Biased Sampling of Early Users and the Direction of Startup Innovation

Ruiqing Cao
Rembrand Koning
Ramana Nanda

Biased Sampling of Early Users and the Direction of Startup Innovation

Ruiqing Cao
Harvard Business School

Rembrand Koning
Harvard Business School

Ramana Nanda
Harvard Business School

Working Paper 21-059

Biased sampling of early users
and the direction of startup innovation[*]

RUIQING CAO

Harvard Business School

REMBRAND KONING

Harvard Business School

RAMANA NANDA

Harvard Business School

November, 2020

Abstract

Using data from a prominent online platform for launching new digital products, we document that the composition of the platform's 'beta testers' on the day a new product is launched has a systematic and persistent impact on the venture's success. Specifically, we use word embedding methods to classify products launched on this platform as more or less focused on the needs of female customers, and show that female-focused products launched on a typical day – when nine-in-ten users on the platform are men – experience up to 40% less growth a year after launch. By isolating variation in the composition of beta testers that is unrelated to the characteristics of launched products, we further show that a startup's user growth and success with VC financing can be traced to variation in the number of female beta-testers on the platform the day it's product is launched. Overall, our findings suggest that the composition of early users can create systematic bias in signals of a startup's market potential, with consequential real effects – including a dearth of innovations aimed at consumers who are underrepresented among early-users.

1 Introduction

Given the irreducible uncertainty associated with founding new ventures (Hayek 1948), entrepreneurship is increasingly practiced and modeled as a process of experimentation. The value of this approach stems from the ability to learn from early tests in order to optimize decision-making for the next stage of a venture's development (Ries 2011; Kerr, Nanda, and Rhodes-Kropf 2014; Gans, Stern, and Wu 2019). A growing body of research in this vein has also highlighted the value of a scientific approach and tools like A/B testing that help startups run more informative tests (Camuffo et al. 2020; Koning, Hasan, and Chatterji 2019; Thomke 2020).

This literature has largely taken for granted an important assumption that makes experimentation valuable: that tests conducted by entrepreneurs will produce unbiased – albeit noisy – signals of how promising the underlying startup idea is. While a single test might overstate or understate the value of an idea, a collection of noisy but unbiased signals will inevitably produce a consistent picture, allowing the entrepreneur to correctly learn about the market potential of her idea. On the other hand, biased signals will lead some types of entrepreneurs to *systematically* abandon promising ideas too early while others will persist for too long. To the extent that a few gatekeepers account for the majority of early feedback to startups, such systematic differences in firm exit due to biased feedback can have wider, society-level consequences. For example, it can lead to a dearth of firms catering to a certain demographic or customer segments, completely independent of drivers such as prejudice.[1]

Why might the results from entrepreneurial experiments be systematically biased? As with medical trials and RCTs in the social sciences, an often overlooked but crucial source of bias stems from the composition of participants in the experiment (Allcott 2015; Deaton

---

[1]For example, in learning models such as Jovanovic (1982), unbiased but noisy signals can lead to inefficiently early (or late) exit by any individual firm, but will not lead to systematic differences in the types of firms that exit inefficiently. Such systematic differences will arise from the systematic bias in the signals received.

and Cartwright 2018). For example, after decades of excluding women from medical trials, the FDA now suggests that trials strive for demographic representativeness to ensure that the medical treatments brought to market serve the needs of the entire population (Office of Research on Women's Health 2016; Dusenbery 2018). Relatedly, recent work shows that speech recognition algorithms fail more often when listening to African Americans, likely because the training data these algorithms were tested on included few black voices (Koenecke et al. 2020).

Returning to entrepreneurial experiments, if early adopters and initial users are not representative of the larger set of consumers, then measures of early 'traction' from even the most internally valid A/B tests will lead entrepreneurs to incorrect inferences about how consumers value their new product or service. Here we focus on gender and test if the over-representation of men amongst early users constrains the growth and success of startups that are especially focused on women users and consumers.

Our setting is the online platform Product Hunt which plays a prominent role in enabling early adopters of digital products to share and vote on new products and venture ideas.[2] Each day, dozens of products are "launched" and voted on by the Product Hunt community. These launches help entrepreneurs discover users for their early stage products, serve as a signal to investors (Cao 2019), and help entrepreneurs get feedback on the potential of their ideas from the notable founders, developers, and designers who use regularly use the platform. Product Hunt has enabled tens-of-thousands of entrepreneurs to quickly test their ideas and served as a launchpad for products like RobinHood, Checkr, Gimlet Media, Cheddar, and Gigster. Yet early users on platforms, like Product Hunt, that serve the needs of early stage ventures are disproportionately dominated by men: 75% of visitors to Kickstarter are men, 67% for Indiegogo, and 79% of visitors to YCombinator's influential platform HackerNews are men.[3] In the case of Product Hunt, 90% of

---

[2]Product Hunt is a a subsidiary of fund raising and hiring platform AngelList.

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

voters are men and nearly 80% of products launched are built by all male teams. Given that Product Hunt users come from the ranks of Silicon Valley engineers, managers, founders, and Venture Capitalists these percentages are to be expected (Gompers and Wang 2017; Guzman and Kacperczyk 2019).

Our analysis of whether the over-representation of male beta testers on the Product Hunt platform impacts the growth and survival of female-focused startups proceeds in four broad steps. First, as described in Section 3, we create a measure of the degree to which a product is likely to cater or appeal to women consumers, using word embedding techniques enabled by recent advances in machine learning. In doing so, we create and validate an *ex ante* measure of a product's appeal to women consumers allowing us to compare whether nascent female-focused products are less likely to gain traction and grow. Instead of relying on realized market-share, which confounds potential appeal with realized success, our ex-ante measure allows us to directly identify female-focused ventures. We believe that our approach–of mapping text to an interpretable underlying dimension–has wider applicability in settings ranging from the study of startup strategy to firm culture to measuring the direction of innovation (Srivastava et al. 2018; Koning, Samila, and Ferguson 2019; Guzman and Li 2019).

Second, we use data from Product Hunt that measures the products users browse, vote for, and pass over to show the presence of gender-based preferences for new products among the early users on the platform. When compared to the most-male products, the most female-focused products are roughly 20% to 25% more likely to be upvoted by women than men. We show that this association remains present after accounting for a whole range of observable covariates, as well as product- and individual-beta-tester fixed effects. This implies that the relationship we observe is not driven by fixed differences in the quality of products or the harshness of particular reviewers.

Importantly, the channel we identify is distinct from one in which early users are biased against female founders (cf. Ewens and Townsend 2020). In fact, we show that the gen-

4

dered differences in preferences for products are also present for all-male entrepreneurial teams launching new female-focused products. The fact that many men prefer products that are predicted to cater to men (and vice-versa) is unsurprising. However, this association not only provides additional validation of our measure but also enables a quantitative assessment of how these preferences can lead to biased feedback from early users.

Third, we show that startups commercializing female-focused products perform at least as well as male-focused products before their Product Hunt launch. However, after launching on Product Hunt, a gender gap emerges: on average, startups launching female-focused products experience roughly 40% lower growth in the year after they launch on Product Hunt. Further, they are 5 percentage points less likely have active users suggesting that female-focused startups are more likely to have failed. These findings continue to hold even after controlling for differences in time-invariant product quality and allowing for time-varying impacts of the team's gender composition.

The systematically worse performance of female focused products post-launch is consistent with a view in which the over-representation of men distorts the direction of product innovation towards more male preferences. By lowering the cost of acquiring male early users Product Hunt might make it easier for male-focused technology startups to grow. Beyond this direct effect, post-launch growth might also serve as a signal to the entrepreneur and to investors about the potential of the idea causing both groups to under-invest in promising female-focused ideas. However, the results above are also consistent with a more pedestrian explanation: namely, that women oriented products just have a systematically harder time gaining traction in the market overall. For example, perhaps the nature of demand or the competitive environment facing these startups makes it more costly for startups aimed at women to gain new users (e.g. Lambrecht and Tucker (2019)). Perhaps this fact – rather than the composition of the beta testers and related biased signals – accounts for the systematic differences we observe in the post-launch outcome of women-oriented products.

We address this concern in our fourth step by isolating exogenous variation in the number of women who engage with the platform that is unrelated to the quality of the startups launched that day. If our findings are driven by entrepreneurs testing on biased samples, then these exogenous 'shocks' that increase the representation of female users on the platform on certain days should reduce the post-launch product-gender gap among startups launched on *those*, but not other, days.

We isolate exogenous variation in female engagement using Product Hunt's daily newsletter, which gives updates on *previously* launched products.[4] Our key insight—which we validate with proprietary data from Product Hunt—is that a daily newsletter that gives an update on particularly female-focused product will be of more interest to women and so will cause disproportionately more women to visit platform that day. Crucially for our research design, the newsletter goes out at 7 a.m. Pacific Time, which is after the vast majority of products are launched for the day.[5] Our identification is predicated on the assumption that newsletters with updates on more female-focused products will increase the female engagement with the platform on those days in a manner that is plausibly exogenous to the types of products launched that day. Indeed, a balance test finds no observable differences in the products launched on days where the newsletter features a very female-focused product as opposed to a very male-focused product.

We find that on days when the newsletter is predicted to drive more female engagement with the platform, the gap we typically see in the traction and survival of female-oriented products shrinks towards zero. On the other hand, the gap is even larger on days when more male beta testers are predicted to be on the platform. This provides strong evidence in favor of the view that our results are driven by the composition of who is on the platform rather than differences in the types of products catering to women.

Beyond the direct impact on post-launch startup growth, we also test if the composi-

---

[4]Note that our sample specifically excludes the featured products from our analysis so as not to contaminate the results.

[5]We exclude the small percentage of products launched after the newsletter is sent from our analysis.

tion of beta testers impacts investor and entrepreneur decision making. A key implication of entrepreneurship as experimentation is that early signals are used to guide future investments. When the results of a new user survey or A/B tests show that there is little demand for an innovation the entrepreneur should be more likely to terminate her venture (Koning, Hasan, and Chatterji 2019). When an early test is an unexpected success she should invest more time and effort into developing the business (Camuffo et al. 2020). Similarly, venture capitalists should be more willing to invest when the startup has stronger signals of traction and user growth.

Consistent with the launch-as-signal view, we find that entrepreneurs who launch female-focused products on days when the platform is especially male dominated are 4 percentage points less likely to raise funding post-launch. Using data on the startup's technology stack we also find that these entrepreneurs put 30% less effort into product development. For entrepreneurs who happen to launch female-focused products on days when the newsletter brings in more women we find that these gaps close. Overall, we find evidence that "sampling bias" shapes startup growth by directly impacting user acquisition and by distorting the signals entrepreneurs and investors have of a startup's potential.

Our results—that there is a product gender gap and that sampling bias appears to be at least partly responsible—are relevant to scholars and practitioners interested in entrepreneurship, innovation, and gender. Firstly, our findings enrich our understanding of the benefits and costs of experimental strategies (Levinthal 2017; Camuffo et al. 2020; Gans, Stern, and Wu 2019; Koning, Hasan, and Chatterji 2019). While prior work has largely focused on the benefits of business experimentation, our findings shed light on when such strategies may fail and highlights how shifting the representativeness of the test sample may mitigate these failures. 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 appear to be oriented towards the needs of men over women (Koning, Samila, and Ferguson 2019; Feng and Jaravel 2019; Koning, Samila, and Ferguson 2020). Our results 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 cannot observe what the founders beliefs and strategic decisions are over this time 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 the Product Hunt signal is strong enough that they abandon the idea? Is it that alternative means of discovering demand are harder to find for female-focused startups? If so, given competition between platforms, why do we observe so few platforms with a majority of women? These and other questions appear important, given the growing importance of entrepreneurial testing as well as the dominance of a few platforms, that de-facto serve as gatekeepers to the types of ideas that are ultimately successful.

The rest of our paper is organized as follows. In the next section, we provide further details about our empirical context. We then describe and validate our text-based measure of a product's gender appeal before turning to our outcome measures and basic descriptive statistics. We then present evidence that there is a product-gender gap and that newsletters which bring in more female users close this gap and in so doing impact entrepreneur and investor behavior. Finally, we conclude with a discussion of limitations and where we see opportunities for further work.

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 with one another 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). While the majority of 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.[6] 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. Occasionally, a product that is posted will be asked to make a few changes to the descriptive texts, images, or videos, and the featuring may be delayed to the next day. Our study focuses on these featured products.

A product submission includes photos, sometimes videos, a detailed text description

---

[6]Products are submitted to the platform by power users called "hunters.", 40% of the time the hunter is the same person who build the product (makers). In the other cases, 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.

of the product, and links to the product's own website. It also includes a profile picture and name of each 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 which receive more votes are displayed prominently at the top of the page. Again, this incentives entrepreneurs to post as early in the day as possible so as 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 at later dates, and success on the platform often piques the interest of journalists and investors.

2.2 Data

Product Hunt We have two years of 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 describing 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 product launched by technology incumbents or posts that are not actually new technology products.

First, we restrict our sample to featured technology products launched on Product Hunt.[7] We make sure to discard an 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

---

[7]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.

consistency in the types of products being examined across our sample period.

Second, as noted above, Product Hunt is also used by large companies such as Stripe, Microsoft, and LinkedIn to feature new technology product releases. However, since we are interested in studying startups we restrict our sample to only include the set of products that are launched by early-stage entrepreneurial firms. 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.[8] Since companies sometimes post the platform multiple times we restrict our analysis to the first post from a given website domain in order to look at new ventures and not subsequent iterations from already established companies. To further ensure our sample consists of early stage 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, and so the best rank, they can on the launch day. Further, by only retaining products submitted before 7AM Pacific Time we ensure the product was submitted before the daily newsletter was sent out. By restricting our sample to products launched before the newsletter goes out, we avoid including products that might strategically decide to launch or not because of the newsletter content.

Fourth, we restrict our analysis to the days in which a "product update" newsletter is sent out to the Product Hunt community. During the week, Product Hunt sends out a daily newsletter just after 7AM Pacific Time. The majority of these newsletters provide

---

[8]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 finance, or those who have only raised Pre-Seed, Seed, Accelerator, Convertible Notes, Angel or Series A financing.

updates on a handful of products that have 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 on those days in a manner that is unrelated to the types of products launched that day.[9]

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 a number of 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 as unknown.

SimilarWeb To measure venture growth we merge in 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, key word effectiveness, and a host of other digital trends. Using data from ISPs and 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 of digital startups' initial traction, and predictor of future investment and revenue (Koning, Hasan, and Chatterji 2019). 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,

---

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

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 (2019) in the appendix we show that financing and page visits are strongly correlated, with the bottom decile of visits in our sample raising no funding and the top decile having a nearly 15% chance of raising venture financing.

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 have 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 (2019) BuiltWith tracks the the technologies startups use to run their websites. Since many of these technologies need to be "client" facing BuiltWith simply scrapes the website to see if it is using Google Analytics, Optimizlely A/B testing, Facebook tracking pixels, Shopify payment tools, and a myriad of other technologies. Using BuiltWith's free API we were able to measure 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 to develop the idea or if they instead halted their efforts. In the appendix we show that products in the top decile of technology stack size have a one-in-five chance of raising venture funding where as those in the bottom decile have essentially no chance of raising venture financing. Product development efforts appear to correlate with more traditional measures of venture growth and success.

3 Product Gender Focus

3.1 Creating a Measure of a Product's Gender Focus

To investigate whether there is a product-gender gap, we require a measure of whether a product is directed more to women, more to men, or both sexes equally – in a manner that is independent of the realized outcome for these products.

We create such a measure by analyzing the product's text description on the day it is launched. Our approach is similar to Koning, Samila, and Ferguson (2019), 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 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 unidimensional 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. Note that the comment is also an ex-ante measures 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. 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 largely for new products made by individual makers and non-public firms. Appendix Figure A.1 shows an example of each piece of text used to construct the measure of gender focus.

Using this text we first remove common stopwords[10] and we keep only nouns, verbs,

---

[10]The stopwords are a union of the following lists: http://www.ranks.nl/stopwords;

and adjectives.[11] For the remaining words we then use a pre-trained word embedding model[12] 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 $\boldsymbol{v}_f$

---

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.

[11]Hoberg and Phillips (2016) keeps only nouns in product descriptions in 10Ks to construct word vectors. We include also 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.

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

and $\boldsymbol{v}_m$:

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

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 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), using texts of all products launched on Product Hunt as the corpus.[13] 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 female keyword. Since each of these three keyword pairs likely capture slightly different and idiosyncratic meanings we take 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 A.6. The histogram is bell-shaped and symmetric around the mean, but the tails are wider than those of a standard normal distribution.

---

[13]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.

3.2 Validating our Gender Focus Measure

3.2.1 Face Validity

Our measure of a product's gender focus is created for all products on the platform and we validate this measure in three main 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 content from around the web" is also consistent with being completely gender neutral. Finally, you see that at the 75th percentile you see 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.

3.2.2 Supply Side

Second, we validate the measure by checking whether, as we might expect from the literature, there is a positive correlation between our measure of a product's gender focus and

other observable correlates of the product. For example, recent work shows that female-focused products are much more likely to come from women inventors and entrepreneurs (Feng and Jaravel 2019; Koning, Samila, and Ferguson 2019). In Figure 1, we document 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. As can be seen from Figure 1, 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.

3.2.3 Demand Side

Third we provide evidence that male and female users respond differently to the same products in a manner that is consistent with gendered preferences. For a subset of the products in our data, we have 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 dataset 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 then he or she is 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. Along with our measure of the product's gender score we also have the imputed gender for each user. Appendix Section B describes how we construct this preference data in detail. Our sample is the larger sample of products we analyze in Figure 1, but because we only have proprietary data for part of our sample's time period we end up with 11,212 products in our sample.

Using this data we test if female users are more likely to vote for female-focused

products even after controlling for user- and product-level fixed effects. User-level fixed effects account for differences in how "harsh" different voters are, including differences by gender. Product-level fixed effects account for observed and unobserved quality differences across products.[14] The left panel in Figure 2 presents a binned scatter of the female-user-minus-male-user residuals plotted against the product's gender focus .The plot also controls for log number of words in product texts, launch year-month fixed effects and day-of-week fixed effects. The strong upward slope indicates that, even after accounting for quality differences, 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 by between female and male voters by 0.3 percentage points. Relative to the baseline, this represents an effect of about 20-25% depending on whether women upvote more or men merely upvote female-focused products less. Again, we find the same pattern holds when we restrict the sample to products in our sample of 5,742 new product-startup launches.

4 Results

4.1 Descriptive Statistics

Table 2 shows basic descriptive statistics for the 5,742 products used in our analysis, and also compares products that are in the quartile of products that are most female focused to those in the quartile of products that are most male focused. Appendix Table A.1 presents additional descriptive statistics.

Product Hunt's topic categories reveal that the products span a large number of topics

---

[14]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.

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 women on them.

Unsurprisingly, products catering 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 is able to 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.

4.2 Post-Launch performance of female-focused vs. male-focused startups

We begin our analysis by estimating the following simple difference-in-differences 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.

Y_t is the the log-pageviews for a product in a month and $POST_t$ is an indicator which takes the value of 1 after the product has launched and and zero otherwise. The model includes product fixed effects, γ_i, and year-month fixed effects, δ_t. 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 particularly high or low. The coefficient $\beta_1 POST_m$, which measures the average increase in visitor growth for products

20

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 in $\beta_2 GenderFocus_i \times POST_t$ which captures the degree to which this visitor growth varies post-launch when a product is more female focused.

We report the results of this regression in Panel A of Table 3. Looking first at the coefficient on the post-launch indicator, 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. Importantly, 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. The size of the relationship is stable when we control for the gender composition of the maker team. This suggests that our findings are not merely capturing bias towards the gender of the entrepreneur, but differences in the appeal of the product. In fact, as seen from column (5), the relationship continues to hold, and if anything becomes slightly stronger when looking only at the subset of startups with all-male founders. It is worth noting that, consistent with prior work, we find products built by women teams perform less well after launch, with growth rates roughly 35% of male only teams. There is both a product and founder gender gap.

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 quartile of female-focused products and the bottom quartile of female-focused products to the middle two quartiles of our product-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.

As can be seen from Panel B, the most male focused products do not experience any difference in web traffic from the products that are more gender neutral However, on average, the most female-focused products experience nearly 45% lower web traffic post launch. As with Panel A, we see this relationship remains robust to the inclusion of further controls and moreover, is equally strong when looking only at teams with all-male entrepreneurship.

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 growth 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 and so while the difference between male and female products is visually small, it actually represents substantive differences in relative growth. Figure 5 shows the 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.

Figures 4 and 6 report the difference between gender neural products and male- and female-focused products. The picture emerging from these plots is consistent with that seen in Table 3, with the added understanding that there doesn't appear to be any difference in growth of web-traffic for these products before the launch, and that the female-focused startups experience a persistently lower growth in web traffic, which is visible right from the very first month on the launch. We find no evidence that the most male-focused products experience less growth post launch. Our findings do not merely seem to reflect limited growth opportunities for strongly gendered products. Rather they are

22

distinctly related to the product being female-focused.

In Table 4 we dichotomize our measure of monthly web traction to examine the impacts of launch on whether the startup has an active user base or not. Startups with no visits are much more likely to have failed. Specifically, we define a startup as active if it has more than zero visitors in a month according to SimilarWeb. In Table 3, we see that months post-launch are positively associated with having an active user base consistent with the idea that many startups use ProductHunt to gain their earliest users. However, the benefits for female-focused startups appear to be smaller, with female-focused startups being 5% percentage points less likely to be active post-launch. As before, these results are present in both the parametric and non-parametric specifications, as well as with all-maker teams. It appears female-focused products are more likely to end up with no zeros suggesting that they are more likely to have failed.

4.3 Are the results driven by the composition of the platform's beta-testers?

Thus far, we have found that after launching on Product Hunt, a gap appears in the traction experienced by female-focused startups relative to other products, which is not merely reducible to the gender of the entrepreneur. It appears there is a gender gap in the product market *independent of who makes the product.*

That said, while the findings are consistent with a gap emerging because of skewed gender representation on the Product Hunt platform, our analysis has thus far only looked at growth before and after launching on the platform. Though suggestive, it does not tell us if a shift in the gender composition of who is active on Product Hunt would lead to different products succeeding. For example, something else – such as the overall market size or amount of competition facing startups addressing this market might systematically lead such ventures to under perform relative to others.

An ideal experiment to convincingly tell these mechanisms apart would entail randomly shifting the gender composition of who is on the platform, in a manner unrelated to the products launching on the different days. If female-focused startups launching on days when there were more women on the platform experienced different growth and survival trajectories from those launched on 'normal' days, this would provide convincing evidence that the gap was driven by "sampling bias" and not other unobserved factors. Below we outline how we use Product Hunt's daily newsletter to isolate plausibly exogenous variation in who is engaging with the platform on a given day.

Specifically, we use the fact that Product Hunt's newsletter regularly features prior products on the platform—allowing us to measure the gender appeal of each newsletter— and is designed to increase engagement, interest, and traffic to the website. Our argument is that the newsletter, which is sent to out between 7 and 11AM Pacific Time, will bring *systematically different types of users* to the website on any given day, due to differences in the types of products that are featured. For example, on a day when the newsletter features content on Fintech, one would imagine more users interested in Fintech would also browse the Product Hunt website that day. In turn, if a Fintech product happened to launch on that day, we would expect they would see a boost in subsequent growth. In our case, when the Newsletter content is likely to draw in more women we would expect that female focused products should experience stronger post-launch growth.

What sort of content might draw more women onto the platform? Building on the word embedding analysis above, newsletters which discuss past products that are more female-focused would be likely to bring more women onto the platform that day. Fortunately, the daily newsletter features both free-form content and updates and information on an average of four products launched on the platform in the past.[15] If one of these products is

---

[15]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.

particularly appealing to female users it should increase engagement by women. Further, since the newsletter comes out after products are launched, it is extremely unlikely that product launches can be systematically related to the content of the newsletter. Indeed, in Appendix Figure A.3 we present two example newsletters that are likely to bring more women onto the platform: a sponsored post (by the birth controls startup Nurx) and coverage of a startup's acquisition (Lululemon's purchase of Mirror).

Building on these examples, we calculate a variable "Female Newsletter" that is the gender score of the most female-focused product listed in the newsletter that day. To aid in interpretation of the triple interactions we will run, we then rescale this variable to have a minimum of 0 and a maximum of 1. Table 5 shows summary statistics for each of the 419 newsletters in our sample.

Consistent with idea that products launching on the platform are not strategically gaming the newsletter in the appendix Table A.2 presents a formal balance table showing that 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. 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. f

Do "Female Newsletters" actually draw in more female users? Ideally, we could instrument female share browsing the website using the female newsletters data. Unfortunately, measurement and limitations prevent us from doing so. First, the browsing data we have from ProductHunt only partially covers the days in our sample. Restricting to the days where we have browsing data significantly reduces our power. Second, even if we had full data from ProductHunt a large number of users visit the website without being logged in. For these users, we are not able to estimate their gender composition. Especially if we think women users may be less engaged, and so more likely to visit Product Hunt as logged out users, we would expect the browsing data underestimate shifts in who visits

the website due to the newsletter. Third and finally, the newsletter might also shift the composition of men coming to the platform as well, brining on men whose preferences are closer to the preferences of female users. Thus while we think there the newsletter shifts who engages with the platform, we don't think its effect only flows through the number of logged in female users on the platform that day (i.e., the exclusion restriction doesn't hold).

That said, in Appendix Table A.3 we validate the idea that the newsletter can shift who is engaging with the platform. On days with female-focused newsletters we find that 800 more female active users visit the homepage, 60 more visit each product's page, and 7 more vote for a product. These reflect relative increases of roughly 25%. Overall, these findings strongly support our argument that the newsletter exogenously shifts who is participating on the platform on any given 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 including our measure of how "female-focused" the newsletter is. Specifically, we run the regression:

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

where $NewsletterShock_i$ is measured as the maximum female focus (after rescaling to between 0 and 1) 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 the female-focus on post-launch growth. 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 estimate difference between a 75th percentile female-focused product and a median product at different quartiles of the newsletter shock. 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 a gender neutral product. Female focused products launched on days where more men are brought onto the platform suffer even larger growth penalties compared to 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.

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 the long term outcomes of the ventures.

4.4   Impact on investor funding and entrepreneur effort

Our findings thus far show that the growth of new female-focused products is stunted when early adopters are disproportionately male. Moreover, this discount is not limited to growth on the platform nor is it fleeting. Instead, the product gender gap is apparent in a startup's overall growth and persists for at least a year after launching on Product Hunt.

An interesting implication of our results is that since launch 'success' is publicly visible to the broader community, it has the potential to be used as a signal of an idea's viability, not just for the entrepreneur, but also other actors in the ecosystem such as investors. A 'biased signal' stemming from the composition of users on the platform could cause investors to underestimate and overlook a promising opportunity. Further, even if entrepreneurs and investors know that female-focused products struggle to gain early users they might still be more likely pass on these ideas. With less growth and fewer users the signal of a startup's potential is inherently less certain. This uncertainty, in turn, might be enough for an investor to overlook a promising female-focused product even if she has "corrected" for the product growth gender gap.

We test if sampling bias impacts investor and entrepreneur decision making using 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? We collected funding data from Preqin and CrunchBase as of October 2020. Our measure of funding covers rounds raised from institutional sources including venture capitalists and accelerators.

Our measure of product development effort comes from BuiltWith's technology stack database. If entrepreneurs are actively developing their product they should be adding new technologies to their website which in turn should increase the size of the website's technology stack. Instead, if the entrepreneur is putting less effort into her idea we should see less development and so likely fewer technologies being used on the website. Again, our data is from October 2020.

Unlike our monthly web traffic measure, our measures of funding and the technology stack 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

28

of product fixed effects, we do control for the number of page visits in the month before the product launched and for whether the product startup had raised venture funding before launching. Though imperfect, these controls do allow us to account for differences in quality between more and less female-focused products. These models do include fixed effects for the year-month of launch to account for the fact that startups launched at the start of our sample period have more time to raise venture funding than those launched at the end.

In Table 8 we test if sampling bias shapes funding decisions. 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 also 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 A.5 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 statistical significant. Column 2 restricts the sample to 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 precent of products raise funding post-launch the magnitudes of these effects are economically significant.

In Table 9 we test if sampling bias shapes an entrepreneur's product development effort as measured by the size of the technology stack.[16] We log the dependent variable to ac-

---

[16]Our technology stack models only include 5,312 products for which BuiltWith had current technology stack information for. In Appendix Table A.4 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.

count for the skewed distribution and to aid in interpretation.[17] Again, in In Column 1, we find that entrepreneurs who launched a standard deviation more female-focused products have nearly 30% smaller technology stacks. In Appendix A.6 we find that the top quartile of female-focused products have 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 just look at products launched by all male teams. Overall, it appears that entrepreneurs put less effort into product development when launching female-focused products on days when the platform is especially dominated by male users.

5   Conclusion

Building on the role of sample bias in driving failures of external validity, we have argued that gender imbalances among early users have the potential to systematically impact the growth and survival of new ventures catering to female customers. We find that after launching on Product Hunt, female-focused products experience 40% less growth and are 5% less likely to have any active users after 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 is not present on days when are more women 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 nature of product development undertaken by the entrepreneurs themselves, suggesting that the signals arising from such biased sampling is not being effectively factored in by participants in the ecosystem. An implication of this finding is that systematic gaps can arise in products catering to specific customer segments, if they are not appropriately

---

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

represented among beta-testers and the 'corrupted' signals from early engagement are widely broadcast to other stakeholders.

More generally, our results contribute to research on both startup strategy and gender inequality. On the startup strategy front, our findings imply that while user-focused, lean, and experimental strategies can 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 2019; 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 on 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 explains this gap?

Methodologically, our paper also contributes to a emerging body of work using text and machine learning to estimate a startup's positioning and strategy (Guzman and Li 2019). Here we show how to embed a product's position in the market onto an underlying gendered dimension ranging from female-focused to male-focused. Given that our technique allows for the mapping of text onto a single-dimension between any pair of opposing words one could extend our methods to study differences beyond gender. While we hope work looks at other sociodemographic differences (e.g. "rich" vs. "poor") there is no reason the technique couldn't be used to measure traditional differences in firm strategy (e.g. "flexibility" vs. "commitment").

Returning to gender, our findings build on recent work highlighting the existence of product-market bias (Koning, Samila, and Ferguson 2019). This work documents that women are more likely to invent for women and suggests that increasing the number of

women inventors results in more female-focused inventions (Koning, Samila, and Ferguson 2020). Our findings suggest that product-market bias can operate even if the inventor gender split is even. If early gatekeepers—early adopters, VCs, buyers—tend to be men then the signals entrepreneurs receive will distort the direction of innovation towards men.

References

Allcott, Hunt. 2015. "Site selection bias in program evaluation." *The Quarterly Journal of Economics* 130 (3):1117–1165.

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.

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.

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

Deaton, Angus and Nancy Cartwright. 2018. "Understanding and misunderstanding randomized controlled trials." *Social Science & Medicine* 210:2–21.

Dusenbery, Maya. 2018. *Doing Harm: The Truth About How Bad Medicine and Lazy Science Leave Women Dismissed, Misdiagnosed, and Sick*. HarperOne.

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." *Available at SSRN 3383703* .

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.

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." *Available at SSRN 3489585* .

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.

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

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica: Journal of the Econometric Society* :649–670.

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

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. 2019. "Experimentation and Startup Performance: Evidence from A/B testing." *NBER Working Paper No. 26278* .

Koning, Rembrand, Sampsa Samila, and John-Paul Ferguson. 2019. "Female Inventors and Inventions." *Harvard Business School Working Paper No. 19-124.* .

———. 2020. "Inventor Gender and the Direction of Invention." *American Economic Association Papers & Proceedings* .

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* .

Lambrecht, Anja and Catherine Tucker. 2019. "Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads." *Management Science* 65 (7):2966–2981.

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

Office of Research on Women's Health. 2016. "Report of the Advisory Committee on Research on Women's Health." Tech. rep., National Institute of Health. NIH Publication No. 17 OD 7995.

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

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.

Thomke, Stefan H. 2020. *Experimentation Works: The Surprising Power of Business Experiments.* Harvard Business Press.

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

Figure 1: Binned scatterplot 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.
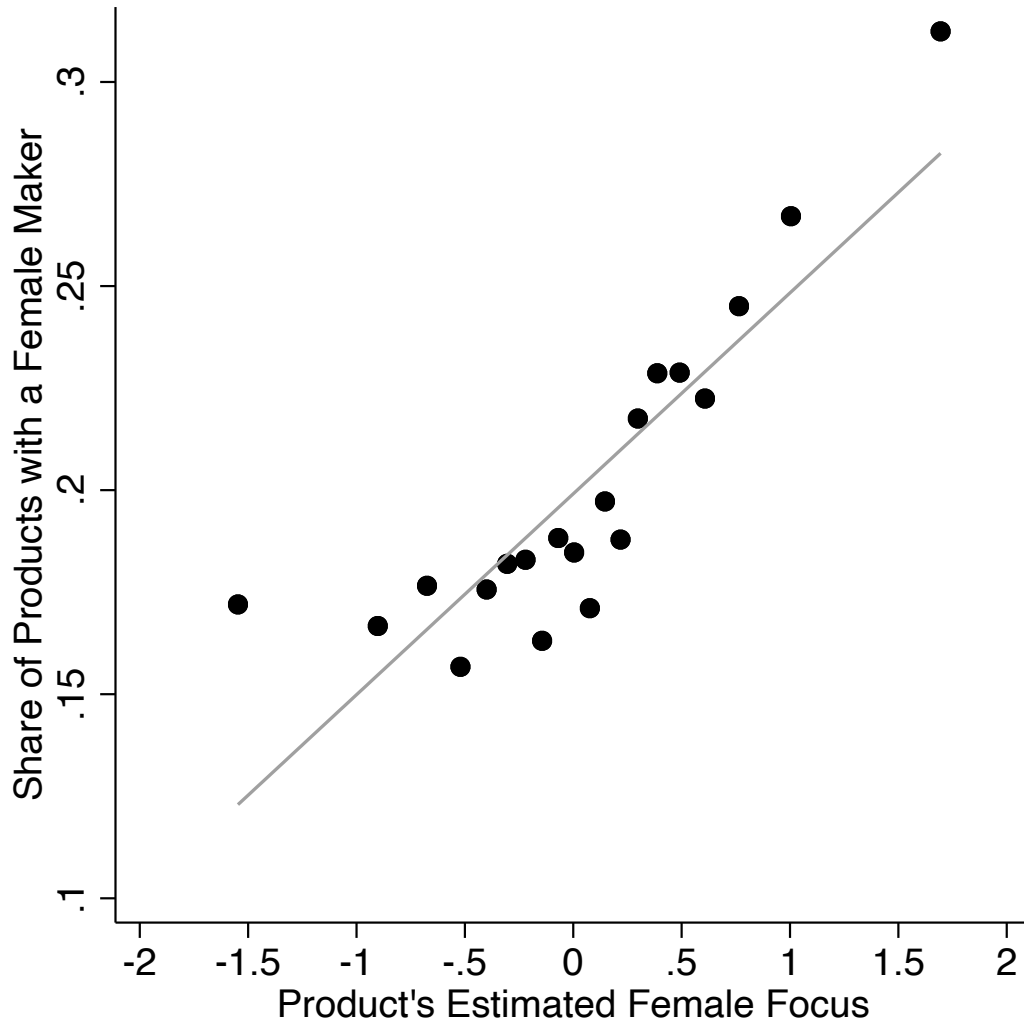
Notes: The Y-axis represents the probability that there is at least one female on the team that made the product. The X-axis is our text-based estimate of the degree to which the product focuses on female users. The binscatter controls for product launch year-month fixed effects, day-of-week fixed effects, and the logarithm of the number of words in product texts. The model includes 19,388 products with non-missing teammember gender data from 1st January 2016 to 31st December 2018.

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 .

Figure 3: Female-focused products (top quartile) have a similar growth trajectory to gender neutral products (2nd and 3rd quartiles) before lauching on Product Hunt but experience nearly 30% less user growth after launching on the platform.
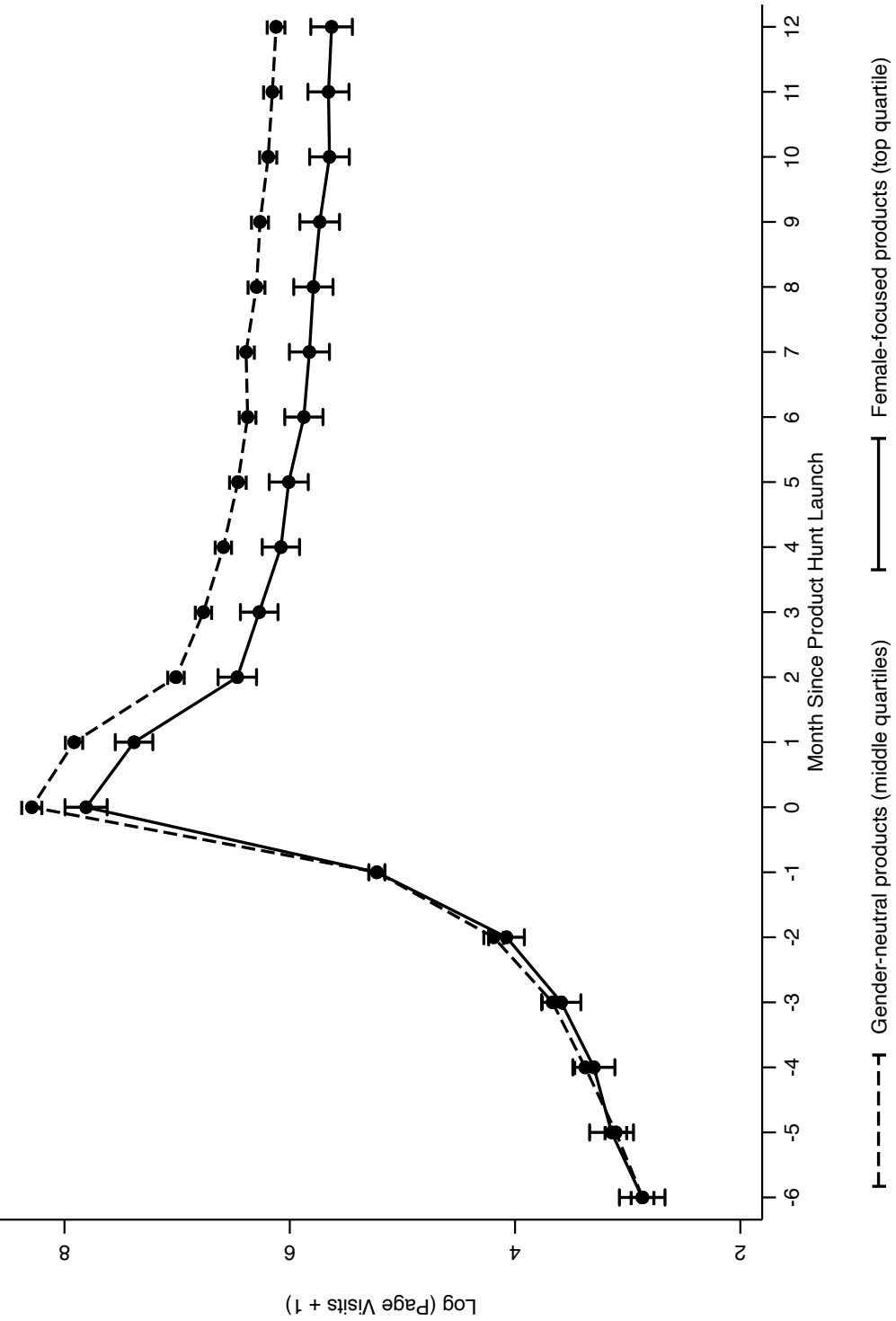
37

Figure 4: Difference in growth trajectories for Female-focused products (top quartile) compared to gender neutral products (2nd and 3rd quartiles) before and after lauching 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.
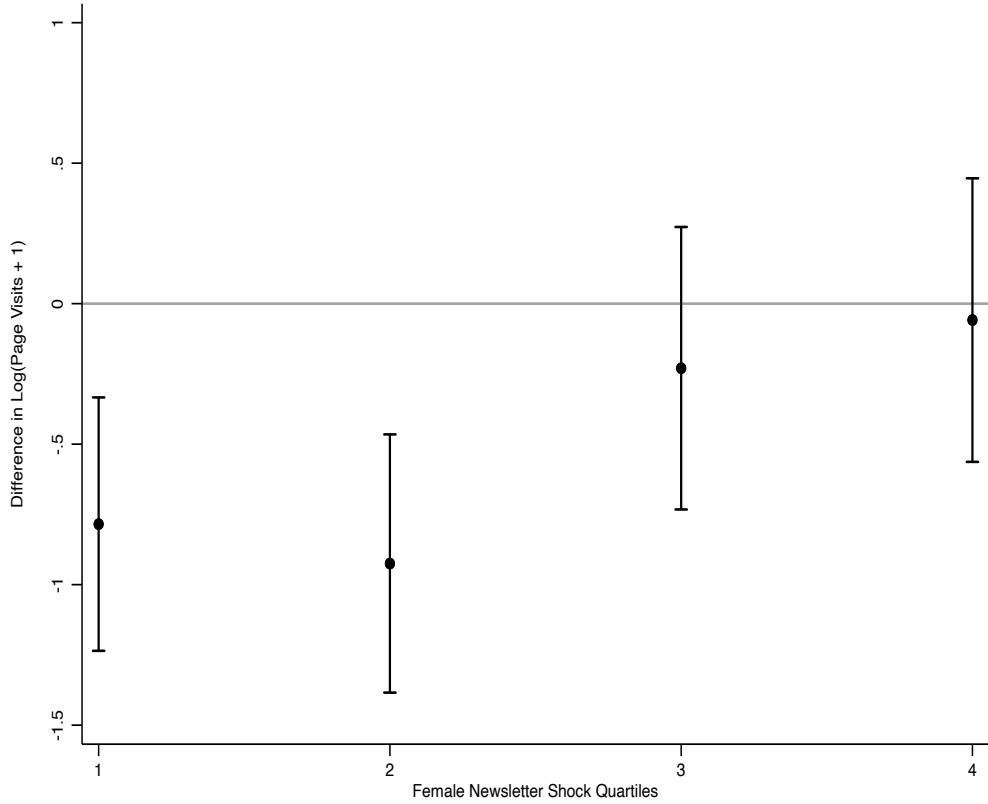
Notes: The estimates and 95% confidence intervals are from a discretized version of model 1 in Table 6 where the "newsletter shock" variable bucketed into quartiles and the product's female focus is bucketed into the top and bottom quartiles. The model includes fixed effects for number of newsletter-suggested 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 unexpecetdly more female focused. If anything, male-focused products do slightly worse when the newsletter is especially female focused.
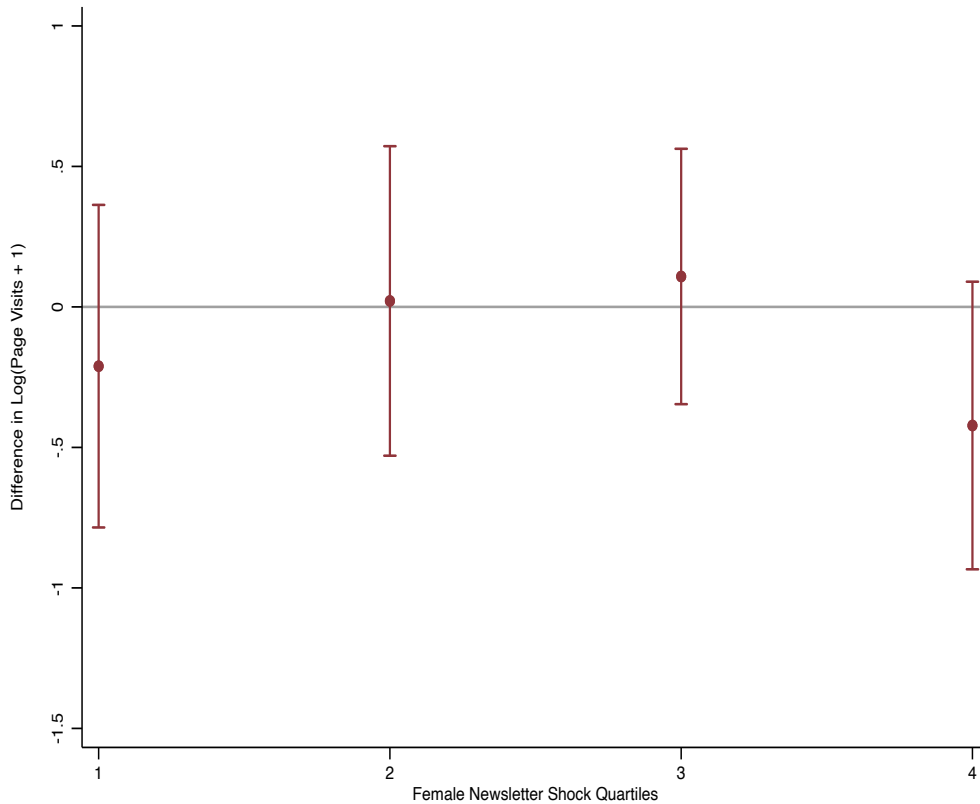
Notes: The estimates and 95% confidence intervals are from a discretized version of model 1 in Table 6 where the "newsletter shock" variable bucketed into quartiles and the product's female focus is bucketed into the top and bottom quartiles. The model includes fixed effects for number of newsletter-suggested 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
P1	ThxBro	Generate deliciously random, jargon-laced e-mails	-5.109
	Ballmetric	Your favorite plays from the NBA	-3.711
	Beard Bib 2.0	Hair clippings catcher from Beard King	-3.653
P5	SPECTRA	The most portable electric skateboard	-1.305
	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
P10	SnapHunt	Product Hunt for Snapchat. Discover new people to follow	-0.903
	Nikola	See your Tesla's battery percentage from your menubar	-0.900
	Hackuna	Secure yourself from all kinds of hackers	-0.898
P25	Morph - PokemonGo Bot	Chatbot to find and report Pokemon around you	-0.404
	Phish.AI	Anti-phishing platform powered by AI & Computer Vision	-0.401
	Sqreen API	A security toolbox for developers	-0.401
P50	Cemtrex Smartdesk	The world's most advanced workstation	0.027
	Yomu	One place to read your favorite content from around the web	0.027
	Adzoola	Hyper-targeted advertising and outreach	0.028
P75	Borsch	The AI app that helps you discover the yummiest dishes	0.495
	Cuddle Mattress	Hug your better half without the arm numbing	0.495
	Joonko	Personal diversity and inclusion AI-coach for managers	0.500
P90	The Silver Post	Do more for grandma or grandpa	0.999
	Kindred	Friends for when you travel	1.001
	Ropazi	Personal shopper for busy parents	1.001
P95	Artwxrk	Curated collection of the world's best contemporary art	1.451
	Chairman Mom	A social, Q&A platform for working moms	1.457
	VINA	Connecting awesome women for fun, for work, for life	1.474
P99	Babee on Board	Pregnant? Request a seat on public transport	5.104
	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

	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 effects of female newsletter shock by 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 effects of female newsletter by product's female focus shock on 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 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	0.025	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 effects of female newsletter by product's female focus shock 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

Who benefits from entrepreneurial experiments?

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. [18] 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 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 (**?**). 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

---

[18]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.

venture capitalists, hedge funds, and private equity firms.

BuiltWith As described by Koning, Hasan, and Chatterji (2019), 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 compeition. 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.

B   Using the viewing data to estimate preferences

We begin by identfying all the users who spent time on the Product Hunt homepage on a given day. For each user, we define users to be *active* on a given day, if they have visited Product Hunt homepage at least once[19]. The user is determined to have viewed 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.

The product-user view data is constructed by combining proprietary data on users' browsing behavior on the platform with their upvoting history.

C   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.[20] In the cases where the total number of names in the database for inferring the gender is small or

---

[19]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

[20]In rare cases registered users are actually organizations, for which we are unable to map the name to a

zero (when the name cannot be parsed), we apply Bayesian updating to a Beta prior $B(31, 71)$[21], 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.

---

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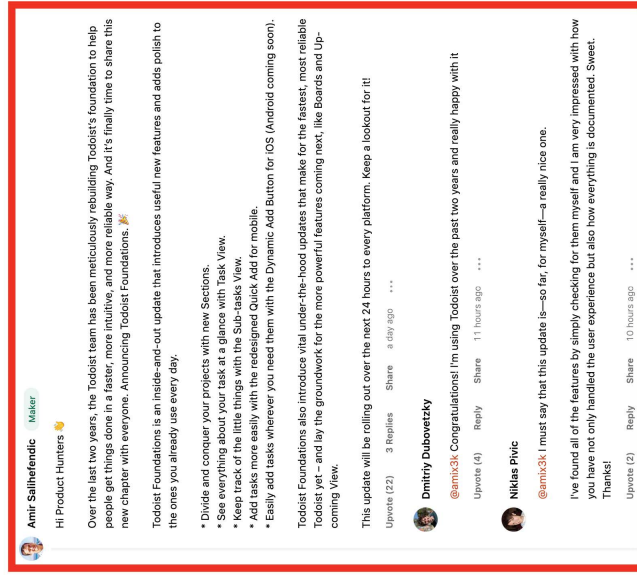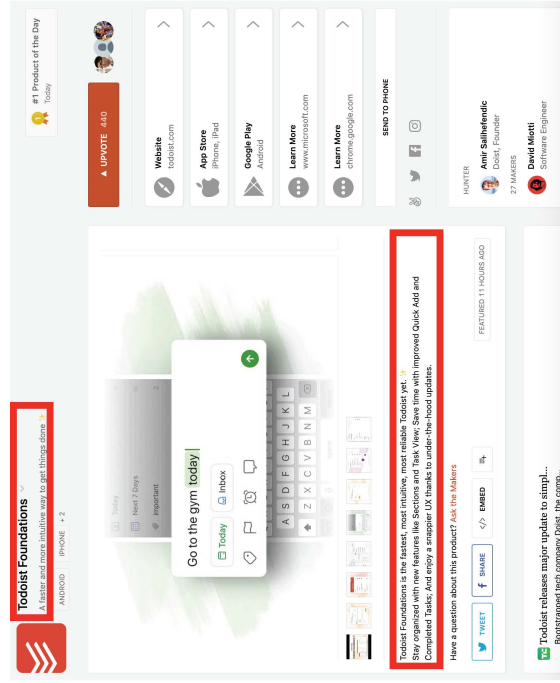
[21] Among 100 individuals, 30% are female.

Figure A.1: 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.
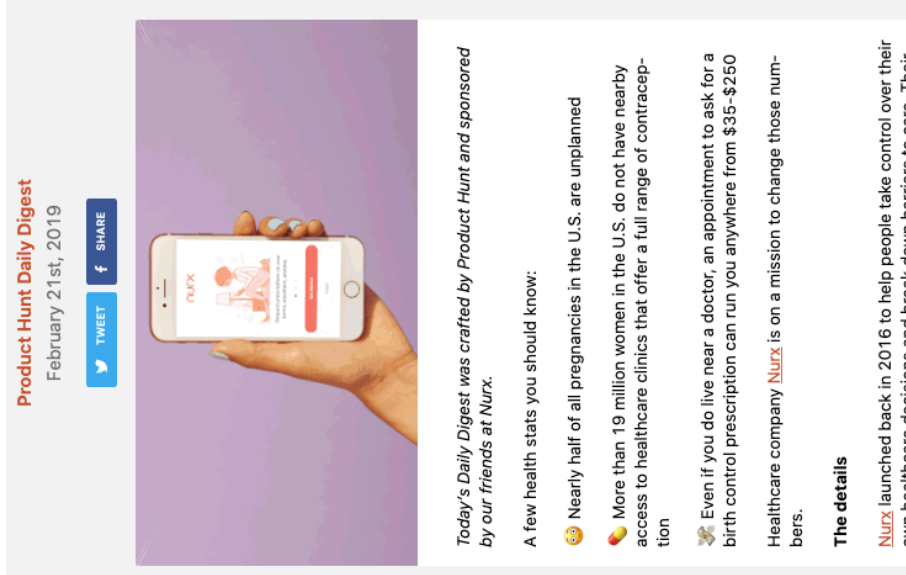
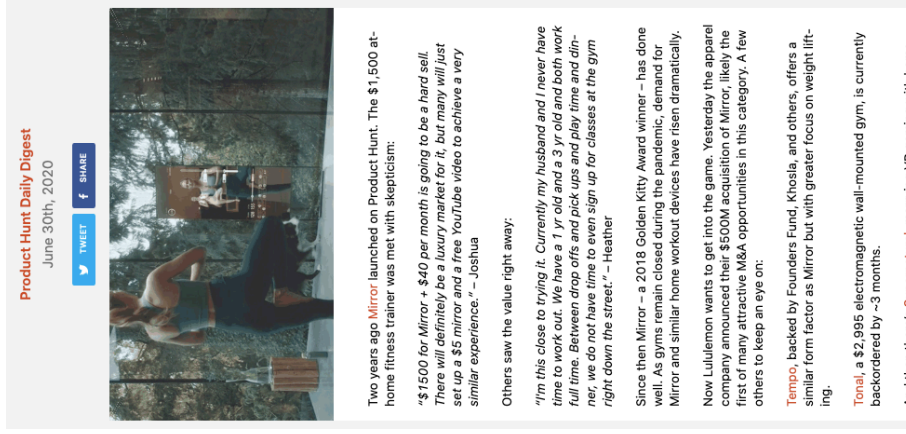Figure A.2: Distribution of the Product Gender Score
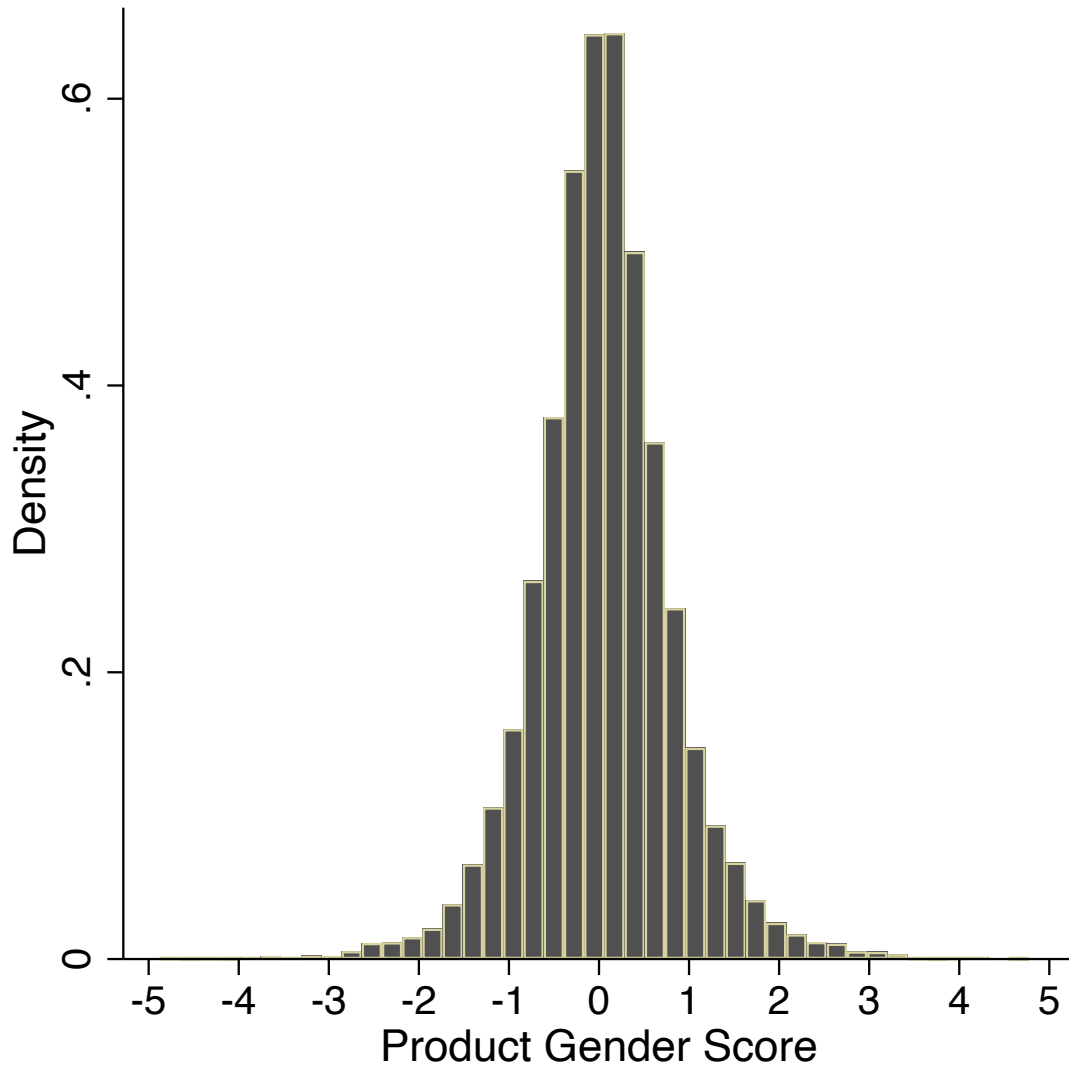
Notes: This figure shows the distribution of the product gender score, on a sample of entrepreneurial products launched on Product Hunt on weekdays from September 2016 to October 2018. The score is estimated using the text of each product's description. See text for further details. Higher scores imply that the product is more likely to serve or appeal to women.

Figure A.3: 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 A.4: Binned scatter plot showing that whether a startup ever raises venture funding is strongly correlated with monthly visits.
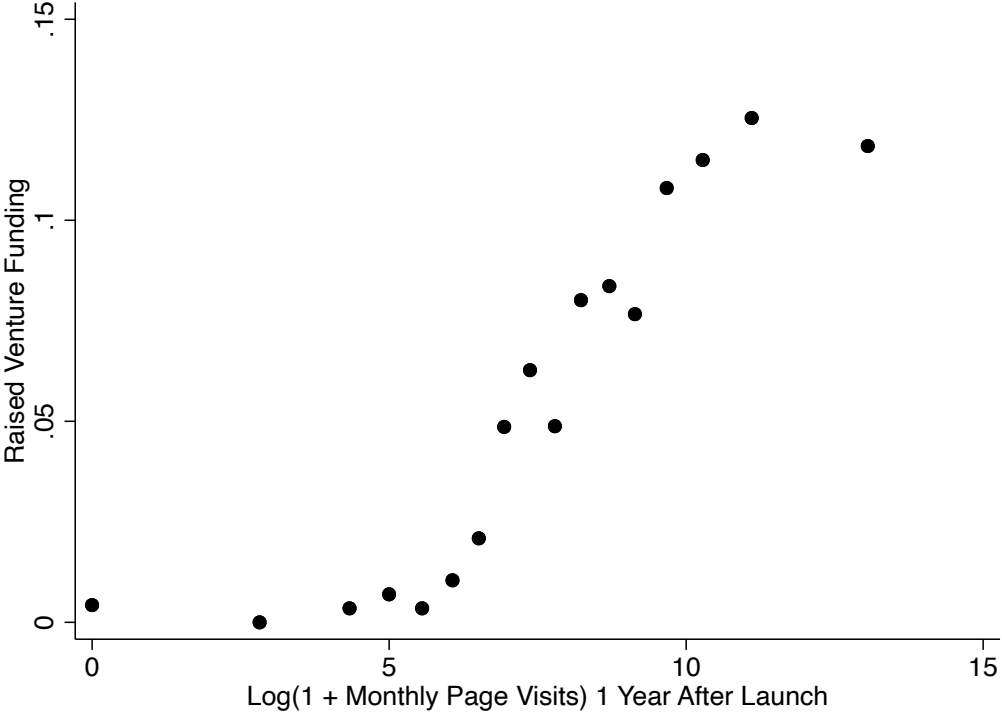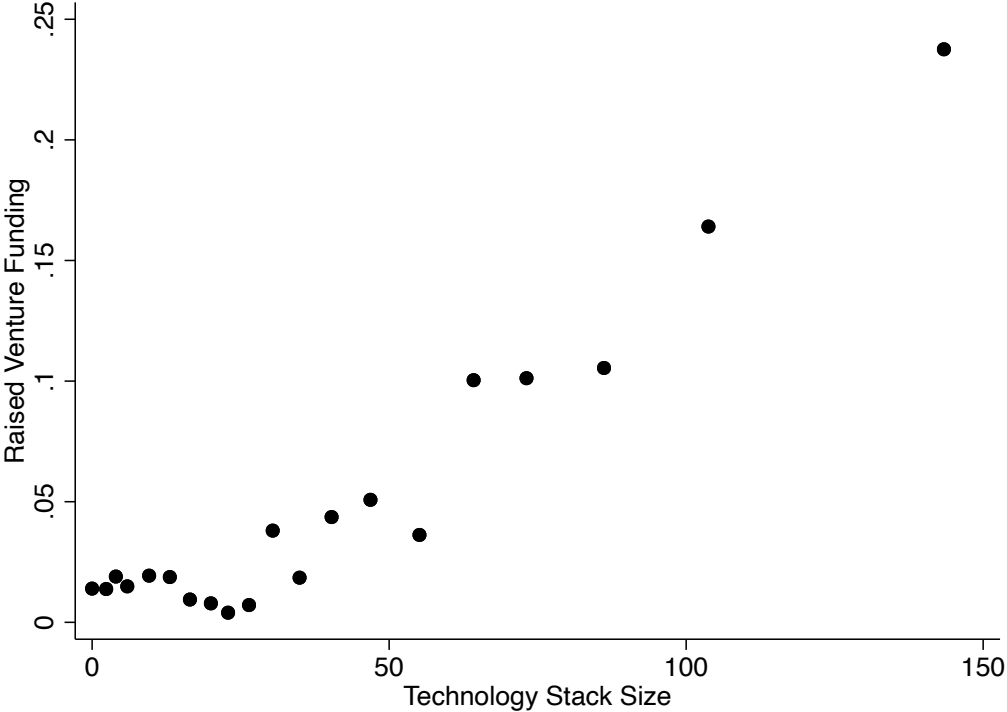

Notes: This figure shows the correlation between a startup's monthly visits a year after Product Hunt launch with whether the startup has ever raised VC funding as of October 2020.

Figure A.5: Binned scatter plot showing that a startup's technology stack size is strongly correlated with monthly visits.
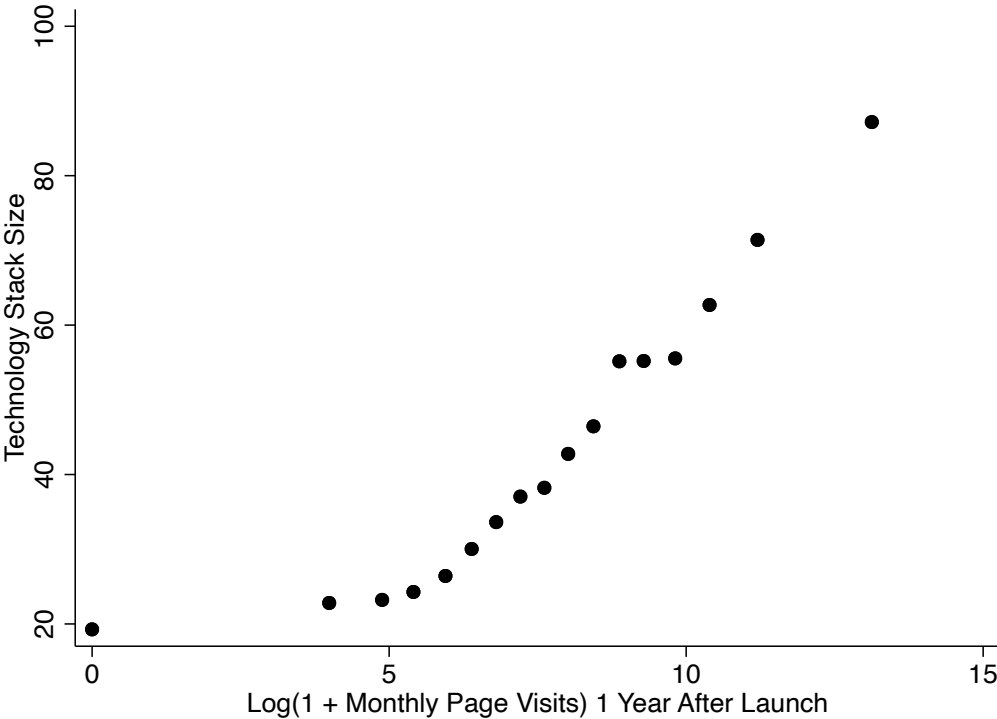
Notes: This figure shows the correlation between a startup's monthly visits a year after Product Hunt launch with whether the the size of a startup's technology stack as of October 2020.

Figure A.6: 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.

Table A.1: 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 A.2: 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 A.3: Estimated Effects of Newsletter Shock on Active Users' Platform Participation

| | Number of active female users who: | | | | | |
| | Visit ProductHunt | | Visit a Product's Page | | Vote for the Product | |
	(1)	(2)	(3)	(4)	(5)	(6)
Female Newsletter	799.030*	0.220**	59.256**	0.211**	6.819*	0.278*
	(413.080)	(0.111)	(23.353)	(0.083)	(3.830)	(0.163)
Constant	3,179.390***	8.096***	246.493***	5.531***	21.083***	3.082***
	(217.596)	(0.059)	(12.387)	(0.044)	(1.977)	(0.084)
Year-Month FE	Y	Y	Y	Y	Y	Y
Model	OLS	QPML	OLS	QPML	OLS	QPML
Observations	4,261	4,261	4,261	4,261	5,742	5,742
R-Squared	0.529		0.572		0.032	

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. 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, columns 3 and 4 if more female users visits the product's page, and columns 5 and 6 if female users vote for the product. 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 A.4: 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 A.5: 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 A.6: 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. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.