HARVARD | BUSINESS | SCHOOL

Judging Foreign Startups

Nataliya Langburd Wright
Rembrand Koning
Tarun Khanna

Working Paper 21-097

Judging Foreign Startups

Nataliya Langburd Wright
Harvard Business School

Rembrand Koning
Harvard Business School

Tarun Khanna
Harvard Business School

Working Paper 21-097

Funding for this research was provided in part by Harvard Business School.

Judging foreign startups*

March 12, 2021

Nataliya Langburd Wright (Harvard Business School)

Rembrand Koning (Harvard Business School)

Tarun Khanna (Harvard Business School)

ABSTRACT

Accelerators and investors evaluate entrepreneurial ideas in an increasingly global context. Can they pick the most promising startup ideas no matter their provenance? Discerning the potential of early-stage ideas is challenging and may be particularly so when evaluating foreign ideas, for which judges lack contextual expertise. Furthermore, biases may interfere, leading judges to systematically boost or discount foreign startup ideas. We investigate this question using unique data from the global round of a large accelerator in which multiple judges from different regions are randomly assigned to evaluate startups from across the globe. Our analysis of this natural experiment shows that judges are informed about the quality of both foreign and local startups, though they are biased against foreign startup ideas. Judges are less likely to recommend startups headquartered outside their home region by 4 percentage points. Additional information about a startup does not appear to attenuate this bias. Back-of-the-envelope calculations suggest that judges pass over 1 in 10 high-potential foreign startups. Together, our results reveal that judges are able to discern high from low quality startups regardless of their geographic provenance, but they systematically discount foreign ideas relative to local ones.

*nlangburdwright@hbs.edu, rkoning@hbs.edu, tkhanna@hbs.edu. We thank Susan Cohen, Jorge Guzman, Ramana Nanda, and Maria Roche for valuable comments. We also would like to thank Julia Comeau for excellent copy-editing. All errors remain our own.

I. Introduction

Startups, like corporations, are increasingly globalized in terms of their markets, investments, and workforce participation (Kerr, 2016; Ghemawat and Altman, 2019; Lu and Beamish, 2001; Oviatt and McDougall, 2005), partially due to the advent of technology that reduces the cost of expanding internationally (Brynjolfsson, Hui, Liu, 2019). As a result, entrepreneurial gatekeepers, ranging from investors to accelerators, increasingly evaluate a global pool of startups and must choose the most promising to provide support and funding. For example, Silicon Valley-based Y Combinator funded Ukraine-based Petcube, an interactive pet monitor startup that went on to become a unicorn, valued at over $1 billion (Y Combinator, 2020; X1 Group, 2018). At the same time, gatekeepers have missed out on promising international startup opportunities; for example, Silicon Valley-based Bessemer Venture Partners passed over Australian-based Atlassian, a project management software company, which is now worth over $43 billion (Bessemer Venture Partners, 2020).[1]

Can accelerators and investors choose the most promising startups from this increasingly global pool? Judges may not be able to discern the quality of startups overall, given the highly uncertain nature of early stage ideas (Kerr, Nanda, and Rhodes-Kropf, 2014; Gans, Hsu, and Stern, 2008; Luo, 2014). They may be particularly inaccurate in discerning the potential of foreign startups because they lack the contextual expertise and information necessary to sort winners from losers. Moreover, judges may carry a bias against foreign startups, similar to the gender, race, and expertise biases documented across a range of entrepreneurial settings (e.g. Lee and Huang, 2018; Niessen-Ruenzi and Ruenzi, 2019; Hegde and Tumlinson, 2014; Li, 2017). Overall, judges may be either uninformed about the quality of all startups, only informed about

[1] The valuation is in terms of market capitalization.

the quality of local startups, or informed about foreign and local startups equally. They also may be simply biased.

While prior research shows that trade partners, financial analysts, and investors are more likely to select companies that are nearby, research documenting such “home bias” often conflates crucial differences in the mechanisms underlying the effect (French and Poterba, 1991; Disdier and Head, 2008; Dziuda and Mondria, 2012; Coval and Moskowitz, 1999; 2001). As mentioned above, “home bias” by accelerators and startup investors could be the result of a simple preference for home-grown startups irrespective of each startup’s potential. Under this mechanism, an accelerator could simply counter its bias by lowering its threshold for accepting or investing in foreign firms. However, such a move will backfire if the underlying “home bias” mechanism is instead rooted in the inability of judges to distinguish foreign winners from losers. In this situation, judges pick the most promising local ventures whereas their evaluations of foreign firms are potentially no better than random draws. No matter the threshold, judges will still end up selecting lower quality foreign ventures than local ones. In this case, remedying the underlying “bias” requires fundamentally shifting the composition of who evaluates; redesigning how scores are aggregated into decisions will not matter. In short, the underlying mechanism that leads to “home bias” effects has strong implications for how accelerators and investors should design their selection processes.

However, teasing apart these mechanisms is non-trivial. First, estimating judge “home bias” effects, in and of itself, is not easy. Estimates that rely on the location of selected startups and the investors and accelerators who select these firms will nearly always confound supply-side (the judge’s choice of who to pick) and demand-side (the founder’s choice of where to apply) concerns. Further, even with data where the risk set of startups to be selected is fully observed (e.g. venture competitions), startups may selectively choose which local or foreign competitions to enter, and judges are often non-randomly assigned which startups to assess. In these cases, estimates are again biased because higher-quality startups might disproportionally select into local competitions, or harsher judges might be assigned to foreign ventures. Finally, even if judges and startups from different countries are randomly assigned to one another, finding evidence that judges discount foreign startups is insufficient to reveal the underlying mechanism. Indeed, teasing apart whether "home bias" is rooted in uniform discounting or differences in a judge's ability to evaluate requires not just random assignment of judges but also measures of each startup's quality.

Here we analyze data from an accelerator's global venture competition in 2017 and 2018 that meet these criteria and so allow us not only to causally identify if judges exhibit home bias, but also pinpoint the mechanisms underlying this effect.[2] In the first round of this competition, judges from North America (the United States and Canada), Latin America, Europe, and Israel evaluated startups from across the globe. Crucially, in the first round, judges were randomly assigned to evaluate startups no matter their origin, and no startups could "opt out" of being evaluated by judges from particular regions. In total, we analyze the decisions of 1,063 judges who are assigned to evaluate 3,784 startups. Further, we pair this data with information on financing and proxies for user growth that allow us to measure the ex-ante potential and ex-post success of startups in our sample irrespective of whether they were accepted into the accelerator or not. These measures of quality allow us to identify the mechanisms underlying any "home bias" effect.

[2] Accelerators are defined as "fixed-term, cohort-based program for startups, including mentorship and/or educational components, that culminates in a graduation event" (Cohen, Fehder, Hochberg, and Murray, 2019).

We find that judges are less likely to recommend startups from a foreign region by 4 percentage points after accounting for observed and unobserved differences in startup quality with startup-level fixed effects. The magnitude is meaningful. It is nearly half of the effect of a startup going from having no users to some user traction and a tenth of the size of a startup going from no funding to raising venture financing. These magnitudes are consistent with prior work documenting "home bias" in other settings ranging from financial markets to trade (Coval and Moskowitz, 1999; Disdier and Head, 2008).

This effect is driven by a consistent discounting of foreign startups by local judges and not by differences in the ability of judges to better pick winners from losers amongst local firms. Surprisingly, we find judges are equally good at evaluating startup quality if the startup is from their home region or is foreign. Startups given higher scores by judges – irrespective of being accepted into the accelerator – are more likely to raise financing and experience more user growth. Again, this correlation holds within local and foreign startups, contrary to prior work showing judges struggle to pick winners from losers in technology businesses (e.g. Scott, Shu, and Lubynsky, 2020). Further, this foreign discount does not diminish when stronger signals of startup quality are present, suggesting that the judge's "home bias" is unlikely to be rooted in some form of statistical discrimination. Indeed, when we conduct back-of-the-envelope calculations, we find that judges passed over 521 promising foreign startups, equating to roughly 1 in 10 foreign startups in our sample. Overall, we find strong evidence that judges can evaluate startup quality no matter the firm's provenance, but they uniformly discount foreign firms, suggesting that simple changes to how accelerators and incubators aggregate judges' evaluations might mitigate the impact of "home bias" on outcomes.

This study makes three primary contributions. First, our findings suggest that entrepreneurial gatekeepers are adept at evaluating startup quality, in contrast to prior work showing that accelerators, investors, and judges struggle to separate the most from the least promising startups (Nanda, Samila, and Sorenson, 2020; Scott, Shu, and Lubynsky, 2020; Kerr, Lerner, and Schoar, 2014). Crucially, we find little evidence that gatekeepers have a local advantage in screening startups, as judges were no better able to choose winners from losers among local startups as against among foreign startups. To be clear, local judges might hold other advantages when it comes to investing in or nurturing these startups, but our findings suggest screening is unlikely to explain such differences. Further, the fact that judges were equally good at screening foreign and local startups suggests that increasingly standardized business models[3], management practices (Bloom and Reenen, 2007; Chatterji, Delecourt, Hasan, and Koning, 2019), and technology (Haefliger, Von Krogh, and Spaeth, 2008) might be making it easier to evaluate startups no matter their origin.

Second, we show that geographic bias may distort the composition and direction of entrepreneurship and innovation in ways that research has shown in terms of gender and race (e.g. Lee and Huang, 2018; Koning, Samila, and Ferguson, 2020). If gatekeepers are biased against foreign startups, and if the majority of these gatekeepers still reside in entrepreneurial hubs like in the U.S., this may potentially result in a gap in startups from non-hub regions. And this bias does not just impact which startups succeed, but also may impact who benefits from their innovations (Koning, Samila, and Ferguson, 2020). Indeed, if accelerators overlook ideas from these non-hub markets, then there may be too few startups serving the needs of customers in those foreign markets.

[3] For example, Silicon Valley-based Y Combinator highlights "nine business models and metrics that investors want:" https://www.ycombinator.com/library/8E-nine-business-models-and-the-metrics-investors-want-sus-2019.

Third, we show the potential limitations of accelerators in their ability to help startups grow because judges select out foreign startups from gaining access to the resources of the accelerator's ecosystem. While various studies find that accelerator programs result in positive performance gains for startups (Cohen, Bingham, and Hallen, 2019; Hallen, Cohen, and Bingham, 2020; Yu, 2020; Chatterji, Delecourt, Hasan, and Koning 2019; Howell, 2017; Gonzalez-Uribe and Leatherbee, 2018; Fehder and Hochberg, 2014; Yin and Luo, 2018), our results suggest that their impact may be muted for foreign startups because these organizations discount them. That said, our results also suggest that fixing such bias requires relatively minor tweaks to how a firm aggregates decisions. Indeed, given that judges can evaluate foreign startups as well as local ones, accelerators and investors need merely to lower their threshold for selecting a foreign firm versus a local firm.

The rest of this paper is structured as follows: Section II discusses the theoretical framework. Section III describes the context. Section IV presents the data. Section V discusses the empirical specification. Section VI presents the results. Section VII concludes.

II. Theoretical Framework

Evaluating Startup Quality

Predicting startup quality is difficult because of at least three information challenges. First, the success of startup ideas hinges on a host of complex factors, including the viability of the technology itself, the quality of the business model, the demand of customers, competition, and the ability of the founding team to execute (Gompers, Gornall, Kaplan, and Strebulaev, 2020; Sorensen, 2007; Kaplan, Sensoy, and Strömberg, 2009; Aggarwal, Kryscynski, and Singh, 2015; Hoenig and Henkel, 2015). It is difficult to understand each individual factor, and importantly, their interactions.

Second, there are few precedents on which to base startup predictions. Startup ideas may be novel, and there are very few startup success cases relative to failed ones (Hall and Woodward, 2010). Just as supervised machine learning models fail to predict accurately when their training data are insufficient, evaluators too may fail in their predictions without appropriate historical cases.

Third, entrepreneurs may provide incomplete information about their ideas, resulting from Arrow's paradox that the value of an idea requires disclosure, but disclosure eliminates incentive to "pay" for an idea (Gans, Hsu, and Stern, 2008; Luo, 2014; Arrow, 1962). As a result, the information that evaluators do observe may only be a small portion of the full information needed to make an accurate prediction.

The combination of complexity, lack of precedents, and incomplete disclosure creates a context of imperfect information that makes early-stage ideas highly uncertain and difficult to evaluate. Indeed, research shows that entrepreneurial gatekeepers, including investors and mentors, may lack the ability to predict quality of startups ex-ante (Nanda, Samila, and Sorenson, 2020; Scott, Shu, and Lubynsky, 2020; Kerr, Lerner, and Schoar, 2014). Rather, access to deal flow allows some investors to have a persistent advantage over others (Nanda, Samila, and Sorenson, 2020).

Contextual Intelligence

Given these challenges in discerning startup quality, when can evaluators distinguish winners from losers? Evaluators may be able to do so when they have expertise (Li, 2017) or intuition (Huang and Pearce, 2015) that may compensate for the imperfect information provided about startups. They may build this expertise in their local region where they live and work. However, the expertise may not be transferable to a foreign context because of differences in institutional, cultural, linguistic, and related factors (Khanna, 2014). Evaluators, therefore, may only be able to use this expertise to reduce the uncertainty in quality of local, but not of foreign startups. For example, an Israeli investor would be able to use his/her expertise of Israel's military structure to reduce uncertainty around founders of an Israeli company with military experience. The same investor may not be able to use his/her expertise to reduce uncertainty when evaluating a U.S. company. Consistent with this view, prior work has shown that financial analysts are worse at picking foreign stock winners, relative to local stock winners (Malloy, 2005; Coval and Moskowitz, 1999), and the information value of an acquisition is highest for foreign acquirers (Conti, Guzman, and Rabi, 2020). Figures A1a-b show how expertise may compensate for imperfect (unobserved) information for local startups more so than for foreign startups. This situation would result in judges being better able to choose winners from losers among local versus among foreign startups.

[Insert Figures A1a-b]

Bias in Evaluations

However, reliance on local expertise may come with a cost of bias. Evaluators may systematically score lower or higher some startups over others for reasons other than quality. In contrast to measurement error that may result in deviations from objective quality in multiple directions because of an inability to predict with accuracy, biases manifest as deviations from objective quality measures in a consistent direction because of a preference, as defined in other evaluation studies (e.g. Li, 2017). Research has documented such biases particularly in the gender and race domains.

The direction of this biases may be either positive (favoring) or negative (penalizing) for foreign startups. They may be positive if judges more similar to startups are harsher to them and therefore those foreign to them are relatively less harsh. This may occur because evaluators oversample on factors of startups closer to them in geography, which then distorts their ability to detect quality of those startups (Boudreau, Lakhani, and Riedl, 2016) or because of varying perceptions of stereotypes as seen in investments of early-stage female-led ventures (Johnson, Stevenson, and Letwin, 2018).

The more likely case, based on existing research, is that the bias is negative, such that judges discount foreign startups. Judges may prefer what is more "familiar" (Huberman, 2002; Franke, Gruber, Harhoff, and Henkel, 2006; Lin, Prabhala, and Viswanathan, 2013) or have an interest in promoting the success of local startups no matter the quality. Such a direction would reflect homophily as evaluators prefer startups more similar to themselves, as seen among startup founders (Ruef and Aldrich, 2003). Similarly, prior research has found evidence of bias against entrepreneurs from different genders, and races (Lee and Huang, 2018; Niessen-Ruenzi and Ruenzi, 2019; Hegde and Tumlinson, 2014). Studies in financial and trade markets relatedly have detected a home bias for geographically proximate portfolio stocks or trade partners (French and Poterba, 1991; Disdier and Head, 2008; Dziuda and Mondria, 2012; Coval and Moskowitz, 1999; 2001).

Hypothesis Development

The combination of uncertainty, expertise, and bias generates six scenarios that each call for different strategic responses by organizations. These scenarios depend on whether judges are uninformed about the quality of all startups, only informed about the quality of local startups, or informed about the quality of all startups. Figure 1 sketches the relationship between measures of startup quality (x-axis) and judge scores (y-axis) for startups foreign to the judge (red dotted line) and local to the judge (blue solid line). Each combination of mechanisms results in a different relationship between startup quality and judge scores, corresponding to different organizational responses to improve the quality of selected startups.

[Insert Figure 1]

In the first row of Figure 1, we show cases when judges are uninformed about startup quality. They cannot pick winners from losers among startups because of the high uncertainties associated with early stage ventures, as shown in several entrepreneurship settings (Nanda, Samila, and Sorenson, 2020; Scott, Shu, and Lubynsky, 2020; Kerr, Lerner, and Schoar, 2014). No matter whether judges are biased (cell B) – systematically preferring local or foreign startups — or unbiased (cell A), the same performance outcome results: the selected startups are random on the basis of their quality. While the level of quality of startups may not vary in the biased or unbiased case, the distribution does. The case in which judges are unbiased results in the selection of a higher diversity of local and foreign startups. Thus, organizations with a pure performance objective would be better off allocating resources to monitoring selected startups, rather than improving the selection process (including rooting out bias). Organizations with a diversity objective, however, would have incentive to reduce bias in the selection process, albeit without improvements to the quality of selected startups.

Yet, the deep experience judges gain in their local contexts gives reason to believe that judges may be informed about the quality of local startups, as shown in the second row of Figure 1. In this case, judges' bias impacts both the performance and diversity outcomes. The quality of the ventures selected may be higher if judges are biased against foreign startups because they will disproportionately choose among local startups from which they can pick winners from losers (cell D). The diversity of the selected startups, however, will be lower in the biased case.

In this case, organizations may be better off assigning local judges to startups, which they are better fit to evaluate.

Lastly, the standardization of technologies and business practices may enable judges to be informed about the quality of both local and foreign startups, as shown in the third row of Figure 1. Startups may follow a similar enough playbook that separating good from bad investments across countries is no harder than within countries. For example, work has shown the benefits of good management appear universal for corporations and startups across the globe (Bloom and Reenen, 2007; Chatterji, Delecourt, Hasan, and Koning, 2019). The same appears to be true for coding practices (Haefliger, Von Krogh, and Spaeth, 2008). The standardization of such business and technology practices means that judges might be able to evaluate a startup's potential with similar clarity no matter its country origin. In this case, bias interferes with picking the most promising startups. If judges are biased against foreign startups, they may pass over some higher quality foreign startups for lower quality local startups (cell F). Given their ability to detect quality of local and foreign startups alike, organizations may leverage the opportunity to select from startups globally. Rather than changing the composition of judges as would be the case in the locally informed-biased scenario (cell D), organizations would be better off revising their processes to reduce bias either in aggregate (e.g. by lowering the threshold for selecting a foreign firm versus a local firm) or at an individual judge level (e.g. by introducing nudges).

Taking these cases together, it becomes apparent that simply evaluating whether judges give lower scores to foreign startups does not allow us to disaggregate cases when judges are biased and uninformed (cell B), informed about local startups and biased (cell D), or informed about both local and foreign startups and biased (cell F), though the cases yield vastly different strategic responses. Our empirical strategy, discussed further below, disaggregates these six cases to reveal optimal organizational strategies.

Among these scenarios, we can infer that judges can pick promising startups only when they are informed about all startups, whether unbiased (E) or biased (F). The biased case (F), however, demonstrates that judges may choose not to pick the most promising startups – though they can discern them – because of their preferences. Thus, their prediction error is a matter of choice (conscious or sub-conscious) rather than lack of ability.

This framework of scenarios extends information-bias frameworks presented in other evaluation studies in the science context (e.g. Li, 2017; Boudreau, Guinan, Lakhani, and Riedl, 2016) by de-coupling bias, not only from information, but also partial information enabled by local expertise. We proceed to test these hypotheses. In particular, we investigate whether judges exhibit a bias for or against foreign startups and whether they are able to detect the quality of foreign startups to understand the extent to which judges can pick promising startup ideas in global evaluations.

III. Context: Global Accelerator Competition

To unbundle these scenarios, we use data from a large global accelerator's new venture competition. The accelerator operates in four regions around the world: the U.S., Europe, Israel, and Latin America. There are four rounds in the accelerator program. In the first round (the global round), startups apply to several of the regional locations of the accelerator program. In the latter rounds, the accelerator assigns startups (based on their preferences and judge scores) to a particular regional location of the accelerator, and judges generally local to that area evaluate the startups. Roughly a third of startups make it from the initial applicant pool into the second round, a third from the second to the third round, and a quarter from the third to the final round.

Startups who made it to the third round (approximately 10 percent of the initial applicant pool) participate in the full accelerator program, including the educational curriculum, mentorship program, and other networking events. The top 10-20 rated startups across the globe at the conclusion of the last round gain both credibility and monetary prizes worth $10-$110K.

We focus on the global round of the competition where judges – representing executives (61%), investors (14%), and other professionals (25%) – across these international regions initially screen startups from across the globe. Applications ask startups to list the company's background and funding, industry & competitors, and business model & financials. We show the full application template in Appendix A7. All applications are in English.[4] While the applications do not specifically inform judges of the startup's location, judges may infer it fairly easily through the description of the startup, founder(s), and market. Judges review these applications online. Each judge evaluates roughly 20 startups, and each startup receives evaluations from 5 judges on average. Judges provide a score on whether they recommend the startup to the next round of the competition, which the accelerator uses to select startup cohorts for the next round. Judges also provide subscores on a scale of 1-10 on the following criteria: startup team, industry & competitors, and business model & financials.

To infer judges' location, we use data on the location of the accelerator the judge is affiliated with.[5] As judges need to evaluate startups in person during the later rounds of the competition, judges tend to be assigned to an accelerator to which they are physically proximate. We therefore categorize judge locations as corresponding to the accelerator's locations: Europe, Latin America, Northern America (U.S. & Canada), and Israel. We manually checked a random

[4] While English applications may mask quality of startups whose founders have a different native language with different writing styles, such a language requirement is common for startup accelerator program applications.
[5] The accelerator does not collect data on judges' location of residence. It only collects the home accelerator program of each judge.

subsample of 136 judges from our data to test if the residence regions of the judges matches those of their home programs. Of this subsample, 76 percent (103/136) of judges resided in their home program region. Crucially, this error should bias our estimates towards zero because the broad regional categories could underestimate biases within regions. For example, a U.K. judge evaluating a Latvian startup would appear as a regional match in our data, though we can imagine that the judge would consider the startup foreign (and potentially discount it).

IV. Data

Our data come from the 2017 and 2018 cycles. During these two years, judges were randomly assigned to startups during the initial global round. This random assignment allows us to overcome selection by startups into programs in particular locations. Such selection would confound our ability to measure whether judges discount foreign startups because we would not be able to disaggregate judges' evaluations of startups from startups' underlying quality that may have caused them to choose certain accelerator locations over others. Our overall data across the 2013-2019 four rounds consist of 88,715 startup-judge level observations, including 11,197 unique startups and 3,747 unique judges. When we narrow down our sample to the 2017 and 2018 cycles, we end up with 21,017 startup-judge level observations, including 4,425 unique startups and 1,067 unique judges. We then remove startups whose headquarter regions do not match any of the judges' home programs to exclude the possibility that a startup is foreign to all judges in our sample and therefore lacks a local judge score as a basis of comparison.[6] This brings our final sample to 17,959 startup-judge level observations, including 3,784 unique startups and 1,063 unique judges.

Dependent Variables

[6] Our results are robust to including or excluding startups whose headquarter regions do not match those of any of the judges' home programs.

***Score* –** Our first dependent variable is a composite z-score created from the subscores judges give to startups. These underlying subscores include: customer pain and solutions, customer needs and acquisition, financial/business model, industry competition, overall impact, regulations and intellectual property, team (including advisors and investors), and the overall recommendation. These subscores correspond to the sections in the applications startups initially complete. They range from a scale of 1-10. While not all judges complete every subscore evaluation, the vast majority do.

***Recommend* –** Our second dependent variable is a binary dependent variable indicating whether a judge recommended the startup to advance to the next round of the competition. The accelerator only considers this overall recommendation from each judge to determine whether a startup makes it to the next round of the competition.

Independent Variables

***Foreign Startup* -** Our key covariate captures whether the judge and startup are from the same region (e.g. both from Europe, the U.S./Canada, Israel, or Latin America). We construct a binary variable indicating whether a judge is evaluating a foreign startup ("1" indicates a foreign startup, "0" indicates a local startup). These broad geographic categories likely bias downward foreign discounting because judges may perceive geographic differences within our currently constructed regions that would increase the home bias effect.

Post-Accelerator Logged Financing Value - To understand whether judges discount foreign startups because they struggle to evaluate their future potential, we measure the post-accelerator performance of startups. Specifically, we test if judges give higher ratings to higher performing local startups but not to higher performing foreign ones. We use logged financing value six months after the program.[7] This variable indicates the logged amount of USD startups received from investors six months after the program.

Post-Accelerator Logged Website Page Visits - We also use logged monthly page visits after the accelerator program in 2019 (the latest data we have available). Website page visits are a common performance indicator for early stage (pre-revenue) startups because they reveal engagement with users and are predictive of venture funding (Koning, Hasan, and Chatterji, 2019; Cao, Koning, and Nanda, 2020; Hallen, Cohen, and Bingham, 2020). We confirm that website page visits are positively correlated with financing value on the basis of pre-accelerator and post-accelerator measures (Figures A5 and A6).

Pre-Accelerator Financing - We use pre-accelerator quality measures of the startups as proxies of the startup quality at the time of application. These measures allow us to assess whether judges can evaluate the quality of startups at the time of evaluation. We use two versions of the pre-accelerator financing variable: a continuous measure of logged financing value before the program and a binary variable indicating whether a startup received financing before the program to indicate financing traction.

Pre-Accelerator Page Visits - We also use two measures of pre-accelerator website page visits as a proxy of startup quality at the time of evaluation. These measures include (1) a continuous logged average website page visits 3 months before the initial application review period of the accelerator and (2) a binary variable on whether a startup reached above the first decile (relative to our overall sample) of website page visits three months before the program to indicate user traction.

[7] All logged values are of (1+x) because of frequency of zeros in our dataset.

If no page visit or financing data are available for startups, we impute them to be zero. In robustness checks, we confirm that whether a startup has financing and page view data are positively correlated with their evaluations, suggesting that the missing values are the result of startup shutdown or slow maturity.[8]

Accelerator Participation - We also account for whether a startup participated in the accelerator. This variable allows us to control for the potential treatment effects of the accelerator that may confound our ability to assess whether judges are able to detect the post-accelerator performance quality of startups. We include it in specifications involving post-accelerator financing and page visit variables.

Descriptive Statistics for Evaluations - Table 1 shows summary statistics for our main sample from the global round of the competition, including 17,959 startup-judge level observations, 3,784 unique startups, and 1,063 unique judges. These summary statistics break up our main dependent variables (judge score measures) and independent variables (startup quality measures) by whether a startup is local or foreign to the judge in a given evaluation. The raw data comparing means of scores given to foreign and local startups show that for the most part there is no difference in the quality measures between local and foreign startups with the exception of pre-page traction, where local startups have a higher value on average by 0.07 point. This may be because regional differences in target market distributions: U.S. and Canadian startups, which are more likely to be local to judges and have consumer-facing ideas, have higher user traction. Table 1 also reveals that judges are less likely to recommend foreign startups and rate them as lower quality.

[Insert Table 1]

[8] Our results are robust to imputation or lack of imputation of zeros in the page visits data. We do not have sufficient sample size to evaluate results without imputation of zeros for the financing data.

V. Empirical Specification

To assess whether judges systematically give lower or higher scores to foreign startups, we fit the following model:

$$(1)\ SCORE_{ijt} = \propto + \beta FOREIGN_{ij} + JUDGE_j xYEAR_t + STARTUP_i xYEAR_t + \epsilon_{ijt}$$

Where $SCORE_{ijt}$ is either a z-scored average or a binary variable on whether judge j recommends startup i to the next round. $FOREIGN_{ij}$ is our binary variable indicating whether the region of startup i is different from that of judge j. Our main coefficient of interest is β, indicating whether judges discount startups from outside their home region. If β is statistically significant, then judges consistently discount or boost foreign startups, allowing us to reject that judges are informed and unbiased (Figure 1, cell E), as well as uninformed and unbiased (Figure 1, cell A).

We include a battery of fixed effects to identify judge effects from differences in startup quality. We account for judge harshness and judges participating across multiple years of the program through judge-year fixed effects ($JUDGE_j xYEAR_t$), so that our analysis focuses on judge evaluations of startups within the same year.

We phase in several fixed effects to account for startup-level differences. The most stringent version – startup fixed effects by year ($STARTUP_i xYEAR_t$) – are shown in equation 1. Before moving to this version, we first account for startup region-year fixed effects ($STARTUPREGION_i xYEAR_t$) to measure startup evaluations within a particular region (e.g. Europe, Latin America, Israel, and Northern America) in a given year to account for differences across startup headquarters regions over time. These fixed effects allow us to account for differences in quality among, for example, European-based startups and Israeli-based startups. This is the loosest interpretation of geography as it does not account for country differences within regions. We then tighten our specification to startup country-year fixed effects ($STARTUPCOUNTRY_i xYEAR_t$) to focus our analysis on startup evaluations within a particular country in a year to account for differences across startup headquarters countries (within regions) over time. These fixed effects allow us to account for quality differences between, for example, a U.K.-based startup and a Latvia-based startup within Europe. In our most stringent specification, we focus on evaluations at the startup level in a given year (across multiple judge evaluators) using startup-year fixed effects ($STARTUP_i xYEAR_t$), enabling us to account for differences in individual startup quality within countries and across years. We cluster robust standard errors at the judge and startup levels.

To assess whether foreign discounting is driven by judges being better at evaluating local startups, we estimate a model similar to Li (2017) that measures the sensitivity of judges' scores to local vs. foreign startups' performance measures. This model allows us to discern the remaining scenarios in Figure 1, including whether judges are informed and biased (Figure 1, cell F), informed only about local startups and biased (Figure 1, cell D), informed only about local startups and unbiased (Figure 1, cell C), or uninformed about all startups and biased (Figure 1, cell B).

$$(2)\ SCORE_{ijt} = \propto + \beta FOREIGN_{ij} + \delta POSTACCELERATOR\_PERFORM_i + \phi FOREIGN_{ij} xPOSTACCELERATOR\_PERFORM_i + JUDGE_j xYEAR_t + STARTUPCOUNTRY_i xYEAR_t + \epsilon_{ijt}$$

Where $POSTACCELERATOR\_PERFORM_i$ indicates the logged page visits of startups after the accelerator application review cycle begins (in March 2017 and 2018) until July 2019 (our latest available data). Our main coefficients of interest are β, which indicates whether judges discount foreign startups, δ, which indicates whether judges can discern winners from losers overall, and ϕ, which shows the extent to which judges can discern winners from losers among foreign startups versus among local startups. A positive and significant β indicates that, given a certain level of quality among startups, judges discount foreign versus local startups. A positive and significant δ indicates that judges are able to discern winners from losers among startups overall. A negative and significant ϕ indicates that judges are less sensitive to the quality of foreign versus local startups, and therefore judges are better able to discern winners from losers among local versus foreign startups. A concern with this approach is that the accelerator itself is impacting the post-accelerator performance of startups, which confounds our ability to proxy the potential of startups at the time of judging. We account for such possible treatment effects by controlling for startups' participation in the accelerator program. Prior to accelerator participation, startups get little recognition or feedback for participating in the competition.

VI. Results

Are foreign judges actually randomly assigned?

Our ability to measure the presence and impact of foreign discounting hinges on the assumptions that startups and judges are randomly assigned. For this assumption to hold, we need to net out the effect of startups selectively choosing where to apply (impacting their distribution of judges) and judges choosing which startups to assess. In normal in-person rounds of the competition, this random assignment is difficult to confirm because it is less feasible for foreign judges to travel to international locations to judge startups. Because startups choose where to apply, and there are mostly local judges in those locations, by design, startups choose which judges (in geographic terms) evaluate them.

To check random assignment, we use chi-squared tests (Tables 2a-b). These chi-squared tests allow us to measure whether there is a statistically significant difference between a predicted distribution of startup-judge regions under random assignment versus the actual distribution of pairs observed in the data. In 2017, there is no statistically significant difference between the predicted distribution of startup-judge region assignments under random allocation and the observed distribution. Thus, we cannot reject the null hypothesis that startup-judge assignments on basis of geography are random. In 2018, we see that we can reject this null hypothesis because of the perhaps non-random assignment of Israeli judges to European startups, a fairly small share (0.14%) of our sample, representing 15 startups. However, when we take out Israeli judges, we see a similar situation as in 2017. The distribution is again consistent with random assignment. No matter, our results hold if we include or exclude these non-randomly evaluated European startups.

[Insert Tables 2a-b]

Is there foreign discounting of startups?

We now turn to whether judges discount foreign startups. Summary statistics of scores for startups that match the geography of the judge show that, on average, the main composite score, recommend, and subscores are lower for startups that do not match geographies (that is, for foreign startups) versus those that do (Table 1).

Figure 2 also reveals that the distribution of scores from judge evaluations of foreign startups are lower on average than those of local startups. We confirm in a two-sample Kolmogorov-Smirnov test that the two distributions are statistically different from one another (p=0.000). However, this graph may reflect the fact that most judges in our sample are U.S.-based. Startups that are therefore foreign are likely to be those that are non-U.S. based, and non-U.S. based startups may be worse quality on average than are U.S.-based startups.

[Insert Figure 2]

We account for these regional quality differences in our regression models. Table 3a (column 1) shows that when only controlling for judge-year fixed effects (and not differences in region or startup quality), judges give 0.2 standard deviation lower scores to foreign vs. local startups. Table 3a (column 2) adds startup region-year fixed effects to account for regional variations among startups. This estimate also reveals judges' discounting of foreign startups, though lower at -0.063 standard deviation, indicating that regional-level startup quality differences account for about 70 percent of the effect. Table 3a (column 3) adds more restrictive startup country-year fixed effects. We find foreign discounting in this specification as well at -0.063. Table 3a (column 4) adds startup-year fixed effects and finds a similar effect of -0.059. This suggests that there is little in the way of systematic differences between startups within regions. The coefficient that accounts for regional quality is about 29 percent of the original effect, suggesting that regional differences in startup quality account for about two-thirds of the foreign discounting effect, and judges account for one-third.

Column 5 in Table 3a includes measures for whether a startup has user traction and financing at the time of the application. Controlling for these pre-accelerator quality measures allows us to compare the coefficients on the variables to our estimate of home bias. The home bias effect (-0.06) is about 30 percent of the size of a startup having user traction (0.20) and about 8 percent of the effect of a startup having raised a round of financing (0.71) at the time of the application. This effect appears non-trivial.

Table 3b is similar to Table 3a, but uses our binary measure of whether a judge recommended a startup to the next round of the competition as the dependent variable. Judges are less likely to recommend foreign vs. local startups to the next round by 9 percentage points before accounting for quality. This coefficient remains significant and negative, but falls to 4 percentage points when accounting for region-year fixed effects (column 2), country-year fixed effects (column 3), and startup-year fixed effects (column 4), indicating that judge preferences account for over 40 percent of the foreign bias effect.

[Insert Tables 3a-b]

Together, these results reveal that judges consistently give a lower evaluation to foreign versus local startups. This result rules out that judges are informed and unbiased (Figure 1, cell E) and that judges are uninformed and unbiased (Figure 1, cell A).

Does foreign discounting result from judges better evaluating local startups?

We next proceed to analyses that allows us to discern whether the foreign discounting effect reflects that judges are informed and biased (Figure 1, cell F), informed only about local startups and biased (Figure1 1, cell D), informed only about local startups and unbiased (Figure 1, cell C), or uninformed about all startups and biased (Figure 1, cell B). We disentangle these alternatives in two ways. First, we evaluate whether judges are better able to select winners from losers among local startups, revealing whether judges are more informed about local versus foreign startups, either isolating or ruling out scenarios D and F. Second, we assess whether judges discount foreign startup ideas only when there are few quality signals available about a startup, allowing us to isolate whether judges are biased, thereby confirming or ruling out scenarios B, D, and F.

We first assess whether judges are better able to select winners from losers among local versus among foreign startups. Figure 3 shows a binscatter graph depicting the relationship between startups' website page visits 1-2 years after the program (x-axis) and the scores given by judges (y-axis), after incorporating judge-year and startup country-year fixed effects. The slope of the best fit line depicting the relationship between these two variables indicates the sensitivity of judge scores to a measure of startup quality. We break this best fit line into two buckets of evaluations akin to our regression specification: when judges evaluate startups local to their regions (blue solid line) and when judges evaluate startups foreign to their regions (red dashed line). The steeper the slope of these lines, the better are judges at discerning the potential of these startups. By comparing the differences in the slope between these two lines, we visually can test if judges are better at evaluating local versus foreign startups. Both lines have a positive slope, suggesting that judges can separate high potential startups from those destined to fail among both local and foreign startups. The fact that the blue solid line depicting local startup evaluations is above the red dashed line across the quality spectrum suggests that judges give an across-the-board penalty to foreign startups no matter their quality. Further, the blue and red lines are similarly sloped. It does not appear that judges are better able to pick winners and losers among local versus among foreign startups.

[Insert Figure 3]

We confirm in regressions that judges are not any better at evaluating local startups relative to foreign startups. Table 4 (column 1) reveals that the coefficient on the interaction term between foreign startups and logged post-page visits ($FOREIGN_{ij}xPOSTACCELERATOR\_PERFORM_i$) is not significant, suggesting that there is no difference in the relationship between startup quality and judge scores among foreign versus local startup evaluations. Consistent with Figure 3, we do find judge scores predict startup quality. To overcome concerns that the accelerator program itself impacts post-accelerator performance results, we control for accelerator participation in Table 4 column 2. While accelerator participation has a positive effect on post-accelerator startup page visits, it does not meaningfully account for the foreign discounting effect.

These results suggest that judges are able to detect quality of all startups with relatively equal precision. Thus, if U.S. startups are on average better quality than are European startups, judges scores across the board would be lower for European startups than for U.S. startups. However, in contrast to work showing a local information advantage (Malloy, 2005; Coval and Moskowitz, 2001), we do not find that judges are more sensitive to quality of startups within their home regions than of those from other regions. This analysis allows us to rule out remaining scenarios in which judges are only informed about local startups or uninformed all-together (scenarios C, D, and B). This leaves the possibility that judges are informed about all startups and biased against foreign startups (scenario F) as the remaining explanation. Indeed, Figure 3 closely reflects the predicted situation in Figure 1 cell F.

[Insert Table 4]

Next, we confirm that scenario F holds by showing that judges' discounting is driven by bias rather than expertise or information to compensate for uncertainty around startup quality. If it were the case that foreign discounting reflected judges' use of expertise about their own geography to compensate for uncertainty around startup quality, then when uncertainty declined, judges would not rely as much on geography as a heuristic for quality. Therefore, foreign discounting would decline. The provision of information about startups, such as whether they received financing, should reduce uncertainty about a startup's quality (Lerner, Schoar, Sokolinski, and Wilson, 2018). We therefore test this channel by comparing the foreign discounting effect among startups that have stronger versus weaker quality signals – in particular whether they have raised financing and user traction. To do so, we restrict equation (1) to startups that raised some financing prior to the program and separately to those who have not to see whether foreign discounting persists when there are stronger quality signals provided about the startups.

Figure A3 shows the relationship between judge scores on the y-axis and pre-accelerator financing of startups on the x-axis. It reveals that evaluations of foreign startups (red dashed line) always get a lower score than those of local startups (blue solid line) across the distribution of startup quality. This illustration suggests that judges do not change their reliance on geographic signals even when given more information about a startup's potential.

We confirm these results in regressions in Table 5a. In this table, we assess whether judges discount foreign startups using the specification from equation 1. We apply this specification to two sub-samples within our data: (1) evaluations involving startups who have no financing at the time of application (column 1) and (2) evaluations involving startups who have some financing indicating stronger quality signals (column 2). We find that judges' bias against foreign startups is stronger among startups with financing (-0.2) than those without financing (-0.04). We confirm that the magnitude of the foreign bias is larger in column 3 by including an interaction term between whether a startup is foreign and whether it has raised financing.[9]

We find a similar effect when assessing home bias on the basis of whether quality signals around user traction are available at the time of the application in Table 5b. We first confirm that judges discount foreign startups without user traction (column 1). While not significant at the 5 percent level, this effect is a similar magnitude to earlier results (-0.05). Foreign discounting not only persists among startups with user traction, but actually gets bigger (-0.07) (column 2), suggesting that judges rely even more on geography when stronger quality signals are available

[9] We also check whether the adjusted R-squared differs across the financed and non-financed samples. If judges rely less on common geography when other quality signals are available, then the adjusted R-squared should be lower for the financed than for the non-financed sample. We find that the adjusted R-squared is similar for financed (0.58) and for non-financed (0.63) samples.

about the startups.[10] The foreign discounting effect among startups without user traction is not statistically significantly different from those with user traction (column 3). This may be the case because website page visits are a proxy for user traction, which judges may not observe directly in the startups' application forms, as they do with financing. Therefore, page visits may not be as obvious of a quality signal at the time of evaluation.

[Insert Tables 5a-b]

Together these results confirm that foreign discounting reflects a bias rather than expertise or information compensating for uncertainty. The effect may grow larger in magnitude with the provision of more information, rather than remaining constant, because the information reduces some noise around the quality of startups. Therefore, some of the foreign discounting effect in the lower information scenario may have been a result of measurement error around startup quality that was in an opposite direction of the bias.

Another possible explanation of why we may see a higher magnitude foreign bias effect for financed startups versus non-financed startups is because of the identity of the investors. In particular, if these early stage startups received pre-accelerator investment from investors local to them, then the home bias effect may reflect discounting against foreign investors. This might happen for several reasons. First, it could be more complicated to allocate prize money to startups if there are pre-existing foreign financing arrangements. Second, judges may perceive the foreign investors to be lower quality, inhibiting the future performance of the startups. Third, judges may not be aware of the quality of the foreign investors and therefore discount them on the basis of uncertainty. Another reason may be that judges give a bigger boost to local vs.

[10] The adjusted R-squared, like in the financing case, is similar for startups with page traction (0.62) and for those without (0.65).

foreign startups when they have stronger quality signals. Alternatively, judges may be more comfortable acting on their bias when they perceive foreign startups to have outside options.

Overall, our results are consistent with a model in which judges are informed about the quality of startups and are biased against foreign startups (Figure 1, cell F). While judges' expertise does account for enough of the unobserved information to accurately choose winners from losers among foreign and local startups, they systematically discount foreign startups relative to local ones.

Robustness

We run a number of robustness checks. We first confirm that our foreign bias result holds in the broader sample of data from the accelerator to ensure the result is not specific to our particular data sample in the global rounds. Specifically, we use data from 2013-2019 across all four rounds of the competition. We find that judges give a 0.06 standard deviation discount to foreign startups and are 4 percentage points less likely to recommend foreign startups to the next round of the competition in this larger sample (Table A1). That said, these rounds do not exhibit random assignment of judges to startups on the basis of location, raising selection concerns. Despite these concerns, the fact that this estimate from the full data on 75,917 evaluations, including 9,415 startups and 3,696 judges[11], is consistent with earlier results on foreign bias lends credence to the idea that judges discount foreign startups.

We also confirm that our estimate on the ability of judges to pick winners and losers among foreign versus local startups holds with different measures of startup pre-accelerator and post-accelerator performance. This robustness check ensures that our results in Table 4 are not specific to the post-accelerator page visits metric of quality. We see a similar pattern between

[11] These evaluations only include startups who come from the same regions as judges' home regions to ensure that startups are both local and foreign to judges in our sample.

startup performance and judge scores holds using measures of logged page visits 3 months before the program, logged financing raised before the program, and logged financing raised six months after the accelerator (Figures A2-A4). We also confirm the results hold in regression models (Table A2). While the pre-accelerator indicators reveal judges' ability to predict current quality of startups, post-accelerator indicators reveal judges' ability to predict the future potential of startups, both of which are key considerations in the entrepreneurial evaluation process.

A concern with our results may be that judges are in the process of learning how to detect quality of startups, and over time, seeing results in the market should reduce their biases, as seen in the baseball card exchange market (List, 2006). If judges have sufficient prior experience evaluating startups, it is unlikely that their evaluations reflect this learning lag. In our global round data, each judge, on average, evaluated 56 startups in prior years/rounds of the program, which is the majority (77 percent) of their evaluations to date. While they did not receive direct feedback on their evaluation accuracy, they had sufficient startup outcomes to observe in the market, so it is unlikely that learning behavior would confound our results.

Lastly, we confirm that the foreign discounting effect exists across all judge-startup foreign pairings. This analysis allows us to address concerns that any particular judge pairing (e.g. U.S. judges evaluating Latin American startups) dominates the effect. While we lack the power to conduct our regression specification (in equation 1) isolated to each judge-startup foreign pairing, we are able to show in summary statistics that evaluations involving judges and startups from different regions tend to have a lower recommendation probabilities than evaluations involving judges and startups from the same region. Figure A8 shows average recommendation probabilities across startup-judge regional pairings. The lowest recommendation probabilities are given by U.S. judges to Latin American startups, European judges to Latin American startups, and U.S. judges to Israeli startups. Indeed, U.S. judges are particularly harsh (which we account for in our judge fixed effects).

Does foreign discounting cause judges to pass on promising foreign startups?

Our results above show that across the quality distribution, judges give lower scores to foreign startups. While our graphs suggest this effect holds across the quality distribution, it is possible that this discounting has little impact on which startups judges select for the next round. For example, perhaps judges discount high quality foreign startups who, though moved down in the ranking, still would be make it to the next round. Conversely, judges may discount low quality foreign startups who would not make it to the next round regardless. In these extreme cases, foreign discounting would not impact the marginal difference of making it to the next round. However, for startups in the middle of the quality distribution, this foreign discounting may make a marginal difference on whether they make it to the next round of the competition.

In Table 6, we estimate the number of "missed foreign startups," for whom foreign discounting does make a marginal difference, by comparing judges' actual recommendation decisions to the judges' decision if they only relied on the "quality dependent" component of their score. We isolate the "quality dependent" portion of the judges' scores by regressing judge decisions on our startup quality measures. This reveals the judges' weights on different measures of startup quality and allows us to construct counterfactual rankings as if judges were not foreign biased.[12] Specifically, we use these estimated weights to predict for each judge a new ranking for the startups they would be most likely to recommend. We then select the same number of recommendations for the next round as the judges actually selected. Appendix A9 provides further information about this calculation.

[12] If foreign startups are lower quality, then judge could still discount them. However, our argument is that judges have a direct bias against foreign startups that is not mediated by quality.

We find that 3,024 startup evaluations (comprising 16.8 percent of all evaluations) are not recommended to the next round of the competition, though they would be in our counterfactual analysis. We find that 521 startups, comprising over 13 percent of all startups, would have made it to the next round of the competition in the counterfactual, but do not in actual judge recommendations.

[Insert Table 6]

VII. Conclusion and Implications

We find that judges can discern the quality of local and foreign startups with similar ability. However, they discount foreign startups no matter their potential. Judges are less likely to recommend foreign startups by 4 percentage points, equivalent to roughly one half of the effect of having some user traction, or a tenth of the effect of going from no financing to some venture capital or angel financing. Stronger quality signals (e.g. financing) do not appear to attenuate this bias. Back-of-the-envelope estimates suggest that this bias results in the potential exclusion of over 1 in 10 highly promising foreign entrepreneurial ideas. These results reveal that judges are informed about quality of both local and foreign startups, but they are biased against foreign startups. Therefore, while they have an ability to select the most promising startups, they systematically discount foreign startups relative to local ones.

Our results suggest that startups from remote locations may fail when they try to move or scale to “hubs” – despite lower costs to such movements because of the advent of IT and cloud technology. Indeed, recent work reveals such home bias in online investment platforms, as California-based non-expert investors prefer California-based ventures (Lin and Viswanathan, 2016). The bias may contribute to explanations to why ventures tend to perform better when located in the native regions of founders (Dahl and Sorenson, 2012) and why investment from "high-status" venture capital firms is most beneficial when the home country of the venture capital firm is more connected to that of the venture (Alvarez-Garrido and Guler, 2018). The bias of judges may therefore constrain the ability of accelerators to be vehicles for startups from remote locations to expand into "hub" markets. Consequently, accelerators may not be able to help startups scale globally as much as they intend, despite their benefits for startup performance found in other studies (Hallen, Cohen, and Bingham, 2020; Cohen, Bingham, and Hallen, 2019; Yu, 2020; Howell, 2017; Gonzalez-Uribe and Leatherbee, 2018; Fehder and Hochberg, 2014).

Foreign bias also may impact the direction of innovation. If accelerators select out startups from remote regions, which are more likely to be foreign to accelerators or investors, they reduce the probability that innovations addressing the needs of those markets will survive and grow. This distortion is similar to effects seen in studies of bias in gender and race contexts where innovations are targeted at the most represented populations and are therefore less likely to benefit less represented populations (e.g. Koning, Samila, and Ferguson, 2020). Similarly, in our geographic context, if accelerator programs overlook promising foreign startups from remote locations that lack sufficient financing and other support resources nearby, then we may see a gap in startup innovations that benefits the needs of those underrepresented remote markets.

While we find that startups face a "liability of foreignness" (Zaheer, 1995), with the across-the-board discount given by foreign judges, we notably do not find that judges face a disadvantage in evaluating foreign startups. Instead, we find that judges can discern quality of startups across regions. This may be because startup practices have standardized into a "playbook" that is comparable across countries, for example, with the proliferation of codified management (Bloom and Reenen, 2007; Chatterji, Delecourt, Hasan, and Koning, 2019) and technology practices (Haefliger, Von Krogh, and Spaeth, 2008). The existence of such a playbook may reduce the need for private information (Malloy, 2005) or "contextual intelligence" (Khanna, 2014) to evaluate local opportunities. While our results reveal that gatekeepers may lack a local information advantage to screen startups, we cannot reject that they may lack a local advantage with respect to deal flow and monitoring, as shown by other research. The deal flow advantage may come from localized networks that provide venture capitalists a higher quality pool of local startups ex-ante (Sorenson and Stuart, 2001). Furthermore, gatekeepers may have a local advantage in adding value to startups because of their closer proximity to monitor startups and intervene when necessary (Bernstein, Giroud, and Townsend, 2016).

Together, our results show that gatekeepers may benefit from looking more globally for startups, given their ability to discern startup quality. This global search need not be done in localized programs, as is the current standard in accelerator and venture capital models with local subsidiaries around the world (e.g. California-based venture capital firm Sequoia has offices in eight other cities around the world). Rather, it may occur directly out of the headquarters location. This global approach, however, can only yield a high quality pool of startups if organizations reduce bias against foreign startups. This bias both reduces the overall quality of selected startups and their diversity.

There are at least two interventions that accelerators, investors, and other entrepreneurial gatekeepers may pursue to reduce the distortionary effects of foreign bias on startup selection. First, interventions that reduce the extent to which individual-level biases impact the future opportunities of startups may alleviate the repercussions of home bias on entrepreneurial growth and innovation. For example, globally distributed accelerators, venture capital firms, and other gatekeepers may lower their threshold for selecting a foreign firm versus a local firm. They also may aggregate biases across a diverse set of gatekeepers to average them out. These global models also may enable gatekeepers to select startups from around the world – a viable option given that we do not find a local informational advantage that would constrain the ability of gatekeepers to evaluate foreign startups. An investor from Europe, for example, is just as good as an investor from the U.S. in sorting winners from losers among U.S. startups. Committee-based judging that includes a diverse set of evaluators similarly may wash away these biases and reduce future distortions.

Second, interventions that directly reduce individual-level biases also may be valuable for reducing distortionary implications. For example, accelerators and investors may build up the experience of judges in evaluating internationally-diverse startups or select judges with such experience. Nudges reminding judges of their unconscious biases with regards to geography also may reduce the individual incidences of bias. Lastly, interventions that either make home country location more salient to avoid subconscious biases or limit the provision of such information all together that elicit biased responses, such as the name of origin country or native language, further may reduce distortion in individual evaluations.

There are also at least two approaches that startups may pursue to alleviate the adverse implications of foreign bias. One approach is to diversify the geographic affiliation of startup teams in order to appeal to a wider set of gatekeepers. For example, a startup team with U.S. and European founders may appeal to both U.S. and European gatekeepers. A second approach is to vary information about the startup's geography to gatekeepers. For example, startups either may include information to make their home region even more salient to reduce subconscious foreign biases or exclude it all together to avoid triggering home bias in evaluations. Startups also may get translation support to effectively express their application responses in a non-native language

(English), as often required in international investor and accelerator programs. Future research may precisely measure the effect of these interventions on the prevalence of foreign bias and the future performance of startups.

Tables

Table 1: Summary Statistics

	Local Startup					Foreign Startup					Local-Foreign
	No. Obs.	Mean	SD	Min	Max	No. Obs.	Mean	SD	Min	Max	Difference in Means
Judge Score Measures											
Composite Score	7232	0.01	1.01	-3.31	2.36	9425.00	-0.12	1.05	-3.31	2.36	0.13***
Recommend	7706	0.61	0.49	0.00	1.00	10247.00	0.56	0.50	0.00	1.00	0.05***
Subscore: Customer Needs and Acquisition	7692	6.25	1.86	1.00	10.00	10173.00	6.08	1.91	1.00	10.00	0.17***
Subscore: Customer Pain and Solution	7694	6.82	1.84	1.00	10.00	10179.00	6.64	1.95	1.00	10.00	0.18***
Subscore: Financial Business Model	7675	5.72	1.98	1.00	10.00	10125.00	5.54	2.07	1.00	10.00	0.18***
Subscore: Industry and Competitor	7690	6.11	1.85	1.00	10.00	10166.00	5.94	1.93	1.00	10.00	0.17***
Subscore: Overall Impact	7686	6.21	1.93	1.00	10.00	10160.00	6.04	2.00	1.00	10.00	0.17***
Subscore: Regulation and IP	7261	5.91	2.15	1.00	10.00	9498.00	5.65	2.25	1.00	10.00	0.26***
Subscore: Team and Advisors Investors	7678	6.51	2.01	1.00	10.00	10143.00	6.32	2.08	1.00	10.00	0.19***
Startup Quality Measures											
Log Pre-Accelerator Total Page Visits	3916	1.36	2.75	0.00	12.50	6117.00	1.45	2.87	0.00	12.50	-0.09
Log Pre-Accelerator Financing	7706	0.47	1.44	0.00	6.12	10247.00	0.46	1.41	0.00	6.12	0.01
Log Post-Total Page Visits	7706	2.87	3.52	0.00	12.82	10247.00	2.93	3.60	0.00	12.82	-0.06
Log Post-Accelerator Financing	7706	0.56	1.63	0.00	6.68	10247.00	0.54	1.58	0.00	6.68	0.02
Has User Traction	7706	0.60	0.49	0.00	1.00	10247.00	0.53	0.50	0.00	1.00	0.07***
Has Financing	7706	0.12	0.33	0.00	1.00	10247.00	0.13	0.33	0.00	1.00	0.00

Notes: The table reports descriptive statistics of scores provided by judges on the sample of 17,959 startup-judge pairings from the 2017 and 2018 global round.

Table 2a: Chi-Squared - 2017

	Pearson chi2(4) = 1.4494 Pr = 0.836			
		Judge Subregion		
Startup Subregion	Europe	U.S. & Canada	Israel	Total
Europe	230	791	206	1,227
	239.4	*784.3*	*203.3*	
U.S. & Canada	1,008	3,325	860	5,193
	1,013.30	*3,319.30*	*860.4*	
Israel	300	922	240	1,462
	285.3	*934.5*	*242.2*	
Total	1,538	5,038	1,306	7,882
	1,538.00	*5,038.00*	*1,306.00*	

Non-italicized numbers indicate frequency. Italicized numbers indicate the expected frequency of the cell counts if they were randomly assigned based on the marginal distributions.

Table 2b: Chi-Squared – 2018

With Israeli Judges: Pearson chi2(9) = 23.1737 Pr = 0.006

Without Israeli Judges = Pearson chi2(6) = 7.9187 Pr = 0.244

	Judge Subregion			
Startup Subregion	Europe	Latin America	U.S. & Canada	Israel
Europe	566	153	1,389	**25**
	594.8	*177.5*	*1,347.30*	***13.4***
Latin America	705	213	1,539	11
	688.3	*205.4*	*1,558.90*	*15.5*
U.S. & Canada	1,406	432	3,135	23
	1,393.30	*415.8*	*3,155.60*	*31.3*
Israel	37	12	84	2
	37.6	*11.2*	*85.3*	*0.8*
Total	2,714	810	6,147	61

Non-italicized numbers indicate frequency. Italicized numbers indicate the expected frequency of the cell counts if they were randomly assigned based on the marginal distributions.

Table 3a: Foreign Bias – Score

	(1)	(2)	(3)	(4)	(5)
			Judge's Total Score		
Foreign Startup	-0.204***	-0.063**	-0.063**	-0.059***	-0.059***
	(0.021)	(0.020)	(0.020)	(0.016)	(0.018)
Has User Traction					0.201***
					(0.028)
Has Financing					0.714***
					(0.022)
Observations	16,639	16,639	16,639	16,593	16,639
Judge x Year	Yes	Yes	Yes	Yes	Yes
Startup Region x Year	No	Yes	No	No	No
Startup Country x Year	No	No	Yes	No	Yes
Startup x Year	No	No	No	Yes	No

Standard errors (shown in parentheses) are clustered at the judge and startup levels.
* p<0.05 ** p<0.01 *** p<0.001

Table 3b: Foreign Bias – Recommend

	(1)	(2)	(3)	(4)	(5)
	Judge Recommends Startup?				
Foreign Startup	-0.091***	-0.038***	-0.039***	-0.038***	-0.037***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Has User Traction					0.085***
					(0.014)
Has Financing					0.342***
					(0.010)
Observations	17,938	17,938	17,938	17,938	17,938
Judge x Year	Yes	Yes	Yes	Yes	Yes
Startup Region x Year	No	Yes	No	No	No
Startup Country x Year	No	No	Yes	No	Yes
Startup x Year	No	No	No	Yes	No

Standard errors (in parentheses) are clustered at the judge and startup levels.
* p<0.05 ** p<0.01 *** p<0.001

Table 4: Quality Interaction Using Ex-Post-Indicators

	(1)	(2)
	Judge's Total Score	
Foreign Startup	-0.068**	-0.066**
	(0.025)	(0.024)
Log Post-Accelerator Page Visits	0.049***	0.037***
	(0.004)	(0.004)
Foreign Startup * Log Post-Accelerator Page Visits	0.001	0.001
	(0.005)	(0.004)
Accelerator Participation		0.625***
		(0.024)
Observations	16,639	16,639
Judge x Year	Yes	Yes
Startup Country x Year	Yes	Yes
Startup x Year	No	No

Standard errors (in parentheses) are clustered at the judge and startup level.

* p<0.05 ** p<0.01 *** p<0.001

Table 5a: Foreign Bias – Score for Pre-Accelerator Financed vs. Non-Financed Startups

	(1)	(2)	(3)
		Judge's Total Score	
Foreign Startup	-0.035*	-0.195***	-0.040*
	(0.017)	(0.041)	(0.017)
Foreign Startup * Has Financing			-0.143***
			(0.038)
Observations	14,404	1,782	16,598
Financed Startups	No	Yes	Yes
Non-Financed Startups	Yes	No	Yes
Judge x Year	Yes	Yes	Yes
Startup x Year	Yes	Yes	Yes

Standard errors (shown in parentheses) are clustered at the judge and startup levels.

* $p<0.05$ ** $p<0.01$ *** $p<0.001$

Table 5b: Foreign Bias – Score for Startups With and Without Pre-Accelerator Page Traction

	(1)	(2)	(3)
	Judge's Total Score		
Foreign Startup	-0.046	-0.074**	-0.047*
	(0.024)	(0.023)	(0.023)
Foreign Startup * Has User Traction			-0.022
			(0.031)
Observations	7,348	9,088	16,598
Startups w/ Page Traction	No	Yes	Yes
Startups w/ No Page Traction	Yes	No	Yes
Judge x Year	Yes	Yes	Yes
Startup x Year	Yes	Yes	Yes

Standard errors (shown in parentheses) are clustered at the judge and startup levels.
* p<0.05 ** p<0.01 *** p<0.001

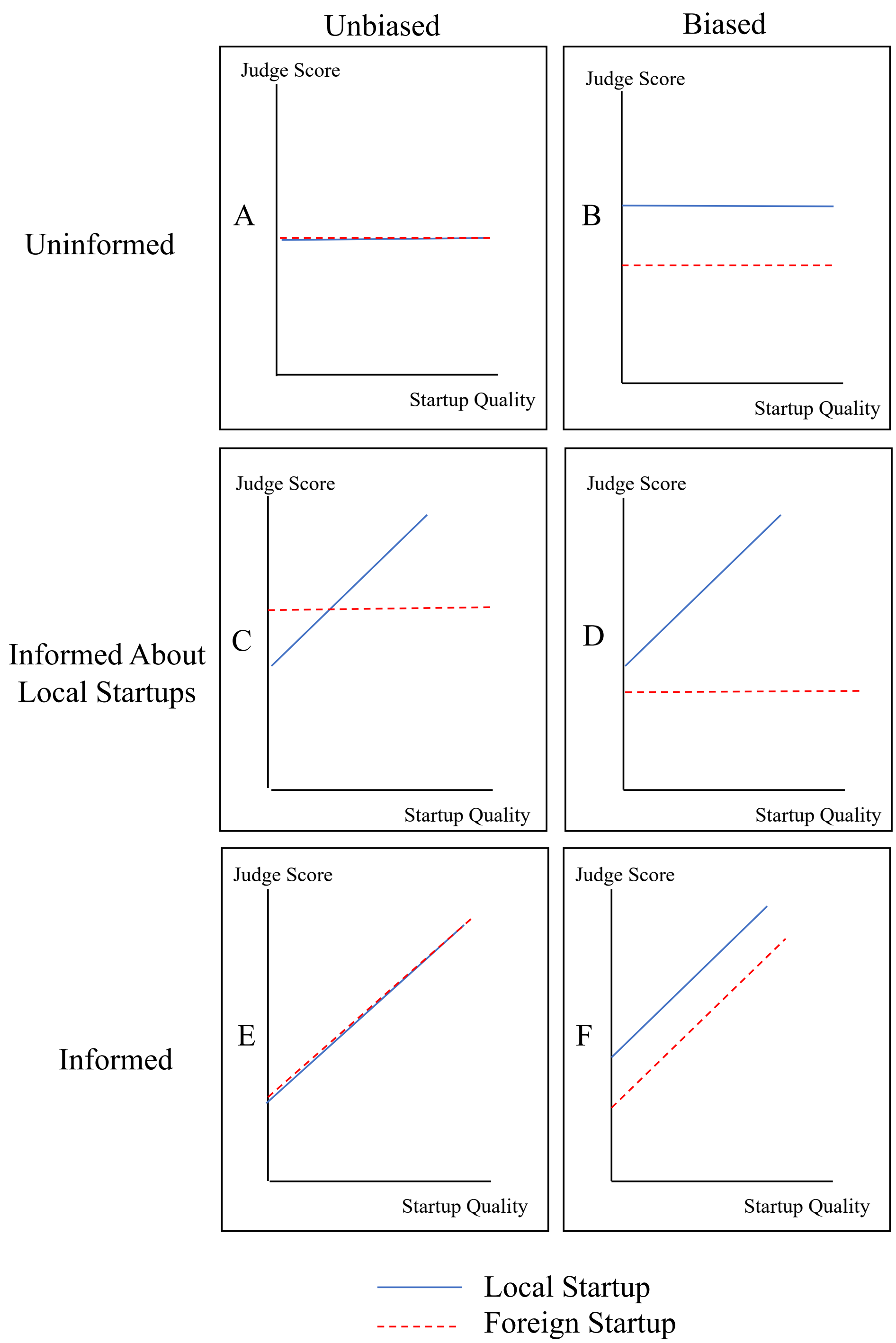
Table 6: Predicted "Missed Out" Startups

Num. Promising Foreign Startup Evaluations Not Recommended	Percent of Evaluations	Implied Num. of Startups	Percent of Startups Evaluated
3,024	16.8	521.0	13.8

The table shows the number of foreign startup evaluations that are predicted to be recommended to the next round of the competition based on pre- and post- metrics of quality, but were not actually recommended to the next round of the competition in the 2017-2018 global rounds. Composite quality includes log pre-accelerator financing, log post-accelerator financing, and log post-accelerator page visits. We calculated the number of implied startups passed over as the number of startups who are recommended to the next round when counterfactual recommendations are averaged across judges, but are not recommended to the next round based on averaged actual recommendations across judges. The regression to generate predictions includes startup country x year and judge x year fixed effects and a control for participation in the accelerator. It clusters robust standard errors at the startup and judge levels.

Figures

Figure 1: Predicted Relationships Between Judge Scores and Startup Quality



Note: The level shifts in cells B, D, and F show a negative bias against foreign startups for illustrative purposes, but conceptually this bias also may be positive.

Figure 2: Kernel Density Plot of Scores

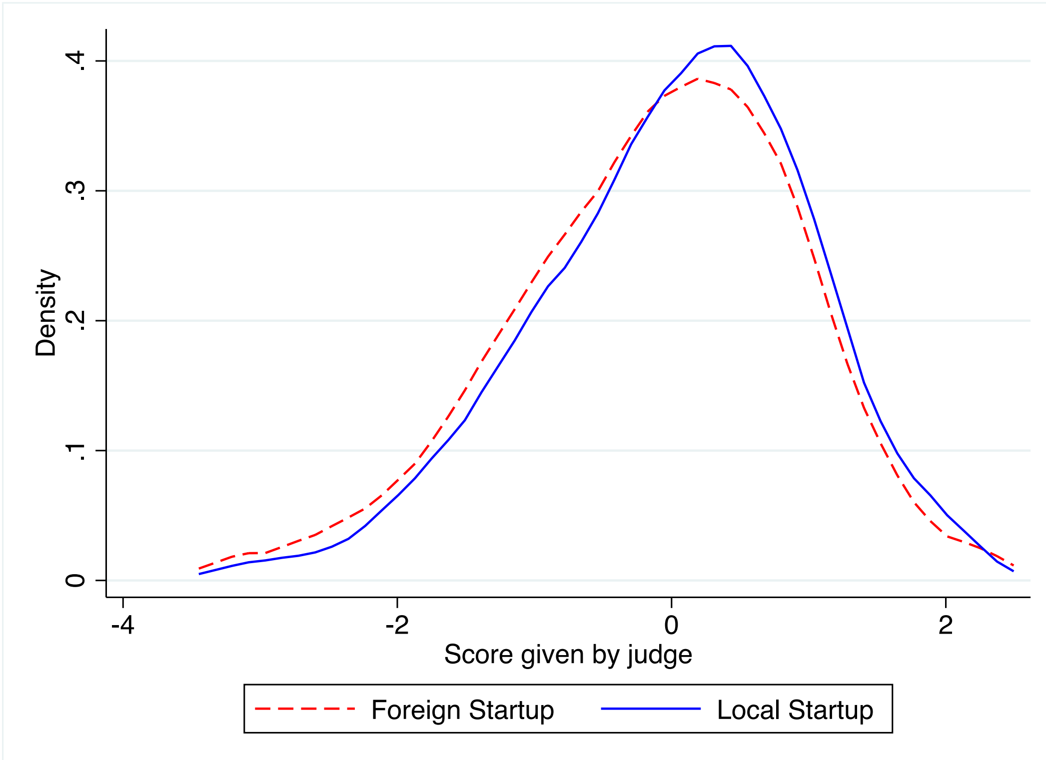
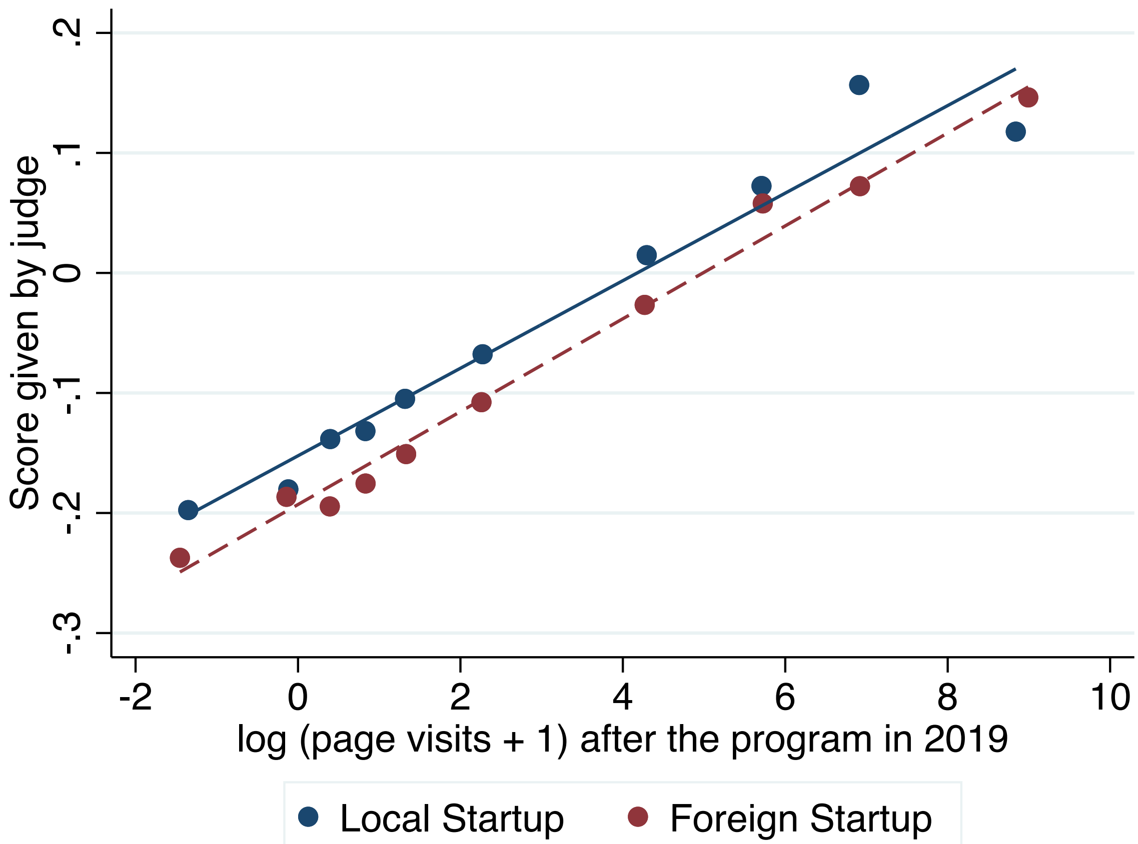


Figure 3: Judge Ability to Predict Future Performance By Whether a Startup is Foreign or Local Using Post-Accelerator Page Visits



The graph includes judgexyear and startup countryxyear fixed effects.
The y-axis indicates log (page visits + 1) after the program in 2019.
The graph shows 2017-2018 global round data.
We use 10 bins.
We control for participation in the accelerator program.

Appendix

Tables

Table A1: Foreign Bias – Score (Full Sample)

	(1)	(2)
	Judge's Total Score	Judge Recommends Startup?
Foreign Startup	-0.064***	-0.041***
	(0.012)	(0.007)
Observations	70184	75732
Judge x Year	Yes	Yes
Startup x Year	Yes	Yes
Program x Year	Yes	Yes

The table shows evaluations across rounds in 2013-2019. Standard errors (shown in parentheses) are clustered at the judge and startup levels.
* p<0.05 ** p<0.01 *** p<0.001

Table A2: Quality Interaction Using Additional Quality Indicators

	(1)	(2)	(3)	(4)	(5)
			Score		
Foreign Startup	-0.044*	-0.044	-0.054**	-0.042	-0.054**
	(0.019)	(0.027)	(0.019)	(0.027)	(0.019)
Has Financing	0.784***				
	(0.029)				
Foreign Startup * Has Financing	-0.111**				
	(0.036)				
Has User Traction		0.257***			
		(0.036)			
Foreign Startup * Has User Traction		-0.035			
		(0.035)			
Log Pre-Accelerator Financing			0.154***		
			(0.007)		
Foreign Startup *Log Pre-Accelerator Financing			-0.014		
			(0.008)		
Log Pre-Accelerator Page Visits				0.036***	
				(0.006)	
Foreign Startup * Log Pre-Accelerator Page Visits				-0.006	
				(0.007)	
Log Post-Accelerator Financing					0.080***
					(0.009)
Foreign Startup *Log Post-Accelerator Financing					-0.009
					(0.007)
Accelerator Participation					0.394***
					(0.043)
Observations	16,645	16,645	16,645	9,347	16,645
Judge x Year	Yes	Yes	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes	Yes	Yes
Startup x Year	No	No	No	No	No

Standard errors (in parentheses) are clustered at the judge and startup level.
* p<0.05 ** p<0.01 *** p<0.001

Figures

Figure A1a: Diagram of Imperfect Information in Local Startup Quality

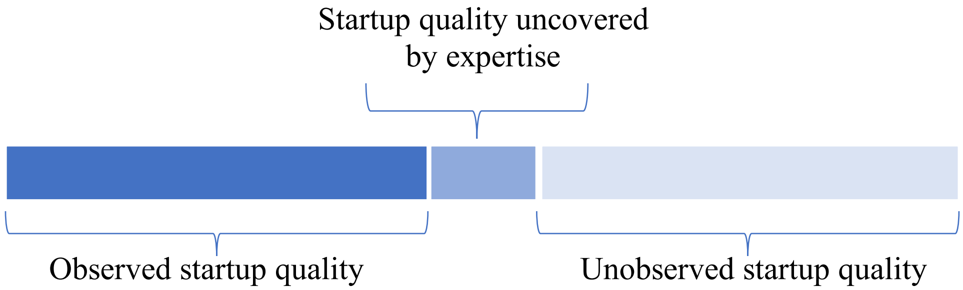


Figure A1b: Diagram of Imperfect Information in Foreign Startup Quality

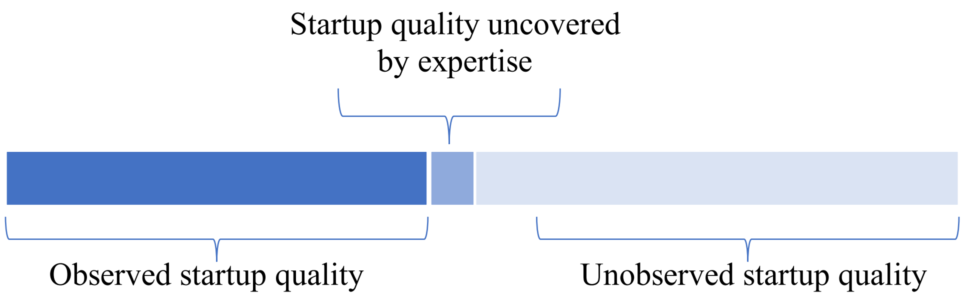
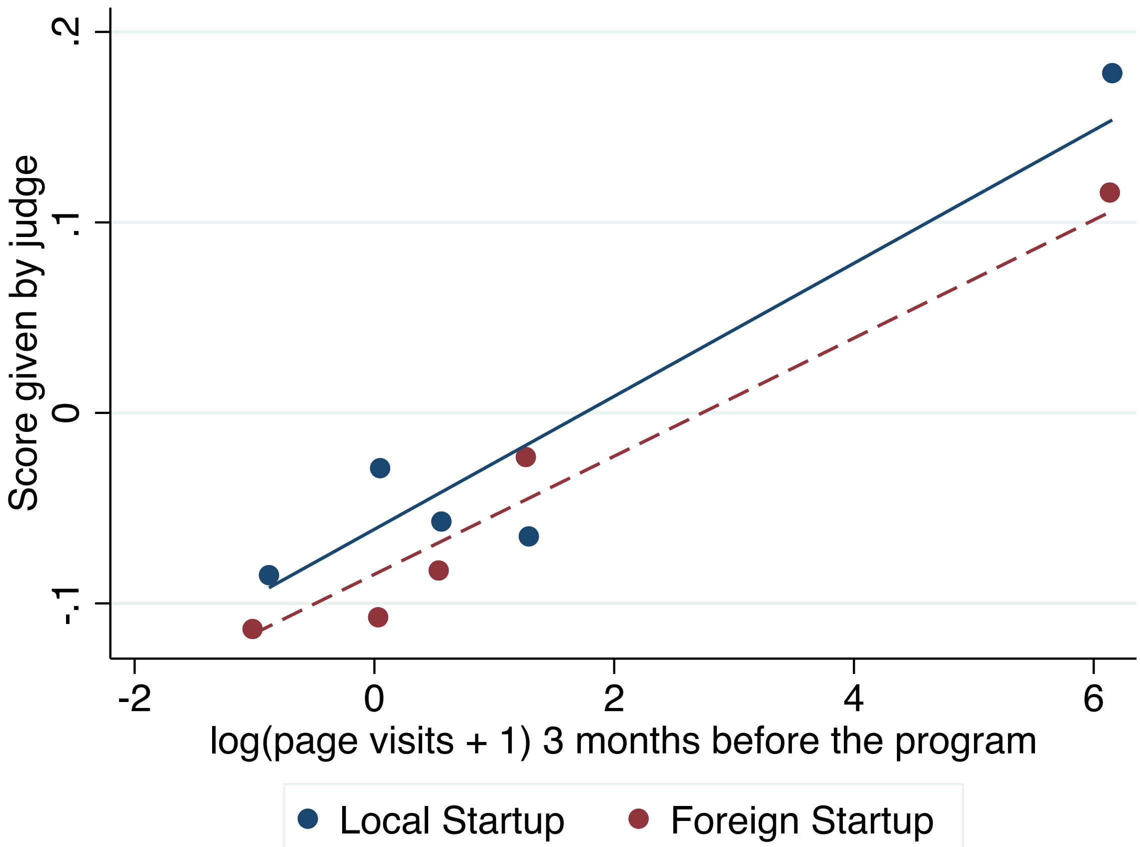
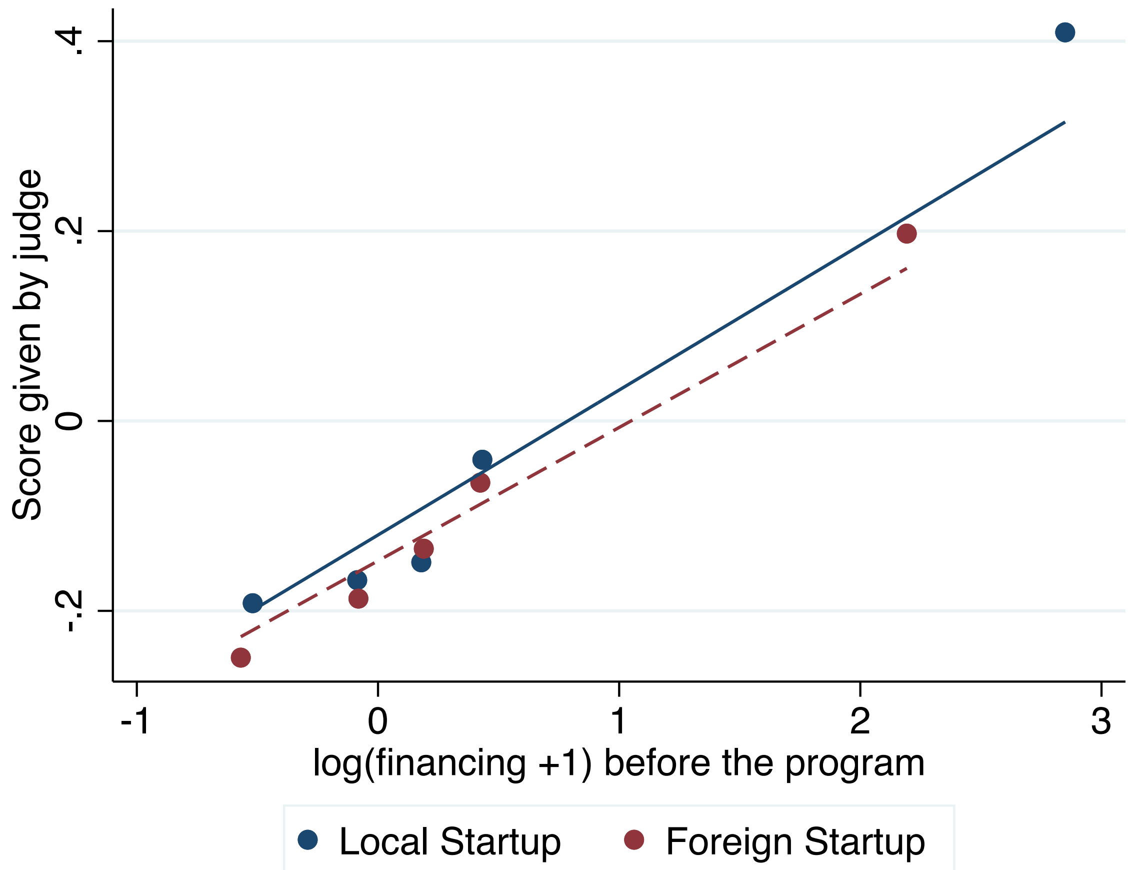


Figure A2: Judge Ability to Predict Future Performance by Whether a Startup is Foreign or Local Using Pre-Accelerator Page Visits



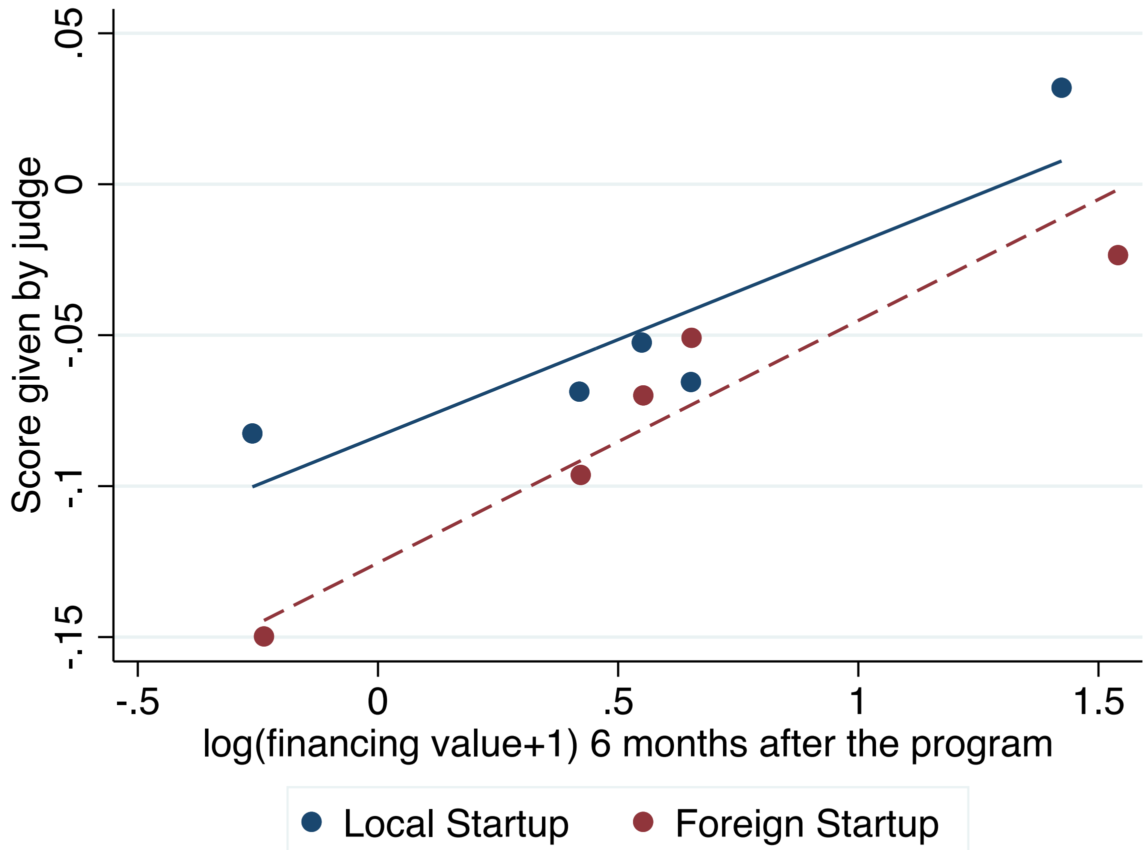
The graph includes judgexyear and startup countryxyear fixed effects.
The y-axis indicates log(page visits+1) 3 months before the program.
We include a zero for startups without page visit data.
The graph shows 2017-2018 global round data.
We use 5 bins.

Figure A3: Judge Ability to Predict Future Performance by Whether a Startup is Foreign or Local Using Pre-Accelerator Financing



The graph includes judgexyear and startup countryxyear fixed effects.
The y-axis indicates log(financing value +1) before the program.
The graph shows 2017-2018 global round data.
We use 5 bins.

Figure A4: Judge Ability to Predict Future Performance by Whether a Startup is Foreign or Local Using Post-Accelerator Financing



The graph includes judgexyear and startup countryxyear fixed effects.
The y-axis indicates log(financing value+1) 6 months after the program.
The graph shows 2017-2018 global round data.
We use 5 bins.
We control for participation in the accelerator program.

Figure A5: Correlation Between Pre-Accelerator Page Visits and Financing

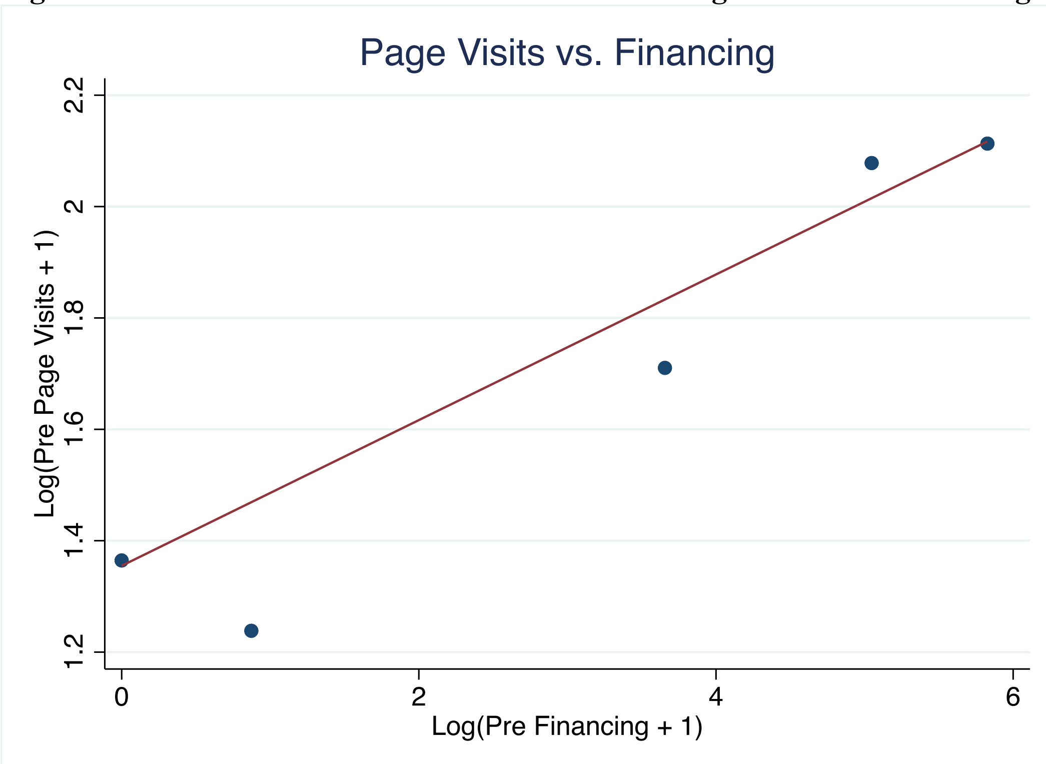
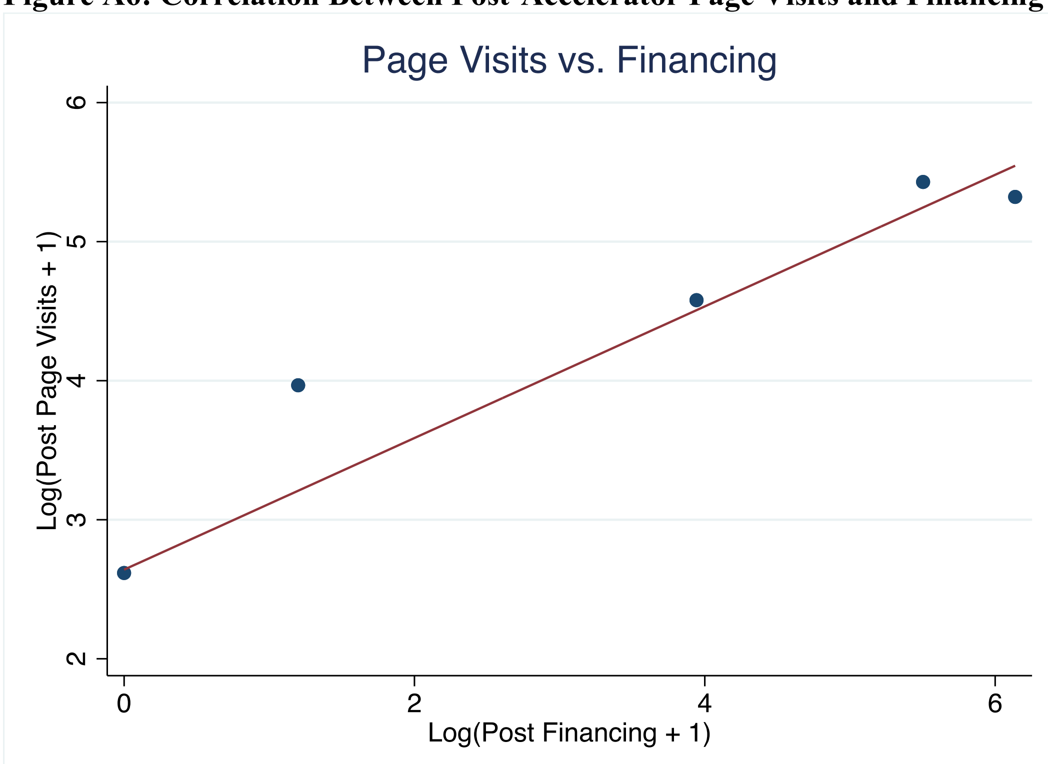


Figure A6: Correlation Between Post-Accelerator Page Visits and Financing



A7: Application Questions

Company Background

1. Full time employees – *the number of full time employees currently in your company.*
2. Part time employees – *the number of part time employees currently in your company.*
3. Interns/volunteers – *the total number of interns or volunteers in your company.*

Customer Pain and Solution

1. Problem – *please describe what problem (customer pain point) you are trying to solve.*
2. Solution – *what is your solution? What is innovative about your solution, technology, business model, etc.?*

Overall Impact

1. Define the 1 year and 5 year impact that you hope to accomplish – *use whatever metrics are most appropriate for you (e.g. revenue, profit, jobs, societal benefits)*

Customer Needs and Acquisition

1. How would you define your potential market, and what is the addressable market size?
2. What traction have you made to date with market validation?
3. Marketing – *what will be your messaging to users and customers, and how do you plan to spread that message?*
4. Sales and distribution – *how will you reach your customers? Via which channels will you likely reach your customers/users?*

Industry and Competitors

1. Which organizations compete with your value offering now, and who might do so in the future?
2. Which organizations complement your offering in the market? Do you know of or anticipate any value chain partners?
3. What are the primary advantages relative to existing or potential competitors? *i.e. Why will you win?*

Business Model/Financials

1. What are the key drivers of business economics (price points, margins, etc.)?

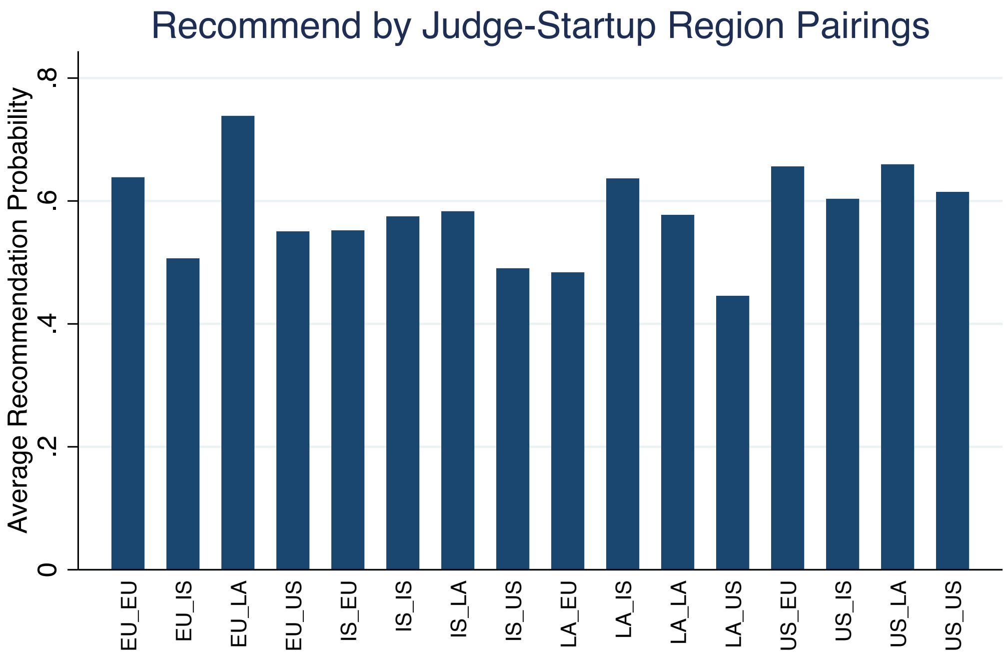
Regulation and IP

1. What IP (intellectual property) or regulatory requirements exist for your business or in your industry?

Founding Team and Advisors/Investors

1. Please share some background information on your team members, and tell us what makes your team special.
2. Please tell us about current or anticipated advisors and investors.

Figure A8: Average Recommendation Probabilities Across Judge-Startup Region Pairings



The graph shows the average recommendation probability by judge-startup region pairing.
The pairings are in the format startup region - judge region.

A9: Back-of-the-Envelope Calculation on Passed Over Startups

We follow a methodology similar to Li (2017) to assess the efficiency of judges' recommendations, assuming that judges are optimizing for the highest performing startups. We compare the cohort of startups selected into the next round of the competition based on judges' actual recommendations compared to a counterfactual algorithm in which judges only consider the startup quality component.

To construct the latter approach, we create a counterfactual decision-making algorithm that only relies on the quality of startups based on the realized outcomes from our data. In this approach, we predict the probability that a startup is recommended to the next round by regressing the realized outcomes – startups' pre-accelerator financing[13] and post-accelerator page visit and financing indicators – on judge scores in the 2017-2018 global rounds of data. Just as in our specification in equation (2), we include judge-year and startup country-year fixed effects to be able to assess evaluations for a given judge, in a given year, within a given country of the startup. We also control for whether startups participated in the accelerator program to account for treatment effects of the accelerator itself. This approach allows us to maintain the weights that judges place on the quality component of startups unlike other machine learning models which would fit new weights on those components.

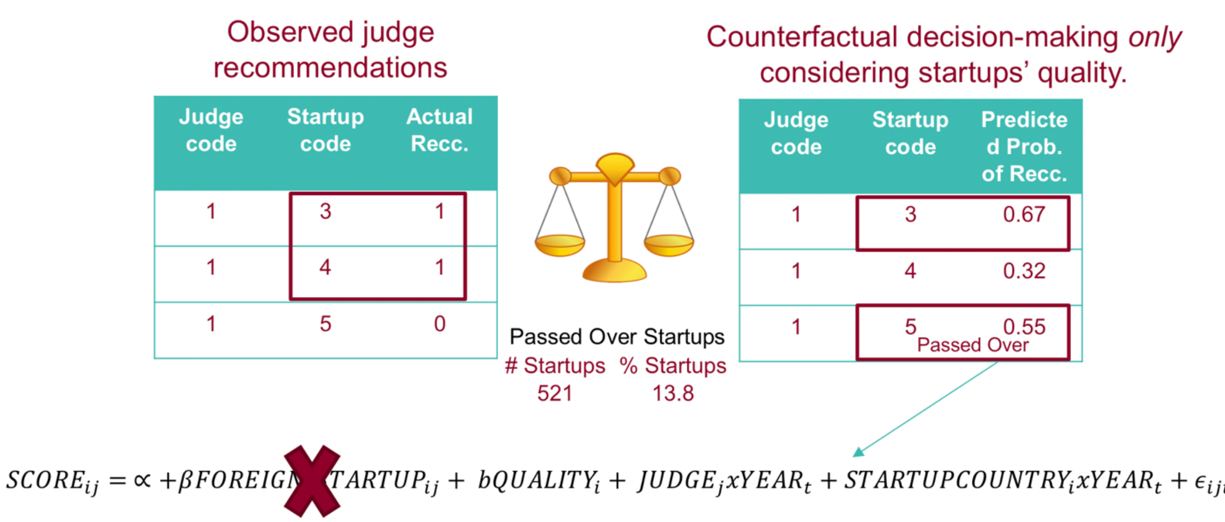
We then rank the recommended probabilities of startups for each judge. We assign a recommendation to the top startups within the number of startups that judges recommended out of their actual evaluated pool. In Figure A10 shown below, if judge 1 recommended two of three startups it evaluated, then in our counterfactual analysis, we would assign recommendation to the top two startups (#3 and 5) on the basis of predicted recommendation probability. We may then

[13] We do not include pre-accelerator website page visits because we lack pre-accelerator website data for startups in the 2017 cohort.

identify which startups would be recommended in our counterfactual analysis versus those that were not in observed outcomes. In Figure A10, startup #5 would fall into this category, as judge 1 would have recommended it in our counterfactual analysis, but had not in actual outcomes. We find that 3,024 foreign startup evaluations are not recommended by judges, though they would be in our counterfactual analysis.

We then estimate how these passed over evaluations may translate into passed over startups. To do so, we average the assigned recommendations in our counterfactual and actual observed outcomes across judges for each startup to evaluate how many startups would be passed over if the accelerator took the average of judge scores across startups.[14] This approach assumes that overall recommendation to the next round of the accelerator would be based on a pure average of judge scores. These calculations bring us to 521 startups – the implied number of startups passed over in the evaluation process (Table 6). These 521 startups comprise 13.8 percent of total startups in these rounds, equating to roughly 1 in 10 startups.

Figure A10: Visual Depiction of Back-of-Envelope Counterfactual Analysis



[14] In actuality, the accelerator applies discretion on top of average judge scores.

References

Aggarwal, R., Kryscynski, D., & Singh, H. (2015). Evaluating venture technical competence in venture capitalist investment decisions. *Management Science*, *61*(11), 2685-2706.

Alvarez-Garrido, E., & Guler, I. (2018). Status in a strange land? Context-dependent value of status in cross-border venture capital. *Strategic Management Journal*, *39*(7), 1887-1911.

Arrow, K. J. (1962). Economic welfare and the allocation of resource for inventions, in the rate and direction of inventive activity: economic and social factors. *RR Nelson (red.), Princeton University, Princeton*.

Bernstein, S., Giroud, X., & Townsend, R. R. (2016). The impact of venture capital monitoring. *The Journal of Finance*, *71*(4), 1591-1622.

Bessemer Venture Partners. (2020). Anti-Portfolio. Retrieved at: https://www.bvp.com/antiportfolio/.

Bloom, N., & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The quarterly journal of Economics*, *122*(4), 1351-1408.

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., & Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, *62*(10), 2765-2783.

Brynjolfsson, E., Hui, X., & Liu, M. (2019). Does machine translation affect international trade? Evidence from a large digital platform. *Management Science*, *65*(12), 5449-5460.

Cao, R., Koning, R., & Nanda, R. (2020). Biased sampling of early users and the direction of startup innovation. *Harvard Business School Entrepreneurial Management Working Paper*, (21-059), 21-059.

Chatterji, A., Delecourt, S., Hasan, S., & Koning, R. (2019). When does advice impact startup performance?. *Strategic Management Journal*, *40*(3), 331-356.

Cohen, S. L., Bingham, C. B., & Hallen, B. L. (2019). The role of accelerator designs in mitigating bounded rationality in new ventures. *Administrative Science Quarterly*, *64*(4), 810-854.

Cohen, S., Fehder, D. C., Hochberg, Y. V., & Murray, F. (2019). The design of startup accelerators. *Research Policy*, *48*(7), 1781-1797.

Conti, A., Guzman, J., & Rabi, R. (2020). Information Frictions in the Market for Startup Acquisitions. *Available at SSRN 3678676*.

Coval, J. D., & Moskowitz, T. J. (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance*, *54*(6), 2045-2073.

Coval, J. D., & Moskowitz, T. J. (2001). The geography of investment: Informed trading and asset prices. *Journal of political Economy*, *109*(4), 811-841.

Dahl, M. S., & Sorenson, O. (2012). Home sweet home: Entrepreneurs' location choices and the performance of their ventures. *Management science*, *58*(6), 1059-1071.

Disdier, A. C., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and statistics*, *90*(1), 37-48.

Dziuda, W., & Mondria, J. (2012). Asymmetric information, portfolio managers, and home bias. *The Review of Financial Studies*, *25*(7), 2109-2154.

Fehder, D. C., & Hochberg, Y. V. (2014). Accelerators and the regional supply of venture capital investment. *Available at SSRN 2518668*.

Franke, N., Gruber, M., Harhoff, D., & Henkel, J. (2006). What you are is what you like similarity biases in venture capitalists' evaluations of start-up teams. *Journal of Business Venturing*, *21*(6), 802-826.

French, K. R., & Poterba, J. M. (1991). Investor diversification and international equity markets. The American Economic Review, 81(2), 222-226.Gans, Joshua S., David H. Hsu, and

Gans, J. S., Hsu, D. H., & Stern, S. (2008). The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. *Management science*, *54*(5), 982-997.

Ghemawat, P., & Altman, S. A. (2019). The state of globalization in 2019, and what it means for strategists. *Harvard Business Review*, 2-8.

Gompers, P. A., Gornall, W., Kaplan, S. N., & Strebulaev, I. A. (2020). How do venture capitalists make decisions?. *Journal of Financial Economics*, *135*(1), 169-190.

Gonzalez-Uribe, J., & Leatherbee, M. (2018). The effects of business accelerators on venture performance: Evidence from start-up chile. *The Review of Financial Studies*, *31*(4), 1566-1603.

Haefliger, S., Von Krogh, G., & Spaeth, S. (2008). Code reuse in open source software. *Management science*, *54*(1), 180-193.

Hall, R. E., & Woodward, S. E. (2010). The burden of the nondiversifiable risk of entrepreneurship. *American Economic Review*, *100*(3), 1163-94.

Hallen, B. L., Cohen, S. L., & Bingham, C. B. (2020). Do accelerators work? If so, how?. *Organization Science*, *31*(2), 378-414.

Hegde, D., & Tumlinson, J. (2014). Does social proximity enhance business partnerships? Theory and evidence from ethnicity's role in US venture capital. *Management Science*, *60*(9), 2355-2380.

Hoenig, D., & Henkel, J. (2015). Quality signals? The role of patents, alliances, and team experience in venture capital financing. *Research Policy*

Howell, S. T. (2017). Financing innovation: Evidence from R&D grants. *American Economic Review*, *107*(4), 1136-64.

Huang, L., & Pearce, J. L. (2015). Managing the unknowable: The effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. *Administrative Science Quarterly*, *60*(4), 634-670.

Huberman, G. (2001). Familiarity breeds investment. *The Review of Financial Studies*, *14*(3), 659-680.

Johnson, M. A., Stevenson, R. M., & Letwin, C. R. (2018). A woman's place is in the… startup! Crowdfunder judgments, implicit bias, and the stereotype content model. *Journal of Business Venturing*, *33*(6), 813-831.

Kaplan, S. N., Sensoy, B. A., & Strömberg, P. (2009). Should investors bet on the jockey or the horse? Evidence from the evolution of firms from early business plans to public companies. *The Journal of Finance*, *64*(1), 75-115.

Kerr, W. R. (2016). Harnessing the best of globalization. *MIT Sloan Management Review*, *58*(1), 59.

Kerr, W. R., Lerner, J., & Schoar, A. (2014). The consequences of entrepreneurial finance: Evidence from angel financings. *The Review of Financial Studies*, *27*(1), 20-55.

Kerr, W. R., Nanda, R., & Rhodes-Kropf, M. (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives*, *28*(3), 25-48.

Khanna, T. (2014). Contextual intelligence. Harvard business review, 92(9), 58-68.

Koning, R., Samila, S., & Ferguson, J. P. (2020, May). Inventor Gender and the Direction of Invention. In *AEA Papers and Proceedings* (Vol. 110, pp. 250-54).

Koning, R., Hasan, S., & Chatterji, A. (2019). *Experimentation and startup performance:*

Evidence from A/B testing (No. w26278). National Bureau of Economic Research.

Lee, M., & Huang, L. (2018). Gender bias, social impact framing, and evaluation of entrepreneurial ventures. *Organization Science*, *29*(1), 1-16.

Lerner, J., Schoar, A., Sokolinski, S., & Wilson, K. (2018). The globalization of angel investments: Evidence across countries. *Journal of Financial Economics*, *127*(1), 1-20.

Li, D. (2017). Expertise versus Bias in Evaluation: Evidence from the NIH. *American Economic Journal: Applied Economics*, *9*(2), 60-92.

Lin, M., Prabhala, N. R., & Viswanathan, S. (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science*, *59*(1), 17-35.

Lin, M., & Viswanathan, S. (2016). Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science*, *62*(5), 1393-1414.

List, J. A. (2006). The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions. *Journal of political Economy*, *114*(1), 1-37.

Lu, J. W., & Beamish, P. W. (2001). The internationalization and performance of SMEs. *Strategic management journal*, *22*(6-7), 565-586.

Luo, H. (2014). When to sell your idea: Theory and evidence from the movie industry. *Management Science*, *60*(12), 3067-3086.

Malloy, C. J. (2005). The geography of equity analysis. *The Journal of Finance*, *60*(2), 719-755.

Nanda, R., Samila, S., & Sorenson, O. (2020). The persistent effect of initial success: Evidence from venture capital. *Journal of Financial Economics*, *137*(1), 231-248.

Niessen-Ruenzi, A., & Ruenzi, S. (2019). Sex matters: Gender bias in the mutual fund industry. *Management Science*, *65*(7), 3001-3025.

Oviatt, B. M., & McDougall, P. P. (2005). Toward a theory of international new ventures. *Journal of international business studies*, *36*(1), 29-41.

Ruef, M., Aldrich, H. E., & Carter, N. M. (2003). The structure of founding teams: Homophily, strong ties, and isolation among US entrepreneurs. *American sociological review*, 195-222.

Scott, E. L., Shu, P., & Lubynsky, R. M. (2020). Entrepreneurial uncertainty and expert evaluation: An empirical analysis. *Management Science*, *66*(3), 1278-1299.

Sørensen, M. (2007). How smart is smart money? A two-sided matching model of venture capital. *The Journal of Finance*, *62*(6), 2725-2762.

Sorenson, O., & Stuart, T. E. (2001). Syndication networks and the spatial distribution of venture capital investments. *American journal of sociology*, *106*(6), 1546-1588.

X1 Group. (2018). 16 Unicorn Startups with R&D Offices in Ukraine. *Medium*. Retrieved from: https://medium.com/x1group/16-unicorn-startups-with-r-d-offices-in-ukraine-6243cd0ddd6.

Y Combinator. (2020). Top Companies. Retrieved from https://www.ycombinator.com/topcompanies/.

Yin, B., & Luo, J. (2018). How do accelerators select startups? Shifting decision criteria across stages. *IEEE Transactions on Engineering Management*, *65*(4), 574-589.

Yu, S. (2020). How do accelerators impact the performance of high-technology ventures?. *Management Science*, *66*(2), 530-552.

Zaheer, S. (1995). Overcoming the liability of foreignness. *Academy of Management Journal*, *38*(2), 341-363.