguments and findings also suggest that there may be limits to such training. If the sample a "scientific" entrepreneur must learn from is simply tiny, then the evidence produced may never be strong enough to convince decision-makers to continue with the idea. Both arguments help explain why the post-launch growth gap persists for a year, impacts VC investment decisions, and shapes entrepreneurial effort. Indeed, prior work shows judge feedback and peer advice, both good and bad, can impact startup outcomes years later by shaping how and what entrepreneurs and investors learn (Howell 2017; Chatterji et al. 2019).

Our setting also suggests that the rise of online platforms like Product Hunt, Angel List, and Kickstarter might magnify and correlate these entrepreneurial biases and mistakes (e.g. Salganik, Dodds, and Watts 2006). These platforms tend to aggregate the opinions, preferences, and feedback of their users. In turn, these aggregated signals—be it interest on Product Hunt or funding raised on Kickstarter—are then widely broadcast to other stakeholders. If the influential users on these platforms are not representative of the larger market, then these platforms may well be sending "corrupted" signals of new venture potential. Similar to how centralized AI tools can amplify but also shed light on the biases present in training data, there is the potential for online platforms to both reinforce and shed light on existing entrepreneurial inequities (Cowgill and Tucker 2019; Obermeyer et al. 2019; Kleinberg et al. 2020; Koenecke et al. 2020).

Turning to research on gender and entrepreneurship, our first contribution is methodological. Here we show how to embed a product's position in the market onto an underlying gendered dimension ranging from male- to female-focused. While prior work has used word embeddings to understand patterns within cultural space (Srivastava et al. 2018; Kozlowski, Taddy, and Evans 2019), here we show that these text analysis tools can be extended to study economic outcomes like venture growth and investment. Beyond gender, our technique allows for the mapping of text onto a single dimension between any pair of opposing words which should allow researchers to study other sociodemographic differences (e.g. "rich" vs. "poor") along with more traditional differences in firm strategy (e.g. "flexibility" vs. "commitment"). Indeed, an emerging set of strategy and entrepreneurship papers are leveraging such text analysis tools to study everything from startup positioning to gender stereotypes in movie production (Guzman and Li 2019; Luo and Zhang 2021).

Our findings also enrich the emerging literature exploring how a lack of diversity leads to product-market bias (Koning, Samila, and Ferguson 2021). While prior work shows that the underrepresentation of female inventors leads to fewer female inventions (Feng and Jaravel 2019; Koning, Samila, and Ferguson 2020) and that a lack of female representation in medical trials lead to gendered inequities in health research (Michelman and Msall 2022; Gupta 2022), here we show that the diversity of key entrepreneurial gatekeepers also matters. If early gatekeepers—early adopters, VCs, buyers—tend to be men, then the signals entrepreneurs receive will distort the direction of innovation toward men. This suggests that the well-documented homogeneity of technology ecosystems like Silicon Valley might have consequences that go well beyond labor markets (Gompers and Wang 2017). Indeed, the dearth of women and African Americans—to name but two underrepresented demographic groups—might lead entrepreneurs in places like Silicon Valley to overlook the potential of serving female and black consumers. These potentially "lost startups" suggest that sampling bias might not just impact a single startup's chance of success but potentially the direction of startup innovation as well.

References

Bennett, Victor M and Aaron K Chatterji. 2019. "The entrepreneurial process: Evidence from a nationally representative survey." *Strategic Management Journal* .

Bojanowski, Piotr, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. "Enriching word vectors with subword information." *Transactions of the Association for Computational Linguistics* 5:135–146.

Boudreau, Kevin J. 2021. "Promoting Platform Takeoff and Self-Fulfilling Expectations: Field Experimental Evidence." *Management Science* 67 (9):5953–5967.

Camuffo, Arnaldo, Alessandro Cordova, Alfonso Gambardella, and Chiara Spina. 2020. "A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial." *Management Science* 66 (2):564–586.

Camuffo, Arnaldo, Alfonso Gambardella, Danilo Messinese, Elena Novelli, Emilio Paolucci, and Chiara Spina. 2021. "A Scientific Approach to Innovation Management: Evidence from Four Field Experiments." *Working Paper* .

Cao, Ruiqing. 2019. "Information frictions in new venture finance: Evidence from product hunt rankings." *Working Paper* .

Chatterji, Aaron, Solène Delecourt, Sharique Hasan, and Rembrand Koning. 2019. "When does advice impact startup performance?" *Strategic Management Journal* 40 (3):331–356.

Cohen, Susan L, Christopher B Bingham, and Benjamin L Hallen. 2019. "The role of accelerator designs in mitigating bounded rationality in new ventures." *Administrative Science Quarterly* 64 (4):810–854.

Cohen, Wesley M and Daniel A Levinthal. 1994. "Fortune favors the prepared firm." *Management science* 40 (2):227–251.

Cowgill, Bo and Catherine E Tucker. 2019. "Economics, fairness and algorithmic bias." *Working Paper* .

Eisenmann, Tom. 2021. *Why startups fail: A new roadmap for entrepreneurial success.* Currency.

Ewens, Michael and Richard R Townsend. 2020. "Are early stage investors biased against women?" *Journal of Financial Economics* 135 (3):653–677.

Feng, Josh and Xavier Jaravel. 2019. "Innovating for People Like Me: Evidence from Female-Founded Consumer Packaged Goods Startups." *Working Paper* .

Fitzpatrick, Anne and Solène Delecourt. 2020. "Childcare Matters: Female Business Owners and the Baby Profit Gap." *Management Science (Forthcoming)* .

Gans, Joshua S, Scott Stern, and Jane Wu. 2019. "Foundations of entrepreneurial strategy." *Strategic Management Journal* 40 (5):736–756.

Gompers, Paul A and Sophie Q Wang. 2017. "And the children shall lead: Gender diversity and performance in venture capital." Tech. rep., National Bureau of Economic Research.

Greenberg, Jason and Ethan Mollick. 2017. "Activist choice homophily and the crowdfunding of female founders." *Administrative Science Quarterly* 62 (2):341–374.

Gupta, Harsh. 2022. "Do Female Researchers Increase Female Enrollment in Clinical Trials?" *Working Paper* .

Guzman, Jorge and Aleksandra Olenka Kacperczyk. 2019. "Gender gap in entrepreneurship." *Research Policy* 48 (7):1666–1680.

Guzman, Jorge and Aishen Li. 2019. "Measuring Founding Strategy." *Working Paper* .

Hasan, Sharique, John-Paul Ferguson, and Rembrand Koning. 2015. "The lives and deaths of jobs: Technical interdependence and survival in a job structure." *Organization Science* 26 (6):1665–1681.

Hayek, Friedrich August. 1948. *Individualism and economic order*. University of chicago Press.

Hebert, Camille. 2018. "Mind the gap: gender stereotypes and entrepreneur financing." *Working Paper* 3318245.

Hoberg, Gerard and Gordon Phillips. 2016. "Text-based network industries and endogenous product differentiation." *Journal of Political Economy* 124 (5):1423–1465.

Howell, Sabrina T. 2017. "Learning from feedback: Evidence from new ventures." *Working Paper* .

Howell, Sabrina T and Ramana Nanda. 2019. "Networking frictions in venture capital, and the gender gap in entrepreneurship." Tech. rep., National Bureau of Economic Research.

Kaplan, Steven N and Josh Lerner. 2016. "Venture capital data: Opportunities and challenges." *Measuring entrepreneurial businesses: current knowledge and challenges* :413–431.

Kerr, William R, Ramana Nanda, and Matthew Rhodes-Kropf. 2014. "Entrepreneurship as experimentation." *Journal of Economic Perspectives* 28 (3):25–48.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Cass R Sunstein. 2020. "Algorithms as discrimination detectors." *Proceedings of the National Academy of Sciences* 117 (48):30096–30100.

Koenecke, Allison, Andrew Nam, Emily Lake, Joe Nudell, Minnie Quartey, Zion Mengesha, Connor Toups, John R Rickford, Dan Jurafsky, and Sharad Goel. 2020. "Racial disparities in automated speech recognition." *Proceedings of the National Academy of Sciences* .

Koning, Rembrand, Sharique Hasan, and Aaron Chatterji. 2022. "Experimentation and Startup Performance: Evidence from A/B testing." *Management Science* .

Koning, Rembrand and Jacob Model. 2013. "Experimental study of crowdfunding cascades: When nothing is better than something." *Available at SSRN 2308161* .

Koning, Rembrand, Sampsa Samila, and John-Paul Ferguson. 2020. "Inventor Gender and the Direction of Invention." *American Economic Association Papers & Proceedings* .

———. 2021. "Who do we invent for? Patents by women focus more on women's health, but few women get to invent." *Science)* .

Kozlowski, Austin C, Matt Taddy, and James A Evans. 2019. "The geometry of culture: Analyzing the meanings of class through word embeddings." *American Sociological Review* 84 (5):905–949.

Kuppuswamy, Venkat and Peter Younkin. 2019. "Testing the Theory of Consumer Discrimination as an Explanation for the Lack of Minority Hiring in Hollywood Films." *Management Science* .

Levinthal, Daniel A. 2017. "Mendel in the C-Suite: Design and the Evolution of Strategies." *Strategy Science* 2 (4):282–287.

Luo, Hong and Laurina Zhang. 2021. "Gender Orientation and Segregation of Ideas:# Metoo's Impact in Hollywood." *Working Paper* .

Michelman, Valerie and Lucy Msall. 2022. "Sex, Drugs, and RD: Missing Innovation from Regulating Female Enrollment in Clinical Trials." *Working Paper* .

Mollick, Ethan. 2014. "The dynamics of crowdfunding: An exploratory study." *Journal of business venturing* 29 (1):1–16.

Moore, Geoffrey. 1991. *Crossing the Chasm.* Harper Business.

Moretti, Enrico. 2011. "Social learning and peer effects in consumption: Evidence from movie sales." *The Review of Economic Studies* 78 (1):356–393.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting racial bias in an algorithm used to manage the health of populations." *Science* 366 (6464):447–453.

Ries, Eric. 2011. *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses.* Crown Books.

Roche, Maria, Alexander Oettl, and Christian Catalini. 2020. "Entrepreneurs (co-) Working in Close Proximity: Impacts on Technology Adoption and Startup Performance Outcomes." .

Salganik, Matthew J, Peter Sheridan Dodds, and Duncan J Watts. 2006. "Experimental study of inequality and unpredictability in an artificial cultural market." *Science* 311 (5762):854–856.

Scott, Erin L and Pian Shu. 2017. "Gender gap in high-growth ventures: Evidence from a university venture mentoring program." *American Economic Review* 107 (5):308–11.

Srivastava, Sameer B, Amir Goldberg, V Govind Manian, and Christopher Potts. 2018. "Enculturation trajectories: Language, cultural adaptation, and individual outcomes in organizations." *Management Science* 64 (3):1348–1364.

Starr, Evan and Brent Goldfarb. 2020. "Binned scatterplots: A simple tool to make research easier and better." *Strategic Management Journal* 41 (12):2261–2274.

Tak, Elise, Shelley J Correll, and Sarah A Soule. 2019. "Gender inequality in product markets: When and how status beliefs transfer to products." *Social Forces* 98 (2):548–577.

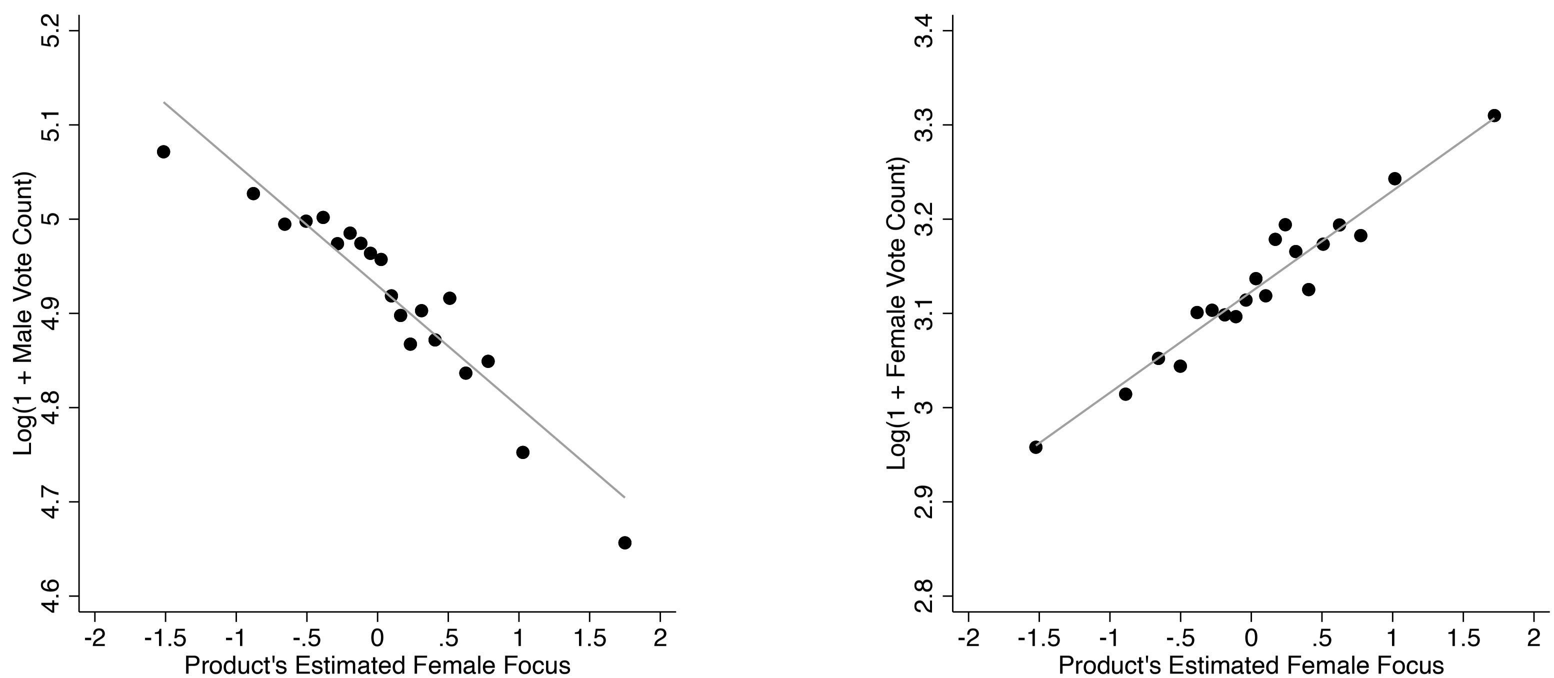
Truffa, Francesca and Ashley Wong. 2022. "Undergraduate Gender Diversity and Direction of Scientific Research." *Working Paper* .

Von Hippel, Eric. 1986. "Lead users: a source of novel product concepts." *Management Science* 32 (7):791–805.

Figure 1: Binned scatter plots showing that as a product's estimated female-focus increases the number of votes from male users decreases (a) while the the number of votes from female users increases (b).

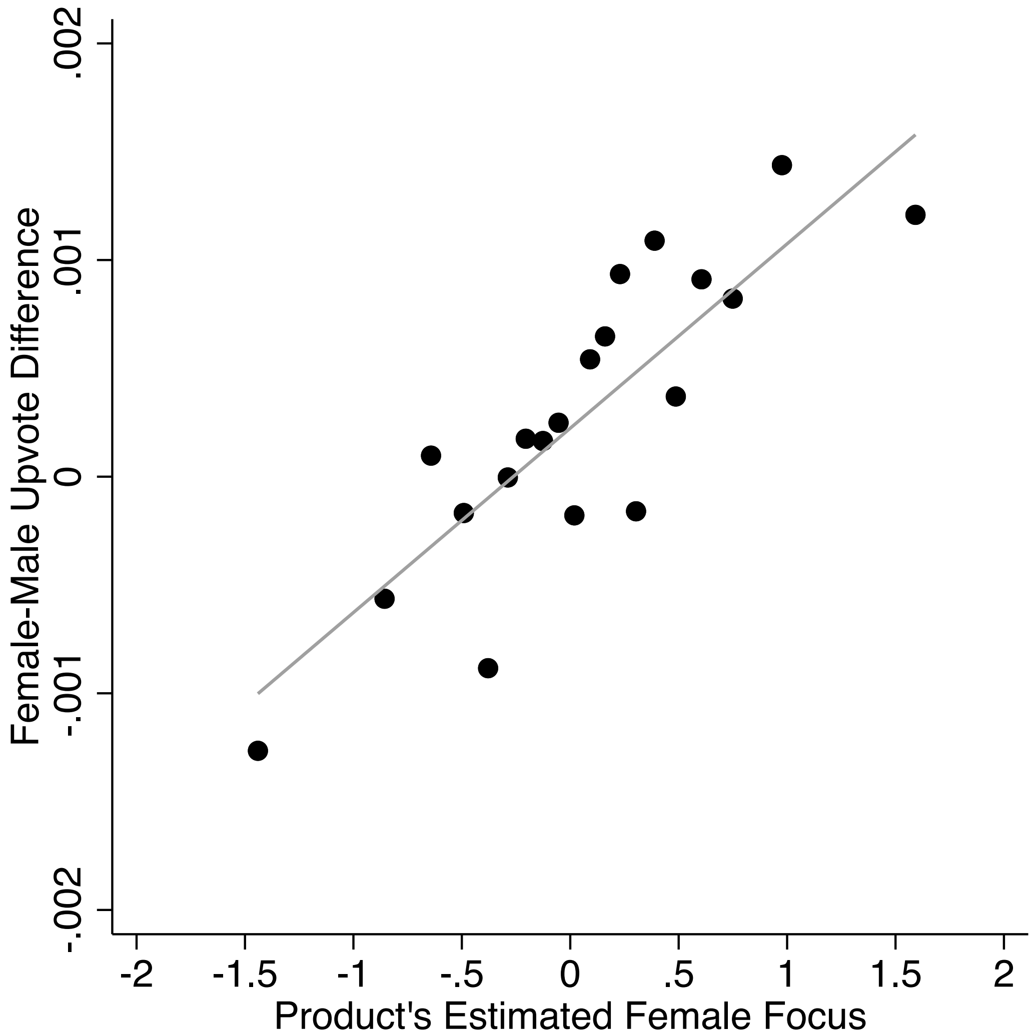
(a) Number of votes from male users

(b) Number of votes from female users



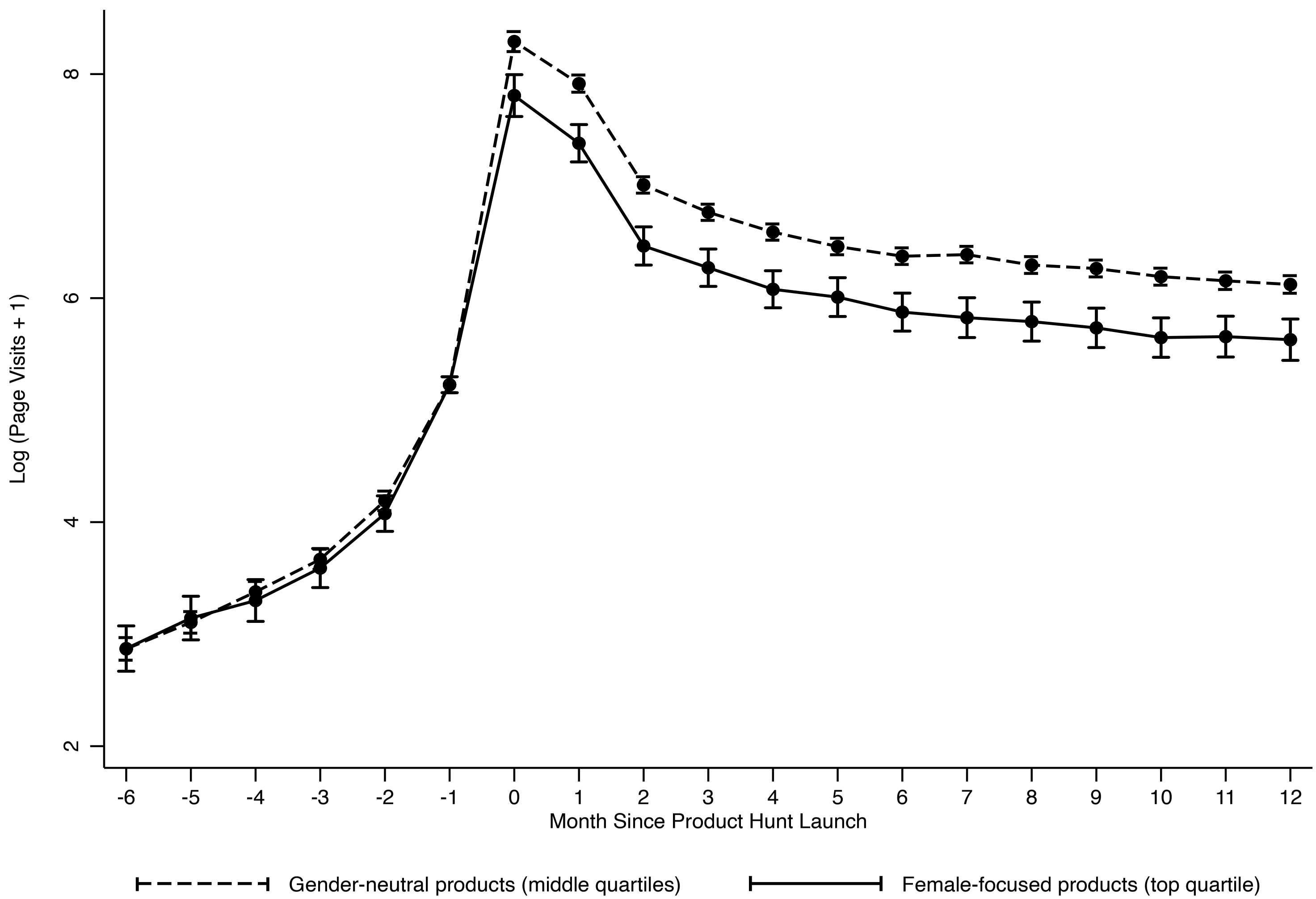
Notes: Binned scatter plots showing the logged number of votes (+1) against the product's estimated female focus. Panel (a) shows logged number of votes from *male* users after controlling for the logged number of votes from female users. Panel (b) shows logged number of votes from *female* users after controlling for the logged number of votes from male users. The model includes the 5,742 products that are included in our newsletter shock analysis.

Figure 2: Binned scatterplot showing that products we estimate as female focused—i.e., more likely to appeal to the needs and preferences of women—are more likely to be preferred ("upvoted") by female users.



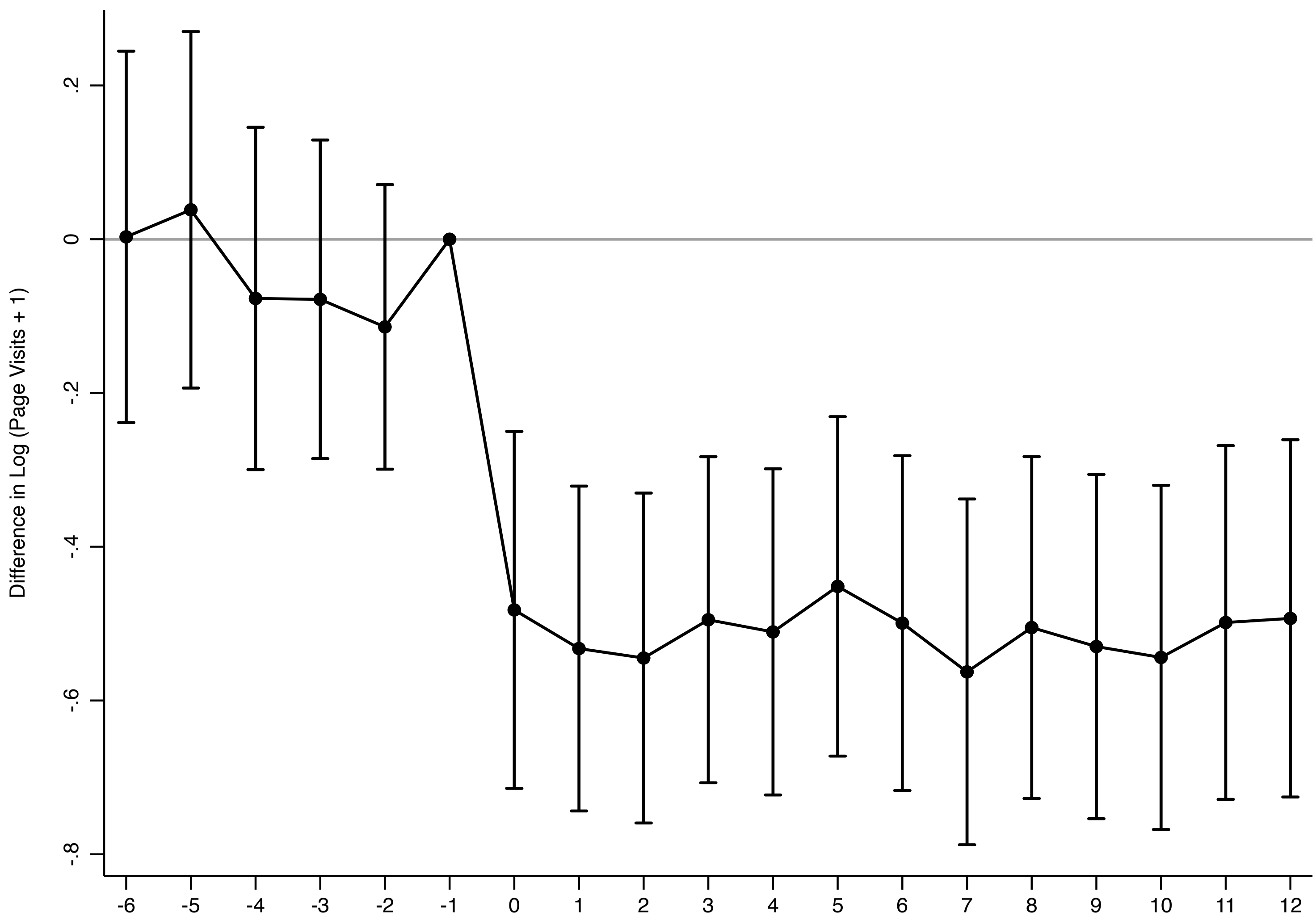
Notes: The Y-axis represents the difference in upvoting behavior between active female and male users who have viewed a product after accounting for voter and product fixed effects. The X-axis is our text-based estimate of the degree to which the product focuses on female users. The binscatter accounts for user and product fixed effects. The model includes 11,212 products launched on weekdays between January 2017 and June 2018, for which the proprietary browsing data on product views are available. See Appendix E for further details.

Figure 3: Female-focused products (top quartile) have a similar growth trajectory to gender-neutral products (2nd and 3rd quartiles) before launching on Product Hunt but experience roughly 45% less user growth after launching on the platform.



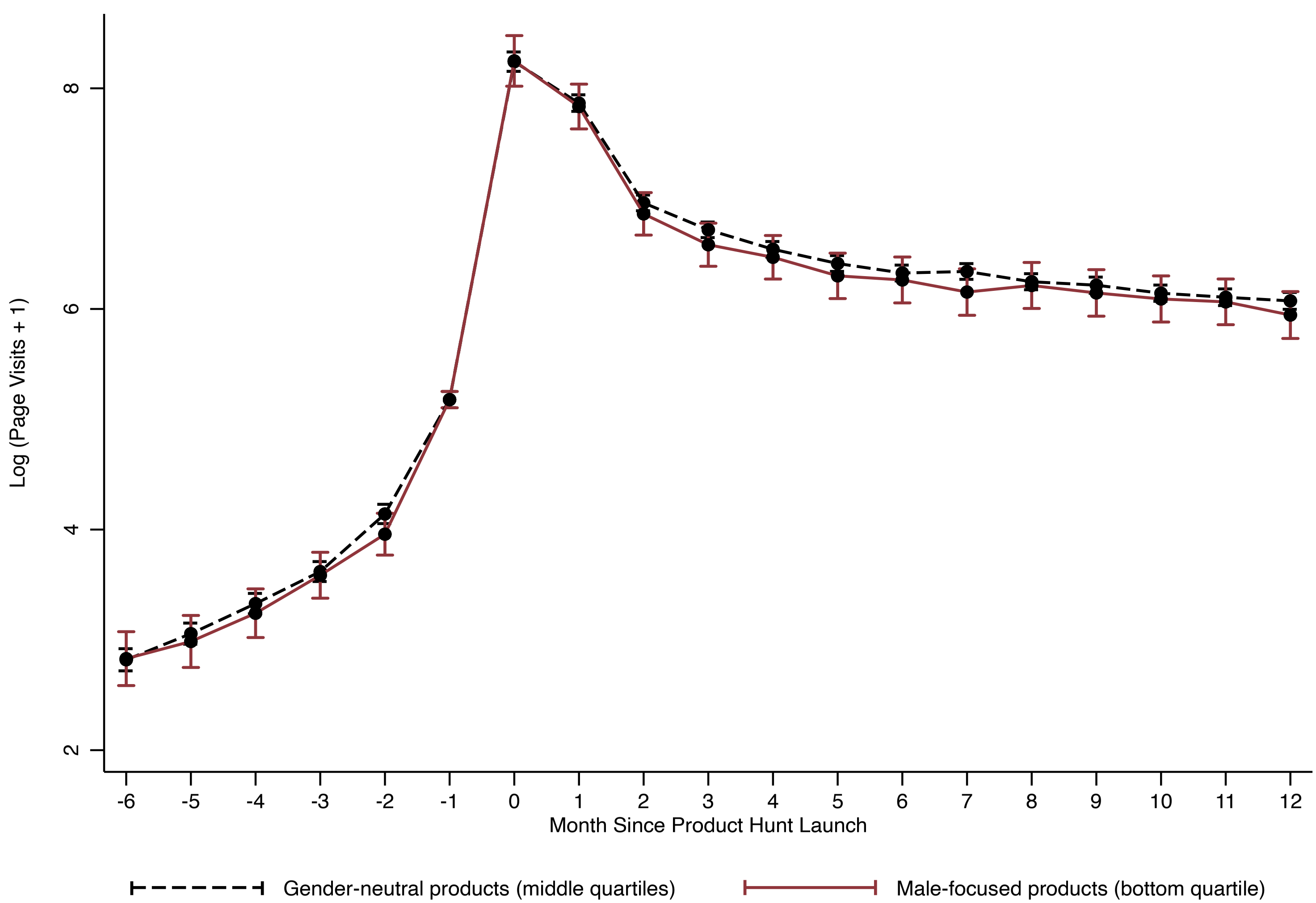
Notes: The plotted estimates and 95% confidence intervals are from an event-study version of Model 3 from Panel B of Table 3. The month before launch serves as the excluded baseline. The model controls for product fixed effects and year-month fixed effects. Standard errors are clustered at the product level. The model includes 5,742 products and 101,803 month-product observations.

Figure 4: Difference in growth trajectories for female-focused products (top quartile) compared to gender-neutral products (2nd and 3rd quartiles) before and after launching on Product Hunt.



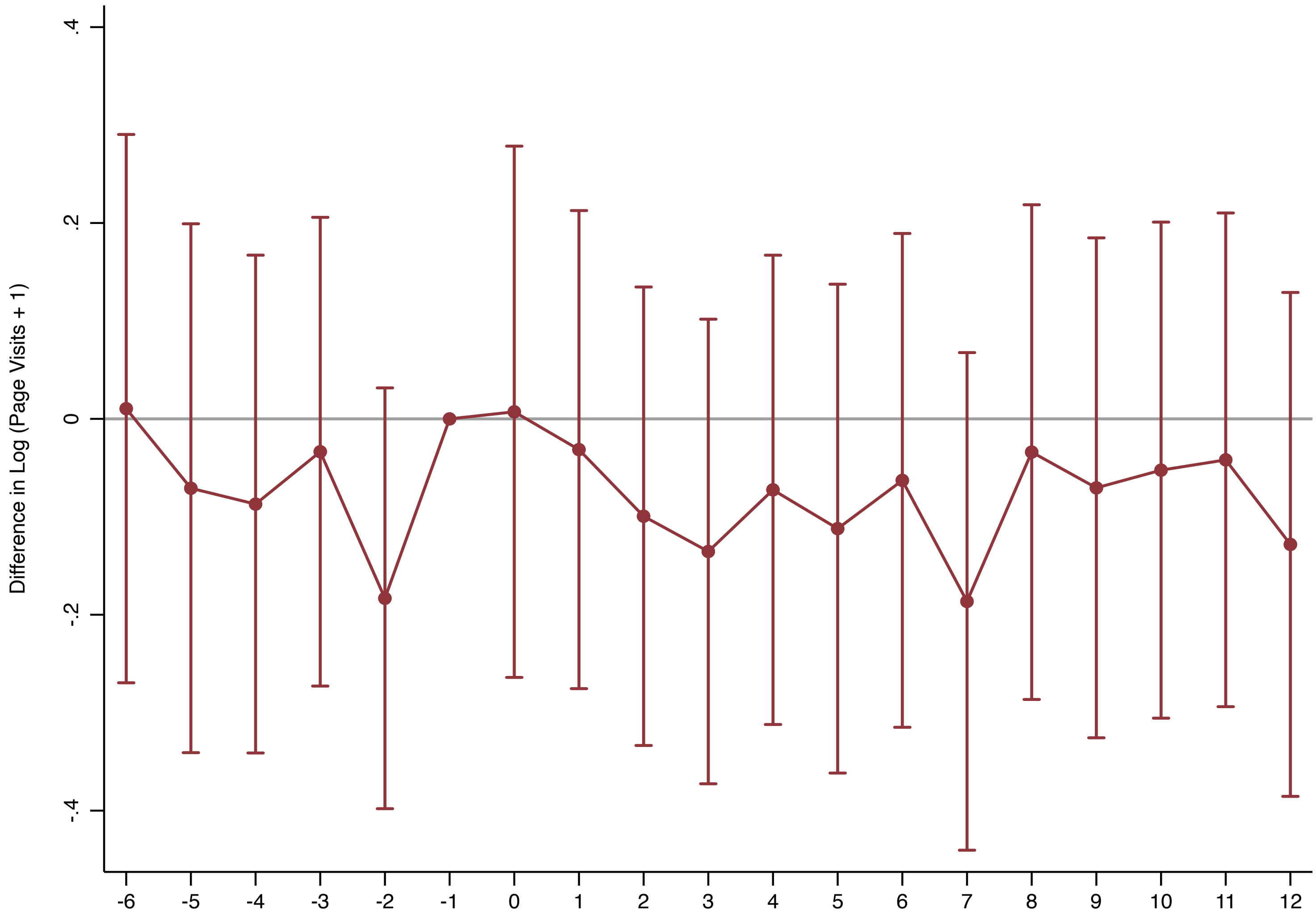
Notes: The plotted estimates and 95% confidence intervals are from an event-study version of Model 3 from Panel B of Table 3. The month before launch serves as the excluded baseline. The model controls for product fixed effects and year-month fixed effects. Standard errors are clustered at the product level. The model includes 5,742 products and 101,803 month-product observations.

Figure 5: Male-focused products (bottom quartile) have a similar growth trajectory to gender-neutral products (2nd and 3rd quartiles) before and after launching on Product Hunt.



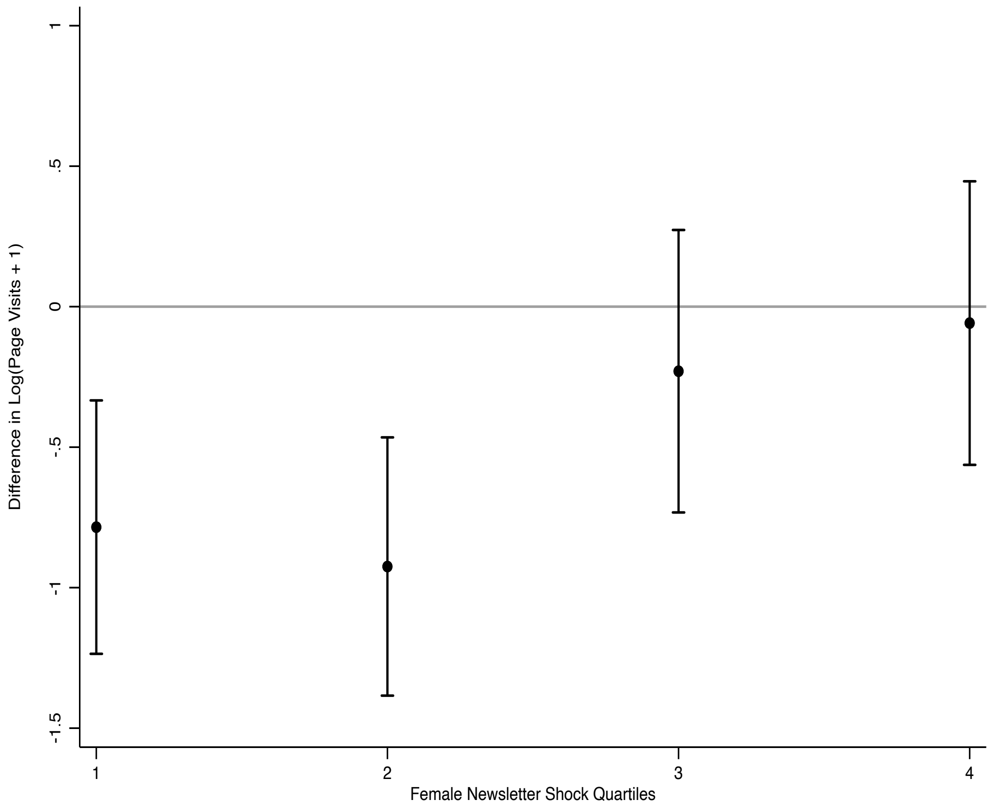
Notes: The plotted estimates and 95% confidence intervals are from an event-study version of Model 3 from Panel B of Table 3. The month before launch serves as the excluded baseline. The model controls for product fixed effects and year-month fixed effects. Standard errors are clustered at the product level. The model includes 5,742 products and 101,803 month-product observations.

Figure 6: Difference in growth trajectories for male-focused products (bottom quartile) compared to gender-neutral products (2nd and 3rd quartiles) before and after launching on Product Hunt.



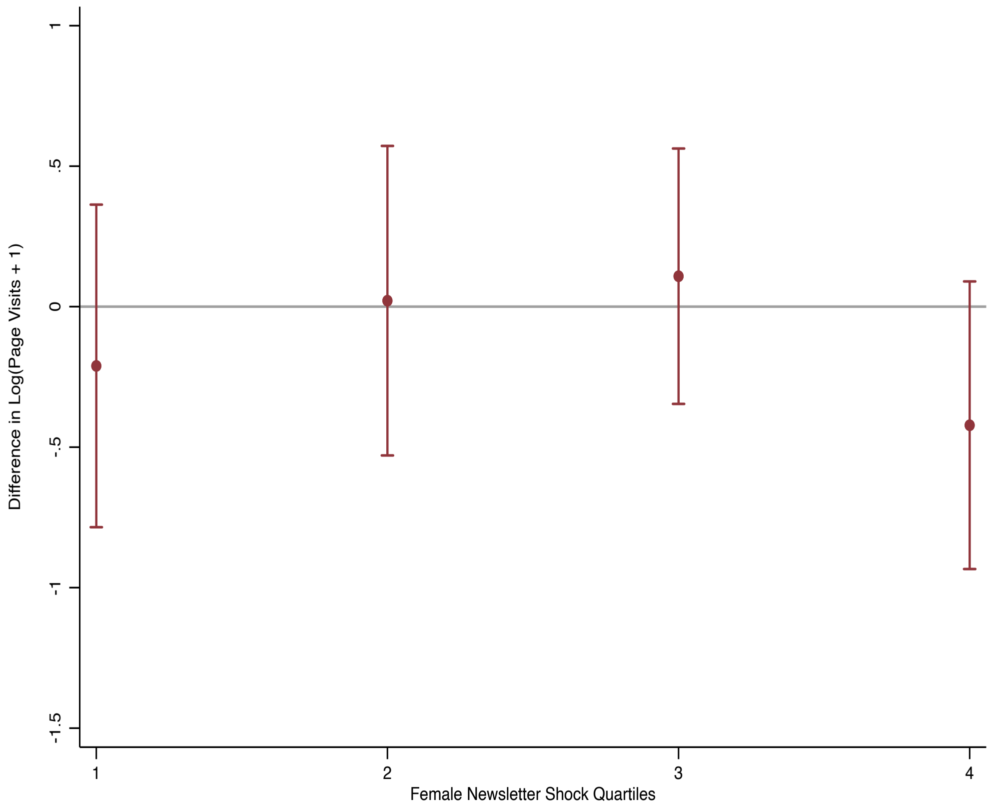
Notes: The plotted estimates and 95% confidence intervals are from an event-study version of Model 3 from Panel B of Table 3. The month before launch serves as the excluded baseline. The model controls for product fixed effects and year-month fixed effects. Standard errors are clustered at the product level. The model includes 5,742 products and 101,803 month-product observations.

Figure 7: The difference in user growth one year after launch between female-focused products (top quartile) and gender-neutral products (middle quartiles) shrinks towards zero when the newsletter is unexpectedly more female focused.



Notes: The estimates and 95% confidence intervals are from a discretized version of model 1 in Table 6 where the "newsletter shock" variable bucketed into quartiles and the product's female focus is bucketed into the top and bottom quartiles. The model includes fixed effects for number of newsletter-suggested products, products, and year-months. Standard errors are clustered at the launch day level. The model includes 5,742 products and 101,803 month-product observations.

Figure 8: There is little change between male-focused products (bottom quartile) and gender-neutral products (middle quartiles) when the newsletter is unexpectedly more female focused. If anything, male-focused products do slightly worse when the newsletter is especially female focused.



Notes: The estimates and 95% confidence intervals are from a discretized version of model 1 in Table 6 where the "newsletter shock" variable bucketed into quartiles and the product's female focus is bucketed into the top and bottom quartiles. The model includes fixed effects for number of newsletter-suggested products, products, and year-months. Standard errors are clustered at the launch day level. The model includes 5,742 products and 101,803 month-product observations.

Table 1: Examples of products by female focus quantile

%tile	Product Name	Tagline	Female Focus
	ThxBro	Generate deliciously random, jargon-laced e-mails	-5.109
P1	Ballmetric	Your favorite plays from the NBA	-3.711
	Beard Bib 2.0	Hair clippings catcher from Beard King	-3.653
	SPECTRA	The most portable electric skateboard	-1.305
P5	Segway Drift W1	The first self balancing e-skate	-1.294
	Keyport Slide 3.0 & Pivot	The swiss army knife of the future	-1.293
	SnapHunt	Product Hunt for Snapchat. Discover new people to follow	-0.903
P10	Nikola	See your Tesla's battery percentage from your menubar	-0.900
	Hackuna	Secure yourself from all kinds of hackers	-0.898
	Morph - PokemonGo Bot	Chatbot to find and report Pokemon around you	-0.404
P25	Phish.AI	Anti-phishing platform powered by AI & Computer Vision	-0.401
	Sqreen API	A security toolbox for developers	-0.401
	Cemtrex Smartdesk	The world's most advanced workstation	0.027
P50	Yomu	One place to read your favorite content from around the web	0.027
	Adzoola	Hyper-targeted advertising and outreach	0.028
	Borsch	The AI app that helps you discover the yummiest dishes	0.495
P75	Cuddle Mattress	Hug your better half without the arm numbing	0.495
	Joonko	Personal diversity and inclusion AI-coach for managers	0.500
	The Silver Post	Do more for grandma or grandpa	0.999
P90	Kindred	Friends for when you travel	1.001
	Ropazi	Personal shopper for busy parents	1.001
	Artwxrk	Curated collection of the world's best contemporary art	1.451
P95	Chairman Mom	A social, Q&A platform for working moms	1.457
	VINA	Connecting awesome women for fun, for work, for life	1.474
	Babee on Board	Pregnant? Request a seat on public transport	5.104
P99	Flo Health	The #1 app for women's menstrual health	5.237
	Wonder	An app for queer & lesbian women to express their uniqueness	6.159

Notes: Table shows examples of products at various points of the distribution of the product's estimated female focus. For each product example, the table includes its name, tagline (short description), and normalized score. Product examples are drawn from the 1th, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 99th percentiles of the distribution of estimated female focus, ranging from the most male-focused (lowest score) to the most female-focused (highest score).

Table 2: Descriptive statistics for the 5,742 products in our sample

	Product Launches Sample, Sep 2016 - Oct 2018				
	All Products (N = 5,742)		**Female Products (Top Quartile)**	**Male Products (Bottom Quartile)**	**T-Test Male - Female**
	Mean	SD	Mean	Mean	Difference
Product Characteristics					
Product Female Focus	0.07	0.72	1.03	-0.99	-2.02***
Log (1 + Monthly Page Visits) 1 Month Before	5.23	4.10	5.43	5.25	-0.18
Pre-Launch Seed or Series A Funding	0.028	0.165	0.026	0.025	-0.001
Topic Category Top #1: Productivity	0.29	0.45	0.24	0.26	0.03
Topic Category Top #2: Developer	0.14	0.35	0.08	0.21	0.13***
Topic Category Top #3: Design	0.10	0.30	0.08	0.08	0.00
Topic Category Top #4: Marketing	0.10	0.30	0.10	0.05	-0.05***
Topic Category Top #5: Artificial Intelligence	0.07	0.26	0.06	0.06	0.00
Topic Category Top #6: User Experience	0.05	0.23	0.04	0.04	0.00
User Statistics					
Hunter is Female	0.10	0.30	0.12	0.07	-0.05***
Maker Team Size	2.02	1.61	1.90	1.88	-0.02
Makers At Least 1 Female	0.19	0.40	0.25	0.16	-0.09***
Active User Votes Female Share	0.14	0.06	0.16	0.12	-0.03***

Notes: Descriptive statistics for the sample of 5,742 products we use in our product lunch and newsletter shock analysis. These product launches take place on the Product Hunt platform between October 4th 2016 and October 19th 2018. The sample includes featured products launched on weekdays and submitted before 7AM Pacific Time on these days – the earliest time of the day at which a newsletter could reach a user's email inbox. The sample only includes products launched on the 347 days on which the newsletter features standard product-list content. The left panel reports summary statistics for the entire sample. The rest of the table reports the same statistics for female-focused (top quartile) products and male-focused (bottom quartile) products, and the differences in means and significance level from two sample T tests on these two groups of products. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Estimated effects of a product's female focus on growth after launching on Product Hunt.

(a) Continuous

	Log (1 + Monthly Page Visits)				
	(1)	(2)	(3)	(4)	(5)
Post-Launch	3.256***	3.270***	3.393***	3.399***	3.472***
	(0.047)	(0.047)	(0.053)	(0.053)	(0.055)
Post-Launch x Female Focus		-0.208***		-0.186***	-0.243***
		(0.054)		(0.054)	(0.070)
Post-Launch x Female Maker			-0.382***	-0.358***	
			(0.102)	(0.102)	
Product FE	Y	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y	Y
Sample	All	All	All	All	Male Makers
# Products	5,742	5,742	5,742	5,742	4,081
Observations	101,803	101,803	101,803	101,803	72,401
R-Squared	0.721	0.721	0.721	0.721	0.709

(b) Quartiles

	Log (1 + Monthly Page Visits)				
	(1)	(2)	(3)	(4)	(5)
Post-Launch	3.256***	3.357***	3.393***	3.465***	3.559***
	(0.047)	(0.056)	(0.053)	(0.060)	(0.064)
Post-Launch x Female Product (Top Quartile)		-0.462***		-0.394***	-0.507***
		(0.098)		(0.098)	(0.120)
Post-Launch x Male Product (Bottom Quartile)		-0.030		-0.002	-0.023
		(0.111)		(0.111)	(0.132)
Post-Launch x Female Maker			-0.382***	-0.356***	
			(0.102)	(0.102)	
Product FE	Y	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y	Y
Sample	All	All	All	All	Male Makers
# Products	5,742	5,742	5,742	5,742	4,081
Observations	101,803	101,803	101,803	101,803	72,401
R-Squared	0.721	0.721	0.721	0.722	0.709

Notes: Estimates from a difference-in-differences model on the sample of 5,742 entrepreneurial product launches. The outcome variable is the launching startup's log monthly website visits. The treatment is whether the product has launched and each of the 101,803 observations corresponds to a product-year-month from 6 months before launch to 12 months after launch. Panel (a) estimates the differential effects by interacting the post-launch dummy with a continuous version of the product's estimated female focus. Panel (b) estimates the differential effects by interacting the post-launch dummy with a top and bottom quartile indicator for the product's estimated female focus. In both panels, column 1 shows the baseline estimates of the effects of the product launch. Column 2 estimates the model after adding the interactions with our estimated female focus measure. Column 3 estimates the model after adding the interaction between the post-launch dummy variable and an indicator of whether at least one maker is female. Column 4 estimates the model with both interaction terms in columns 2 and 3. Column 5 restricts the sample to products launched by all-male makers. All models are estimated in panel regressions with product fixed effects and year-month fixed effects. Standard errors are clustered at the product level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Estimated effects of a product's female focus on whether the startup has an active user base after launching on Product Hunt.

(a) Continuous

	Has Active User Base				
	(1)	(2)	(3)	(4)	(5)
Post-Launch	0.380***	0.381***	0.404***	0.404***	0.411***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Post-Launch x Female Focus		-0.023***		-0.019***	-0.026***
		(0.007)		(0.007)	(0.009)
Post-Launch x Female Maker			-0.074***	-0.071***	
			(0.013)	(0.013)	
Product FE	Y	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y	Y
Sample	All	All	All	All	Male Makers
# Products	5,742	5,742	5,742	5,742	4,081
Observations	101,803	101,803	101,803	101,803	72,401
R-Squared	0.531	0.532	0.533	0.533	0.531

(b) Quartiles

	Has Active User Base				
	(1)	(2)	(3)	(4)	(5)
Post-Launch	0.380***	0.391***	0.404***	0.411***	0.420***
	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)
Post-Launch x Female Product (Top Quartile)		-0.050***		-0.039***	-0.053***
		(0.012)		(0.012)	(0.015)
Post-Launch x Male Product (Bottom Quartile)		-0.005		-0.001	-0.005
		(0.014)		(0.014)	(0.016)
Post-Launch x Female Maker			-0.074***	-0.071***	
			(0.013)	(0.013)	
Product FE	Y	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y	Y
Sample	All	All	All	All	Male Makers
# Products	5,742	5,742	5,742	5,742	4,081
Observations	101,803	101,803	101,803	101,803	72,401
R-Squared	0.531	0.532	0.533	0.533	0.531

Notes: Estimates from a difference-in-differences model on the sample of 5,742 entrepreneurial product launches. The outcome variable measures whether the firm still has more than zero visitors. The treatment is whether the product has launched and each of the 101,803 observations corresponds to a product-year-month from 6 months before launch to 12 months after launch. Panel (a) estimates the differential effects by interacting the post-launch dummy with a continuous version of the product's estimated female focus. Panel (b) estimates the differential effects by interacting the post-launch dummy with a top and bottom quartile indicator for the product's estimated female focus. In both panels, column 1 shows the baseline estimates of the effects of the product launch. Column 2 estimates the model after adding the interactions with our estimated female focus measure. Column 3 estimates the model after adding the interaction between the post-launch dummy variable and an indicator of whether at least one maker is female. Column 4 estimates the model with both interaction terms in columns 2 and 3. Column 5 restricts the sample to products launched by all-male makers. All models are estimated in panel regressions with product fixed effects and year-month fixed effects. Standard errors are clustered at the product level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Daily descriptive statistics suggest that there is no difference between products launched when a newsletter is more or less female focused.

	Product Hunt Daily Newsletters				
	All Days (N = 419)		**Female Newsletter (Top Quartile)**	**Male Newsletter (Bottom Quartile)**	**T-Test Bottom - Top**
	Mean	SD	Mean	Mean	Difference
Female Newsletter Shock	0.52	0.10	0.66	0.42	-0.24***
Product Female Focus	0.07	0.72	0.07	0.07	0.00
Log (1 + Monthly Page Visits) 1 Month Before	5.23	4.10	5.07	5.21	0.14
Pre-Launch Seed or Series A Funding	0.028	0.165	0.029	0.029	0.001
Hunter is Female	0.10	0.30	0.10	0.10	0.01
Maker Team Size	2.02	1.61	2.00	2.01	0.01
Makers At Least 1 Female	0.19	0.40	0.19	0.20	0.01

Notes: Descriptive statistics by newsletter content for the 347 days when a standard product-list newsletter was sent out by Product Hunt. The left panel reports overall summary statistics. The rest of the table reports the same statistics for top quartile days (the most female-focused newsletters) and bottom quartile days (The last female-focused newsletters), and the differences in means and significance level from two sample T tests between products launched on these days. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Estimated effect of the female-newsletter shock by the product's female focus on visit growth after launching on Product Hunt.

	Log (1 + Monthly Page Visits)	
	(1)	(2)
Post-Launch	2.973***	3.435***
	(0.219)	(0.218)
Post-Launch x Newsletter Shock	0.571	0.071
	(0.422)	(0.419)
Post-Launch x Female Focus	-0.878***	-0.918***
	(0.204)	(0.331)
Post-Launch x Newsletter Shock x Female Focus	1.280***	1.303**
	(0.367)	(0.619)
Product FE & Year-Month FE	Y	Y
Sample	All	Male Makers
# Products	5,742	4,081
Observations	101,803	72,401
R-Squared	0.721	0.709

Notes: Estimates from a difference-in-difference-in-differences model on the sample of 5,742 entrepreneurial product launches. The outcome variable is the launching startup's log monthly website visits. The treatment is whether the product has launched and each of the 101,803 observations corresponds to a product-year-month from 6 months before launch to 12 months after launch. The female newsletter shock is measured as the maximum female focus (after rescaling to between 0 and 1) of all suggested products mentioned in the daily newsletter. Column 1 shows coefficient estimates on the main model on our full sample. Column 2 shows coefficient estimates after restricting the sample to all-male made products. All models control for product fixed effects and year-month fixed effects. Standard errors are clustered at the launch day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Estimated effect of the female-newsletter shock by the product's female focus on whether the startup has an active user base after launching on Product Hunt.

	Has Active User Base	
	(1)	(2)
Post-Launch	0.334***	0.389***
	(0.026)	(0.029)
Post-Launch x Newsletter Shock	0.091*	0.042
	(0.050)	(0.055)
Post-Launch x Female Focus	-0.097***	-0.106***
	(0.026)	(0.040)
Post-Launch x Newsletter Shock x Female Focus	0.141***	0.153**
	(0.046)	(0.075)
Product FE & Year-Month FE	Y	Y
Sample	All	Male Makers
# Products	5,742	4,081
Observations	101,803	72,401
R-Squared	0.532	0.531

Notes: Estimates from a difference-in-difference-in-differences model on the sample of 5,742 entrepreneurial product launches. The outcome variable measures whether the firm still has more than zero visitors. The treatment is whether the product has launched and each of the 101,803 observations corresponds to a product-year-month from 6 months before launch to 12 months after launch. The female newsletter shock is measured as the maximum female focus (after rescaling to between 0 and 1) of all suggested products mentioned in the daily newsletter. Column 1 shows coefficient estimates on the main model on our full sample. Column 2 shows coefficient estimates after restricting the sample to all-male made products. All models control for product fixed effects and year-month fixed effects. Standard errors are clustered at the launch day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Estimated effect of the female-newsletter shock by the product's female focus on whether post-launch the team raises venture funding as of October 2020.

	Raises Funding Post-Launch	
	(1)	(2)
Log (1 + Monthly Page Visits) 1 Month Before	0.003***	0.002***
	(0.001)	(0.001)
Pre-Launch Seed or Series A Funding	0.506***	0.540***
	(0.038)	(0.048)
Newsletter Shock	0.025	0.033
	(0.024)	(0.025)
Female Focus	-0.046**	-0.033**
	(0.018)	(0.015)
Newsletter Shock x Female Focus	0.081**	0.051*
	(0.036)	(0.028)
Year-Month FE	Y	Y
Sample	All	Male Makers
Observations	5,742	4,081
R-Squared	0.232	0.245

Notes: Estimates from a linear probability model using our sample of 5,742 entrepreneurial product launches. The outcome variable measures whether the firm raised venture funding between when it launched on ProductHunt and October 2020. All models include fixed effects for the the month-year of launch and the number of products linked to in the newsletter. Standard errors are clustered at the launch day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Estimated effect of the female-newsletter shock by the product's female focus on the product team's web technology investments.

	Log (1 + Technology Stack) Post-Launch	
	(1)	(2)
Log (1 + Monthly Page Visits in Month Before Launch)	0.105***	0.100***
	(0.004)	(0.005)
Pre-Launch Seed or Series A Funding	0.636***	0.755***
	(0.091)	(0.099)
Newsletter Shock	-0.097	-0.297
	(0.181)	(0.203)
Female Focus	-0.282**	-0.301**
	(0.115)	(0.148)
Newsletter Shock x Female Focus	0.567***	0.635**
	(0.213)	(0.280)
Year-Month FE	Y	Y
Sample	All	Male Makers
Observations	5,312	3,748
R-Squared	0.134	0.131

Notes: Estimates from an OLS model using our sample of 5,312 entrepreneurial product launches for which we have technology stack data. The outcome variable measures the (logged) number of active web technologies on the startup's website as of October 2020. All models include fixed effects for the the month-year of launch and the number of products linked to in the newsletter. Standard errors are clustered at the launch day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix

Sampling Bias in Entrepreneurial Experiments

(A) Data sources
(B) Categorizing users by gender
(C) The value of page visits and technology development
(D) Building and validating a measure of a product's gender focus
(E) Using the viewing data to estimate preferences
(F) Validating the newsletter's exogenous impact on female engagement
(G) Robustness checks ruling out alternative newsletter mechanisms
(H) A simple Bayesian example of entrepreneurial sampling
(I) Additional funding and technology stack results
(J) Testing the comments channel
(K) Testing for effect heterogeneity

A Data sources

Product Hunt API. Product Hunt makes its API available to developers, and we obtain all the public data of information displayed on the Product Hunt platform through crawling its Developer API. Product Hunt API Data includes the voting history of each user which covers every product they have upvoted and when. The data also includes all information from the public-facing product profile which includes the name of the product, a catchy tagline that describes the product in brief, as well as media information such as screenshots and marketing videos. Each product is submitted to the platform by a "hunter", often a highly active member on the platform, and in 40% of the cases the Hunter is the lead member of the maker team. Maker (i.e. founder) information is included as well. To engage with the community, the makers often post more information about the product in the comment section to attract attention and feedback to the product from the community. The data also contains user-level information: since users usually register using their real names, we infer the gender of the user to the best extent we can, using their names and Twitter account names to improve the prediction. Data availability is for the entire platform: December 2013 to present.

Product Hunt Proprietary Browsing Data. We augment the public API data using proprietary data on browsing history of users. [21] For each visit to the platform, it is recorded in one of three data sets: (1) homepage viewing (2) viewing of any domain that is not the homepage, and (3) an upvote that was originated from a view event whether it occurred on the homepage or via some other link. Each visit is recorded with a "received at" time stamp, as well as the the URL path of the page that was visited. Data availability is from January 2017 to June 2019. Missing about two months of data from June 2017 to August 2017, because of broken analytics library.

SimilarWeb. We obtain monthly website traffic data from SimilarWeb. SimilarWeb is a market intelligence platform that estimates website and app growth metrics. Using data from a global panel of web browsers, SimilarWeb provides website performance metrics including page views over the last three years at the weekly level. SimilarWeb is used by tech firms for lead generation, to track acquisition targets, and to benchmark performance. We use the SimilarWeb API to pull down weekly website performance metrics for the companies with their products launched on Product Hunt and their website linked to their PH profile in our sample. Data availability is from August

[21] The browsing data only records traffic to the Product Hunt website if the user accesses the platform through a non-mobile device. Therefore, we may be missing users who primarily access the website on their mobile phones.

2016 to August 2019.

CrunchBase API. CrunchBase is a subscription database that tracks technology startups across the globe. The database is used primarily for lead generation, competitor analysis, and investment/acquisition research by industry users. Crunchbase's coverage of internet-focused startups is comparable to other startup data products (Kaplan and Lerner 2016). While the database does include large technology companies such as Google and Microsoft, the majority of firms in its sample are startups. The quality of information about these startups improves significantly after 2008 and includes information on the startups including founding year, firm name, company website, funding raised, and a brief description of the startup's product. Crunchbase data is particularly reliable for companies that have raised funding. Detailed funding data are obtained by querying the CrunchBase API available to researchers.

Preqin is an alternative asset data company. They provide tools to track investments by venture capitalists, hedge funds, and private equity firms.

BuiltWith As described by Koning, Hasan, and Chatterji (2022), BuiltWith is platform for lead-generation and sales intelligence. Companies like Facebook and Optimizely use this database to learn about the adoption of of their tools, generate leads, and monitor competition. BuiltWith indexes more than 30,000 web technologies for over 250 million websites. It tracks these websites' current and prior technology stacks. It's free API, which we use here, provides information on a website's current technology stack.

Table A1 shows descriptive statistics for the 5,742 products in our final sample.s

B Categorizing users by gender

Each user on Product Hunt displays their names on their online profiles. In the majority of cases, these users engage with the platform using their real names. Close to 50% of users link Twitter accounts to their Product Hunt profile as well. They do so primarily to establish a consistent digital presence across online platforms, to build a brand name for their skills to potential investors and employers. As our data set contains all public information displayed on the platform, we can identify the real names of the users, and improve that data when users don't provide their real names but have a linked Twitter account that displays their real names.

We then assign a gender to each name based on the first name using genderizeio API.[22] In the cases where the total number of names in the database for inferring the gender is small or zero (when the name cannot be parsed), we apply Bayesian updating to a Beta prior $B(31, 71)$[23], and classify the gender to be female if the posterior probability that the user if female is at least 50%.

The name-based gender classification upon users are the basis for aggregating preferences for launched products and showing persistent divergence in these preferences across female and male consumers. For each product we use user data to generate measures of how many people viewed and voted for a product. Using the vote totals on each day for each product we calculate a product's rank within that day. We can tag the gender of the users to generate view, vote, and rank estimates for male and female users.

Makers sometimes try to game the Product Hunt platform by recruiting "friends, families, and bots" to one-off vote for their product. Product Hunt doesn't count, and sometimes penalizes, votes from these types of users. We exclude "friends, family, and bots" by filtering out new users who join the platform within a day of a product being launched, vote for this one product, and then are never active on the platform again.

C The value of page visits and technology development

Our primary dependent variable is monthly page visits, a measure of total user growth for a startup that comes from outside of the Product Hunt platform. That our measure is external to the platform is crucial. It could very well be that female-focused startups get fewer Product Hunt votes, after all 90% of users are men, but that this differential has no impact on performance outside of the platform. For example, maybe it is easy and cheap for female-focused startups to find traction through other channels or the that the Product Hunt platform actually doesn't have much real world impact. In the extreme, perhaps garnering fewer male votes on Product Hunt actually correlates with more growth with women through out channels!

Beyond resolving this measurement concern, visits are increasingly used in the entrepreneurial strategy literature to measure the traction and success of early stage ventures (Koning, Hasan, and

[22]In rare cases registered users are actually organizations, for which we are unable to map the name to a gender prediction. For non-western names (e.g. China) the given name cannot be parsed by genderize.io, and hence cannot really extract the gender of the name.

[23]Among 100 individuals, 30% are female.

Chatterji 2022). Similarly, investments in front-end technology stacks are increasingly being used to measure technological capabiltiies and development (Koning, Hasan, and Chatterji 2022; Roche, Oettl, and Catalini 2020).

But do visits and technology stack investments reflect startup traction and success? As simple test of these measures, in Figure A1 we plot the probability a venture raises venture capital funded against the logged number of page views one year after launch. There is a clear positive relationship. Products in the bottom 5% of pageviews (the dot most on the left) essentially never raise any funding. For those in the top 5% (the most right dot) over one-in-ten go on to raise funding. The same is true for technology stack development. In Figure **??** we show that as a product's technology stack gets larger the probability it raises venture funding grows from basically zero to nearly 25%. Finally, in Figure A3 we show that startups with more page visits invest much more in their technology stacks. While far from causal, this pattern is consistent with our argument in Section 5.2 that a lack of user growth may dissuade firms from investing in further technology development.

D Building and validating a measure of a product's gender focus

Here we present additional evidence in support of our measure of the female-focus of a product's target market. Figure A4 shows an example of the product text we use to construct our female-focus measure. We use the product's name, tagline, description, and first comment from the maker team which is used to provide additional product information and details.

Given this text data, what is the distribution of resulting female-focus score? Is it continuous or do we see "clumps" and outliers suggesting that our measure is best treated as dichotomous rather than a smooth score. Figure A5 shows a histogram of the normalized score. The measure is symmetric around zero, smooth, and roughly normally distributed. The distribution has slightly longer tails suggesting there are a handful of extremely male- and female-focused products.

Another potential concern with our female-focus measure is that it might simply capture differences in the product's category. Perhaps all female-focused products fall under the "designer tools" category while all male-focused products come from the "developer tools" category. To check for a lack of overlap in our female-focus measure in Figure A6 we show the share of male- and female-focused products in the five most popular product categories on Product Hunt. While there are notable differences—developer tools are more much more likely to be tagged as male-focused and marketing female-focused—there are male- and female-focused products in each category. This finding is consistent with our results in Table 2 which shows that prior to launch male- and female-focused products are similar in terms of pre-launch growth, funding, and team size.

Finally, extending Koning, Samila, and Ferguson's (2021) finding that female biomedical inventors are more likely to address female diseases and conditions, in Figure A7 we show that as products become more female-focused the probability that the team includes female entrepreneurs/makers increases. While just over 15% of the most male-focused products have a female entrepreneur on the maker team this rate doubles to just over 30% for the most female-focused products.

E Using the viewing data to estimate preferences

Here we describe how we use the viewing data to estimate female user preferences for more female-focused products while accounting for user and product fixed effects. We take an agnostic view to the origins of the gendered preferences we aim to estimate. Female users are likely to have much greater expertise in menstruation products than men and so may well be more likely to vote and support such products than male users. They also may simply enjoy a VR nail polish visualization app more than men. Women could have more career experience in HR roles and thus have more interest in HR software solutions. Women are likely to have more friends who are women and thus might be more interested in new products aimed at nurses because, as an occupation, there are more female than male nurses. There could be social norms around beauty that may cause women to feel compelled to vote for a new cosmetic startup idea. Any of these reasons are likely to lead female users to prefer female-focused startups more than male users even after accounting for product quality differences and differences in in how likely men and women are to vote for a product irrespective of its gender focus.

To estimate these gendered preferences, we begin by identifying all the users who spent time on the Product Hunt homepage on a given day.[24] For each user, we define users to be *active* on

[24]The product-user view data is constructed by combining proprietary data on users' browsing behavior on the platform with their upvoting history. The sample used in this analysis is the larger than the sample of products we analyze through the rest of the paper. Due to a logging error, this data is missing for a year of our core October 2016 to October 2018 sample. In the end, we end up with 11,212 products in our sample.

a given day, if they have visited Product Hunt homepage at least once[25]. The user is marked as viewing any product launched on that day, of which the creation time of its product post preceded the last time stamp indicating that the user accessed the website (any page, and not necessarily the homepage) that day. For each unique pair of user i and product j, where the user viewed the product on its launch day, the user's review of the product is positive (=1) if (s)he upvoted the product, and is zero (=0) otherwise as the user had seen the product but did not upvote it. This allows us to estimate the preference of user i toward each product j in his or her risk set in the following econometric model

$$Y_{ij} = \beta_i D_i + \xi_j D_j + \epsilon_{ij} \tag{4}$$

The residual $\hat{\epsilon}_{ij}$ from this equation measures user i's preference for product j after netting out individual harshness in reviewing products and quality of the product. We then aggregate these residuals over all female and male users for each product. This provides us with a measure of how much male versus female users like a product while ruling out (1) that differences are because men compared to women on average rate products better or worse and (2) that products that appeal to women as against men are lower quality. Figure 2 plots the male minus female difference in the residual votes after accounting for these fixed effects.

F Validating the newsletter's exogenous impact on female engagement

F.1 Sampling bias's causal inference challenge

As discussed in Section 4.5, it is non-trivial to identify sampling bias effects. First and foremost, talented entrepreneurs who are building female-focused products may bring more female users onto the Product Hunt platform on the day they launch. For example, the president of a university's "Women in CS" club is likely to build a great female-focused product *and* bring more female users onto the Product Hunt platform the day she launches. Is it the entrepreneur's talent that drives success or the presence of more female early users? Figure A9 illustrates this classic omitted

[25] The only exception is when the only thing the user did on the given day is visiting the homepage exactly once, in which case we do not consider the user to be active on that day

variable causal inference challenge as a Directed Acyclic Graph (DAG). Given male-dominated Product Hunt, we want to identify if more female engagement increases the success of female-focused products. However, an omitted variable—in this case, entrepreneurial talent—confounds our ability to estimate the causal effect.

This is where the value of the newsletter shock comes in, as illustrated in Figure A10. If the newsletter is exogenous to the products launched on a given day, and if it increases female engagement, then we can use the shock to identify shifts in female engagement while sidestepping potential omitted variables related to the quality of the product or entrepreneur. That said, like nearly all natural experiments, the newsletter shock, if not perfect. Unlike a standard instrumental variables setup, the shock is likely to impact multiple channels of engagement ranging from votes on Product Hunt to social media likes on Facebook to direct usage of the website itself. In fact, many users visit the Product Hunt website without being logged in. For these users, we cannot estimate their gender nor can we directly see if they engage with female-focused products. As such, we cannot assume an exclusion restriction and simply "instrument" the total number of votes with our shock variable.

Our interest, however, as Figure A10 shows, is in the total impact of increasing the number of women in the sample. Our argument is agnostic between whether these additional women choose to give direct feedback on the product's website as against through Product hunt comments. That said, as the red lines in Figure A10 show, the newsletter shock may well impact product success through alternative mechanisms that are fundamentally different than the sampling bias mechanism we propose. For example, the shock could change the preferences of male users leading them to prefer female-focused products more on days when the newsletter is female-focused. Fortunately, the detailed data we have on Product Hunt users, votes, and products lets us rule out this and other alternative channels. A10 lists the set of plausible alternative mechanisms that we think the newsletter might impact. These alternatives are rooted in shifts in men's voting behavior or shifts in women engaging with a larger set of products rather than just female-focused products, as we have argued. As Figure A10 documents, we present evidence—described in detail in the next few sections—that rules out each of these alternative channels.

F.2 Newsletter increases female engagement with female-focused products

As illustrated in Figure A10, for the newsletter to identify the impact of sampling bias on product success, it must impact female engagement with the Product Hunt platform. Consistent with this idea, in Table A2 we show that on "female newsletter" days more female users visit the Product Hunt homepage (Columns 1 and 2). Furthermore, more women visit the pages for products launched that day (Columns 3 and 4). As discussed in the body of the paper, our analysis excludes the handful of products that are featured in the newsletter itself.

Do these women then engage and support female-focused products at higher rates? While we lack data on the full spectrum of possible engagement—direct visits to the product's web page, social media mentions, and so on—we can test if female-focused products receive more votes from female users on days the newsletter is more female-focused. Further, it should be the case that male-focused products see no such gains from days with "female newsletters."

Figure A11 shows that these two patterns hold. On the left, Panel (a) plots the logged number of female votes for male-focused products against our newsletter shock variable. Like in Figure 1 we control for the number of male votes to account for potential differences in product quality and to improve our ability/power to detect gendered preferences versus differences in overall product appeal. It does not appear the newsletter impacts the number of female votes for male-focused products. In contrast, Panel (B) shows that female-focused products receive more votes from women when the newsletter is more female-focused.

Furthermore, the impact of the newsletter does not extend to voting by men. Figure A12 is similar to A11 but shows the impact on the newsletter on the logged number of votes from male users. Again, in Panel (a) we find no impact for male-focused products. If anything, in Panel (b) we find that men may be less likely to vote for female-focused products on "female newsletter" days, though the association is noisy.

To more formally test these patterns, Table A3 regresses the number of male and female votes on the newsletter shock variable for sub-samples that only include male-focused, gender neutral, or female-focused products. Columns 1-3 show that the newsletter does not impact the number of male votes for products no matter the product's gender focus. Columns 4 and 5 show that the newsletter shock does not impact the number of female votes for male-focused and gender-neutral products. However, Column 6 shows that female-focused products received more votes from female users when they launched on days when the newsletter itself features especially female-focused products. Taken together, this table and the figures above strongly suggest the newsletter impacts female engagement with female-focused products.

F.3 The newsletter does not shift male user preferences

As outlined in Figure A10, while the newsletter may well impact the success of female-focused products, it could do so by shifting men's willingness to vote for these products. Perhaps seeing successful female-focused products, like Nurx or Mirror, in the newsletter leads men to be more likely to engage with and vote for these ventures.

To definitively rule out this possibility, Figure A13 builds on Figure 2 but plots the female-male vote gap against the product' female-focus separately for days when the newsletter is especially male-focused (Panel A) and especially female-focused (Panel B). The slopes are nearly identical across the panels, revealing that the newsletter does not change how men vote compared to women. Table A4 builds on this visual evidence by regressing the estimated female-male vote difference on the newsletter shock and the shock interacted with the whether the product is female focused. There is no evidence that the newsletter cause men to shift their voting behavior in favor of female-focused products.

F.4 Evaluating alternative methods of constructing the newsletter shock measure

Given the newsletter features multiple previously launched products, how can we best convert this list of products into a uni-dimensional score that reflects the likelihood the newsletter brings more women onto the Product Hunt platform? Unfortunately, there are many arguments in favor of different functional forms: Perhaps the average is best because readers look at each product and based on the overall "femaleness" then decide to visit Product Hunt and look for new products. Perhaps the maximum because the most female-focused product will "pop out" and drive female user attention. Perhaps the gender-focus of the first few products is all that matters because people don't scroll down when the newsletter is long.

Fortunately, our rich data lets us check if alternative functional forms are more appropriate.

Specifically, in Table A5 we regress the logged number of female votes on the maximum of the female-focused product scores for the products suggested in the newsletter, the mean of these scores, the share of these with scores in the top 90% of the female-focused score distribution (top decile), and in the top 5% of the distribution (top ventile). Since our argument is that the shock should only impact female-focused products launched on the day of the newsletter, we include the interaction between these measures and our measure of the launched products' female-focus score. Column 1 shows that the maximum predicts an increase in votes and Column 2 shows the mean does not. Column 3 shows the share of products with scores in the top decile also does not predict female votes but Column 4 shows that the share of products in the top ventile does predict female engagement.

What explains the fact that the maximum is what matters? While we lack definitive evidence, we think female users are quickly skimming the first few products and when one is especially female focused they click through and then spend time browsing and voting on Product Hunt; when the most female-focused is a "bro" product they simply delete the email and move on with their day. To test this hypothesis, in Table A6 we regress the logged number of female votes on the maximum, median, and minimum of the suggested products' female-focused scores. If users are looking for "especially" female-focused products then the median and minimum should have no effect, female users are looking for the product that is most engaging. Similarly, if the maximum female-score is still is a "bro" product they should be especially unlikely to visit the platform, again suggesting the maximum and not other measures is what matters. Columns 1 shows that the maximum predicts more female votes for female products that are not featured in the newsletter. Column 2 shows the median does not predict future female votes nor does the minimum (Column 3). Column 4 includes all three measures and, if anything, the coefficient on the interaction term with the maximum gets larger.

If women newsletter readers are skimming for especially interesting female-focused products then the impact of the newsletter should be concentrated amongst the first few products in the newsletter and not in products listed later in the email. A very female product described after the few products should have little impact. Indeed, this appears to be the case. If we calculate the maximum female product score only using the first 5 products listed in the newsletter (Column 5) we find effects similar to what we find in Column 1; if we instead only calculate the maximum over products listed after the first 5 we find no impact of the newsletter on female votes for female-focused products.[26]

Overall, these findings suggest that users appear to quickly skim the newsletter looking for the most relevant content and so using the maximum of the suggested products is most appropriate.

F.5 Balance

To further confirm that the content of the daily newsletter is unrelated to product's launched in Table A7 we regress the newsletter shock variable on gender and quality characteristics of the product. We find no evidence that products that are more female focused or that have female makers are more likely to launch when the newsletter is female focused. Similarly, we find no evidence that higher or lower quality projects are more likely to launch on female newsletter days.

F.6 Female-focused newsletter products are no different in their quality than male-focused newsletter products

Beyond balance across days, it could also be that the suggested products in the newsletter that are female focused also differ on other underlying dimensions like product quality or maker popularity. Akin to concerns discussed in the peer effects literature (E.g. Chatterji et al. 2019), our "female newletter" shock may actually reflect a distinct but correlated underlying shock. If, for example, the female-focused products featured in the newsletter are from more popular makers then perhaps the shock simply brings more people onto the Product Hunt platfrom on that day. As a result, female- and male-products receive more votes and see performance gains, which could explain some but not all of our findings.

To rule out this channel we build a dataset of all the suggested products in the newsletters we analyze. We then test if female-focused products featured in the newsletter are higher quality or differ on other confounding dimensions. Unfortunately, unlike the 5,742 products we analyze in the body of the paper, many of the 2,365 featured newsletter products were either launched by large technology companies or are from startups that launched well before our sample window. As a result, we lack page visits from Similar Web that we could use as a measure of featured product of

[26]What we do find is that the main effect of the newsletter shock calculated over suggested products outside of the top 5 is negative, though at the 10% level. Given the number of cuts of the data in this table that one coefficient is marginally significant is not particularly surprising.

quality. That said, for each featured product we do know the number of votes the product received when it launched and the number of Product Hunt followers the makers of the product have at the time of the newsletter. The first gives us a proxy for the product's quality and the second a proxy for the quality of the founding team. We also can measure if the suggested product was described using more words, suggesting the product may be more ambitious or had a larger scope that needed more sentences to describe. Finally, we also have a measure of the gender composition of the maker team.

In Table A8 we regress our normalized female-focus measure of the newsletter suggested products on each of these variables. We find no evidence that suggested female-focused products received more votes, have more popular followers, or have more detailed longer and more detailed descriptions. Consistent with our analysis in Figure A7 we find that the suggested products that are female-focused are more likely to feature female makers. While distinct, we think this measure is, if anything, also likely to lead more women to visit the Product Hunt website.

Finally, we create a "placebo" shock using the logged number of votes each newsletter product received and then calculate the maximum of this measure for each newsletter. Akin to our "female newsletter" shock, this "quality shock" lets us test if newsletters that include especially high quality products drive greater female engagement. Column 6 in Table A8 shows that, unlike our "female newsletter" measure, the "quality newsletter" measure does not predict female votes directly nor does it lead to more female votes for female-focused products.

G Robustness checks ruling out alternative newsletter mechanisms

Again building off of Figure A10, here we present analyses to further rule out alternative causal mechanisms that might explain our newsletter effect. First, in Table A9 we show that the performance effects of the newsletter shock are not universal to female entrepreneurs, but instead only hold for women who are female-focused entrepreneurs. To do so, we restrict our data to teams with at least one female maker and then regress logged monthly page visits on the newsletter shock for three sub-samples of our data: male-focused, gender-neutral, and female-focused startups. While the estimates have large standard errors due to the fact that we are operating on smaller samples of product launches, we find no evidence that the newsletter shock improves the performance of female made male-focused or gender-neutral products. Instead, as Column 3 shows, the benefits of the female newsletter shock only hold for women launching female-focused products.

Second, in Table A10 we show our performance findings are not merely explained by shifts in the total number of votes a product receives. If the total number of votes can explain the impact of the newsletter, this would be suggestive that the newsletter shock is simply serving as a a source of gender agnostic advertising, and not as shifting the gender composition of the sample. To test this channel, in Table A10 we control for the total number of votes and find no change in our estimates on the impact of the newsletter on performance.

H A simple Bayesian example of entrepreneurial sampling

Given that Product Hunt, and the Silicon Valley technology ecosystem, are dominated by men, why don't entrepreneurs focusing on the needs of women simply "debias" the signals they receive, knowing that so much of the early feedback and traction they get will be from men? While there are likely myriad explanations, ranging from the behavioral to the sociological, here we show that even if women know to ignore or discount feedback from men that they may still be more likely to abandon a promising idea. Why? Because by ignoring feedback from men entrepreneurs resort to learning from smaller samples of women. While these smaller samples will be less biased, they will also be, simply by virtue of being smaller, less informative and so less likely to convince an entrepreneur that her idea is worth scaling. The entrepreneur is confronted with a classic trade-off between generating a biased estimate with lower variance, or an unbiased estimate with greater variance and uncertainty.

Here we present a simple numerical example to illustrate the logic. To begin, assume the entrepreneur learns following a standard Bayesian Beta-Binomial model . She starts with a wide beta distributed prior to her chance of success, $\pi \sim Beta(\alpha = 10, \beta = 20)$, consistent with the idea that the entrepreneur is far from confident that her idea will scale. The beta distribution gives us a prior over the probability any given consumer will like her product. If she ever believes that at least 50% (i.e. $\pi > 0.5$) of users will like her product, she will invest in scaling the idea. Otherwise, she shuts down the venture. Given these beliefs, the mean of her prior must start out at less than

50% as otherwise she would simply choose to scale instead of running an experiment with early users to learn if she should invest more time and effort.

She then launches a prototype of her either female- or male-focused idea with 100 users to get feedback and learn. Unfortunately, early users are 80% male. While not ex ante known to the entrepreneur, female-focused ideas appeal to 60% of women, but only 40% of men. Male-focused ideas to 60% of men, but only 40% of women. From the experiment, the entrepreneur learns what fraction of users like her idea and uses this data to update her beliefs about π.

Figure A14 shows how an entrepreneur testing a female-focused idea with 100 early users updates their beliefs. The top row is the entrepreneur's prior. The column on the left is for a "naive" entrepreneur who learns from all 100 users. In the second row on the left we see the binomial "likelihood" she builds from all the data she collects from all 100 users. She learns that 44% of users like her idea. Her posterior beliefs, shown in the third row, move to the right, but the mean is still well below the 50% threshold. She abandons the venture.

However, perhaps the entrepreneur is "sophisticated" and so knows to ignore the advice from the male early users. This scenario is shown in the column on the right. Unfortunately, as the middle row on the right shows, while the data from the sample of 20 women has a mean of 60%—well above the $\pi > 0.5$ threshold—the binomial "likelihood" is also much wider. Learning from a sample of only 20 early female users is less "biased," but it is also simply less informative. As a result, the posterior in the bottom row still updates to the right, but again the mean is below 50%. Even the sophisticated female-focused entrepreneur will incorrectly shut down her venture.

Finally, it is worth noting that while a female-focused entrepreneur can try to "re-weight" data from men to make it informative of the entrepreneur's more female target market, re-weighting will always introduce some noise and uncertainty unless the entrepreneur already knows with certainty the difference in the male versus female appeal of her product. Given the inherent uncertainty associated with entrepreneurship, we find it unlikely that entrepreneurs would not know the appeal to women, but would know with certainty how much more or less men are likely to support the product! Similar to the uncertainty introduced by the use of survey weights, the entrepreneur can try and use observables to make a sample look like a target population of interest, but in doing so the entrepreneur necessarily relies on models and assumptions that introduce additional uncertainty.

While one can quibble with the particular assumptions used in this example, the underlying idea is a general one. If female-focused founders have access to less relevant "data" from early users, then even if they adjust for "sampling bias," the resulting signals are simply weaker. Having to rely on less data is necessarily less informative.

I Additional funding and technology stack results

Here we present a handful of additional robustness checks related to our funding and technology development results. First, Table A11 we show that products missing BuiltWith technology stack data are smaller and are less likely to have raised funding pre-launch, but are no different in terms of their gender focus or the newsletter shock they experience. Missing data is unlikely to explain our technology development results.

Tables A12 and A13 show our funding and technology development results, but swap out our continuous female-focus measure for male- and female-focus quartiles buckets. Our pattern of findings hold when using these discrete models.

J Testing the comments channel

Does the newsletter lead to more comments for female-focused products? More valuable comments? More comments from women? To test these channels, we collected the comments for 5,742 products in our sample. We find that 5,134 of the products in our sample have at least a one comment from someone outside the maker team. For these products we generate four measures: (1) if anyone who commented is a female user, (2) the logged number of female comments, (3) the average length of the logged number of words included in female user comments, and (4) a score of the sentiment of female user comments using the sentiment analysis function from the the NLTK python package.[27]

Table A14 regresses these measures on our measure of a product's female-focus and our newsletter shock measure along with out standard year-month controls. The models also include a control for the number of comments from men. Columns 1 and 2 show that female-focused products are 5 to 10 percentage points more likely to receive a comment from a female user. However, the newsletter shock appears to have zero impact on the probability a women comments. Similarly, Columns

[27] https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664

3 and 4 show female-focused products receive more comments from female users, but again the newsletter has no impact.

Consistent with these null results, in Columns 5-6 we show the average length of comments from female users is longer for female-focused products, but does not shift because of the newsletter. Finally, Columns 7-8 show that the the there is no impact on the sentiment of the comments from female users. FInally, it is worth noting that to isolate potential impacts on comment content from the number of comments Columns 5-8 only include products where at least one female user left a comment.

Overall, perhaps because women are simply less engaged on the Product Hunt platform to begin with, we find no impact of the newsletter on the quantity or quality of the comments. In contrast, in Appendix Section F.2 we show the newsletter does meaningfully shift the number of votes from female users.

K Testing for effect heterogeneity

As discussed in Section 5.3, we use split sample tests to check for heterogeneity in our effects. We test for heterogeneity across two distinct dimensions. First, as discussed in the body of the paper, to check for differences in prior platform engagement we use the fact that we can link makers to their user profiles to see how many products each maker has voted on in the past. Using this linkage, for each team we then create a total number of votes given to other previously launched products by the team members. We then use this "number of votes by team" measure as a measure for the team's engagement with the platform. Teams with 20 or more prior upvotes are marked as those that are the "most engaged" where as teams with fewer than the 20 are marked as less engaged. While the mean team has voted 131 times before their launch, the median team has voted 23 times and 10% of teams have voted for other projects one or fewer times before launching.

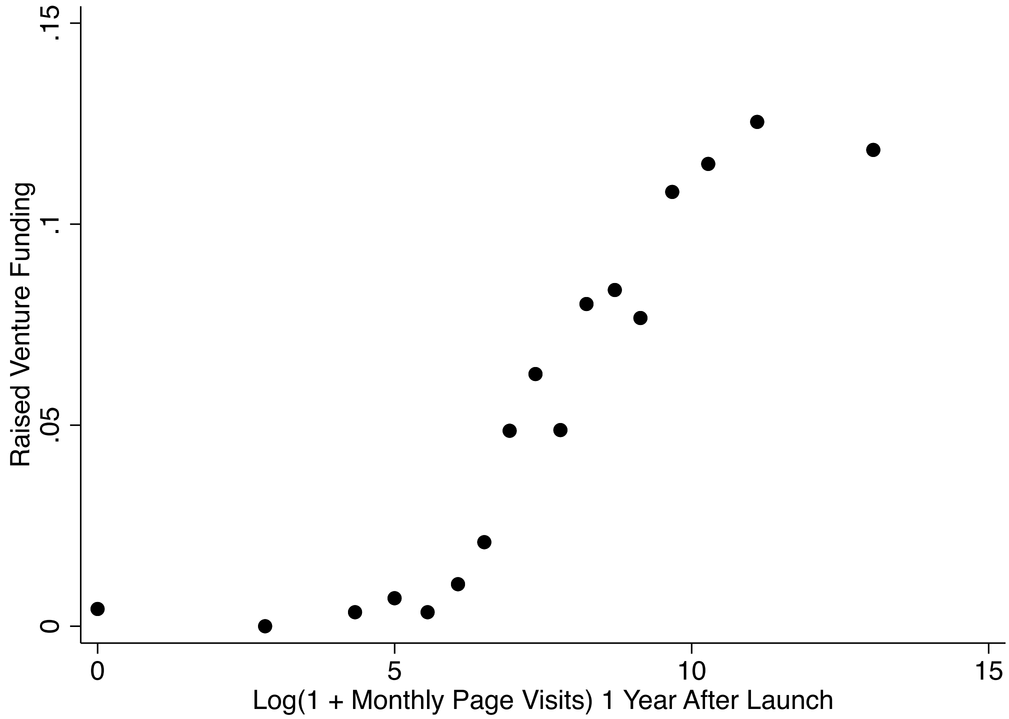
We also check for differences by observed entrepreneurial experience. Perhaps greater entrepreneurial experience leads entrepreneurs to be better informed, have stronger beliefs, and so have a greater set of signals to rely on. For example, entrepreneurs who have already pitched VCs may have already collected a set of useful, if also biased, signals to learn from. Entrepreneurs who have launched other ventures in the past may have access to personal experience that can substitute for the biased and noisy data generated by launching on Product Hunt. These arguments suggest entrepreneurial and fundraising experience may mitigate the impact of sampling bias.

To measure experience, we leverage the fact we know both if a product has already raised venture funding at the time of launch and if someone on the team has launched a different venture on Product Hunt in the past. Roughly 50% of launches feature team members with prior launch experience or that have already raised funding. While there are surely founders with prior founding experience in the group we mark as lacking experience, our argument is that, at least on average, the group that has raised funding or launched on Product Hunt in the past has more experience and so greater access to other signals of an idea's potential.

Table A15 then applies our standard newsletter shock regressions, but split by the two measures described above. Column 1 and 2 show the split by prior engagement with Product Hunt. We find teams that have engaged with Product Hunt in the past see sampling bias effects at least as large as teams that are new to the platform. It does not appear that familiarity, and so awareness, of the platform's composition limits sampling bias. Moreover, in Figure A15 we plot kernel density estimates of our estimated female-focus measure for teams who have been active on Product Hunt (the blue line) and those have been less active (the red line). The distributions are nearly identical, with both groups launching an equal share of female-focused products. This overlap suggests that while teams with more experience on Product Hunt respond more to the platform's signals, they do not appear to be less likely or willing to launch female-focused products.

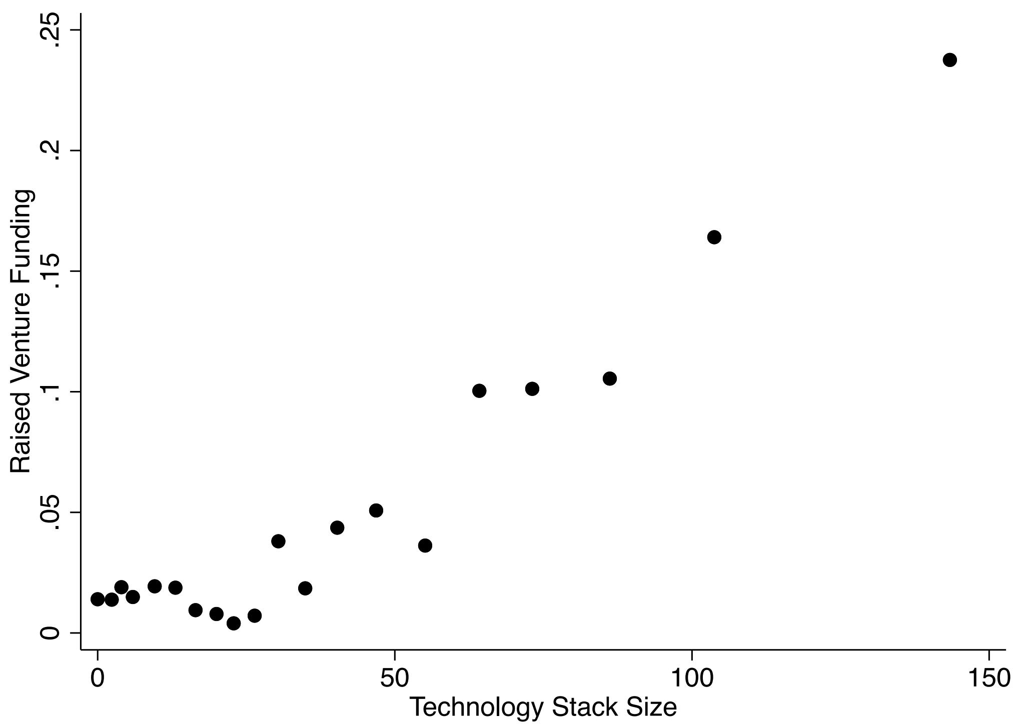
Finally, Columns 3 and 4 in Table A15 show the results for teams with observed entrepreneurial experience and for teams without any such experience. We find little difference across the samples. Entrepreneurs that choose to launch on Product Hunt, no matter their experience, appear equally influenced by sampling bias effects.

Figure A1: Binned scatter plot showing that whether a startup ever raises venture funding is strongly correlated with monthly visits.



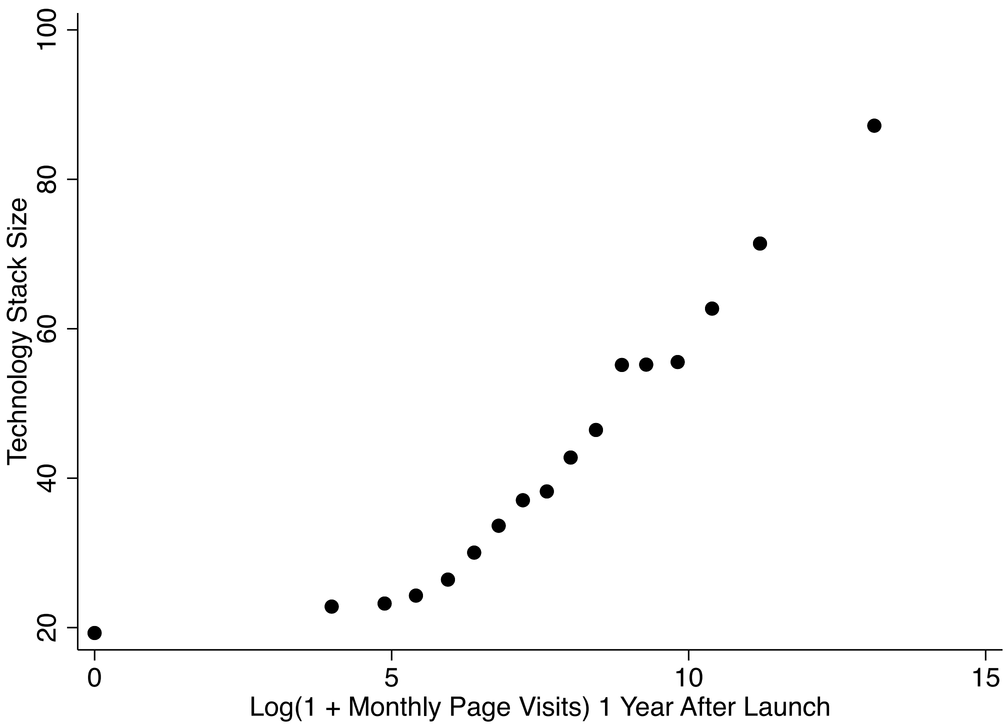
Notes: This figure shows the correlation between a startup's monthly visits a year after Product Hunt launch with whether the startup has ever raised VC funding as of October 2020.

Figure A2: Binned scatter plot showing that venture funding is strongly correlated with technology stack size.



Notes: This figure shows the correlation between the size of a startup's technology stack as of October 2020 and whether the startup has ever raised venture funding as of the same date.

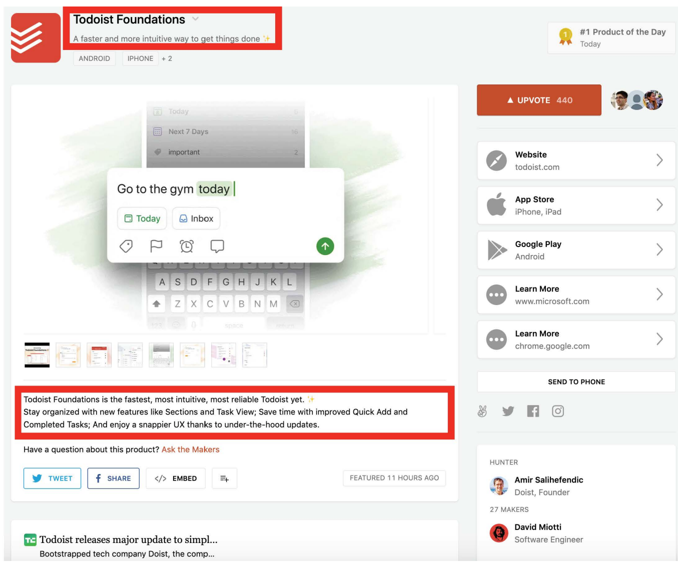
Figure A3: Binned scatter plot showing that a startup's technology stack size is strongly correlated with monthly visits.



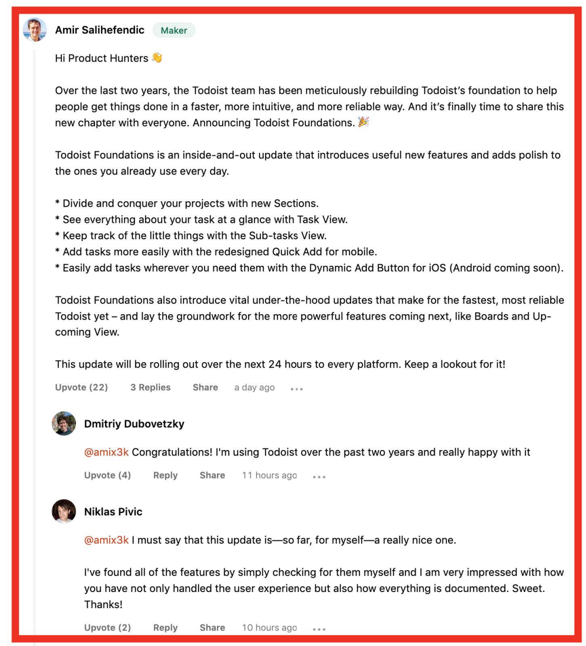
Notes: This figure shows the correlation between a startup's monthly visits a year after Product Hunt launch with whether the the size of a startup's technology stack as of October 2020.

Figure A4: Example of Product Texts: Todoist Foundations

(a) Name, Tagline and Description

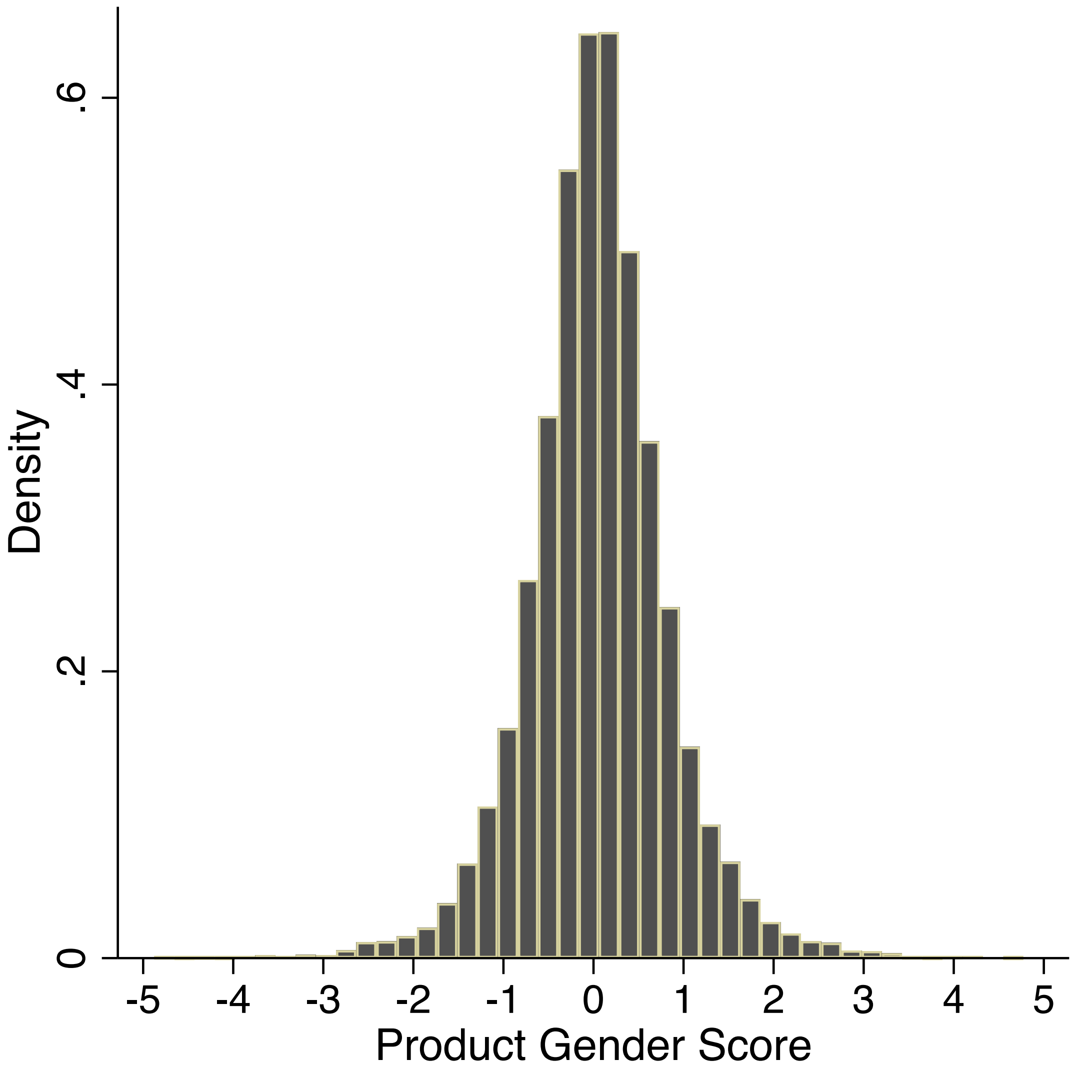


(b) Hunter- and Maker-initiated Comments



Notes: This figure shows the texts in the Product Hunt post for "Todoist Foundations" – a featured product launched on October 23, 2019. The left panel highlights the name and tagline of the product. The right panel highlights comments on the product, initiated by the user posting the product on behalf of the makers (aka hunter) and the makers themselves.

Figure A5: Distribution of the Product Gender Score



Notes: This figure shows the distribution of the product gender score, on a sample of entrepreneurial products launched on Product Hunt on weekdays from September 2016 to October 2018. The score is estimated using the text of each product's description. See text for further details. Higher scores imply that the product is more likely to serve or appeal to women.

Figure A6: Product gender focus for the most popular product categories on Product Hunt. Blue bars show the fraction of products in the category that are male-focused (bottom quartile). Red bars show the fraction of products in the category that are female-focused (top quartile). While there are meaningful differences across product categories there are both male- and female-focused products in each product category. This overlap suggests that when we compare female- to male-focused products there is enough "overlap" across products with different gender scores for meaningful comparison.

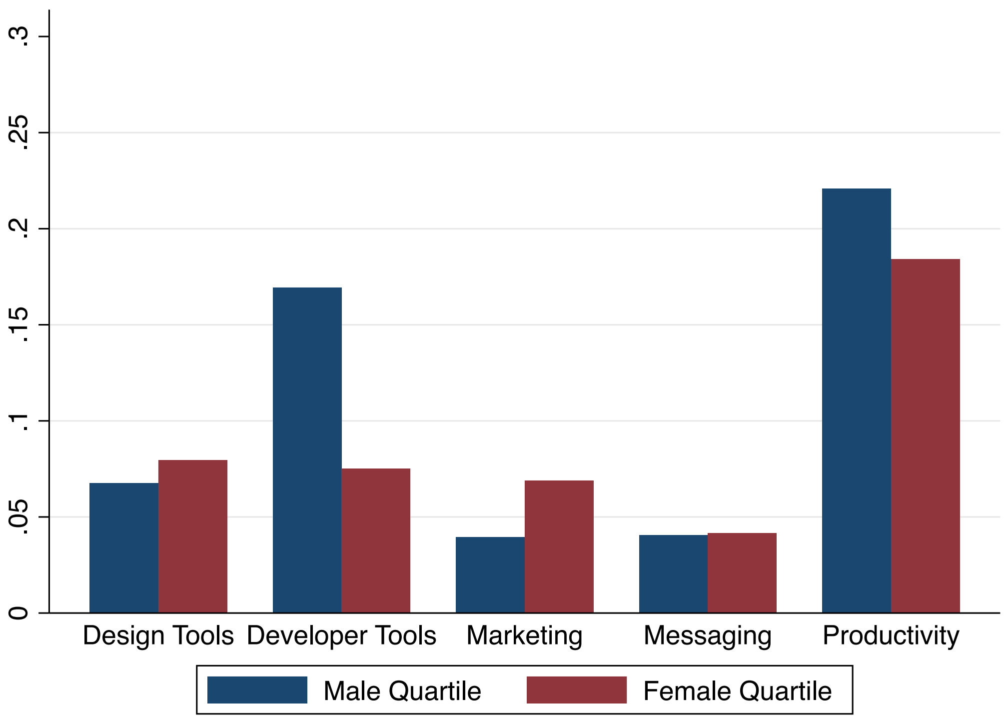
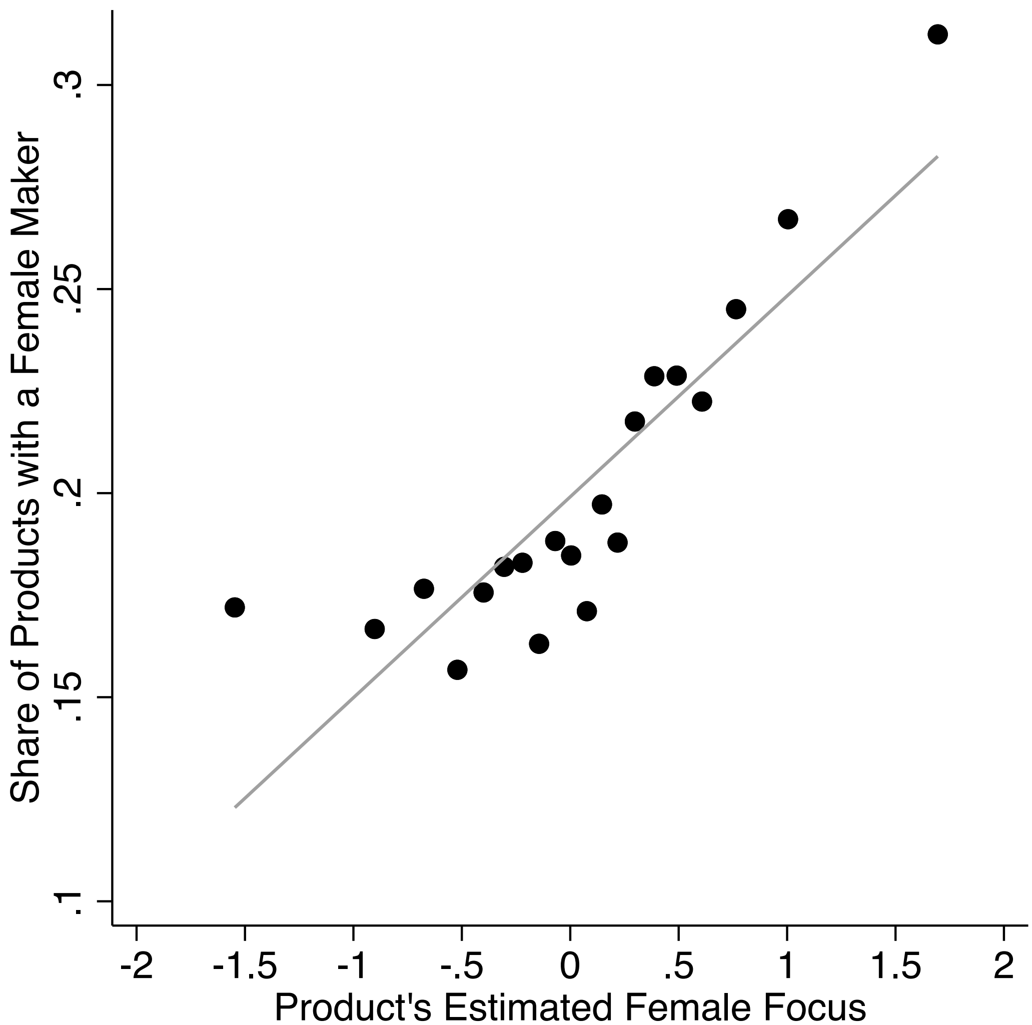


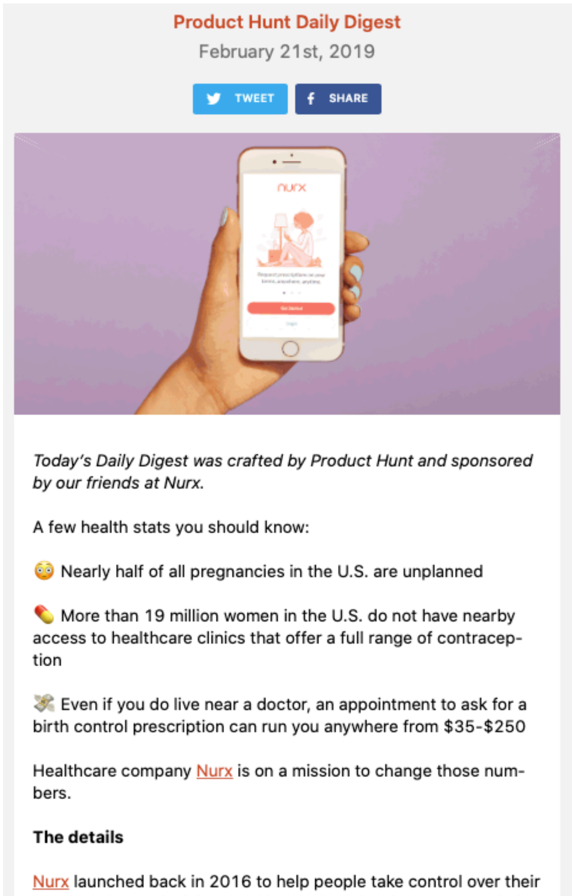
Figure A7: Binned scatter plot showing that products estimated as female focused—i.e., more likely to appeal to the needs and preferences of women—are more likely to be made by a female entrepreneur or inventor.



Notes: The Y-axis represents the probability that there is at least one female on the team that made the product. The X-axis is our text-based estimate of the degree to which the product focuses on female users. The binscatter controls for product launch year-month fixed effects, day-of-week fixed effects, and the logarithm of the number of words in product texts. The model includes 19,388 products with non-missing team member gender data.

Figure A8: Example of female-focused newsletters

(a) Example Newsletter: Nurx sponsored newsletter



Product Hunt Daily Digest

February 21st, 2019

TWEET SHARE

Today's Daily Digest was crafted by Product Hunt and sponsored by our friends at Nurx.

A few health stats you should know:

😳 Nearly half of all pregnancies in the U.S. are unplanned

💊 More than 19 million women in the U.S. do not have nearby access to healthcare clinics that offer a full range of contraception

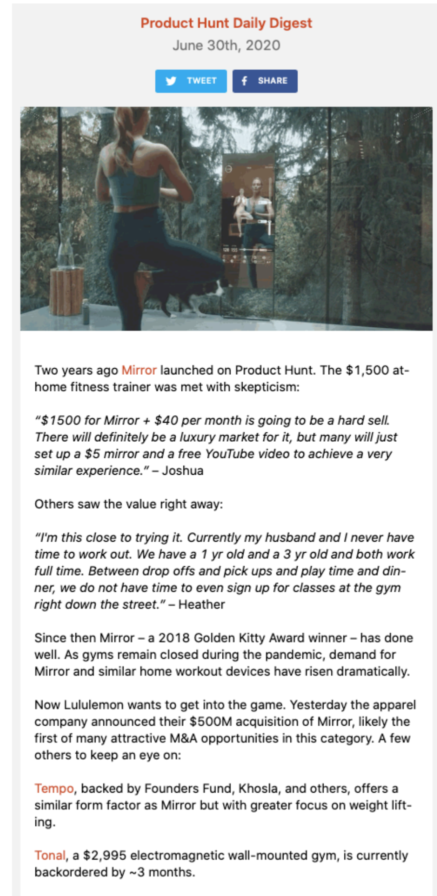
💸 Even if you do live near a doctor, an appointment to ask for a birth control prescription can run you anywhere from $35-$250

Healthcare company Nurx is on a mission to change those numbers.

The details

Nurx launched back in 2016 to help people take control over their own healthcare decisions and break down barriers to care. Their

(b) Example Newsletter: Mirror acquired by Lululemon



Product Hunt Daily Digest

June 30th, 2020

TWEET SHARE

Two years ago Mirror launched on Product Hunt. The $1,500 at-home fitness trainer was met with skepticism:

"$1500 for Mirror + $40 per month is going to be a hard sell. There will definitely be a luxury market for it, but many will just set up a $5 mirror and a free YouTube video to achieve a very similar experience." – Joshua

Others saw the value right away:

"I'm this close to trying it. Currently my husband and I never have time to work out. We have a 1 yr old and a 3 yr old and both work full time. Between drop offs and pick ups and play time and dinner, we do not have time to even sign up for classes at the gym right down the street." – Heather

Since then Mirror – a 2018 Golden Kitty Award winner – has done well. As gyms remain closed during the pandemic, demand for Mirror and similar home workout devices have risen dramatically.

Now Lululemon wants to get into the game. Yesterday the apparel company announced their $500M acquisition of Mirror, likely the first of many attractive M&A opportunities in this category. A few others to keep an eye on:

Tempo, backed by Founders Fund, Khosla, and others, offers a similar form factor as Mirror but with greater focus on weight lifting.

Tonal, a $2,995 electromagnetic wall-mounted gym, is currently backordered by ~3 months.

And then there's Supernatural, marrying VR gaming with home

Notes: Panel (a) which had been featured on ProductHunt in the past, sponsored a newsletter on ProductHunt. Nurx is a digital-first birth control subscription company. The post linked to the original launch. Panel (b) Mirror, which launched on ProdcutHunt two years earlier, was acquired by the apparel company Lululemon. The newsletter covered the acquisition in linked to Mirror and competitors which had also launched on the platform in the past.

Figure A9: A Directed Acyclic Graph illustrating the core causal inference challenge in identifying sampling bias effects.

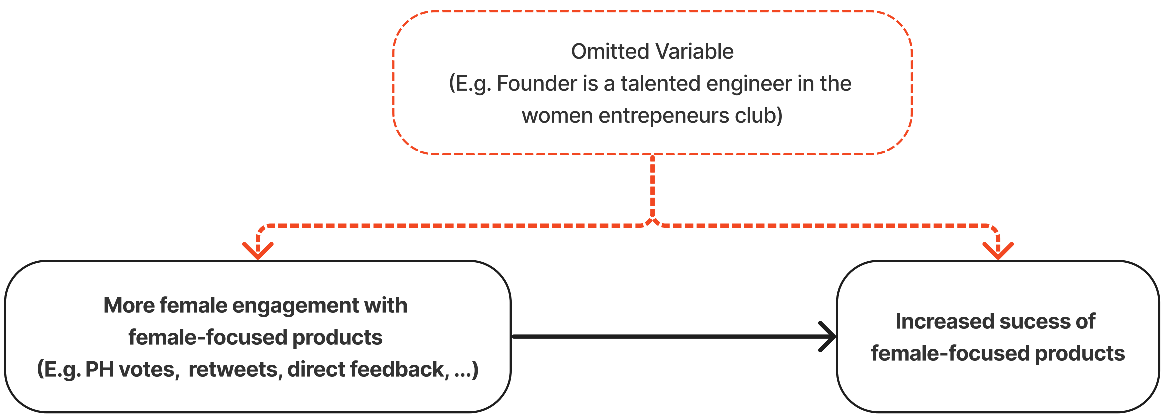


Figure A10: A Directed Acyclic Graph illustrating or study's causal inference challenges and how our results both support our proposed sampling bias channel and rule out potential alternatives.

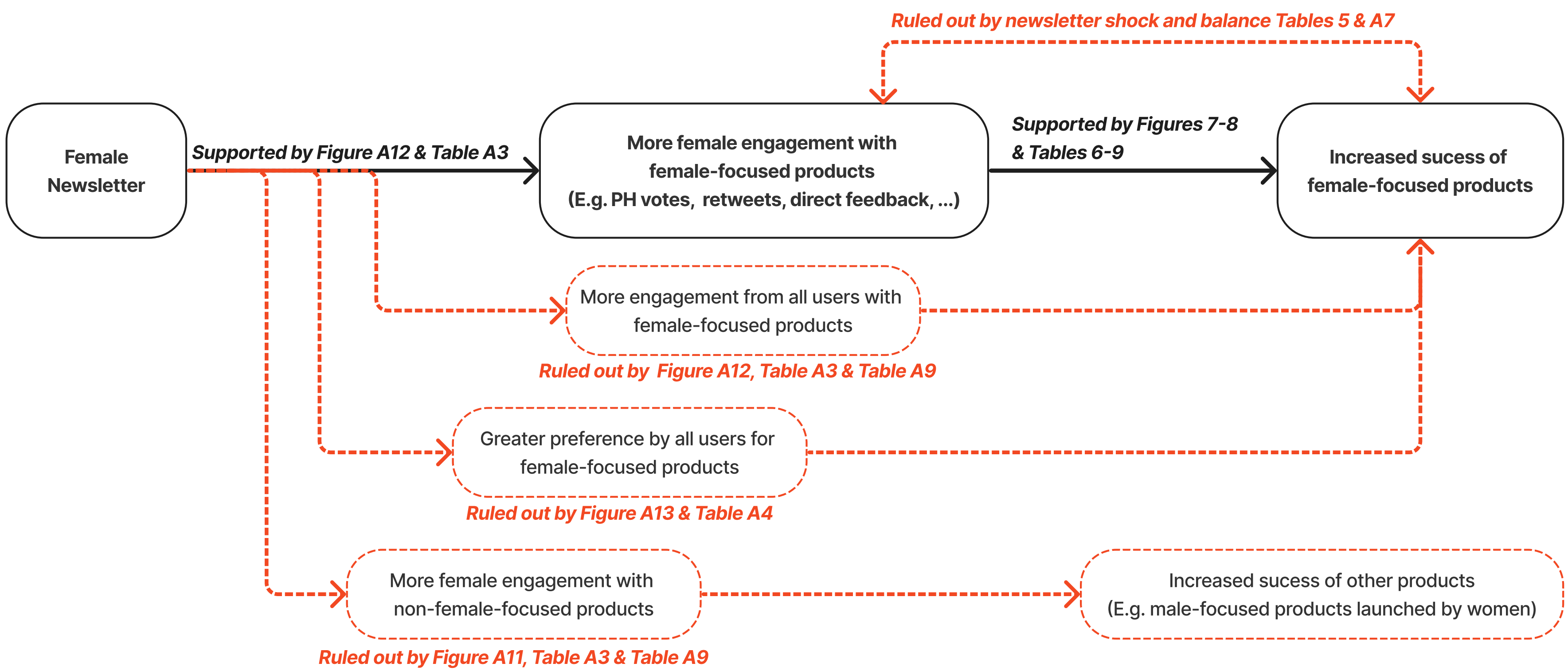


Figure A11: Binned scatter plots showing the impact of the female newsletter shock (x-axis) on the logged number of *female votes* (y-axis) a product receives after controlling for the logged number of votes from male users. Panel (a) is for male-focused products, which reveals the newsletter shock does not impact voting by female users for male-focused products. Panel (b) is for female focused products, which shows that the newsletter shock leads to *more* female votes for female-focused products, consistent with our sampling bias arguments. The sample includes the 5,742 products we analyze in Table 6.

(a) Male-focused products (Bottom Quartile)

(b) Female-focused products (Top Quartile)

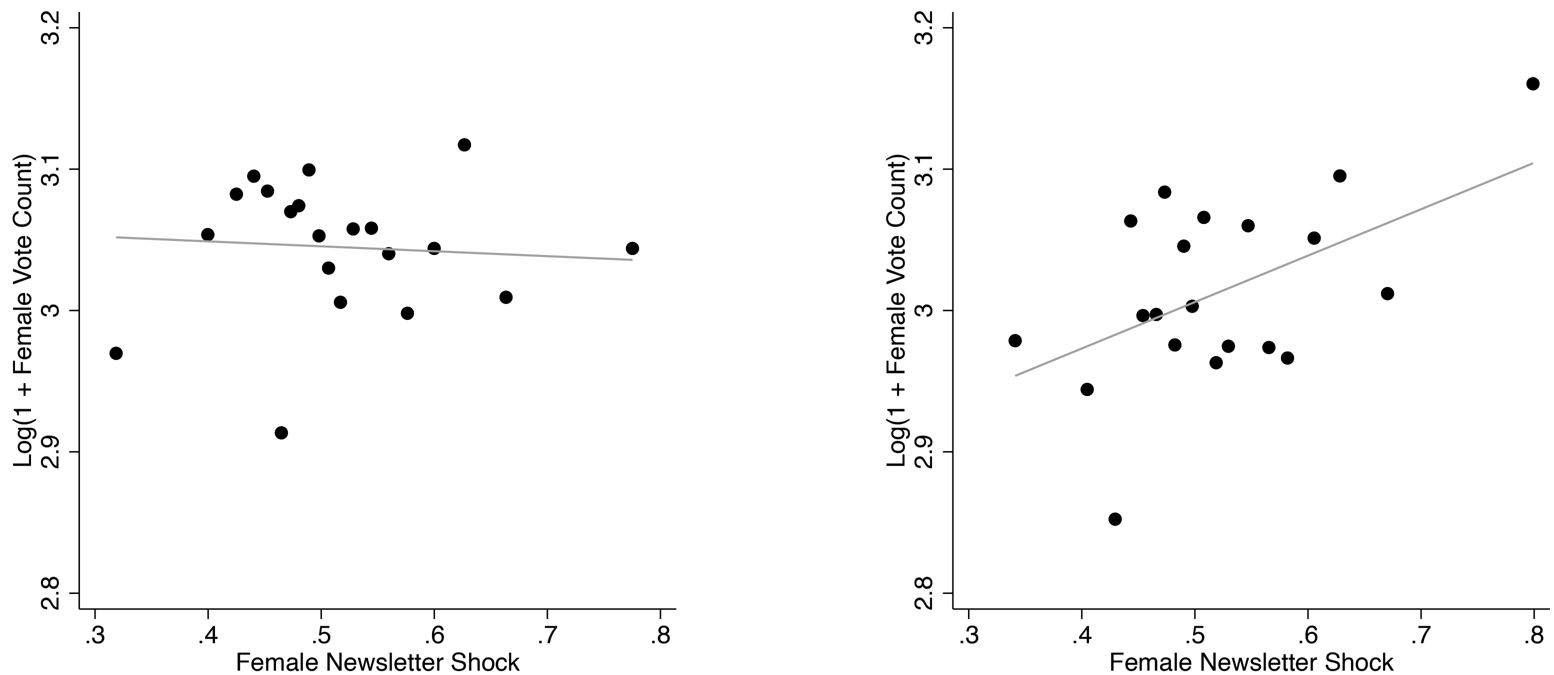


Figure A12: Binned scatter plots showing the impact of the female newsletter shock (x-axis) on the logged number of *male votes* (y-axis) a product receives after controlling for the logged number of votes from female users. Panel (a) is for male-focused products, which reveals the newsletter shock does not impact voting by male users for male-focused products. Panel (b) is for female focused products, which shows that—if anything—the newsletter shock leads to *fewer* male votes for female-focused products. Comparing across the two panels reveals that many more men vote for male-focused as against female-focused products. The sample includes the 5,742 products we analyze in Table 6.

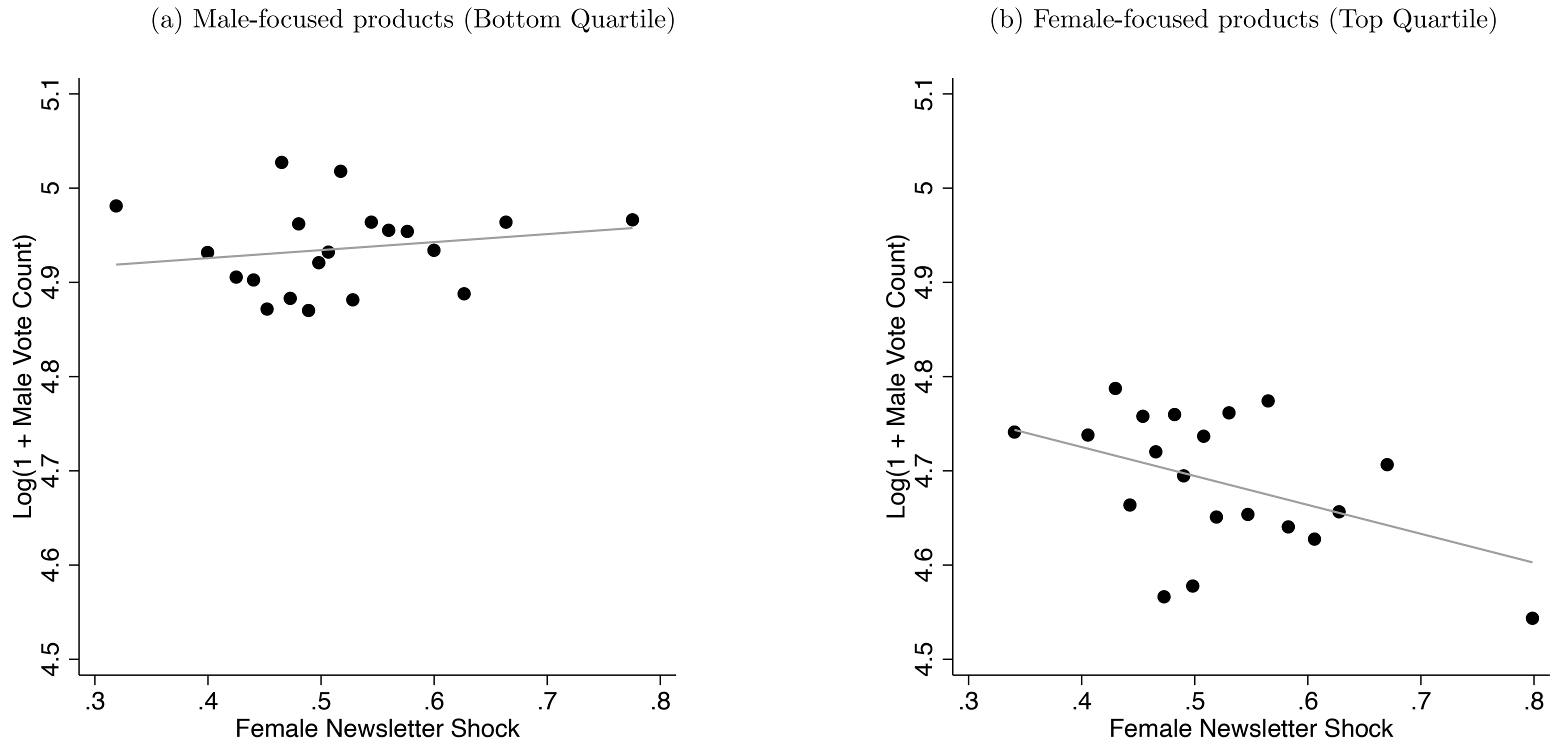
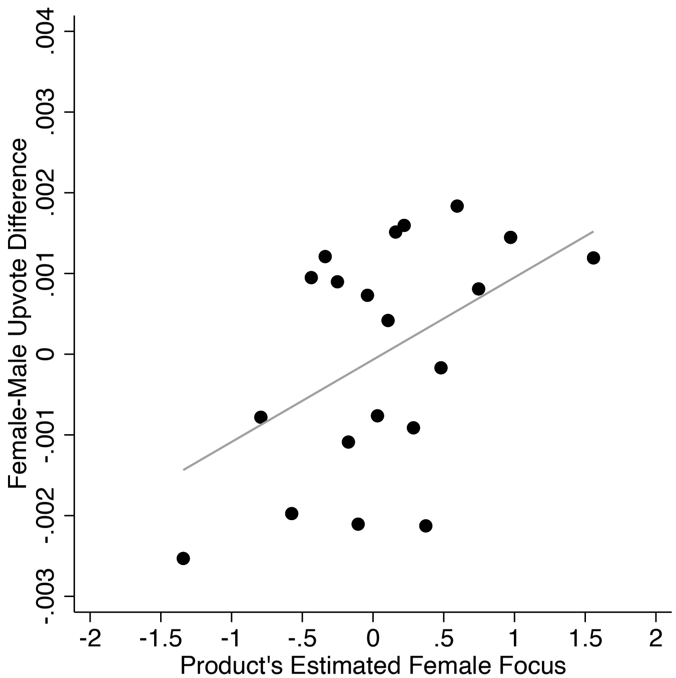


Figure A13: Binned scatter plots showing that increased chance a female user votes for a female-focused product relative to a male user is not different on male- versus female-focused newsletter days. Panel (a) shows the difference in the probability a female versus male user votes for a product given the product's estimated female focus on days when the newsletter is more male focused (bottom quartile). Panel (b) is similar but for days when the newsletter is more female focused (top quartile). Both panels show the difference between female and male users after accounting for product and user fixed effects. As in Table A4, the sample includes the 4,261 products for which we have both newsletter and proprietary viewing data.

(a) Male Newsletter Shock Days (Bottom Quartile)



(b) Female Newsletter Shock Days (Top Quartile)

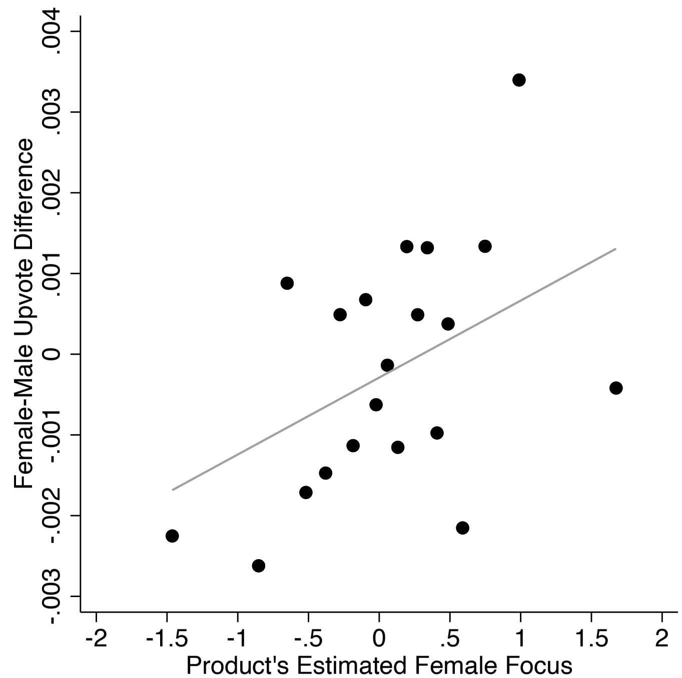
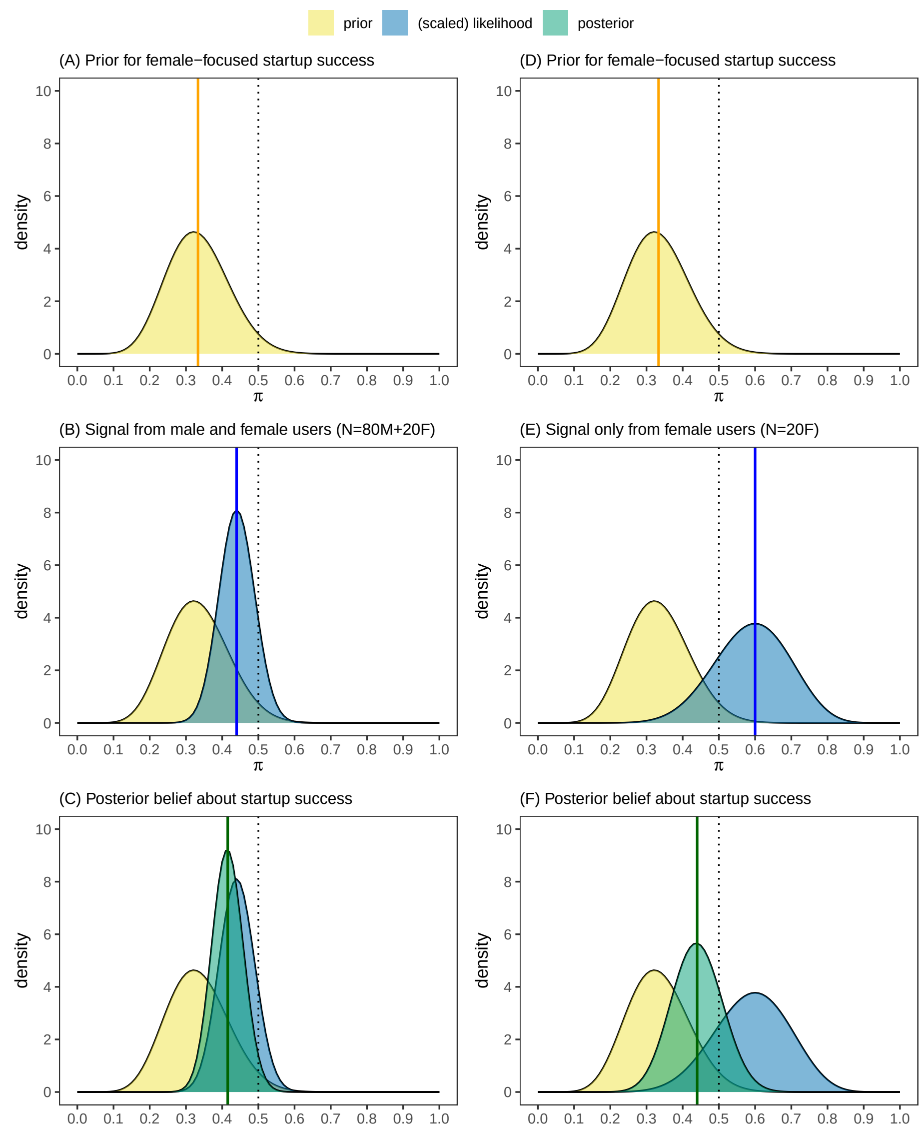


Figure A14: How a Bayesian "beta-binomial" entrepreneur building a female-focused product will update their beliefs about success in the face of an gendered imbalanced sample of early users. The first column is for a "naive" entrepreneur who updates using the full sample. The second column is for a "sophisticated" entrepreneur who only updates based on the female feedback. Even the "sophisticated" entrepreneur fails to update her beliefs above the 50% chance of success threshold. She still erroneously shuts down because the female only signal is simply nosier and so less convincing.



Notes: The first column shows a "naive" entrepreneur who simply aggregates the feedback from the full sample of 80 men and 20 women (Panels A, C, E). The second column shows a "rational" entrepreneur who only learns from the 20 females users and discards the overly pessimistic signal generated by male users (Panels B, D, F). In both columns, the entrepreneur starts with a wide prior with an expected success rate of 33% and will continue with the project if expected success exceeds 50%. Men prefer a male-focused product 60% of the time and female-focused 40% of the time; women like female-focused products 60% of the time and male focused 40%. If the entrepreneur had perfect knowledge, they would scale up both the female- and male-focused products. Though the entrepreneur building a female-focused product knows that women will like her idea more than men will, as is the case with any estimate of startup success, she is does not know by exactly how much. For simplicity, we assume the entrepreneur can either learn from the full sample or the sub-sample, though our logic holds when considering entrepreneurs with priors over the difference in male and female preferences. The second column reveals that when a female-focused entrepreneur only on female early users they receive a less informative signal, which in turn still leads a women with a promising idea to shut down too early.

Figure A15: Kernel density plots of the product's estimated female focus for teams with more versus less engagement on the Product Hunt platform. The blue line shows the distribution for products launched by entrepreneurs that has upvoted other projects 20 or mote times before launching. The red line is for entrepreneurs who have up voted fewer than 20 times before launching. The overlaps shows that, despite Product Hunt's male gender skew, entrepreneurs who have engaged with platform more heavily are equally likely to launch a female- as male-focused product.

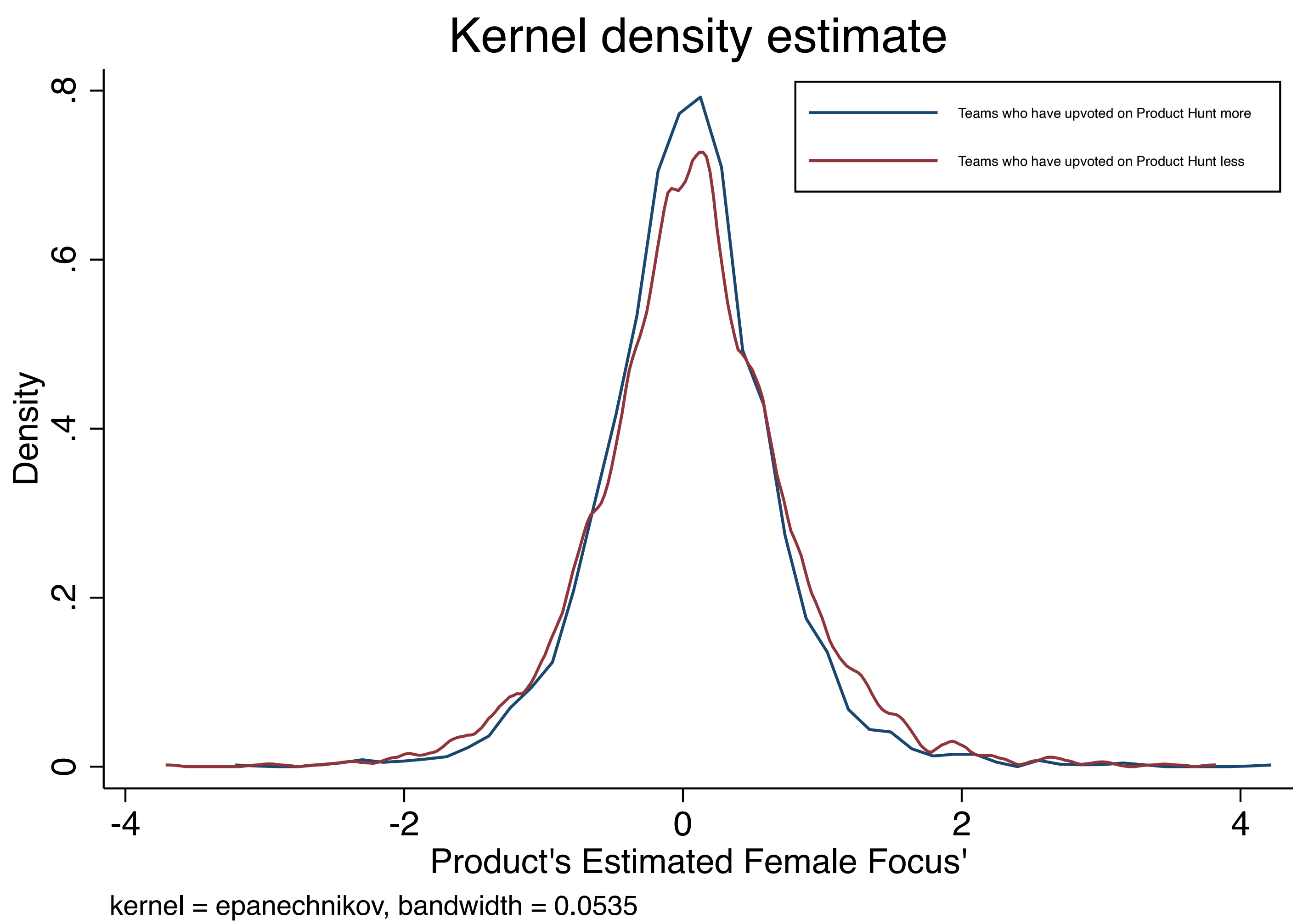


Table A1: Additional descriptive statistics for the 5,742 products in our sample

	Product Launches Sample, Sep 2016 - Oct 2018				
	All Products (N = 5,742)				
	Mean	Median	SD	Min	Max
Product Characteristics					
Product Female Focus	0.07	0.06	0.72	-4.46	5.10
Log (1 + Monthly Page Visits) 1 Month Before	5.23	6.11	4.10	0.00	20.31
Pre-Launch Seed or Series A Funding	0.028	0.00	0.165	0.00	1.00
Topic Category Top #1: Productivity	0.29	0.00	0.45	0.00	1.00
Topic Category Top #2: Developer	0.14	0.00	0.35	0.00	1.00
Topic Category Top #3: Design	0.10	0.00	0.30	0.00	1.00
Topic Category Top #4: Marketing	0.10	0.00	0.30	0.00	1.00
Topic Category Top #5: Artificial Intelligence	0.07	0.00	0.26	0.00	1.00
Topic Category Top #6: User Experience	0.05	0.00	0.23	0.00	1.00
User Statistics					
Hunter is Female	0.10	0.00	0.30	0.00	1.00
Maker Team Size	2.02	1.00	1.61	1.00	20.00
Makers At Least 1 Female	0.19	0.00	0.40	0.00	1.00
Active User Votes Female Share	0.14	0.13	0.06	0.00	0.59
Outcome Variables					
Log (1 + Monthly Page Visits) in 1 Year	6.01	6.73	3.99	0.00	20.10
Has Active User Base in 1 Year	0.757	1.00	0.429	0.00	1.00
Raises Funding Post-Launch	0.034	0.00	0.180	0.00	1.00
Log(1 + Technology Stack) Post-Launch	3.14	3.37	1.25	0.00	5.80

Notes: Descriptive statistics for the sample of 5,742 products we use in our product lunch and newsletter shock analysis. These product launches take place on the Product Hunt platform between October 4th 2016 and October 19th 2018. The sample includes featured products launched on weekdays and submitted before 7AM Pacific Time on these days – the earliest time of the day at which a newsletter could reach a user's email inbox. The sample only includes products launched on the 347 days on which the newsletter features standard product-list content.

Table A2: Estimated Effects of Newsletter Shock on Platform Participation

	Number of active female users who:			
	Visit ProductHunt		Visit a Product's Page	
	(1)	(2)	(3)	(4)
Female Newsletter	799.030*	0.220**	59.256**	0.211**
	(413.080)	(0.111)	(23.353)	(0.083)
Constant	3,179.390***	8.096***	246.493***	5.531***
	(217.596)	(0.059)	(12.387)	(0.044)
Year-Month FE	Y	Y	Y	Y
Model	OLS	QPML	OLS	QPML
Observations	4,261	4,261	4,261	4,261
R-Squared	0.529		0.572	

Notes: This table shows regression results of a linear model that estimates the effects of newsletter shock on engagement by active female users. Active users are defined as those with at least 9 followers on the platform, which excludes about 50% of all users. This allows us to exclude bot and inactive accounts which are especially important to remove when analyzing browsing data. Our data on visiting the ProductHunt platform and viewing a product page is only available for 4,261 products. We have vote data for all 5,742 entrepreneurial products in our sample. Columns 1 and 2 test if more female users visit the ProductHunt on days with female-focused newsletters and columns 3 and 4 if more female users visits the product's page. Odd columns are standard linear models and even columns use quasi-Poisson maximum likelihood models to account for the dispersed nature of the count data. All models control for year-month fixed effects, and total newsletter-suggested products fixed effects. Standard errors are clustered at the launch day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: The newsletter shock only increases the number of female votes for female-focused products. It does not impact male votes for any other product sub-sample nor does it have an impact on female voting for gender-neutral or male-focused products.

	Log (1 + Male Votes)			Log (1 + Female Votes)		
	(1)	(2)	(3)	(4)	(5)	(6)
Newsletter Shock	0.134	0.053	-0.198	-0.034	-0.024	0.334**
	(0.118)	(0.086)	(0.166)	(0.132)	(0.087)	(0.168)
Year-Month FE	Y	Y	Y	Y	Y	Y
Pre-Launch Visits & Funding Controls	Y	Y	Y	Y	Y	Y
Log (1 + Female Votes) Control	Y	Y	Y			
Log (1 + Male Votes) Control				Y	Y	Y
Gender Focus of Sample	Male	Neutral	Female	Male	Neutral	Female
Observations	1,452	2,854	1,436	1,452	2,854	1,436
R-Squared	0.843	0.852	0.796	0.835	0.847	0.787

Notes: All models control for the opposite gender logged vote count to account for non-gendered product quality differences and improve statistical power. Linear regressions with standard errors clustered at the launch date level * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: The newsletter shock does not impact the relative difference a female versus male user will vote for a more female-focused product.

	Female-Male Upvote Difference		
	(1)	(2)	(3)
Female Focus	0.001***		0.001***
	(0.000)		(0.000)
Newsletter Shock		-0.000	-0.000
		(0.001)	(0.001)
Newsletter Shock x Female Focus			0.000
			(0.001)
User FE	Y	Y	Y
Product FE	Y	Y	Y
Observations	4,261	4,261	4,261
R-Squared	0.015	0.010	0.015

Notes: The dependent variable represents the average difference in upvoting behavior between active female and male users who have viewed a product after accounting for voter and product fixed effects. The model includes the 4,261 products launched on weekdays between January 2017 and June 2018, for which the proprietary browsing data on product views are available. Linear regressions with standard errors clustered as the launch date level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: The maximum of the newsletter product gender scores best predicts female upvotes for female-focused products. The mean and share of female-focused products are less informative measures of female user engagement.

	Log (1 + Female Votes)			
	(1)	(2)	(3)	(4)
Female Focus	0.021	0.056	0.103***	0.103***
	(0.038)	(0.046)	(0.009)	(0.009)
Newsletter Shock (Max)	0.063			
	(0.089)			
Newsletter Shock (Max) x Female Focus	0.163**			
	(0.069)			
Newsletter Shock (Mean)		-0.036		
		(0.134)		
Newsletter Shock (Mean) x Female Focus		0.114		
		(0.100)		
Newsletter Shock (Top Decile Share)			-0.036	
			(0.040)	
Newsletter Shock (Top Decile Share) x Female Focus			0.030	
			(0.033)	
Newsletter Shock (Top Ventile Share)				-0.057
				(0.062)
Newsletter Shock (Top Ventile Share) x Female Focus				0.072*
				(0.037)
Log (1 + Male Votes) Control	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y
Observations	5,742	5,742	5,742	5,742
R-Squared	0.845	0.845	0.845	0.845

Notes: See the note under Table 6 and 8 for additional details on how the sample and variables are constructed. Linear regressions with standard errors clustered as the launch date level. * $p < 0.10$, ** $p < 0.05$, ***

Table A6: Regressions exploring why the maximum female-focus score of the products suggested in the newsletter best predicts female-upvotes for female-focused products. That lack of an effect for the median and minimum suggest that users are quickly skimming the newsletter seeking out the most interesting product in the newsletter, which suggests that the maximum score should be most predictive. Consistent with this "skimming" explanation the effects of the maximum are concentrated in the first five products listed in the newsletter and not in products listed sixth or after.

	Log (1 + Female Votes)					
	(1)	(2)	(3)	(4)	(5)	(6)
Female Focus	0.024	0.078*	0.110***	0.055	0.027	0.110***
	(0.038)	(0.046)	(0.029)	(0.048)	(0.037)	(0.010)
Newsletter Shock (Max)	0.009			0.044	0.025	-0.216*
	(0.077)			(0.095)	(0.079)	(0.115)
Newsletter Shock (Max) x Female Focus	0.158**			0.208**	0.152**	-0.029
	(0.069)			(0.088)	(0.068)	(0.037)
Newsletter Shock (Median)		-0.053		-0.099		
		(0.129)		(0.195)		
Newsletter Shock (Median) x Female Focus		0.063		-0.148		
		(0.100)		(0.188)		
Newsletter Shock (Min)			-0.034	0.005		
			(0.089)	(0.113)		
Newsletter Shock (Min) x Female Focus			-0.011	0.021		
			(0.076)	(0.115)		
Log (1 + Male Votes) Control	Y	Y	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y	Y	Y
Newsletter suggest product position	All	All	All	All	Top 5	Outside Top 5
Observations	5,742	5,742	5,742	5,742	5,742	5,742
R-Squared	0.844	0.844	0.844	0.844	0.844	0.844

Notes: See the note under Table 6 and 8 for additional details on how the sample and variables are constructed. Linear regressions with standard errors clustered as the launch date level. * $p < 0.10$, ** $p < 0.05$, ***

Table A7: Balance Test: Daily Newsletter Shock and Product Covariates

Newsletter Shock Balance Test on Product Launches (N = 5,742)		
	Coefficient	P-value
Hunter is Female	-0.048	0.234
Maker Team Size	-0.033	0.895
Makers at least 1 Female	-0.023	0.705
Product Gender Score	0.08	0.432
Log (Web Visits) 1 Month Before	-0.549	0.307
Pre-Launch Seed or Series A Funding	-0.012	0.585

Notes: This table shows the balance test statistics on the newsletter shock. Each row contains the coefficient estimate and p-value from regressing a pre-treatment product covariate on the newsletter shock variable (maximum gender score of newsletter-suggested products after rescaling to between 0 to 1). The regressions control for product launch year-month fixed effects, and number of newsletter-suggested products fixed effects. The product covariates include the gender of the user who submitted the product, maker team size, makers not all men, product gender score, log website visits 3 months before launch, log website visits 1 month before launch, raised funding 1 month before launch, and log total amount raised 1 month before launch. The models include 5,742 products. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Columns 1-5 show that newsletter products that are female focused look similar to male-focused newsletter products, with the exception of the share of female makers. We do not include year-month fixed effects since featured products on a given day can come from a multitude of prior dates. That said, the results are unchanged if we include month-year fixed effects. Columns 4 and 5 which include a variable capturing the share of makers who are female have fewer observations because we drop teams where we are missing a maker gender estimate. Column 6 shows that swapping our "female newsletter" measure for a "newsletter quality" measure leads to null effects. Variation in the quality of listed newsletter products does not account for our female newsletter effects.

	Suggested Product's Normalized Gender Score					Log(1 + Female Votes)
	(1)	(2)	(3)	(4)	(5)	(6)
Log(1+ Upvotes for the suggested product)	-0.001				-0.002	
	(0.002)				(0.002)	
Log(1+ Followers of makers of the suggested product)		-0.002			0.000	
		(0.001)			(0.001)	
Log(1 + Words used to described the suggested prodcut)			-0.002		0.001	
			(0.002)		(0.003)	
Share of makers who are female				0.057***	0.056***	
				(0.011)	(0.011)	
Female Focus						0.183***
						(0.045)
Newsletter Quality Shock						0.01
						-0.009
Newsletter Quality Shock x Female Focus						-0.0142
						-0.008
Year-Month FE	N	N	N	N	N	Y
Observations	2,355	2,365	2,365	1,760	1,755	5,742
R-Squared	0.000	0.002	0.001	0.024	0.024	0.844

Notes: Columns 1-5 includes all products *featured* in the daily newsletters analyzed in Table 5. Column 6 includes all launched products in the sample we analyze in Table 6. Linear regressions with Standard errors are clustered at the launch day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Regressions showing the impact of the newsletter shock on page visits for products with at least one female maker on the team. While female launched male-focused and gender-neutral products appear to benefit from launching on Product Hunt, these products do not benefit from the female newsletter shock. In contrast, female launched female-focused products do see more page visits when launched on days with more female newsletters.

	Log (1 + Monthly Page Visits)		
	(1)	(2)	(3)
Post-Launch	2.337***	1.818***	0.930
	(0.785)	(0.664)	(0.684)
Post-Launch X Newsletter Shock	0.907	1.882	3.216**
	(1.464)	(1.284)	(1.299)
Product FE	Y	Y	Y
Has a female maker?	Y	Y	Y
Gender Focus of Sample	Male	Neutral	Female
Observations	7,016	13,094	9,292
R-Squared	0.741	0.740	0.739

Notes: Observations are month-product observations as in Table 6. Linear regressions with standard errors clustered at the launch date level * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: The impact of the the newsletter on female-focused product page visits, funding, and technology stack size hold even when controlling for the total number of votes the product received.

	Log (1 + Monthly Page Visits)		Raises Funding Post-Launch		Log(1+ Tech Stack)	
	(1)	(2)	(3)	(4)	(5)	(6)
Female Focus	-0.878***	-0.776***	-0.046**	-0.045**	-0.282**	-0.270**
	(0.204)	(0.204)	(0.018)	(0.018)	(0.115)	(0.114)
Newsletter Shock	0.571	0.480	0.025	0.024	-0.097	-0.112
	(0.422)	(0.404)	(0.024)	(0.024)	(0.181)	(0.180)
Newsletter Shock x Female Focus	1.280***	1.291***	0.081**	0.081**	0.567***	0.566***
	(0.367)	(0.370)	(0.036)	(0.036)	(0.213)	(0.211)
Log (1+ All Votes)		0.881***		0.005**		0.090***
		(0.039)		(0.002)		(0.017)
Product FE	Y	Y				
Launch Date FE	Y	Y				
Year-Month FE			Y	Y	Y	Y
Pre-Launch Visits & Funding Controls			Y	Y	Y	Y
Observations	101,803	101,803	5,742	5,742	5,312	5,312
R-Squared	0.721	0.731	0.232	0.232	0.134	0.139

Notes: For additional details on Models 1-2 see Table 6, for Models 3-4 see Table 8, and for Models 5-6 see Table 9. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Products missing BuiltWith technology stack data are smaller and are less likely to have raised funding pre-launch, but are no different in terms of their gender focus or the newsletter shock they experience.

	Technology Stack Data Avaliable		
	(1)	(2)	(3)
Log (1 + Monthly Page Visits) 1 Month Before			0.014***
			(0.001)
Pre-Launch Seed or Series A Funding			0.036***
			(0.009)
Newsletter Shock	-0.058		-0.047
	(0.039)		(0.040)
Female Focus		-0.003	0.015
		(0.006)	(0.037)
Newsletter Shock x Female Focus			-0.036
			(0.072)
Year-Month FE	Y	Y	Y
Sample	All	All	All
Observations	5,742	5,742	5,742
R-Squared	0.010	0.009	0.055

Notes: Estimates from a linear probability model using our sample of 5,742 entrepreneurial product launches. The outcome variable measures if we have technology stack information for the product from BuiltWith. All models include fixed effects for the the month-year of launch and the number of products linked to in the newsletter. Standard errors are clustered at the launch day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Estimated effects of female newsletter by product's female focus shock on on whether post-launch the team raises venture funding as of October 2020.

	Raises Funding Post-Launch	
	(1)	(2)
Log (1 + Monthly Page Visits) 1 Month Before	0.003***	0.002***
	(0.001)	(0.001)
Pre-Launch Seed or Series A Funding	0.506***	0.540***
	(0.038)	(0.048)
Newsletter Shock	0.018	0.028
	(0.029)	(0.032)
Female Product (Top Quartile)	-0.048**	-0.043*
	(0.023)	(0.024)
Male Product (Bottom Quartile)	0.006	0.016
	(0.029)	(0.038)
Newsletter Shock x Female Product (Top Quartile)	0.076*	0.062
	(0.044)	(0.046)
Newsletter Shock x Male Product (Bottom Quartile)	-0.020	-0.028
	(0.055)	(0.073)
Year-Month FE	Y	Y
Sample	All	Male Makers
Observations	5,742	5,742
R-Squared	0.230	0.231

Notes: Estimates from a linear probability model using our sample of 5,742 entrepreneurial product launches. The outcome variable measures whether the firm raised venture funding between when it launched on ProductHunt and October 2020. All models include fixed effects for the the month-year of launch and the number of products linked to in the newsletter. Standard errors are clustered at the launch day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Estimated effects of female newsletter by product's female focus shock on the team's web technology investments.

	Log (1 + Technology Stack) Post-Launch	
	(1)	(2)
Log (1 + Monthly Page Visits) 1 Month Before	0.105***	0.100***
	(0.004)	(0.005)
Pre-Launch Seed or Series A Funding	0.631***	0.754***
	(0.091)	(0.099)
Newsletter Shock	-0.213	-0.257
	(0.244)	(0.288)
Female Product (Top Quartile)	-0.523**	-0.367
	(0.223)	(0.270)
Male Product (Bottom Quartile)	-0.024	0.291
	(0.274)	(0.323)
Newsletter Shock x Female Product (Top Quartile)	0.926**	0.608
	(0.427)	(0.520)
Newsletter Shock x Male Product (Bottom Quartile)	-0.176	-0.749
	(0.507)	(0.601)
Year-Month FE	Y	Y
Sample	All	Male Makers
Observations	5,312	3,748
R-Squared	0.135	0.132

Notes: Estimates from an OLS model using our sample of 5,312 entrepreneurial product launches for which we have technology stack data. The outcome variable measures the (logged) number of active web technologies on the startup's website as of October 2020. All models include fixed effects for the the month-year of launch and the number of products linked to in the newsletter. Standard errors are clustered at the launch day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: The newsletter shock does not shift the quantity nor quality of comments from female users.

	Has Female Comment?		Log(1 + Female Comment)		Log # Words Female Comment (Mean)		Sentiment of Female Comments (Mean)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female Focus	0.049***	0.098**	0.058***	0.122***	0.049*	0.091	0.020	-0.023
	(0.009)	(0.042)	(0.009)	(0.041)	(0.029)	(0.178)	(0.012)	(0.055)
Newsletter Shock		-0.059		-0.095		0.192		-0.010
		(0.076)		(0.069)		(0.181)		(0.069)
Newsletter Shock		-0.094		-0.122		-0.079		0.081
x Female Focus		(0.080)		(0.076)		(0.350)		(0.105)
Year-Month FE	Y	Y	Y	Y	Y	Y	Y	Y
Num Male Comments	Y	Y	Y	Y	Y	Y	Y	Y
Sample	All	All	All	All	Has Female Comment		Has Female Comment	
Observations	5,134	5,134	5,134	5,134	2,216	2,216	2,216	2,216
R-Squared	0.100	0.101	0.148	0.148	0.023	0.023	0.032	0.032

Notes: All dependent variables are measures of comments from female users. Columns 1-4 reflect "extensive" margin effects on the number of comments from female users. Columns 5-8 only include products with at least one female comment and so reflect the "intensive" margin of female comment length and sentiment. Linear regressions with standard errors clustered at the launch date level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A15: The impact of the newsletter shock on entrepreneurs with prior launch experience and who are more active on Product Hunt. Columns 1 and 2 show that having launched on Product Hunt before does not mitigate or strengthen the size of the newsletter shock nor the female-focused growth gap. In contrast, we find stark differences between startups that have been active voters on Product Hunt versus teams that have engaged less with the platform. Teams that have voted on more products in the past show much larger product gender gaps and are much more impacted by the newsletter shock.

	Log (1 + Monthly Page Visits)			
	(1)	(2)	(3)	(4)
Post-Launch	3.096***	2.874***	3.096***	2.874***
	(0.342)	(0.289)	(0.342)	(0.289)
Post-Launch x Female Focus	-1.057***	-0.706***	-1.057***	-0.706***
	(0.356)	(0.269)	(0.356)	(0.269)
Post-Launch x Newsletter Shock	0.649	0.442	0.649	0.442
	(0.655)	(0.554)	(0.655)	(0.554)
Post-Launch x Newsletter Shock x Female Focus	1.602**	0.982**	1.602**	0.982**
	(0.645)	(0.486)	(0.645)	(0.486)
Product FE & Year-Month FE	Y	Y	Y	Y
Teams with more platform engagement (20+ prior votes)	Y	N		
Team has funding or launching experience?			Y	N
Observations	51,920	49,883	51,920	49,883
R-Squared	0.739	0.698	0.739	0.698

Notes: See 6 for further information on the variables and sample construction. Observations are at the product-month level. Linear regressions with standard errors clustered at the launch date level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.