Entrepreneurial Strategy and Scaling in a Global Digital Economy

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the Committee for the PhD in Business Administration have examined a dissertation entitled

Entrepreneurial Strategy and Scaling in a Global Digital Economy

Presented by Nataliya Langburd Wright

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature ______________________________

Tarun Khanna, Chair

Signature ______________________________

Shane Greenstein

Signature ______________________________

Laura Huang

Signature ______________________________

Rembrand Koning

Date: March 1, 2023

Entrepreneurial Strategy and Scaling in a Global Digital Economy

a dissertation presented
by
Nataliya Langburd Wright
to
The Strategy Unit at the Harvard Business School

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Business Administration

Harvard University
Cambridge, Massachusetts
March 1, 2023

Thesis advisor: Professor Tarun Khanna Nataliya Langburd Wright

Entrepreneurial Strategy and Scaling in a Global Digital Economy

Abstract

The number of promising technology startups has increased worldwide, yet few outside the US manage to scale. What accounts for these international scaling disparities? This dissertation assesses the role of entrepreneurial and gatekeeper decisions in these scaling differences. From the perspective of entrepreneurs, the first chapter assesses the role of strategy, finding that strategy matters more in institutional contexts where mistakes are more costly, though it is harder to develop there because entrepreneurs learn from prior mistakes. The second chapter illuminates how such differences in entrepreneurial decisions may emerge from locally-embedded knowledge, showing how geographic exposure shapes entrepreneurs' core experimentation and strategy. From the perspective of gatekeepers, the third chapter finds that judges in accelerator competitions discount startups foreign to them, driven not by a local information advantage, but rather by pure preference. The last chapter explores the role of open-source platforms—another type of gatekeeper for startups to access technical knowledge and coordination. It finds that such platforms increase entry into entrepreneurship overall, but more so for startups already in highly-endowed contexts, suggesting that the platform's design decisions may contribute to entrepreneurial growth differences. Together, these papers reveal how decision-making interacts with institutional contexts to shape the growth trajectories of startups around the world.

Contents

Title Page i

Copyright ii

Abstract iii

Table of Contents iv

List of Figures viii

List of Tables xi

Acknowledgements xvii

0 Introduction 1

1 Where Strategy Matters: Evidence from a Global Startup Field Study 7

1.1 Introduction 7

1.2 Theoretical framework 12

1.3 Methodology 16

1.4 Results 21

1.5 Conclusion 40

2 When the Journey—And Not Just the Destination—Matters: How Internationalization Shapes Entrepreneurial Experimentation 43
2.1 Introduction . . . 43
2.2 Experimentation and internationalization . . . 46
2.3 Methodology . . . 50
2.4 Findings . . . 56
2.5 Discussion . . . 69

3 Judging Foreign Startups 74
3.1 Introduction . . . 74
3.2 Theoretical framework . . . 79
3.3 Context: Global accelerator competition . . . 83
3.4 Data . . . 85
3.5 Empirical specification . . . 89
3.6 Results . . . 91
3.7 Conclusion and implications . . . 107

4 Open Source Software and Global Entrepreneurship 110
4.1 Introduction . . . 110
4.2 Framework and hypotheses . . . 115
4.3 Measurement . . . 120
4.4 Estimates . . . 127
4.5 Discussion . . . 137

Appendix A Chapter 1 Supplementary Material 163
A.1 Interview protocol and scoring . . . 165
A.2 Coding examples . . . 167
A.3 Sampling checks . . . 180
A.4 Machine learning approach to strategy score . . . 184
A.5 Alternative measures of strategy . . . 185

A.6 Robustness checks using LASSO specifications 186
A.7 Results robust to a country index of the ease of recovering from mistakes 188
A.8 Interviews suggest that mistakes are more costly in non-US contexts. 192
A.9 Strategy can help companies avoid mistakes 193
A.10 Content of strategy in US versus non-US firms 195
A.11 US strategies are less geographically oriented. 199
A.12 Organizational approaches are localized. 201
A.13 Mediation analysis using US investor and executives' US work experience 206
A.14 Robustness checks with other outcomes 207

Appendix B Chapter 3 Supplementary Material 208

B.1 Monte Carlo simulations showing that 0–40% judge location measurement error still allows us to detect foreign discounting and local information advantage effects. . . . 211
B.2 Kernel density plots showing that judges give lower scores to foreign startups 216
B.3 Judges across home regions are biased against foreign startups. 217
B.4 Foreign bias results are robust to raw weighted and non-weighted measures of judges final recommendation score. 220
B.5 Judges are equally likely to give incomplete subscores to foreign and local startups. 222
B.6 Alternative foreignness measures produce similar results. 223
B.7 Results hold when excluding investor judges and Latin American startups. 226
B.8 Robustness to check that startups are not strategically disclosing their region in their applications 228
B.9 Robustness to country and founder quality measures 230
B.10 Foreign bias results hold whether the judge-startup industry matches or not. 233
B.11 Hub location does not explain the foreign discounting effect. 236
B.12 The findings generalize to all years and rounds of the accelerator program. 238
B.13 Robustness using alternative measures of startup quality 239
B.14 Judges are equally responsive to quality when we split our sample by foreignness. . 244

B.15 Judges are worse at detecting the quality of startups that have raised financing, are from hubs, or have founders affiliated with elite universities. 246
B.16 Judges are worse at selecting winners from losers among startups in later stages of the competition. 248
B.17 Judges can equally detect the quality of local vs. foreign startups no matter whether startups regions are mentioned in their applications. 249
B.18 Judges have a local information advantage when assessing very localized startups. . 251
B.19 Additional analyses estimating missed startups. 254
B.20 Application questions that startups answer and judges evaluate in the accelerator program 256
B.21 Robustness when including elite university and hub controls 258

Appendix C Chapter 4 Supplementary Material 259
C.1 Example of OSS policies 261
C.2 Impact of OSS on new venture founding: First stage estimates 262
C.3 Correlations 263
C.4 Impact of OSS on hardware vs. software ventures 264
C.5 Predicted increase in number of ventures with 1% increase in 2016 GitHub commits 265

List of Figures

1.1 Strategy weakly predicts performance of firms. 23
1.2 Strategy predicts performance for non-US firms (right), but not US ones (left). 23
1.3 Kernel density plot of strategy scores 26
1.4 US companies see a right shift in the distribution of strategy scores. 27
1.5 Non-US companies are less likely to draw on direct experience, company examples, advisors, and alternatives to inform their strategy. 38

2.1 Theoretical overview 47
2.2 Methods overview 51
2.3 Process model illustration 56

3.1 Predicted relationships between judge scores and startup quality 82
3.2 Binscatter showing that judges give higher scores to startups with more growth one-to- two- years after the accelerator program 98
3.3 Binscatter showing that judges give higher scores to startups with more growth one-to- two- years after the program, but they consistently discount foreign startups no matter their eventual success 99
3.4 Coefficient plot of judge sensitivity to the quality of startups across sub-samples of startups. The bars show 90 percent and 95 percent confidence intervals. 103

3.5 Coefficient plot of judge local information advantage across different local sub-samples of startups. The bars show 90 percent and 95 percent confidence intervals. 105

4.1 GitHub commits vs. new IT venture formation 2016 112

4.2 Possible relationships between OSS participation and entrepreneurship 115

4.3 OSS context illustration 117

A.3.1 Regional distribution of the non-interviewed sample is similar to that of the interviewed sample. 181

A.3.2 Employee distribution of the non-interviewed sample is similar to that of the interviewed sample. 182

A.3.3 First-financing-size (in USD) distribution of the non-interviewed sample is similar to that of the interviewed sample. 183

A.7.1 The ease-of-recovery index by country. 189

A.11.1 North American interviewed companies are more likely to mention geography in their market scope responses. 200

A.12.1 North American interviewed companies' organizational responses tend to mention an objective more than those of companies from other regions. 202

A.12.2 Machine learning models are more accurate at detecting the region of organizational design and culture responses than of other responses of interviewed companies. 203

A.12.3 Organizational structure in terms of breakdown by roles varies across regions. 204

A.12.4 Organizational "flatness" varies across regions. 205

B.1.1 Classical measurement error: Foreign bias coefficients 212

B.1.2 Classical measurement error: Foreign bias coefficients 213

B.1.3 Classical measurement error incorporating information asymmetries: Foreign information coefficients 213

B.1.4 Classical measurement error: Foreign bias coefficients with asymmetrical error 214

B.1.5 Classical measurement error incorporating information asymmetries: Foreign bias coefficients with asymmetrical error 215

B.1.6 Classical measurement error incorporating information asymmetries: Information coefficients with asymmetrical error . . . 215
B.2.1 Kernel density plot of scores by whether the judge and startup are from the same region (local startup) or from different regions (foreign startup) . . . 216
B.3.1 Bar graph showing that US, EU, and Israeli judges are more likely to recommend local over foreign startups to the next round of the competition. Panel A is for all startups and Panel B is only for startups that have raised financing. . . . 218
B.13.1 Binscatter showing that judges can give higher scores to startups with more growth before entering the accelerator program, but consistently discount foreign startups no matter their growth. . . . 242
B.13.2 Binscatter showing that judges can give higher scores to startups with more financing before entering the accelerator program, but consistently discount foreign startups no matter their financing. . . . 243
B.19.1 Distribution of startups actually selected versus estimated to be missed with foreign discounting. . . . 255

List of Tables

1.1 Sample comparison table showing that interviewed companies do not vary from non-interviewed ones based on whether they are US-based, the number of employees, funding, and valuation. 17

1.2 Summary table 22

1.3 Strategy predicts performance more for non-US firms. 25

1.4 Non-US firms have lower strategy scores. 28

1.5 The strategy gap between US and non-US firms remains, even when accounting for the content of strategy. 31

1.6 Non-US firms are less likely to use direct experience, alternatives, company examples, advisors, and investors to inform their strategy. 39

2.1 Interview sample breakdown 52

2.2 Interview coding example 55

3.1 Summary statistics at the evaluation level 89

3.2 Chi-squared table for the 2017 global round showing distribution of judges to startups is no different than what we would expect from random chance 92

3.3 Chi-squared table for the 2018 global round showing distribution of judges to startups is no different than what we would expect from random chance when excluding outliers 93

3.4 Regressions showing that judges give lower scores to startups from outside their home region even when we control for judge and startup fixed effects . . . 95
3.5 Regressions showing that judges are less likely to recommend startups from outside their home region even when we control for judge and startup fixed effects . . . 96
3.6 Regressions showing judges (1) give higher scores to more successful startups, (2) are equally good at evaluating success for local and foreign startups alike, and (3) still discount foreign startups . . . 101

4.1 Summary statistics . . . 128
4.2 Impact of OSS on new venture founding . . . 130
4.3 Impact of OSS on new venture founding by GDP per capita and human capital . . . 131
4.4 Impact of OSS on new venture founding by GDP per capita: 2SLS . . . 132
4.5 Impact of OSS on new venture founding by human capital: 2SLS . . . 134
4.6 Impact of OSS on global and mission-oriented venture founding . . . 135
4.7 Impact of OSS on quality of ventures . . . 136
4.8 Difference-in-difference: Impact of OSS on new venture founding . . . 138
4.9 Summary of hypotheses, implementation, and results . . . 139

A.2.1 Market scope examples . . . 169
A.2.2 Moat examples . . . 172
A.2.3 Organizational design examples . . . 175
A.2.4 Organizational culture examples . . . 178
A.4.1 Human and NLP calculations of strategy are positively correlated. . . . 184
A.5.1 Alternative measures of strategy . . . 185
A.6.1 The strategy gap between US and non-US firms holds when using a double-selection LASSO model controlling for the strategy's content. . . . 186
A.6.2 Strategy predicts performance more for non-US firms even when controlling for the strategy's content using a double-selection LASSO model. . . . 187
A.7.1 Strategy predicts performance more where it is harder to recover from mistakes. . . . 190
A.7.2 Strategy scores are higher where it is harder to recover from mistakes. . . . 191

A.8.1 Perception of mistakes and failure as learning opportunities by whether US HQ . . . 192
A.9.1 Strategy and office closures . . . 194
A.10.1 US versus non-US objectives . . . 195
A.10.2 US versus non-US market scope approaches . . . 195
A.10.3 US versus non-US moat approaches . . . 196
A.10.4 US versus non-US organizational design approaches . . . 196
A.10.5 US versus non-US organizational culture approaches . . . 197
A.10.6 US versus non-US next three actions . . . 197
A.10.7 US versus non-US use of Series A funding . . . 198
A.10.8 US versus non-US biggest uncertainty . . . 198
A.11.1 US firms are less likely to focus on geography in their strategy. . . . 199
A.11.2 Global strategies in pitch decks . . . 199
A.13.1 Gap between US and non-US strategy scores declines when accounting for US investors and executive work experience in the US. . . . 206
A.14.1 Strategy predicts performance more for non-US firms . . . 207

B.3.1 Israeli, European, and Latin American judges are similarly discounting foreign startups as are other judges. . . . 219
B.4.1 Regressions showing that judges score lower startups from outside their home region even when we control for judge and startup fixed effects: Weighted raw recommendation score 220
B.4.2 Regressions showing that judges score lower startups from outside their home region even when we control for judge and startup fixed effects: Non-weighted raw recommendation score . . . 221
B.5.1 Probability that judges give incomplete subscores . . . 222
B.6.1 Example localized words by startup home region . . . 224
B.6.2 Foreign discounting using other measures of foreign startup . . . 225
B.6.3 Foreign discounting weakens when the region is not explicit in the application text . . . 225
B.7.1 Foreign discounting effect excluding investor judges . . . 226
B.7.2 Foreign discounting effect excluding Latin American startups . . . 227

B.8.1 Startup quality does not predict whether region is explicitly mentioned in applications 229
B.9.1 Foreign discounting effect controlling for country HQ quality 231
B.9.2 Foreign discounting effect controlling for founder quality 231
B.9.3 Distribution of startups in our sample by hub 232
B.10.1 Judges are similarly biased against foreign startups no matter whether their industry matches that of the startups. 234
B.10.2 Judges can tell the quality of foreign vs. local startups with similar precision no matter whether their industry matches that of the startups. 234
B.10.3 Judges do not discount startups from outside of their industry. 235
B.11.1 Whether a startup is in a hub does not remove or moderate the foreign discounting effect. 237
B.12.1 Foreign discounting effect across all years and rounds of the accelerator program . . . 238
B.13.1 The ability of judges to evaluate startup quality: Pre-accelerator quality measures . . 240
B.13.2 The ability of judges to evaluate startup quality: Post-accelerator quality measures . 241
B.14.1 The R-squared values for the extent to which judges are informed are similar when we split our sample into startups foreign to the judge and startups local to the judge. . . 245
B.14.2 Regressions showing that judges scores predict future startup performance. 245
B.15.1 Foreign discounting effect across different quality sub-samples of startups 247
B.16.1 Judges ability to detect startup quality among startups accepted into the accelerator program declines. 248
B.17.1 Judges can equally detect the quality of local vs. foreign startups no matter whether startups regions are mentioned in their applications. 250
B.18.1 Top words in application by startup home region 252
B.18.2 As we restrict the sample to more localized startups, we see that judges become better at evaluating startup quality and worse at evaluating foreign startup quality. . . . 253
B.21.1 Regressions showing judges (1) give higher scores to more successful startups, (2) are equally good at evaluating success for local and foreign startups alike, and (3) still discount foreign startups when controlling for elite university affiliation and hub location 258

C.1.1 Example of OSS policies 261
C.2.1 Impact of OSS on new venture founding: First stage estimates 262
C.3.1 Correlations 263
C.4.1 Impact of OSS on hardware vs. software ventures 264
C.5.1 Predicted increase in number of ventures with 1% increase in 2016 GitHub commits . 265

To my family.

Acknowledgments

I am deeply grateful to my dissertation committee for their guidance and wisdom. Tarun Khanna inspired me to tackle research questions that can make an impact. Shane Greenstein pushed me to not lose sight of the forest for the trees. Laura Huang helped me see and convey my ideas across disciplinary boundaries. Rem Koning encouraged me to harness my passion for international startups and fieldwork into action.

I was incredibly fortunate to have such wonderful mentors and colleagues, whose advice, feedback, and encouragement were instrumental throughout my PhD journey. I am grateful to Juan Alcacer, Victor Bennett, Daniel Brown, Natalie Carlson, Innessa Colaiacovo, Chris Eaglin, Michael Impink, Geoffrey Jones, Bill Kerr, Hyunjin Kim, Josh Lerner, Hong Luo, Frank Nagle, Jan Rivkin, Laura Weimer, Andy Wu, Dennis Yao, and countless others. And I could not have done it without the support of Amisha Miller, Ashley Palmarozzo, Emily Prinsloo, and Kala Viswanathan, devoted friends who have accompanied me through the ups and downs of the journey.

Thank you as well to the incredible HBS Doctoral Programs team for guidance in each step of this PhD journey. I am also thankful for the SMS SRF Dissertation Research Grant and to the participating companies that made this research possible.

A special thanks to my family—the Benins, Langburds, and Wrights—and friends, without whom I would not be here. Your understanding and encouragement made even the tallest mountains seem like hills. I am forever indebted to my parents for inspiring me to shoot for the moon. On a journey that took us from Kyiv to the Bronx to Philly, I've learned so much from your resilience, determination, and selflessness. I am thankful to my brother, Michael, for always making me smile. And finally, to Billy, for being there for me through thick and thin and always believing in me. *Thank you*.

0

Introduction

The number of promising technology startups has skyrocketed worldwide, yet few outside the US manage to scale. Venture-backed US technology firms have 20-percent higher valuations than those from other countries.[1] In 2021, the US produced nearly 10 times the number of billion-dollar-valued unicorns than France, Israel, Canada, and Brazil combined.[2] Indeed, there are few Airbnbs, WhatsApps, or Slacks of Europe, Latin America, Asia, or Africa. What accounts for these international scaling disparities? Why do some startups scale while others are left behind?

A rich body of research discusses how differences in external resources across geographies—whether it be related to the availability of venture capital, customers, talent, or suppliers—can explain differences in entrepreneurial growth. For example, Conti and Guzman (2021) find that access to venture capital and market size help account

[1]PitchBook (2022). Company Profile. Retrieved from Pitchbook database.

[2]Statista. (2022). Number of unicorns worldwide as of 2021, by country.

for the US comparative advantage in entrepreneurship. Further, the flexibility of labor markets and supply of knowledge workers offer startups in hubs like Silicon Valley access to cutting-edge technology that helps them gain and sustain a competitive advantage (Kerr, 2018; Kerr and Robert-Nicoud, 2020; Tambe and Hitt, 2014). This research generally shows that resource constraints outside of the hubs make it difficult for entrepreneurs and gatekeepers to implement their pre-set choices, which are, in essence, similar to those of actors in the hubs. For example, hub and non-hub entrepreneurs might have the same type of experiments they want to run. However, resource constraints facing non-hub entrepreneurs make it more costly for them to do so, limiting their growth. In this case, choices are fixed, and external resources impact whether entrepreneurs can execute them.

But these external resource differences may actually shape the choices of these key actors. The hub and non-hub entrepreneurs might conceive fundamentally different types of experiments or even view the value of experiments differently as a result of their geographic circumstances. Indeed, research shows that firms update their choices to address institutional voids (Khanna and Palepu, 1997; 2010; Khanna and Rivkin, 2001). These same institutional voids—for example, under-developed talent or capital markets (Khanna and Palepu, 1997; 2010)—might limit the availability of experienced local investors and advisors who can inform the decisions of entrepreneurs. Limited access to this expertise might constrain the choice set that startups perceive in the first place (Chatterji et al., 2019; Gavetti and Rivkin, 2007; Shane, 2000; Vissa and Chacar, 2009). Startups' choices might be particularly contingent on these environmental signals and knowledge inputs because of the high uncertainty entrepreneurs face and their relative lack of organizational inertia (Kerr et al., 2014b; McDonald and Gao, 2019; Zhang et al., 2016). Ultimately, these choices—whether related to which market positions to pursue, what capabilities to invest in, or what experiments to run—impact performance (Barney, 1991; Camuffo et al., 2020; Koning et al., 2022; Porter, 1991; Wernerfelt, 1984). Systematic differences in these choices across geographies would therefore result in international differences in entrepreneurial growth.

In light of the influence that external resources across geographies can have on choices, this dissertation explores the role of gatekeeper and startup decisions in international scaling differences. While much prior work in entrepreneurship has assumed a shared choice framework among actors across international contexts—at least partly because it is difficult to directly measure these choice frameworks—my work seeks to uncover the heterogeneities in the decision-making of entrepreneurs and gatekeepers across these contexts. It does so through leveraging large-scale field data, for example, involving interviews with nearly 340 startups from over 40 countries. Understanding the geographic variance in these decisions may shed light on why we see international differences in scaling and reveal capital-efficient interventions—for instance, related to knowledge exchange—that may nurture entrepreneurial growth around the world.

The first two chapters of my dissertation assess the role of entrepreneurs' decisions, and the latter two assess the role of gatekeepers' decisions to understand international differences in scaling. The first chapter, "Where Strategy Matters: Evidence from a Global Startup Field Study," assesses the role of strategy in startup scaling. It proposes that institutional context can reconcile conflicting views on the role of strategy for entrepreneurs. To detect whether startups have a strategy, I interviewed executives of 253 scaling software ventures from 34 countries and scored the alignment of their market and organizational choices—following from a key prediction that strategy, the smallest set of choices to optimally guide other choices (Van den Steen, 2017), creates alignment across firm decisions—developing the first dataset of its kind. Having a strategy predicts performance more for non-US startups, for which a one standard deviation increase in the strategy score is associated with an increase in valuation by over a third. Yet, non-US startups are less likely to develop a strategy; they have a 0.3 standard deviation lower strategy score than do others. Additional analyses suggest that mistakes are more costly in non-US contexts because of financial, talent, and cultural differences, penalizing firms there without a strategy that helps anticipate sources of failure. Creating a strategy, however, is more difficult without the ability to learn from prior mistakes. Together, this research suggests that in institutional contexts where mistakes are more costly, strategy matters more, but is also harder to develop. The varying role of strategy across international contexts may therefore contribute to differences in scaling.

Why do differences in strategy persist across international contexts? My second chapter, "When the Journey—And Not Just the Destination—Matters: How Internationalization Shapes Entrepreneurial Experimentation" (co-authored with Laura Huang), offers a knowledge-based explanation for why these entrepreneurial decision-making differences emerge and how cross-border exposure can overcome them. It shows how exposure to both near and distant geographies shapes the way early-stage ventures engage in experimentation and, in turn, how they define their business. Scholars have shown how entrepreneurs develop products or services and, after achieving some traction, turn to international markets to help them continue growing and scaling their businesses. Yet, what may be neglected in this prior work is how internationalization may not always be the result of—but instead the catalyst to—new information, entrepreneurial discovery and sensemaking, and important developments in an entrepreneur's core business. Through an inductive field study of 84 entrepreneurs across 27 countries in the global technology industry, we examine how internationalization influences the entrepreneurial process and the profound effect that it has on how entrepreneurs identify and exploit opportunities. We find that internationalization shapes the way these entrepreneurs engage in experimentation and, in turn, how they define their business. We propose a process model that sheds light on how entrepreneurs (a) define, (b) scope, and (c) externally validate ideas based on their international exposure—a process that aids them in creating and capturing value. Our findings show how cross-border markets may offer unanticipated information to entrepreneurs, which adjusts their perceived choice sets and helps them ultimately define their business, contributing to research on experimentation and internationalization and new theory on how cross-border exposure shapes innovation and growth. In doing so, this research reveals how geographic context can fundamentally influence the nature of entrepreneurs' key decisions. Because entrepreneurs are often embedded in local geographies, the decision-making of entrepreneurs across countries may systematically vary—as documented in Chapter 1—contributing to differences in scaling.

I next turn to the role of gatekeeper decisions to understand international differences in entrepreneurial scaling. The third chapter, "Judging Foreign Startups" (co-authored with Rem Koning and Tarun Khanna), shows how accelerator judges discount foreign startups, reducing their access to the accelerator's resources and knowledge for growth. Using unique data from a global accelerator where judges are randomly assigned to evaluate startups headquartered across the globe, we show that judges are less likely to recommend startups headquartered outside their home region by 4 percentage points. Back-of-the-envelope calculations suggest this discount leads judges to pass over 1 in 20 promising startups. While prior research in trade and finance shows that such foreign discounting results from judges' local information advantage, our research shows that preference underlies their behavior. Judges are, in fact, no better at detecting the quality of local startups than of foreign ones because these startups pursue globally standardized business models that require less local knowledge to understand. These findings show how accelerator judges' decision-making—which may penalize startups from remote areas—can contribute to international differences in startup scaling.

Not only accelerators but also digital platforms are important gatekeepers for entrepreneurs to access the knowledge and resources needed to grow. The design decisions of these platforms may also contribute to international scaling differences. To shed light on the role of these latter actors, the last chapter, "Open Source Software and Global Entrepreneurship" (co-authored with Frank Nagle and Shane Greenstein), reveals how open source platforms may also heterogeneously shape entry into technology entrepreneurship across countries. As the first study to consider the relationship between open source software (OSS) and entrepreneurship around the globe, this study measures whether country-level participation on the GitHub OSS platform affects the founding of innovative ventures and, where it does so, for what types of ventures. We estimate these effects using cross-country variation in new venture founding and OSS participation. We propose an approach using instrumental variables and cannot reject a causal interpretation. The study finds that an increase in GitHub participation in a given country generates an increase in the number of new technology ventures within that country in the subsequent year. The evidence suggests this relationship is complementary to a country's endowments and does not substitute for them. In addition to this positive change in the rate of entrepreneurship, we also find a change in direction—OSS contributions lead to new ventures that are more mission- and global-oriented and are of a higher quality. Together, the results suggest that OSS can boost entrepreneurial activity, albeit with a human capital prerequisite. We consider the implications for policies that encourage OSS as a lever for stimulating entrepreneurial growth. Ultimately, the design decisions of OSS platforms may also contribute to international differences in startup scaling.

Together, this research reveals how the decisions of entrepreneurs and gatekeepers—shaped by their geographic contexts—may contribute to international differences in startup scaling. In doing so, it reveals how institutional context can shape when deliberate decision-making—whether related to positioning in the market (Porter, 1996) or building internal capabilities (Barney, 1991; Wernerfelt, 1984)—matters for startups, shedding light on a long-standing debate in entrepreneurship and strategy (e.g., Bhide, 2000; Delmar and Shane, 2003; Dencker et al., 2009; Mintzberg and Waters, 1985; Ott et al., 2017; Rivkin, 2000). Institutional factors, therefore, not only impact which strategies firms choose (Gao et al., 2017; Khanna and Palepu, 1997; 2010) but also whether strategy development capabilities are rare and valuable resources (Barney, 1991; Wernerfelt, 1984).

The research further sheds light on knowledge mechanisms that may constrain entrepreneurial scaling. Entrepreneurs gain knowledge to inform their experimentation and strategy from their often locally embedded experiences and experimentation or those of their investors, advisors, and mentors. However, ventures outside of the historically successful US hubs—where we see a concentration of experienced entrepreneurs, investors, and advisors—may face higher hurdles to accessing this knowledge. Therefore, startups outside these US hubs need to be knowledge mobilizers—and not only resource mobilizers (Clough et al., 2019)—to grow.

Lastly, this research helps bridge the gap between the global emergence of innovative startups and entrepreneurship research that is mainly US-focused. By leveraging global data on startups—whether hand-collected or acquired through partnerships with multinational organizations—this research systematically assesses how decisions and outcomes of entrepreneurs and gatekeepers vary across international contexts, revealing how such contexts can fundamentally shape the decisions of these actors.

My future work will assess the role of entrepreneurial decision-making and platforms in global startup scaling, leveraging rich field panel and experimental data. First, I seek to use large-scale field methods to assess entrepreneurial decision-making longitudinally, specifically, how startups around the world iteratively balance market and organizational commitments to grow. I plan to do so by "scaling" the strategy database from Chapter 1 through more touch points with existing companies and broader coverage of companies worldwide. This scaled data would allow me to assess how strategy varies over time and to conduct field experiments to test knowledge interventions that affect strategy and performance. Second, I seek to delve longitudinally into the role of platform gatekeepers in startup scaling. Panel data from digital product, open source, corporate, and related platforms will enable me to assess how such gatekeepers influence the direction of innovation, markets, and growth of startups around the world over time. These two future avenues of research will shed light on how entrepreneurial and gatekeeper decisions contribute to persistent international differences in scaling and interventions that can overcome these differences.

1

Where Strategy Matters: Evidence from a Global Startup Field Study

1.1 Introduction

Is strategy only for large bureaucratic corporations, or is it also valuable for innovative startups? This question is much debated. Some argue that the value of strategy—"the smallest set of choices to guide other choices" (Van den Steen, 2017)—is limited at best. Practitioner frameworks like the Lean Startup movement consider such strategic planning a distraction from experimentation. As the movement's founder explains: "Rather than engaging in months of planning…entrepreneurs accept that all they have on day one is a series of untested hypotheses—basically, good guesses" (Blank, 2013). Research, too, suggests that strategy can hurt entrepreneurs' performance by making them less able to adapt and learn in their highly uncertain environments (Bhidé, 2000; Delmar and Shane, 2003; Mintzberg and Waters, 1985).

Yet others suggest that strategy can boost entrepreneurial performance. The holistic nature of strategy—in contrast to that of modular experiments—enables startups to anticipate inter-dependencies and create "fit" across company activities (Dencker et al., 2009; Porter, 1996; Rivkin, 2000; Siggelkow, 2001; Sørensen and Carroll, 2021; Van den Steen, 2017). This strategy can therefore help startups avoid costly mistakes that would occur were these activities to clash. Such anticipation might be particularly valuable as startups mature and face inter-connected decisions like expanding their markets (Bingham and Eisenhardt, 2011), creating moats relative to competitors (Guzman and Li, 2022), designing their organization (Lee, 2022), and formalizing an organizational culture (DeSantola and Gulati, 2017). For example, hiring the wrong sales talent can destroy the company's culture and its ability to expand into new markets. While experiments can test a particular approach to hiring, they might miss how this approach impacts the firm's market expansion and culture. By anticipating such contingencies, strategy can help startups achieve their goals (Dencker et al., 2009).

To help reconcile these views, this paper argues that the value of strategy depends on the startup's institutional context. Resource and cultural factors may shape how crucial it is to anticipate interdependent decisions. For example, failing to anticipate that new sales hires would destroy the company culture may be fatal in contexts where it is hard to find the money and talent to replace those hires and where there is a cultural stigma associated with failure. Prior studies in this debate generally assess firms within single-country contexts (Bhidé, 2000; Delmar and Shane, 2003; Dencker et al., 2009) and therefore do not capture how the value of strategy depends on such context in the first place.

To test this argument, this paper assesses: how does strategy's relationship with performance vary across institutional contexts? Answering this question requires measuring whether companies have a strategy. One way to detect strategy is by observing the alignment of entrepreneurs' choices, building on prior strategy theory that shows that strategy creates alignment, and so the lack of alignment suggests a lack of strategy (Porter, 1996; Van den Steen, 2017). This approach to detecting strategy allows comparing startups across countries. Comparing whether startups have a strategy using alternative approaches, such as measuring particular positioning choices (Porter, 1996) or resource investments (Barney, 1991; Wernerfelt, 1984), is difficult in international contexts where no one choice fits all (Khanna and Palepu, 1997). But using this flexible approach is nontrivial. Existing databases can reveal the talent that startups hired or the new country offices they built, but not whether entrepreneurs thought through the fit of these hires with their new country markets and their objectives. Neither do the surveys used in prior work to proxy strategic planning (Delmar and Shane, 2003; Dencker et al., 2009) capture how startups reason across choices before acting.

To overcome this empirical challenge, this paper pursues a field methodology to measure whether a global sample of scaling startups has a strategy. This field methodology enables measuring alignment across startups' choices to detect whether executives have a strategy. Structured interviews ask executives about their market scope, moat, organizational design, and organizational culture choices. Like the World Management Survey, the study quantitatively scores executives' interview responses according to a rubric (Bloom and Van Reenen, 2007). Rather than scoring the use of specific practices, however, evaluators measure the alignment of these choices with the objective of the executive and with the other choices. These measures then aggregate to a numerical strategy score. Matching the strategy scores—validated with natural language processing (NLP) techniques—with startups' geographic backgrounds and subsequent performance outcomes makes it possible to measure variance in the value of strategy across institutional contexts.

The resulting dataset covers the strategies of 253 software companies from 34 countries and six continents. The sample includes companies that received Series A funding ($5–20 million) from January 2019–September 2021. These high-growth companies, which received investments from top venture capital firms like Sequoia Capital, Y Combinator, Andreessen Horowitz, and the Founders Fund, make up over 12 percent of such software Series A deals in this time frame.[1] They raised about $30 million in funding and employed 90 people on average to date. This interviewed sample generally looks like the rest of the Series A population in observables like headquarters region, employee count, and initial financing amount, suggesting that it is a globally representative sample of firms in this phase. This dataset consists of roughly 190 hours of interviewing, a million words, and 63,000 coded observations. It is the first to systematically capture the strategic alignment of a globally representative sample of scaling startups.

Consistent with the idea that the institutional context shapes the value of strategy for startups, the paper finds substantial variance in strategy's relationship with performance across geographies. This variance particularly emerges between US and non-US contexts, where we also see stark historical disparities in startup scaling (Conti and Guzman, 2021; Kerr and Robert-Nicoud, 2020; PitchBook, 2022). While strategy weakly predicts the performance of firms, it strongly predicts subsequent performance for non-US firms. For the median non-US firm,

[1] Excluding China, where it is difficult to get performance data.

a one standard deviation increase in the strategy score is associated with an increase in valuation from about $32 million to $44 million and an increase in the probability of a successful exit by over four percentage points.

If strategy predicts performance more for non-US startups, then we would expect them to be more likely to develop a strategy to realize its seemingly high returns. Surprisingly, the study finds the opposite. Non-US firms have a 0.3 standard deviation lower strategy score, even when controlling for the strategy's content and readability. Non-US firms are less likely to develop a strategy.

Why does strategy predict performance more for non-US startups but is less likely to be developed by them? Qualitative analysis of the interview data reveals that due to financial, talent, and cultural constraints, mistakes—for example, hiring the wrong sales talent, which then impairs the company's ability to enter new markets—are more costly in non-US contexts, penalizing firms there without a strategy that can help anticipate such sources of failure. Yet, it is often through their own mistakes and those of their peers, advisors, and investors that entrepreneurs gain the knowledge to develop a strategy in the first place. Consistent with this explanation, executives outside the US—that is, where mistakes are more costly and therefore less likely to be sources of learning—rely less on direct experience, investors, and advisors to inform their strategy.

These results generalize beyond US and non-US contexts. Additional analyses—using a continuous country index of the ease of recovering from mistakes—yield similar results. This index is a composite of World Economic Forum indicators reflecting financial, talent, and cultural constraints that interviews reveal make mistakes more costly in non-US contexts. Using this index, it is possible to compare strategy scores and their relationship with performance across countries where mistakes are more versus less costly; for example, comparing Israel to France or the UK to South Korea. Consistent with prior results, strategy is more predictive of performance for firms headquartered in countries where it is harder to recover from mistakes. Yet this is also where strategy scores are lower.

Together, these findings suggest that in institutional contexts where mistakes are more costly, strategy matters more, but these contexts are also where it is harder to find the knowledge—gained from prior mistakes—to develop this strategy. The results shed light on why studies such as Bhidé (2000), focusing on US firms, find that planning has limited value. Others—like Delmar and Shane (2003), focusing on Swedish firms—find a higher value. By assessing entrepreneurs across countries, this paper reveals the institutional factors that may condition both the value and development of strategy.

This study makes several contributions to strategy and entrepreneurship research. First, it contributes to the debate on whether strategy matters for entrepreneurs (e.g., Bhidé, 2000; Delmar and Shane, 2003; Dencker et al., 2009; Mintzberg and Waters, 1985; Ott et al., 2017; Rivkin, 2000) by showing that institutional context shapes the value of strategic planning. Specifically, resource and cultural differences across geographies influence the cost of mistakes. Strategy, by helping startups anticipate conflicts related to fit, may particularly matter in contexts where the cost of mistakes is high. In contributing to this debate, the study sheds light on a potential cost of experimentation: the propensity to make mistakes related to fit between company activities. The modular nature of experiments makes it difficult to test interactions across company activities. Complementing experiments with strategic planning that holistically considers these activities can mitigate this risk.

Second, this study shows that not only the substance of choices but also their alignment varies across firms and geographies. While prior work focuses on the substance of startups' market, technology, and human capital choices (e.g., Eisenhardt and Schnoohoven, 1990; Gans et al., 2021), this study reveals that understanding how choices fit with a company's objective, assumptions, and other choices is also essential. This understanding is particularly important in contexts where mistakes are more costly to help firms avoid clashing market and organizational commitments that could be fatal. Consequently, interventions to change specific strategic choices in isolation might not always be sufficient to boost entrepreneurial performance. Interventions that increase the alignment of those choices could also be necessary.

Third, the research reveals that entrepreneurs in non-US contexts face not only resource frictions (Clough et al., 2019) but also knowledge frictions that may constrain growth. Knowledge from experimentation and experience—often involving mistakes—helps shape strategy. However, the concentration of experienced entrepreneurs, investors, and advisors who can afford to learn from prior mistakes in historically successful US hubs may make it harder for ventures outside these hubs to access this knowledge. Therefore, startups outside these US hubs need to be knowledge mobilizers to grow.

Fourth, the research extends theories on institutional voids and the resource-based view of the firm. Institutional factors—including constrained capital and talent markets (Khanna and Palepu, 1997)—that shape the cost of mistakes not only influence which strategies firms choose (Gao et al., 2017; Khanna and Palepu, 1997), but also whether strategy development capabilities are rare and valuable resources (Barney, 1991; Wernerfelt, 1984). Specifically, investing in the capability to develop strategy may yield dividends for firms in institutional contexts where mistakes are costly.

Lastly, the research helps bridge the gap between the global emergence of innovative startups and entrepreneurship research that is still primarily US-focused. Through interviews with startup executives from 34 countries and six continents, supplemented with human coding and NLP, the paper captures otherwise tacit strategy development among growing ventures across international contexts. The approach combines the depth of qualitative methods with the generalizability of quantitative methods to expand our geographic lens of entrepreneurial strategy and scaling.

1.2 Theoretical framework

The section proposes that strategy may matter for performance more in some geographic contexts than in others, depending on how costly it is to recover from mistakes. Yet, the cost of mistakes—by shaping access to knowledge—may also influence the development of strategy. These two trends may result in a perfect storm: strategy may matter more where it is more costly to develop.

1.2.1 The value of strategy

A rich body of scholarship suggests that strategy enables companies to develop a competitive advantage, in particular, by creating fit across company activities (Porter, 1996; Van den Steen, 2017). The interconnected nature of activities makes it difficult for rivals to duplicate efforts, even for seemingly public strategies (Rivkin, 2000; Porter, 1996). By creating fit across company activities, strategy can also improve the efficiency of internal processes, whether related to resource allocation or search (Rivkin and Siggelkow, 2003; Siggelkow, 2001). Lastly, strategy can help companies adapt to external changes (Sørensen and Carroll, 2021; Weick, 1976). The interconnected nature of activities means that any changes in the environment will come to the radar of multiple parts of the firm at once (Siggelkow, 2001; Weick, 1976).

In scaling ventures, which are less studied in this scholarship, failing to adopt a strategy that can anticipate fit can result in inconsistent commitments that fracture the company. This problem is unique to adolescent ventures because they face high uncertainty (Ozcan and Eisenhardt, 2009) and a relative lack of organizational knowledge and inertia (Gavetti and Rivkin, 2007). The result is extreme sensitivity and opportunism with regard to signals and changes in the external environment, with only a thin safety blanket of knowledge and resources. While this sensitivity helps these ventures escape the trap of inertia that sometimes cripples mature organizations with good fit among obsolete activities (Siggelkow, 2001), it presents a new risk: piecemeal changes that tear away at the fabric of the company. For example, tackling a new market opportunity in the consumer space while hiring enterprise sales talent salient in the local labor market may drive the company in two incompatible directions. Given entrepreneurs' scarce organizational knowledge and resources, recovery from such a fracture may be impossible.

Avoiding such misalignment can be particularly important as ventures pursue both market and organizational growth. In the scaling phase, ventures simultaneously face choices regarding growing their customer base and growing their team that can be incompatible (Eisenmann and Wagonfeld, 2014). Indeed, commitments to expand market scope (Bingham and Eisenhardt, 2011; Santos and Eisenhardt, 2009), create a moat relative to competitors (Gans et al., 2021; Guzman and Li, 2022), design the organization (DeSantola et al., 2022; Lee, 2022; Lee and Kim, 2022), and formalize an organizational culture (DeSantola and Gulati, 2017) may not go hand in hand. For example, founding team characteristics may imprint a firm's organizational culture (Beckman and Burton, 2008; Nelson, 2003; Stinchcombe, 1965), while external signals and market dynamics might influence its market scope and professionalization (DeSantola et al., 2022; Eisenhardt and Schoonhoven, 1990) in quite a different way. When the natural tide does not result in synchronization (DeSantola and Gulati, 2017; Eisenmann and Wagonfeld, 2014), active thinking about aligning these company building blocks becomes necessary. Strategy enables an increasingly multi-faceted venture to make commitments that complement rather than conflict with one another (Porter, 1996; Rivkin, 2000; Siggelkow, 2001; Van den Steen, 2017).

Still, strategy might not be equally valuable for all ventures. Avoiding misaligned commitments might be more important in contexts where it is costly to recover from mistakes, such as bad hires that destroy a company's ability to expand into new markets. Replacing a sales manager who proves to be a poor fit is harder where there are few sales candidates locally and limited venture capital money to cover the cost of recruiting and on-boarding a replacement, not to mention where the company will be stigmatized for having made the hiring mistake in the first place. Elsewhere, where mistakes are less costly, companies failing to adopt a strategy might not face as high of a penalty because it is easier to try again; for example, to fire and replace a bad hire. The value of getting alignment right ex-ante—which strategy enables doing (Van den Steen, 2017)—may not be as crucial for these firms because—with sufficient resources at hand—they can achieve alignment ex-post through, for example, trial-and-error.

The cost of such mistakes might arise from scarce resources—money, talent, and customers—as well as from a cultural aversion to failure. The differences in these resource and cultural factors influencing the cost of mistakes might especially fall on geographic lines. The agglomeration of venture capital (Chen et al., 2010; Sorenson and Stuart, 2001), talent (Glaeser et al., 2015; Fallick et al., 2006; Kerr, 2018; Kerr and Robert-Nicoid, 2020; Tambe and Hitt, 2014), and a culture embracing failure (Saxenian, 1996) in hubs like Silicon Valley makes it systematically easier to recover from mistakes there relative to other locations. Therefore, startups outside of hubs like Silicon Valley may face a particularly high penalty for lacking a strategy that helps them avoid mistakes that are so costly for them.

1.2.2 The development of strategy

The same institutional factors—such as the cost of mistakes—that systematically vary across geographies may influence not only the value of strategy but also its development. Strategy is a product of prior knowledge. Like scientific theory, it requires primary or secondary sources to inform hypotheses. This knowledge often forms from prior mistakes. For example, the recognition that a customer success team is vital for anticipating feedback from mainstream enterprise customers might emerge from having observed such customers fail to give feedback organically in the past. Similarly, the recognition that creating a user-friendly product for engineers requires formalizing a culture of autonomy and creativity might emerge from having earlier tried out a hierarchical culture that failed to meet engineer users' immediate needs. While a final strategy is straightforward, the process of getting to it might be anything but.

The knowledge needed to develop strategy comes from the experience—often including mistakes—of a variety of locally embedded sources. Entrepreneurs' local experience and experimentation may shape how well their choices fit together (Dahl and Sorenson, 2012; Gans et al., 2019; Gavetti et al., 2005; Gavetti and Porac, 2018; Michelacci and Silva, 2007; Shane, 2000; Wang, 2015). For example, directly experiencing the value of customer success teams for enterprise customers in a past venture helps entrepreneurs understand how well their subsequent hiring and market choices do or do not fit together. In place of direct experience, local advisors, investors, and peer entrepreneurs may convey similar lessons from their own experiences or those of their portfolio companies (Chatterji et al., 2019; Gavetti and Rivkin, 2007; Schilling, 2018; Vissa and Chacar, 2009). No matter the source, this knowledge conveys not only best practices but also fundamental ways of thinking about a concert of choices—mental models, so to speak (Baron and Hannan, 2002). Thus, developing a strategy requires access to knowledge drawn from locally embedded experience.

But entrepreneurs in contexts where mistakes are more costly may have less access to knowledge from their own experience or that of their peers, investors, and advisors to inform their strategy. The perceived cost of mistakes may disincentivize entrepreneurs from trying new approaches. Doing so risks making mistakes, from which they may not survive long enough to learn (Cahn et al., 2021; Kerr et al., 2014b; Landier, 2005). The same holds true for their local peers and the portfolio companies of local investors and advisors. Instead, accessing knowledge from remote sources may prove difficult. Research shows that companies can more easily mobilize resources and the knowledge that comes with them in their local context because resource holders may be biased against foreign startups or rely on local networks to source startups (Clough et al., 2019; Lin and Viswanathan, 2016; Shane and Cable, 2002; Vissa, 2011; Wright et al., 2022).

The resulting unequal access to knowledge may result in systematic differences in the development of strategy across geographies. Where knowledge from direct experience—including mistakes—is rare, startups may struggle to understand the holistic nature of the choices they make. Imitating another company's specific choice, after all, is far easier than recognizing, understanding, and imitating a complex system of choices following a guiding principle (Rivkin, 2000; Van den Steen, 2017). For example, Silicon Valley entrepreneurs can rely on their own scaling experiences or those of nearby peers, investors, and advisors. These prior experiences have taught them what types of talent best fit an enterprise sales model or which types of cultural values to formalize for millennial customers. Their direct experience conveys a way of approaching strategic choices in concert rather than in isolation. Compare this to an entrepreneur in Munich who can more easily find a former corporate executive than an entrepreneur who has scaled a business. Because it is harder to find scaling knowledge locally, the Munich entrepreneur might look to the specific strategic choices made by far-away companies, say in Silicon Valley. For example, the entrepreneur might borrow the software-as-a-service (SaaS) business model that many successful companies in Silicon Valley adopted. But that entrepreneur might fail to recognize the need to adopt a culture prioritizing feedback and transparency or to create an organizational structure based on industry divisions—choices that may have enabled the success of the SaaS model in the first place. By "mimicking" specific choices, the entrepreneur might lose sight of how these choices interact and what guiding principles drive them, inhibiting their ability to develop a strategy.

This theory suggests that strategy may be particularly valuable to help ventures avoid incompatible market and organizational commitments. However, not all companies may face the same penalty for such misalignment. The penalty might be higher where mistakes are more costly. Yet, it is in these same contexts where it may be harder to develop a strategy because entrepreneurs often accumulate knowledge to inform their strategy from prior mistakes. Thus, in institutional contexts where mistakes are more costly, strategy might matter more, but also be harder to create.

1.3 Methodology

The methodology for this study leverages structured interviews with startup executives that elicit their market scope, moat, organizational design, and organizational culture choices, along with the reasoning behind them. Measuring the internal and external alignment of each of these choices creates a numerical measure of the extent to which a company has a strategy. This measurement approach builds upon prior strategy scholarship that contends that strategy creates alignment among companies' choices (Porter, 1996; Rivkin, 2000; Sørensen & Carroll, 2021; Van den Steen, 2017). Connecting this measure with firms' financial and team data allows for measuring the relationship between strategy and performance across contexts.

1.3.1 Structured interviews to measure strategy

The field methodology to create a strategy measure uses data from interviews with executives of a globally representative sample of software startups beginning to scale. These interviews elicited executives' market scope, moat, organizational design, and organizational culture choices.

The interviews capture how startup executives in the scaling phase think about key market and organizational choices ahead of execution in the scaling phase and how their knowledge shapes their strategy. Deriving this information from existing databases or third-party sources is otherwise virtually impossible. While such sources may show the commitments made by firms—perhaps intentionally or not—they generally do not allow one to capture the thinking that led to those commitments to test whether companies had a strategy in the first place.

The interviews targeted software companies that had raised a Series A round since 2019. They focused on the software industry because companies in this sector often pursue standardized business models, such as software-as-a-service, that make cross-country comparisons feasible. This sector also drives high-growth entrepreneurship, accounting for most of the billion-dollar-valued unicorns that have emerged worldwide. The interviews focused on the Series A phase because ventures generally have reached product-market fit within their early adopter market and are now actively thinking about scaling to a broader market. In this scaling phase, companies simultaneously face market and organizational choices that can be aligned or misaligned, allowing one to detect whether a company has a strategy (Eisenmann and Wagonfeld, 2014). Further, sampling on this funding stage allows controlling for a quality threshold, as companies undergo rigorous due diligence to get the Series A. Indeed, these firms are highly promising, with investments from prestigious venture capitalists like Sequoia Capital and Andreessen Horowitz.

These interviews resulted from directly reaching out to startup executives and getting a positive response from a representative sample. From July to November 2021, executives—generally CEOs and co-founders—of startups in the software sector who raised a Series A ($5–20 million) round of funding since 2019, as listed in PitchBook, received a standardized email template inviting them to participate in a 45-minute interview as part of an academic project assessing how startups scale. These emails excluded companies from China because of the difficulty of getting performance data. Overall, 12 percent of such startups (253 companies) agreed to interview, exceeding the five-to-ten percent response rate seen in other research involving private sector surveys or interviews (Ben-David et al., 2013; Bloom et al., 2012a). The startups that agreed to interview do not appear to differ systematically from those that did not. For example, performance indicators do not predict whether a startup agreed to an interview (Table 1.1). The interviewed sample generally looks like the non-interviewed sample in terms of factors such as headquarters region (Figure A.3.1), employee count (Figure A.3.2), and first financing amount (Figure A.3.3). The interviews occurred from July to November 2021.

Table 1.1: Sample comparison table showing that interviewed companies do not vary from non-interviewed ones based on whether they are US-based, the number of employees, funding, and valuation.

	(1) Whether Interviewed
Whether US-Based	-0.004
	(0.034)
Log(Employees at Time of Interview + 1)	-0.016
	(0.014)
Log(First Funding Value + 1)	-0.005
	(0.006)
Log(Valuation at Time of Interview +1)	-0.004
	(0.010)
Primarily English-Speaking Country	0.056
	(0.049)
N	1106

The table compares interviewed versus non-interviewed companies in the sampling frame of software companies that raised a 5–20M USD Series A Jan. 1, 2019–Sep. 30, 2021 not including China. Robust standard errors (in parentheses) are clustered at the company level.
$+\ p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Interview questions were open-ended and future-oriented. They were open-ended to ensure the accuracy of responses by minimizing "social desirability bias" and the leading of interviewees, consistent with studies measuring management and strategy in mature firms (Bloom and Van Reenen, 2007; Yang et al., 2020). The questions were future-oriented to capture how executives think about scaling choices yet made rather than capturing retrospective accounts of decisions, which may be fraught with measurement error. The questions elicited executives' objectives over the next 3–5 years, how they planned to expand their markets, what they saw as their biggest moat against competitors, how they planned to expand their organization, and how they defined their company's culture. Additional questions asked executives what they surmised to be their next three action items to reach their objective, how they planned to use their Series A funding, and what they saw as their biggest uncertainties. The interviews also captured the sources of information that executives used to develop their strategy and why they did not pursue particular alternative approaches. The Appendix shows questions from the structured portions of the interview.

Independent evaluators coded the interviews in a double-blind manner. Five coders—MBA students and those with similar experiences—coded each interview transcript independently. To quantify the strategy, coding was based on a rubric[2] that measured how well each interview response fit with executives' assumptions, objectives, and other responses on a scale of one to five. Evaluators also provided binary codes (0/1) to indicate the presence or absence of particular aspects of the strategy—such as expanding across geographies or verticals—to measure the strategy's content. The interviews were double-blind, consistent with other research scoring management and strategic practices across organizations (Bloom and Van Reenen, 2007; Yang et al., 2020). Interviewed executives did not know that their responses would get quantitative scores. Interview coders also did not have performance information about the interviewed firms. Thus, neither the interviewees nor those doing the evaluation knew the relationship between a firm's strategy and its performance.

Using multiple independent evaluators validated the strategy coding. The coding rubric achieves a relatively high and stable inter-coder reliability: 0.9 correlation across all questions and 0.5 correlation among the questions coded one to five. This reliability is similar to correlations seen in past research (Bloom et al., 2012a). In the final dataset, about one-fifth of the interviews received independent evaluations from two coders, and the final strategy score was the average of the two. Due to resource limitations, the rest of the interviews received evaluations from one coder each.

[2]This rubric was the consequence of (a) pilot interviews with colleagues who were startup executives but not in the final sample and (b) feedback sessions with the evaluators.

1.3.2 Calculating strategy scores

Measuring how well executives' market scope, moat, organizational design, and organizational culture choices align with executives' assumptions (external alignment), as well as their objective and other choices (internal alignment), creates a numerical strategy score. This score indicates the extent to which a company has a strategy. Like other strategy measures, it captures a company's market and organizational choices, but unlike other measures, it also considers their fit with a common objective and assumptions.

The strategy score reflects the alignment of executives' market scope, moat, organizational design, and organizational culture choices. The final score sums the external alignment of each choice weighted by its internal alignment (Rivkin, 2000; Siggelkow, 2001; Sørensen and Carroll, 2021; Van den Steen, 2017). Equation 1 shows the algorithm that calculates the strategy score across these four choices:

$$strategy_i = \sum_{i}^{4} externalalignment_i * internalalignment_i \quad (1)$$

The dependent variable $strategy_i$ is the weighted overall strategy score of a company. $externalalignment_i$ reflects how well choice i fits with the executive's assumptions. $internalalignment_i$ refers to how well it fits with the executive's objective and other choices.

Why not consider only external alignment? No matter how well a choice fits with an executive's assumptions, it is ineffective if it does not move the company toward a broader goal or if it clashes with the other choices. To demonstrate, Table A.2.3 shows a company building a "one-stop-shop" for retail with a plan for designing the organization based on hiring experienced talent from around the world. While the choice fits well with the executive's assumption that there is a talent gap in the local labor market, it does not align with the company's objective to be a "super regional app with multiple services" in the Middle East. There is a discrepancy between the local nature of that objective and the global nature of the organizational design choice. As a result, this organizational design response is weighted lower in the final strategy score because it will be less likely to get the company to where it wants to be in the next three-to-five years.

The strategy score is robust to alternative coding and aggregation approaches. One alternative approach uses an SBERT NLP model, a word-embedding model that can capture the semantic meaning of the text at the sentence level (Carlson, 2022; Devlin et al., 2018; Reimers and Gurevych, 2019). This model measures the similarity between each of the market scope, moat, organizational design, and organizational culture responses and the objective to estimate external alignment. The similarity of the sentences within each of these four responses allows for proxying internal alignment. Multiplying these two NLP-based measures creates an alternative strategy composite score. Table A.4.1 shows that these NLP-based measures correlate with the human-generated scores. The results are also robust to different ways of summing the coding variables, as shown in Table A.5.1. These alternative measures do not fill in missing values, aggregate the scores based on simple averages, exclude fit with other choices, or double-weight the fit with the objective. The measures have at least a 0.93 correlation with the main score and with one another.

Robustness checks also show that the scores withstand differences in speaking styles. The length of responses does not predict the strategy score. This analysis reduces the concern that time constraints or speaking styles confound this measure. Later specifications also control for the English readability of the transcribed responses to ensure that language differences do not meaningfully affect score comparisons.

The strategy measure builds on a rich body of strategy and entrepreneurship research. As in Kaplan and Norton (1992), this measure brings together a view of both the customer-facing (market scope and moat) and internal business (organizational structure and culture) perspectives. While Kaplan and Norton (1992) measure the quality of each of these perspectives based on its objectives—new products from the customer perspective and technological capability from the internal business perspective—the measure presented in this paper also assesses how these choices address each other and the firm's broader aim. It thus captures interdependence among the firm's core decisions (Rivkin, 2000; Van den Steen, 2017).

1.3.3 Data to measure variance in strategy

The final dataset connects startups' strategy scores and other interview responses with their team characteristics and post-interview performance outcomes. Specifically, it contains data on the firms' financials from PitchBook, Crunchbase, and PrivCo; employee backgrounds from online resume data; website visits from SimilarWeb; technology tools from BuiltWith; and websites over time from the Internet Archives, for which summary statistics are shown in Table 1.2. It also includes the pitch decks of a third of the firms and the organizational charts of a fourth of them to corroborate interview information. Below are the primary variables used in this study.

- **Strategy** indicates the 1–100 strategy score that is then standardized using interview data.

- **Logged post-money valuation** indicates a company's post-money valuation, reflecting its expected value at the time of investment, using PitchBook's database.

- **Logged valuation per employee count** indicates valuation dollars per employee as a rough metric of productivity for these young ventures, using PitchBook data.
- **Exit or additional funding** indicates whether a company achieved a successful exit in the form of an acquisition or initial public offering, as well as whether it received additional funding rounds (Series B+), using PitchBook data.
- **Logged employees** indicates the logged employee count using PitchBook data.
- **Reached 150 or 200 employees** indicates an alternative measure of employee count, that is, by whether a company has reached 150 or 200 employees—the top percentile (10-15 percent)—of the employee distribution across sampled companies.
- **Logged page visits** indicates the number of page visits to a company's homepage, reflecting user growth, using SimilarWeb's database. Entrepreneurship studies increasingly use website page visits as a proxy for startup performance (Cao et al., 2021; Hallen et al., 2020; Koning et al., 2022).
- **Logged technology tool count** indicates the number of technology tools installed on the company's homepage since the interview period, using BuiltWith's database. This variable indicates technological sophistication.

1.4 Results

1.4.1 How does strategy's relationship with performance vary?

The study begins with assessing the baseline relationship between strategy and performance. Figure 1.1 shows a weak positive association between strategy on the x-axis and logged valuations—as one metric of performance—on the y-axis. However, this graph masks substantial variance across firms. Figure 1.2 breaks down this same relationship between US and non-US firms. Surprisingly, valuations appear to be sensitive to strategy scores for non-US firms, but not for US ones. Non-US firms (right panel) see a steeper slope between strategy and log valuations than do US firms (left panel). This figure suggests that strategy predicts valuations only for non-US firms.

Table 1.2: Summary table

	(1)				
	Non-US Obs	Non-US Mean	US Obs	US Mean	Difference
Revenue-Generating Status	129	0.98	124	0.98	0.00
Age (Years)	129	5.76	124	5.48	0.28
Num. Employees	127	107.35	124	71.92	35.43*
Log(Employees+1)	127	4.32	124	3.90	0.42***
Page Visits (Thousands) Since Oct. 2021	129	131.78	124	76.92	54.87
Log(Page Visits +1)	129	9.46	124	8.91	0.55*
Funding Amount (Millions USD)	129	34.29	124	28.08	6.20
Log(Funding Amount + 1)	129	3.01	124	3.03	-0.02
Post-Money Valuation (Millions USD)	75	76.04	120	118.01	-41.97
Log(Valuation+1)	75	3.43	120	4.12	-0.69***
Whether Acquired	129	0.05	124	0.06	-0.01
Whether Profitable	129	0.02	124	0.01	0.02
Revenue (Millions USD)	45	8.42	70	6.80	1.63
Num. Tech Tools on Website (Thousands) Since Oct. 2021	129	0.27	124	0.29	-0.02
Log(Total Tools on Website+1)	129	5.47	124	5.58	-0.10
Whether CEO/Founders Have PhD	129	0.03	124	0.05	-0.02
Whether CEO/Founders Have MBA	129	0.18	124	0.35	-0.18**
Whether CEO/Founders Have Law Degree	129	0.00	124	0.08	-0.08***
Whether CEO/Founders Have Masters	129	0.33	124	0.40	-0.07
Whether CEO/Founders Have Worked Outside of HQ	129	0.89	124	1.00	-0.11***
Whether CEO/Founders Have Worked in the US	129	0.61	124	1.00	-0.39***
Whether CEO/Founders Worked in Hub Ecosystem	129	0.79	124	0.83	-0.04
Whether CEO/Founders Worked in US Hub Ecosystem	129	0.22	124	0.87	-0.65***
Whether CEO/Founders Attended Elite Global University	129	0.34	124	0.39	-0.05
Whether CEO/Founders Attended US University	129	0.27	124	0.83	-0.56***
Whether CEO/Founders Were Previously CEO/Founders	129	0.53	124	0.54	-0.01
Whether CEO/Founders Were Previously Investors	129	0.05	124	0.08	-0.03
Whether CEO/Founders Were Previously Consultants	129	0.19	124	0.19	0.00
Whether Have US Investors	129	0.69	124	0.98	-0.29***
N	253				

The table shows summary statistics for interviewed firms, broken up by whether firms are headquartered in the US (right) or not (left).

Figure 1.1: Strategy weakly predicts performance of firms.

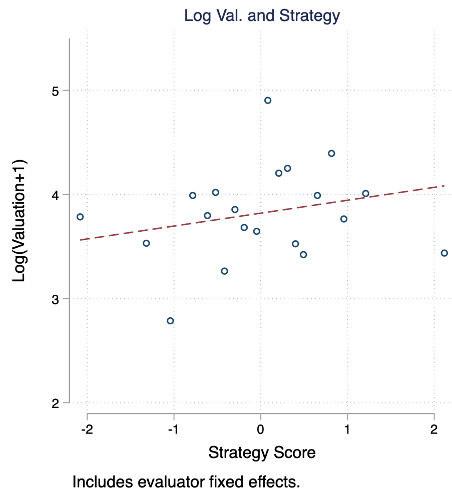
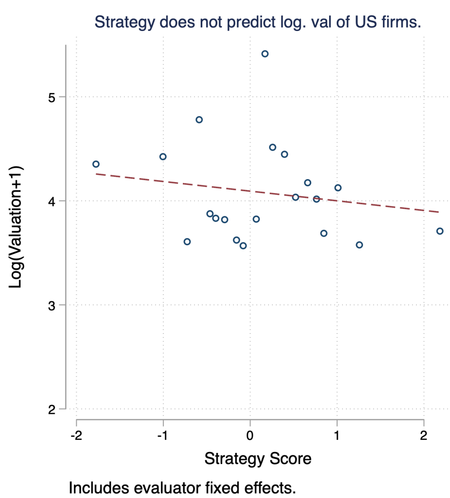
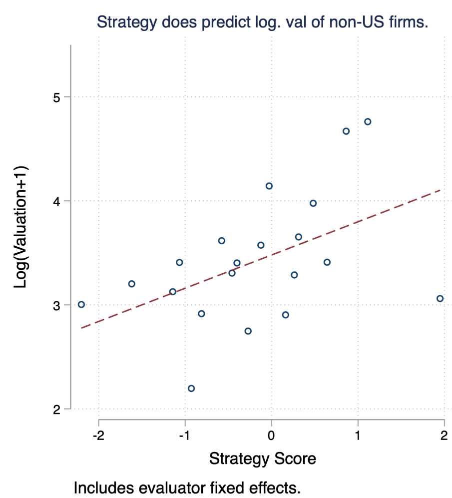


Figure 1.2: Strategy predicts performance for non-US firms (right), but not US ones (left).





Equation (2) tests this visual relationship more rigorously:

$$y_i = \beta_1 us_i + \beta_2 strategy_{ij} + \beta_3 us_i x strategy_{ij} + foundedyear_i + industry_i + evaluator_j + readability_{ij} + gdpcapita_i + firstfinancing_i + \gamma_{ij} + \varepsilon_{ij} \quad (2)$$

The dependent variable y_i is a vector of post-interview performance outcomes, including logged valuation, valuation per employee, future funding, and exit. This vector also includes a performance index that is the normalized average of the normalized transformations of the previously mentioned performance outcomes along with other funding, employee, page visit, and technology tool measures individually shown in Table A.14.1. $strategy_{ij}$ indicates the standardized strategy score calculated using Equation 1 by a given evaluator j for company i. us_i indicates whether the firm is headquartered in the US. $us_i x strategy_{ij}$ indicates whether the relationship between strategy and performance varies between US and non-US firms. $foundedyear_i$ indicates firm i's founding year to control for differences in firm maturity. $industry_i$ indicates the industry cluster of firm i generated from a k-means clustering (unsupervised) machine learning model using the company's keywords. $evaluator_j$ reflects evaluator fixed effects. $readability_{ij}$ reflects (a) the English language quality of responses, taking into account the evaluator's attested understanding of the interview transcripts due to language barriers (irrespective of the content) and (b) the Flesch Reading Ease Score using an NLP technique from the Python textstat library. This algorithm allows for more objectively measuring how feasible it is to read a body of text. γ_{ij} reflects whether the analysis filled in missing values for the strategy score of firm i with evaluator j's average evaluations for firm i. These missing values made up less than two percent of the codes. $gdpcapita_i$ accounts for logged GDP per capita differences across countries, as GDP per capita may impact both the strategy score and performance. $firstfinancing_i$ indicates the logged initial financing amount (in USD) that company i received, which may affect the strategy score and current performance. The coefficient of interest is β_3. This coefficient shows how strategy scores predict performance for US versus non-US companies. A negative coefficient indicates that the strategy score is more predictive of subsequent performance for non-US firms.

Table 1.3 shows results applying Equation 2. Specifically, it shows that strategy scores are more predictive of performance for non-US firms than for US firms. The coefficients on the interaction terms between whether a company has a headquarters in the US and the strategy score (third row) across outcomes is negative. The binscatter and regressions suggest that strategy matters more outside the US. This effect is economically significant: for the median non-US firm, a one standard deviation increase in the strategy score is associated with over a third increase in valuation—roughly from $32 million to $44 million—and a four percentage point increase in the probability of a successful exit.

Table 1.3: Strategy predicts performance more for non-US firms.

	(1) Log Val.	(2) Log Val. Per Employee	(3) Exited	(4) Exited/Raised Future Funding	(5) Performance Index
US HQ	0.614*	0.376**	-0.009	0.164*	0.358*
	(0.258)	(0.119)	(0.042)	(0.074)	(0.163)
Strategy	0.501**	0.145+	0.061*	-0.013	0.288**
	(0.189)	(0.075)	(0.030)	(0.048)	(0.110)
US HQ x Strategy	-0.669**	-0.275**	-0.081*	-0.116+	-0.463**
	(0.220)	(0.090)	(0.040)	(0.059)	(0.141)
N	184	183	230	230	230
Evaluator FE	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Log GDP Capita	Yes	Yes	Yes	Yes	Yes
Log First Financing	Yes	Yes	Yes	Yes	No
Filled-In FE	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes

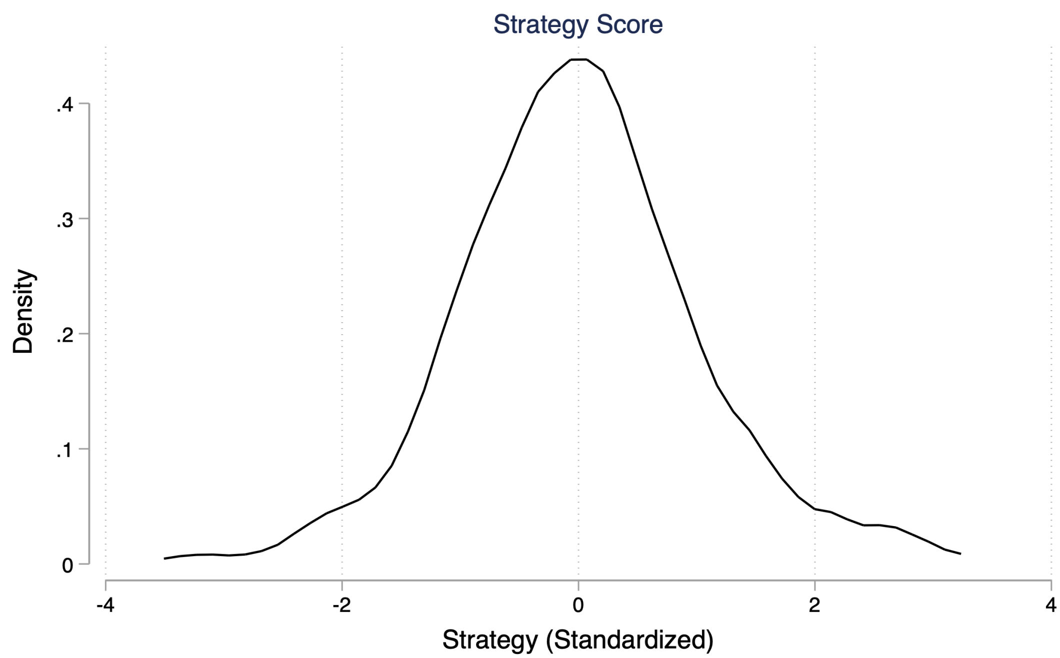
The table shows how the relationship between the strategy score and performance varies for US versus non-US firms. The models control for the English readability of the text, firms' founding year and industry, evaluator fixed effects, whether any responses were filled in, the GDP per capita of the headquarters country, and ventures' logged initial financing amount in USD. The sample size drops because of missing valuation and first financing size data from PitchBook. The results are robust to excluding first financing size as a control. Robust standard errors (in parentheses) are clustered at the company level.+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

One concern with this result is that it may simply reflect that US and non-US firms are in different phases of development. US firms might not have reached product-market fit yet because they can get a Series A at an earlier stage. Thus, the results could reflect that strategy matters when companies have reached product-market fit. This would suggest that the strategy depends on the firm's maturity rather than on its institutional context. To test this confounding variable, Table 1.2 compares the development phase of US and non-US companies at the time of the interview. It shows that according to PitchBook data, US and non-US firms are equally likely to be generating revenue at the time of the interview. They are also of similar age—at about five years old. Further, US companies mention in interviews that they have reached product-market fit. For example, one US company mentioned that as part of its objective: "is a simple goal of scaling and growing...We've got a good product-market fit." Together, this quantitative and qualitative evidence suggests that US and non-US companies are at a similar lifecycle phase at the time of the interview, so firm maturity is unlikely to be driving the results.

1.4.2 How does the development of strategy vary?

If strategy is more predictive of performance for non-US startups—suggesting it may yield higher returns for them—then we would expect that these firms would be more likely to develop a strategy. To test this prediction, the study assesses variance in the development of strategy among US and non-US firms. It begins by assessing the baseline score across all firms. Figure 1.3 shows a kernel density plot of the standardized strategy scores, which are scaled from a 1 to 100 range to a -3.2 to 3.0 range. They approach a normal distribution with a standard deviation of 0.93. A similar variance emerges when looking within particular objectives or choices. For example, firms that have a social impact objective or market scope approach focused on geographic expansion see a wide spread of strategy scores.

Figure 1.3: Kernel density plot of strategy scores

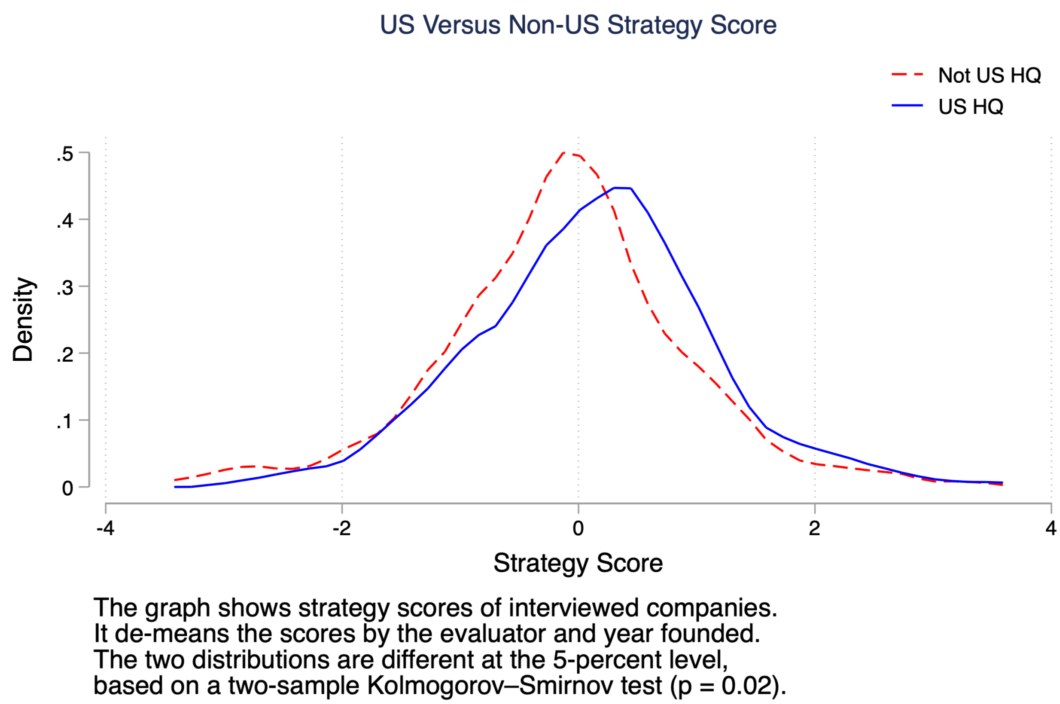


The graph shows strategy scores of interviewed companies.

In contrast to the prediction that non-US startups would be more likely to develop a strategy, the kernel density plot in Figure 1.4 shows that the distribution of strategy scores of US firms relative to others is shifted to the right. This suggests that non-US firms have lower scores and, thus, are less likely to develop a strategy. Kolmogorov-Smirnov tests show that the distributions differ statistically from one another (p=0.02).

Equation (3) measures the variance in the development of strategy between US and non-US firms more rigor-

Figure 1.4: US companies see a right shift in the distribution of strategy scores.



ously:

$$strategy_{ij} = \beta_1 us_i + foundedyear_i + industry_i + evaluator_j + readability_{ij} + \gamma_{ij} + \varepsilon_{ij} \quad (3)$$

The variable of interest is β_1, indicating how strategy varies across US and other firms.

Table 1.4 shows the results from Equation 3. US firms have a 0.3 standard deviation higher strategy score than do non-US firms (Column 1). This result suggests that US firms have a higher strategy score than other firms, consistent with the kernel density plot shown in Figure 1.4. Specifically, it shows that US companies pursue moats that better fit other choices and assumptions by 0.2-0.3 standard deviation (Columns 6–7), have organizational designs that better fit their objective by 0.2 standard deviation (Column 8), and have organizational cultures that better fit their assumptions by 0.2 standard deviation (Column 13). The coefficients on the remaining subscores are generally positive, though not significant at the five-percent level.

This geographic variance is partially consistent with prior work on management differences across countries. The gap between US and non-US companies is similar to differences in management practices (Bloom et al., 2012a). However, this difference does not appear to be a function of GDP, as seen in the management studies. Logged GDP per capita does not predict the strategy score at the 5–10-percent–significance level.

Table 1.4: Non-US firms have lower strategy scores.

	(1) Strategy Composite Score	(2) Market Internal Align.-Obj.	(3) Market Internal Align.-Other	(4) Market External Align.	(5) Moat Internal Align.-Obj.	(6) Moat Internal Align.-Other	(7) Moat External Align.	(8) Org. Internal Align.-Obj.	(9) Org. Internal Align.-Other	(10) Org. External Align.	(11) Culture Internal Align.-Obj.	(12) Culture Internal Align.-Other	(13) Culture External Align.
US HQ	0.269*	-0.045	0.110	0.091	0.163	0.189+	0.283*	0.244*	0.104	0.152	0.068	0.157	0.212+
	(0.128)	(0.115)	(0.117)	(0.118)	(0.104)	(0.114)	(0.117)	(0.111)	(0.125)	(0.118)	(0.116)	(0.124)	(0.128)
_cons	0.322	-0.671	0.847	-0.214	0.988*	0.910	0.494	-0.001	-0.163	-0.343	0.914	0.155	0.094
	(0.569)	(0.476)	(0.687)	(0.547)	(0.480)	(0.570)	(0.516)	(0.526)	(0.665)	(0.630)	(0.566)	(0.584)	(0.682)
N	304	304	304	304	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Filled-In	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table shows how being headquartered in the US associates with the strategy score and its subcomponents, including the fit of executives' market scope, moat, organizational design, and organizational culture choices with their assumptions (external alignment), as well as with their objectives and other choices (internal alignment). The models control for the English readability of the text, firms' founding year and industry, evaluator fixed effects, and whether any missing responses were filled in. The sample size is 304 because about a fifth (51) of the interviews received two evaluators to ensure the robustness of the scoring. Robust standard errors (in parentheses) are clustered at the company level.

$+\ p < 0.1$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Assessing specific subscores that vary across US and non-US firms illustrates the nature of these differences. US firms have a higher score on their moat, driven by its external alignment (Table 1.4 Column 7). For example, a US-based company with a recruiting platform logically laid out why its moat would be enduring:

> Data...The problem with talent data is it's huge. And it is prohibitively expensive for even a multi-billion-dollar organization to try and structure data that way. So we instead were...going to build [our own data], which was a risk. And once we pulled that off, we really started to see that we could identify people in this much better way...We could also match people much better for organizations based on capability. And the people that we were putting there were so much more diverse.

The company articulated that its data moat can withstand competitors because of how difficult it was to collect and create meaningful insights from the available human resource data.

In contrast to this US company's logic of how data can create barriers to entry, a Singapore-based company creating a mobile credit-scoring platform had difficulty conveying why others would not be able to replicate its data moat:

> So there are companies who are developing scorecards based on telco data. We don't consider them as a direct competitor because we have clients who are using both, right? Scorecards developed based on device, mobile device data. That is [our company]. So we can coexist together.

The company was not able to articulate why other players would not be able to build their own mobile data moat. While the company argued that it could coexist with companies providing telco data because consumers used both, it is not clear why other players would not have an incentive to produce mobile-derived data if they would achieve economies of scope providing both types of data to consumers. Table A.2.2 shows additional examples of moat responses and scores.

US firms' higher scores on organizational design are driven by their internal alignment, specifically related to fit with the objective (Table 1.4 Column 8). The ability of US firms to recruit and structure their organization in closer alignment with their objective is consistent with "hiring ahead of the curve," which prior work suggests is important for scaling (Eisenmann and Wagonfeld, 2014). For example, one US company spoke of deliberately hiring for new roles and positions—particularly in sales—ahead of current demand consistent with its objective:

> There's proactive hiring and then there's kind of like reactive hiring. So proactive hiring [includes] sales. And the engineering org is proactive hiring. We can't build a thing and we can't sell an incremental customer unless we have someone staffed and trained and ready to go ahead of time.

This forward-looking approach is consistent with the company's objective, focused on "not just the absolute place that we get, but the speed...$100 million top-line business, that's kind of like the North Star." Proactive hiring enables the company to prioritize increasing revenue and doing so quickly.

On the other hand, a UK-based company desired to defer hiring as long as possible. However, this measured approach makes it difficult to achieve the company's ambitious objective.

> And one of the things we want to do is stay small...So I see a lot of founders just use whatever cash they can, just really increase headcount way ahead...We're trying to stay lean. And I think what we'll do is, we will grow mainly in design and product and engineering. Basically at the rate that you see.

While reasonable on its own, this frugal approach does not align closely with the company's ambitious objective to "recreate the success of something like Slack....this means tens of millions of users, likely across maybe a few 100,000 companies that are using this." Such a high-reach objective suggests the need for robust marketing and customer success teams to acquire and retain this large customer base. Additional examples of organizational design responses and scores are shown in Table A.2.3.

A concern with this strategy gap is that it is capturing differences in the content of the strategies. For example, US and non-US companies might have different objectives and, therefore, might pursue them by different means. The methodology would then be comparing not only the internal consistency of planning but also its content. The calculation would then confound the results.

To address such concerns, additional analyses assess differences in the content of strategy among US and non-US firms and then control for them. US and non-US firms pursue similar financial objectives (Table A.10.1). However, non-US firms are more likely to pursue a geographically oriented strategy in terms of seeking a global objective, expanding markets by international expansion, and hiring geographically oriented talent (Table A.11.1). To account for this, Equation (3) controls for these geographically oriented factors. Table 1.5 shows the results: the strategy gap remains when controlling for the geographically oriented elements of strategy that vary among the US and non-US firms—using OLS (Column 1) and LASSO (Column 2) models—as well as for all contents of strategy (Column 3). These results suggest that the strategy gap between US and non-US firms is not driven by their content but by how aligned they are.

Another concern with these results is that underlying quality differences between US and non-US firms may be driving the results; that is, non-US firms in the final sample might be of lower quality than the US firms. Con-

Table 1.5: The strategy gap between US and non-US firms remains, even when accounting for the content of strategy.

	(1) Strategy (OLS)	(2) Strategy (LASSO)	(3) Strategy (LASSO)
US HQ	0.252^{+}	0.263^{*}	0.228^{*}
	(0.130)	(0.122)	(0.109)
N	295	304	304
Evaluator FE	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Filled-In FE	Yes	Yes	Yes
Readability	Yes	Yes	Yes
Geo Oriented Content	Yes	Yes	Yes
All Strategy Content	No	No	Yes

The table shows how being headquartered in the US associates with the strategy score. Column 1 shows the same OLS model as used in Table 3. Column 2 uses an OLS model that controls for the geographically oriented content of strategy that varies across US and non-US firms. Column 3 uses a LASSO model to control for all contents of strategy. The models also control for the English readability of the text, firms' founding year and industry, evaluator fixed effects, and whether any missing responses were filled in. The sample size is 304 because about a fifth (51) of the interviews received two evaluators to ensure the robustness of the scoring. Robust standard errors (in parentheses) are clustered at the company level.+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

ceptually, we might expect the opposite. Financial resources are more scarce outside the US, making it harder to get Series A funding. Non-US firms that do manage to get this funding might therefore be of higher quality. However, the study mitigates that concern by controlling for various quality differences. First, the analysis controls for a threshold of quality by focusing on companies that raised a Series A round of funding, generally requiring rigorous due diligence. Second, the regressions control for the English readability of the text and for the firm's founding year, which reflects its maturity. Together, these conceptual and empirical checks suggest that underlying quality inherently unrelated to strategy is unlikely to be driving the results.

Together, these results show that strategy varies among US and other firms. Despite strategy being more predictive of performance for them, non-US firms have lower strategy scores—they are less likely to develop a strategy—irrespective of the content of their strategy.

1.4.3 Why are strategy scores lower where they matter more? A cost-of-mistakes mechanism.

Why do non-US firms have lower strategy scores if these scores are more predictive of performance? Qualitative analysis of the interviews reveals a cost-of-mistakes mechanism. Non-US firms discussed how strategic planning was important to avoid mistakes that were costly in their local environments. Yet, mistakes also helped entrepreneurs build knowledge to develop their strategy in the first place. Additional analyses using a country index of the ease of recovering from mistakes support these qualitative findings.

Cost of mistakes and the value of strategy

Why does strategy associate with performance more in non-US versus US contexts? Figure 1.2 shows that US and non-US startups get a similarly high valuation when they do have a strategy (at a higher strategy score), but non-US companies get a substantially lower valuation when they lack a strategy (at a lower strategy score). Based on the data and interviews, it appears that it is not that non-US companies get a higher reward for having a strategy, but rather that they face a higher penalty for lacking a strategy.

To understand why this penalty for lacking a strategy is higher for non-US firms, it is important to understand what it means to lack a strategy in the first place. Why does it matter at all? Turning back to the conceptual framework in Section 2, strategy can help companies anticipate mistakes, particularly those related to fit (Porter, 1996; Siggelkow, 2001). Conversely, without a strategy, startups might be more likely to hire the wrong talent, enter unfruitful markets, or adopt imitable technology. To illustrate, a South Korean startup discussed how lacking strategy—not being "smarter earlier," as the executive put it—resulted in hiring talent that did not fit with its system:

> One of the main mistakes that we actually made...right after Series A is that we went on a hiring spree. And the team went from 20 to actually 70 people in less than six months. And that rapid growth unexpectedly damaged our culture because a lot of new people with different ideas and different working habits came in too fast...And so there was a lot of back progress that we have to have gone through. And that's just time and money being spent not on the company but rather fixing mistakes that, if we were smarter earlier, then we probably wouldn't have made in the first place.

The South Korean startup shows how this hiring mistake—which could have been avoided had they been "smarter earlier,"—that is, had they planned ahead through a strategy—was costly in terms of "time and money." The company faced a steep penalty for not having a strategy. Consistent with this view that lacking a strategy makes companies more likely to make mistakes, Table A.9.1 provides suggestive quantitative evidence that a higher strategy score is associated with lower regional office or subsidiary closures—as a proxy for mistakes.

In contrast, US companies discussed how their mistakes did not incur as much of a penalty. If anything, they served as a learning opportunity. The availability of money and capital enabled them to try again, for example, by replacing bad hires. Without a strategy, these companies could end up making mistakes but could also afford a second chance. For example, one US company noted being able to afford to make mistakes—such as hiring the wrong talent—and viewed such mistakes as a learning opportunity:

> Because you have access to money and capital, then you make a lot of mistakes. And those can be expensive mistakes...Some of the biggest mistakes and learnings for me is hiring people—the right role...and who you're really hiring for, cultural fit, and things like that.

The US company paid a relatively low penalty for making a mistake related to "hiring for fit." While a similar mistake was detrimental to the South Korean startup, the US startup perceived its own mistakes as harmless and even helpful for learning about its talent needs. This startup was not an exception in the US. Other US startup executives discussed the relatively harmless nature of mistakes, embracing them as learning opportunities. One executive noted that he "intentionally hired very smart people. So I let them be smart and execute. And I also let them make mistakes because that's how you learn." Another US executive explained: "If a mistake happens, or like a true mistake...it's fine...What are you going to do to fix it? Great—fix it. How to prevent it? Cool, I'm not going to ask you about it again." Indeed, Table A.8.1 shows that US companies were more likely to discuss mistakes as learning opportunities rather than as something to avoid, suggesting that mistakes are less costly for them.

Now, why might the penalty for such a hiring mistake be higher for the South Korean startup than for the US one? Executives in South Korea and elsewhere outside the US discussed how difficult it was for them to try again because of scarce financial, talent, and customer resources and of a cultural aversion to failure. Mistakes were, therefore, costly and strategic planning critical in order to avoid them. One Brazilian company noted the financial constraints that made trying again difficult and strategy crucial:

> In the Valley, you can validate an idea really fast. You can fail fast, and you can try again...And you will find a path of funding pretty—not easy—but you have many options...In Brazil, you have, like, 10 VCs...And so you need to [consider], can I extract value from this product in 24 months or maybe 36 months...and use this to self-fund the next step?

In addition to financial constraints, executives outside the US discussed the talent constraints that made mistakes more costly. For example, one Swedish company mentioned:

> I believe that the best people already have jobs, so you would have to rip people away from great positions with vested equity...So talent is absolutely the hardest...I believe that the cost of a failed

> recruitment outweighs the benefit of speed, probably by a factor 10 or more. So hence, I'm involved in every single recruitment.

The Swedish example demonstrates how talent resources, like financial resources, were hard to find locally. As a result, companies needed to think ahead about each hire to avoid mistakes that carried a high price tag.

Like financial and talent constraints, having limited customers locally made it that much harder to try again if one of the customer relationships failed. An Australian company discussed the need to plan market expansion early because of the scarcity of local customers:

> If you originate in San Francisco, as long as you focus from [there to] New York, you're pretty good to cover most of the world. But the difference in Australian companies...when you're talking about a population base of 20 million...that's not really enough to build...a billion-dollar business. And so you immediately have this perspective of, how will I scale my product into other geographies?

Making a mistake for this company in any one country market is consequential. They would need to pay the fees associated with closing an operation in the "failed" country market and then pay high fees to open up operations in a new country. Compare this approach to a US company that fails in Pennsylvania. The cost to close up operations in Pennsylvania and open up an office in New York likely would be far lower than in the international case.

Beyond financial, talent, and customer resource constraints, non-US startups discussed how a broader cultural aversion to failure made mistakes quite costly. For example, one German startup mentioned: "What we see currently in Europe is that many mistakes are very much punished. There's not a culture of [making] mistakes." Similarly, a UK startup discussed how averse customers were to disruptions. This reality contrasted with what the interviewee considered to be the case in Silicon Valley, where customers might be more forgiving:

> The approach to the development of technology differs between Europe and Silicon Valley. In Silicon Valley, the key approach is to fail fast but move on quickly. So yes, we're losing out if something breaks, not a big problem. Resolve it in the next release, and you're fine. Both my last business and the current one focus on providing services to institutional clients in financial services. [For us], failing fast is not an option. So if you fail and your bank account is not accessible for whatever reason, the payment is not successful for some reason. Then that creates a headache in itself. And you may not have that customer anymore.

Like the financial and talent constraints, the cultural aversion to failure that the German company saw in its customers made it difficult to try again after a mistake.

These examples show how non-US companies face higher penalties for not having a strategy, consistent with the steeper relationship between the strategy score and valuations that we observe for these companies relative to US ones in Figure 1.2. Without a strategy to anticipate sources of failure, these companies are more likely to make bad hires, enter the wrong markets, or make other mistakes. And these mistakes carry steep penalties outside of the US. Trying again is difficult amid scarce financial, talent, and customer resources and a cultural aversion to failure. Having a strategy to avoid these costly mistakes becomes essential.

To what extent is the cost of mistakes a reflection of investor preferences rather than the startup's institutional context? For example, perhaps US investors with their Lean Startup playbooks prefer their portfolio companies—no matter where they are in the world—to experiment and learn from mistakes. These investors condition their future funding on observing such experimental behavior. Investors from other parts of the world might prefer planning ahead to avoid mistakes and condition their future funding and support on having a strategy. If this were the case, then we would see similar results as in Table 1.3 if we replaced the variable reflecting whether startups are headquartered in the US with whether they have US investors. Interestingly, the results do not hold with this measure: the relationship between strategy and performance does not vary based on whether startups have US investors. This suggests that the cost of mistakes is not operating through the geography of the investors but rather through the geography of the startups.

Cost of mistakes and the development of strategy

While the cost of mistakes might make it more important for startups to have a strategy, it may also make it harder for them to develop the strategy in the first place. Interviews reveal that entrepreneurs learn from their prior experience—and specifically from mistakes—to inform their strategy. As one US startup noted:

> So, nowadays, I can better plan because of those mistakes I have made in the past and because of all the feedback I have received. But for me, it has been more roll up your sleeves and just get into the work and just do it. Just because I don't know how to do it, it's not a good excuse. Go ahead and just do it and then ask for input and people will be there to help you.

But startups outside the US—where mistakes are more costly—discuss how they cannot afford to experiment and potentially make mistakes that are disruptive, even if these mistakes can inform the startup's strategy. For example, one UK executive noted how he preferred to learn from others' mistakes, given how costly it had been for his company to deal with prior hiring mistakes.

> [We] went through...mistakes like that. And you think, oh, it's easy. We'll fire them. It's not easy. It always creates collateral damage. People can get toxic. Other people get disoriented...Firing the people that don't seem to fit can also be quite disruptive to your organization...I really do not like first-principle thinking because I hate making mistakes that other people know the answers to.

Yet, relying on knowledge from others is also difficult in these contexts. When mistakes are costly, peer entrepreneurs have just as hard a time experimenting, which constrains their ability to develop a strategy that enables them to scale. Thus, there are fewer local company examples to rely on to inform a strategy (Gavetti and Rivkin, 2007). One South Korean startup recounted the difficulty of finding appropriate company examples locally:

> Our situation is very unique. So I can't really come up with one [company example] in 2010 because all the well-known companies, their success formula was: scale fast, burn money fast, scale fast, become too big to fail. That was not our situation. All the players, Korean ones, you make money first and then bring your cronies together and leave in all the interests and launder your image...That wasn't the case for us. Interesting, I don't think I have ever had any role model.

This executive revealed the struggle that non-US companies faced in accessing appropriate knowledge to inform their strategy. The company could not easily turn to either foreign or local company examples. Foreign companies "burned money" in a way that could be fatal locally. Local companies pursued approaches that were well adapted to the local context but which conflicted with the focal company's ideals.

Where there are few companies historically that have experimented, made mistakes, and learned from them to develop their strategy and scale, there are also few advisors and investors who worked with such companies. Therefore, not only does the cost of mistakes make it harder for entrepreneurs to experiment themselves or to learn from other local companies that have done so, but they also have a hard time finding experienced investors and advisors locally, whom prior work shows are important inputs into entrepreneurs' strategy development (Bernstein et al., 2016; Chatterji et al., 2019; Vissa and Chacar, 2009). Consistent with this view, interviews reveal that non-US firms found it difficult to find competent advisors and investors locally as they approached scaling. An Australian company noted: "In our part of the world, there wasn't really a lot of people that we could learn from in the earliest phases of the company." Similarly, a UK-based company acknowledged that local advisors lacked the experience to inform the company's strategy: "[There is] not much [reliance on advisors]—for sure, not on the organizational setup—mostly because I don't think our investors had that experience, unfortunately." A German company described that such experience is important because it helps advisors identify where they actually have the expertise to offer valuable advice on strategy:

> Most of the mentors will advise you on the rest of the 50 percent, where they don't have the expertise...At least in Europe...Possibly [the] US is different. You have a high number of experienced founders in the market. They know exactly where they are good, [and] where they are not so good...That's where we see a more mature market in the US.

That example shows how experience in scaling—as, for example, seen in the US—equips advisors with the ability to screen good from bad advice based on what worked in the past. These advisors can therefore offer wisdom to entrepreneurs that can improve strategy development.

Conversely, the companies that could access US advisors and their direct experience used the resulting knowledge to improve their strategy. For example, an Indian startup discussed the influence of US investors on its organizational structure:

> One of my big advisors is...the head of product of [a successful US startup], previously, [another successful US startup]. And [a third successful US startup] is one of my investors. So we follow what these companies have historically set up, which is very par-driven structures focused on customer objectives.

Specifically, the Indian startup's US advisors with experience in successfully scaled companies inspired it to pursue a more customer-oriented organizational design. This structure aligned well with its customer-oriented objective to be "the best partner to merchants [in Southeast Asia] in their ever-changing world."

Similarly, an Indonesian company noted adopting a customer-centric culture based on the US experiences of its co-founders:

> My co-founder and I came from [a successful US startup], another co-founder came from [a US company], and the fourth one came from [an international e-commerce company]. All organizations, all four of us, are very customer-centric organizations. So from our cultural standpoint, it all starts with the customer.

The Indonesian company's customer-oriented culture—shaped by the co-founders' direct experiences in successfully scaled US companies—was well aligned to its user-oriented objective of "creating a new form of e-commerce [in] the smartphone era" targeted at "mom-and-pop stores." These entrepreneurs' experiences helped improve their strategy.

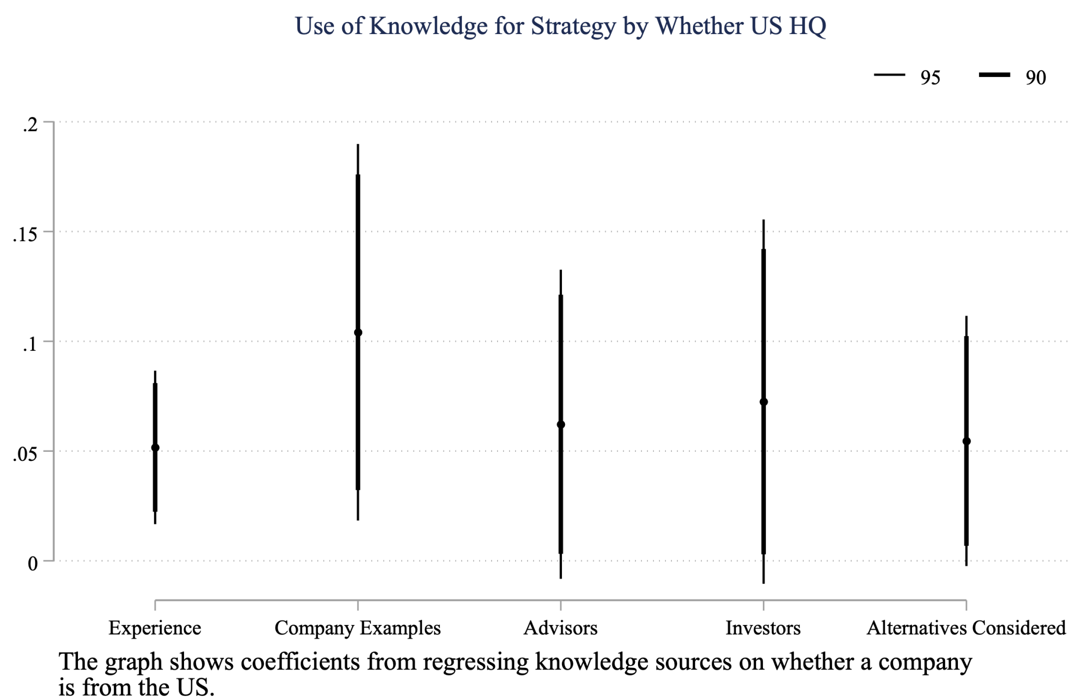
Additional quantitative analyses of the interview data corroborate these qualitative insights that the higher cost of mistakes in non-US contexts makes it harder to access scaling knowledge that can inform strategy. If access to knowledge varied across US and non-US contexts, contributing to differences in their strategy scores, we would expect US executives to use more knowledge inputs than non-US ones to develop this strategy. To measure this variance, Equation (4) switches out the strategy-dependent variable in Equation (3) with several knowledge inputs:

$$knowledge_{ij} = \beta_1 us_i + foundedyear_i + industry_i + evaluator_j + readability_{ij} + \gamma_{ij} + \varepsilon_{ij} \quad (4)$$

The dependent variable $knowledge_{ij}$ is an array of inputs that inform strategy, including (a) direct experience, (b) advisors, (c) investors, and (d) company examples. These sources of knowledge may expose founders to alternative approaches. Consequently, considering such alternatives is also a dependent variable in this specification. The coefficient of interest is β_1, which indicates whether US executives are more likely to use such knowledge inputs to inform their strategy.

Table 1.6 and the coefficient plot in Figure 1.5 show the results from this regression model. US firms are indeed more likely to use direct experience (Column 1), company examples (Column 2), advisors (Column 3), and investors (Column 4)—by 0.1 standard deviation—to inform their strategy. They also are more likely to consider alternatives (Column 5).

Figure 1.5: Non-US companies are less likely to draw on direct experience, company examples, advisors, and alternatives to inform their strategy.



Together, these results reveal that the high cost of mistakes that non-US startups face makes it more costly

Table 1.6: Non-US firms are less likely to use direct experience, alternatives, company examples, advisors, and investors to inform their strategy.

	(1) Uses Direct Experience	(2) Uses Other Company Examples	(3) Listens to Advisors	(4) Listens to Investors	(5) Considers Alternative Options
US HQ	0.052**	0.104*	0.062+	0.073+	0.055+
	(0.018)	(0.044)	(0.036)	(0.042)	(0.029)
_cons	0.545***	1.133***	1.217***	1.112***	0.916***
	(0.086)	(0.190)	(0.201)	(0.213)	(0.131)
N	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes

The table shows how being headquartered in the US associates with using different inputs to inform strategy. These inputs result from coding executives' interview responses. The models control for the English readability of the text, firms' founding year and industry, evaluator fixed effects, and whether any missing responses were filled in. The sample size is 304 because about a fifth (51) of the interviews received two evaluators to ensure the robustness of the scoring. Robust standard errors (in parentheses) are clustered at the company level.+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

to experiment and scale, look to other companies that have done so, and find advisors and investors who have worked with such companies. As discussed in the conceptual framework, scaling experience helps companies develop a strategy. Without the ability to acquire this experience directly—or indirectly through company examples, investors, and advisors—companies have a harder time developing a strategy, even if it could help them avoid future mistakes, as discussed in the previous section. The cost of mistakes, therefore, is one explanation that can shed light on why non-US startups are less likely to develop a strategy, though doing so is more predictive of their performance.

Generalizing beyond US versus non-US contexts

Do these results generalize beyond US and non-US contexts? The qualitative evidence suggests that in non-US contexts, a combination of financial, talent, and customer resources, along with cultural factors, make mistakes more costly. Strategy that can help anticipate these mistakes is, therefore, more important, though it is harder to develop without knowledge drawn from prior mistakes. These constraints may vary within non-US contexts to enable a more granular analysis.

In fact, a more granular analysis measuring the cost of mistakes across countries triangulates the main findings. For example, if mistakes were more costly in Japan than in Israel because of cultural and resource differences, we would expect strategy to predict performance more for Japanese companies than for Israeli ones. To measure this variance, Table A.7.1 uses additional data from the World Economic Forum to construct an "ease-of-recovery" country index.

The "ease-of-recovery" index is an aggregate of variables measuring financial, talent, customer, and cultural factors—inductively drawn from the interview data—that make mistakes more costly. Specifically, countries' access to venture capital availability proxies financial constraints. Countries' local supply of skilled employees, ability to attract talent, and flexibility to set wages reflect talent constraints. The domestic market size reflects customer constraints. Countries' openness to entrepreneurial risk reflects cultural orientations toward failure. Figure A.7.1 shows the index values across countries.

The results hold when using this continuous "ease-of-recovery" index instead of the US versus non-US binary measure in Equation 2 across the same dependent variables as in Table 1.3. Strategy predicts performance more in contexts where it is harder to recover from mistakes (Table A.7.1). Further, consistent with the idea that mistakes help startups build knowledge to develop their strategy, Table A.7.2 shows that strategy scores are lower in contexts where mistakes are more costly.

Together, these results suggest that in institutional contexts where mistakes are more costly, strategy matters more, but is harder to develop. These costs come from financial, talent, customer, and cultural constraints. Strategy may help companies avoid mistakes as they scale. Yet mistakes offer entrepreneurs knowledge to develop their strategy. Thus, where mistakes are more costly, entrepreneurs are less able to afford this avenue of accumulating knowledge, making it more difficult to develop a strategy.

1.5 Conclusion

This paper shows that strategy and its relationship with performance vary across firms and geographies. Strategy predicts the performance of firms outside the US, where mistakes are more costly, but not of US firms. Yet, despite the sensitivity of their performance to strategy, non-US firms have lower strategy scores. This strategy gap seems to be shaped by access to knowledge, which often accumulates from prior mistakes: non-US firms are less likely to rely on investors, advisors, company examples, and direct experience to inform their strategy. Together, these results suggest that the same institutional contexts where mistakes are more costly—making strategy more important—also make it harder to get the knowledge to develop this strategy in the first place.

These results reveal how institutional context helps determine when strategy matters in entrepreneurship (e.g., Mintzberg and Waters, 1985; Ott et al., 2017). Specifically, this research shows how strategy is especially valuable in geographic contexts where it is more costly to make mistakes. By enabling companies to achieve alignment, strategy may help startups avoid such costly mistakes.

The findings also show how not only specific strategic choices but also their alignment with one another vary across entrepreneurial firms. Research suggests that firms choose different market entry approaches (Alvarez-Garrido and Guler, 2018; Bingham and Eisenhardt, 2011), ways to position relative to competitors (Guzman and Li, 2022), organizational designs (DeSantola et al., 2022; Lee, 2022), and organizational cultures (DeSantola and Gulati, 2017). This study shows that how these choices fit together also varies across firms, with implications for performance.

Furthermore, the results suggest that knowledge of how to build a strategy may compensate for difficult venture capital conditions (Chen et al., 2010; Sorenson and Stuart, 2001), talent conditions (Fallick et al., 2006; Glaeser et al., 2015; Kerr, 2018; Kerr and Robert-Nicoid, 2020; Tambe and Hitt, 2014), institutional conditions (Delacroix and Carroll, 1983; Khanna et al., 2005), and industry conditions (Delgado et al., 2016; Eisenhardt and Schoonhoven, 1990). As a result, sending additional material resources to countries may not be the only way to stimulate entrepreneurial growth. Support for accessing knowledge to form a strategy may also be a crucial intervention.

The study also reveals the intertwined nature of knowledge and resources in entrepreneurial contexts. It shows that the knowledge to inform strategy comes from the same investors, advisors, and direct experience that provide resources to grow in the first place. This endogeneity in mechanisms may emerge because young ventures' responsiveness to external signals (Ozcan and Eisenhardt, 2009) and relative lack of organizational inertia (Gavetti and Rivkin, 2007) make the advice they get particularly influential in their strategic decisions. Ventures in geographies that lack a rich history of entrepreneurial successes and resources might be able to compensate with a strategy but may struggle to get the knowledge—often a byproduct of those same resources—to form that strategy. Thus, this research suggests that firms need to be not only resource mobilizers (Clough et al., 2019), but also knowledge mobilizers to grow.

Lastly, this research contributes to understanding the "digital divide" around the world. While digitization has helped reduce knowledge barriers among businesses globally in some ways—for example, in terms of market or coding practices (Brynjolfsson et al., 2019; Dushnitsky and Stroube, 2021; Wright et al., 2020; 2023)—knowledge frictions remain. This research suggests that firsthand experience is a source of knowledge that is difficult to codify on online platforms or other digital sources. Such experience helps entrepreneurs understand the alignment between their choices. Ultimately, these knowledge frictions may inhibit the benefits of digitization from spreading equally across firms internationally.

2

When the Journey—And Not Just the Destination—Matters: How Internationalization Shapes Entrepreneurial Experimentation

2.1 Introduction

Entrepreneurs are more exposed than ever to international markets, connected by the click of a button to potential customers, suppliers, partners and peer companies, investors, and employees across the globe (Kerr, 2016; Lu and Beamish, 2001).[1] Given the rapid advancements in digital and physical technology, as well as increasingly accepted norms for global outsourcing and internationalization, entrepreneurs can more readily execute their

[1] Co-authored with Laura Huang

strategic plans and business models. They more easily reach new customers in international markets (Knight, 1996; Knight and Cavusgil, 2004; 2005; Wormald et al., 2021). They have unique access to international financial capital and advice to fuel their growth (Alvarez-Garrido and Guler, 2018; Balachandran and Hernandez, 2020; Lin and Viswanathan, 2016; Sørensen and Sorenson, 2003; Stuart and Sorenson, 2003), and they are afforded the opportunity to hire talent (Balachandran and Hernandez, 2018; Cullen and Farronato, 2021; Kulchina and Hernandez, 2016) and gain access to new technologies (Alcácer and Chung, 2007; Chung and Alcácer, 2002) from across the world, to name a few. Put simply, increased access to international markets allows entrepreneurs to innovate and build their businesses by scaling their existing resource bases and offerings and, in turn, boosting performance to ultimately create and capture more value.

Yet, despite all the benefits that internationalization affords to entrepreneurs as they are building their ventures, internationalization may also play an alternate role in the life cycle of an entrepreneurial venture. That is, despite internationalization being seen as the result of the strategic progress of a business, as prior literature often suggests (Knight, 1996; Knight and Cavusgil, 2004; 2005; Symeonidou et al., 2017; Wormald et al., 2021), it may also be a critical part of its initial trajectory. Even in a venture's earliest days, where entrepreneurs are still trying to establish product-market fit (Bingham and Eisenhardt, 2011) or gain an initial foothold with their product or service, internationalization is helpful. Rather than merely serving as an expansion mechanism (allowing entrepreneurs to expand their existing opportunities to new markets), exposure to international markets may serve also to shape the processes through which entrepreneurs later identify and subsequently exploit opportunities. Internationalization may be importantly influencing how entrepreneurs think through the value they can create and capture—especially as they are adapting to particular new markets, for example, through trial and error (Bingham, 2009; Bingham and Davis, 2012; Bingham and Eisenhardt, 2011; Gupta and Khanna, 2019; Shaheer and Li, 2020). In this way, we argue that internationalization exposes them to more options, more ideas, and more frameworks and schemas (McDonald and Eisenhardt, 2020), which helps founders redefine the opportunities they observe in valuable ways. Indeed, entrepreneurs update their strategy whenever confronted with new information—such as the type that each international market may provide—in a way that validates or conflicts with their prior beliefs (Kirtley and O'Mahony, 2020). The local knowledge gained from each market may ultimately influence how entrepreneurs perceive their opportunities, as well as provide important information on where to focus their attention and how to think about their ultimate strategy (Gavetti et al., 2005; Gavetti and Rivkin, 2007; Ocasio and Joseph, 2018; Wang, 2015).

Therefore, in our current investigation, we examine if, and how internationalization influences the process through which entrepreneurs realize the value they can create and capture and, ultimately, how they define their businesses. We conduct an inductive field study of 84 entrepreneurs across 27 countries to explore how, rather than being merely a post-hoc strategy or an objective to be implemented, internationalization helps entrepreneurs (unintentionally) decide what to implement in the first place. We find that internationalization shapes early-stage entrepreneurs' experimentation efforts and helps them define their business value proposition and strategy. Specifically, internationalization shifts entrepreneurs' perceived choice sets by making salient the similarities and differences across cross-border markets. In doing so, this process helps entrepreneurs avoid hyper-focusing prematurely at the expense of more optimal opportunities. It also helps prevent over-diversifying before successfully tapping into any of their existing opportunities. We contribute a process model that demonstrates how entrepreneurs do so, showing how international markets help entrepreneurs define, scope, and externally validate ideas that can create and capture value—specifically through mechanisms that allow them to both expand and focus the scope of opportunities they consider in their experimentation (and ultimately in their developed strategy). Our focus is specifically on entrepreneurs in early-stage ventures, rather than more mature firms, because of the high uncertainty under which young companies operate (Eisenhardt and Bingham, 2017; Kerr et al., 2014b) and the importance they place on experimentation. These early-stage ventures also lack organizational and network inertia, which makes them especially responsive to the external environments they are exposed to (McDonald and Gao, 2019; Zhang et al., 2016).

Our findings contribute to scholarly work in numerous areas, including entrepreneurial experimentation, internationalization, and early-stage venture strategy. First, we contribute to theories on the distinct relationship between internationalization and experimentation and the various ways that directionality and trajectory can influence entrepreneurs and their ventures. While prior research shows how entrepreneurs experiment to subsequently enter particular international markets (Bingham 2009; Bingham and Davis, 2012; Bingham and Eisenhardt, 2011; Bruneel et al., 2010; Gupta and Khanna, 2019; Wang, 2020), we show how international exposure can also be a precursor to firms' core experimentation efforts and opportunity identification. As a result, internationalization may help entrepreneurs overcome uncertainty rather than only creating uncertainty through regulations, cultural differences, language, and other cross-country factors (Berry et al., 2010; Ghemawat, 2001).

Second, we contribute the idea that internationalization and geography can help entrepreneurs both search for, and optimize, opportunities. While research has shown that markets are part of the opportunity itself (Bingham et al., 2007; Gans et al., 2019; Sørensen and Sorenson, 2003), we reveal that there are benefits to having them be inputs in the decision-making process and choosing opportunities. Our findings validate that context is important to consider when assessing entrepreneurial decisions, and specifically demonstrate how entrepreneurs' perceived choice sets are shaped by, and endogenous to, their environment.

Finally, we show how geography plays an important role in the strategy and shaping of innovation. Gaining exposure to international markets influences the nature of the new ideas that entrepreneurs pursue, specifically which markets benefit from their innovations, similar to what has been seen in the gender context (Cao et al., 2021; Koning et al., 2020; 2021). We contribute the idea that internationalization affects both the industry and geographic markets that entrepreneurs target with their innovations. These industry and geographic lenses are dimensions by which we may observe shifts in the direction of innovation that profoundly impact who benefits from new breakthroughs.

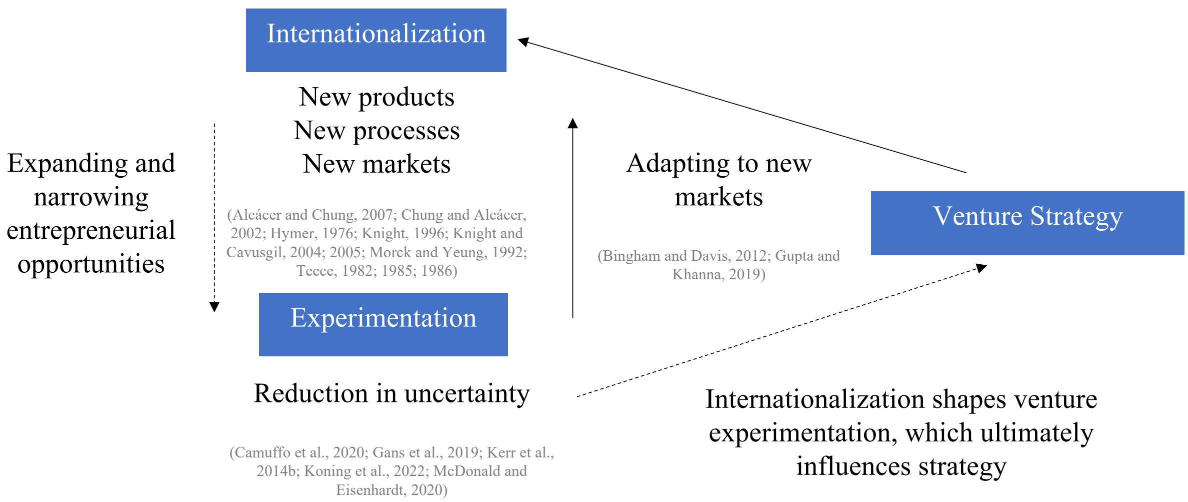
2.2 Experimentation and internationalization

Entrepreneurs engage in experimentation due to the innate uncertainty involved in early-stage venturing (Bingham, 2009; Bingham and Davis, 2012; Bingham and Eisenhardt, 2011; Gupta and Khanna, 2019). At the earliest stages, there is often little information on which markets to enter and how to enter them. As a result, rather than making a large commitment to entering a new market right away, experimentation helps entrepreneurs reduce uncertainty about both their execution abilities as well as the environmental conditions in a new market (Bhidé, 2000; Gans et al., 2019; Kerr et al., 2014b). In this way, entrepreneurs provide a partial up-front investment into experiments—whether it be scientific hypothesis testing (Camuffo et. al., 2020), product tests (Brown and Eisenhardt, 1997; Pisano, 1994), "parallel play" (McDonald and Eisenhardt, 2020), or exploratory alliances (Cui et al., 2018; Santos and Eisenhardt, 2009)—to gain information on the likelihood of success of different ideas in a market. This information helps entrepreneurs proceed in a more knowledgeable and confident manner (Camuffo et. al., 2020; Gans et al., 2019; Kerr et al., 2014b) and ultimately see boosts in performance (Koning et al., 2022; McDonald and Eisenhardt, 2020).

Yet, while a multitude of studies reveal how experimentation helps ventures address uncertainty in entering new markets, the new information that entrepreneurs observe when entering international markets may shape their subsequent experimentation. Indeed, young ventures may be highly responsive to signals in their external environment to inform future experiments given the high level of uncertainty that they face (Eisenhardt and

Bingham, 2017; Kerr et al., 2014b) and their malleable organizational structures, as seen in ventures' adaptation in nascent industries (McDonald and Gao, 2019) and amid institutional change (Zhang et al., 2016). We, therefore, examine how internationalization—a salient dimension of the external environment—may influence experimentation. Figure 2.1 summarizes how our analysis fits with prior literature.

Figure 2.1: Theoretical overview



2.2.1 Exposure to international markets builds knowledge

To unpack how internationalization may shape experimentation, we begin by examining the knowledge sources that scholarship has shown shape entrepreneurial experimentation and strategy development. Entrepreneurs learn from their direct experiences. For example, prior work experiences shape what strategic paths (Gavetti and Rivkin, 2007) or innovations (Wang, 2015) entrepreneurs perceive and ultimately pursue. The experiences of investors and advisors, who themselves have gained knowledge and expertise from working with other companies, also inform entrepreneurs' experimentation (Chatterji et al., 2019; Gavetti and Rivkin, 2007; Vissa and Chacar, 2009). Case studies, too, are an input, allowing entrepreneurs to observe and assess how other companies pursue experimentation (Gavetti et al., 2005).

We extend this literature to show how entrepreneurs can access these sources of knowledge through international exposure. Whether it be through their own experiences or those of their advisors, investors, or peers, entrepreneurs gain knowledge that seemingly bounds the nature of experiments they subsequently pursue. Because their personal experiences or those of their advisors and investors are often locally embedded, these learnings capture the cultural particularities of a region. In a new market, when entrepreneurs gain access to these locally-embedded knowledge sources, they can more readily observe what is similar and different about the environment relative to what they observed from similar knowledge sources in prior markets. These comparisons may resonate when exposed to markets across borders because the cultural, institutional, and related differences are particularly large relative to those within-country jurisdictions (Berry et al., 2010; Ghemawat, 2001).

These comparisons help entrepreneurs understand the underlying logic behind phenomena in their environment, which they might have taken for granted in any single market. For example, observing that customers in one high-income market have a taste for the most cutting-edge features, but in another high-income market, seek a few simple features would reveal to the entrepreneur that income alone does not predict taste for innovation or complexity. Thus, from a pure monetary perspective, investing in more features might not be the optimal opportunity in the high-income total addressable market. Rather, the opportunities may lie in going deeper into a few of the features. Observing any one of these markets in isolation might obscure this "null" relationship between income and features and lead the entrepreneur astray. In the first market, the entrepreneur may form a biased belief that income predicts such innovative taste, leading the entrepreneur to sink more time and money into developing more features. In the second market, the entrepreneur might perceive the opposite relationship and merely remove features without going deeper into any particular ones. Thus, the differences between the markets give the entrepreneur a more balanced view of the opportunity set.

Entrepreneurs in young organizations—as opposed to those in mature organizations—may also be particularly likely to build this knowledge. The high uncertainty that young companies face (Eisenhardt and Bingham, 2017; Kerr et al., 2014b) and lack of organizational inertia make them especially responsive to signals in their external environment and, therefore, more likely to observe differences in international markets that shape their knowledge (McDonald and Gao, 2019; Zhang et al., 2016).

2.2.2 International markets shape experimentation

By virtue of helping entrepreneurs build knowledge by revealing the differences and similarities relative to entrepreneurs' priors, exposure to international markets may shift their experimentation efforts. They do so by shifting the set of opportunities that they perceive in the market and, therefore, what is possible to test. For example, an entrepreneur that targets a particular country market for its commercial potential may observe unanticipated differences or similarities in that market relative to her priors from previous markets. This comparative observation may enable her to realize new opportunities, as well as rule out previously developed opportunities, ultimately shaping her opportunity choice set that allows her to identify new gaps in that market (Shane, 2000). The realization of these gaps may then trigger her to devise new solutions to target in experiments and strategy.

Because both geography and experimentation occur dynamically, they may shift entrepreneurs' ultimate strategy substantially from its starting place. Each change creates ripple effects for subsequent stages. In this way, entrepreneurs iteratively adapt not only their market entry (Bingham, 2009; Bingham and Davis, 2012) and management (Gupta and Khanna, 2019) approaches in particular international markets, but also their subsequent experimentation to uncover their core value that eventually translates into their overall strategy.

2.2.3 Boundary conditions for internationalization

Internationalization, rather than geographic expansion more broadly, may particularly shape entrepreneurial experimentation because cultural, institutional, and other differences between countries may be larger than within countries (Berry et al., 2010; Ghemawat, 2001), providing more room for learning to guide experimentation. That being said, our theory may generalize to within-country geographies when there are sufficient differences between them.

We focus on entrepreneurs in young ventures rather than in mature firms because the high uncertainty that young companies face (Eisenhardt and Bingham, 2017; Kerr et al. 2014b) and organizational inertia make them particularly responsive to external signals (McDonald and Gao, 2019; Zhang et al., 2016). However, there may be cases in which internationalization may shape the core experimentation and strategy development of more mature firms, particularly if they are operating in an environment with high uncertainty and have malleable organizational structures.

Taken together, we argue that internationalization—as the context of entrepreneurial opportunity and knowledge sources— shapes which ideas entrepreneurs choose to test and ultimately commit to in their strategy. Therefore, internationalization may be an input into entrepreneurial experimentation and strategy development, in addition to being an outcome of it. Internationalization exposes ventures to new entrepreneurial opportunities that are tested in experimentation processes, which reduce uncertainty around these opportunities to ultimately inform venture strategy.

2.3 Methodology

The prior literature we discussed earlier provides us with the theoretical handholds from which to study the role of internationalization in entrepreneurial experimentation and strategy (see Figure 2.1 for a summary). Our field study subsequently relied upon field methods to understand how internationalization influences the entrepreneurial process (Charmaz, 2014; Eisenhardt, 1989). Field methods are particularly useful for understanding this experimentation process because, conceptually and quantitatively, it is difficult to track the role of internationalization in the entrepreneurial process, as well as the motivations for these geographic decisions.

Our sample consisted of 84 entrepreneurs. We primarily relied on interviews for this analysis, supplemented with third-party databases and survey data. A summary of our methods process, adapted from Huang (2018), is shown in Figure 2.2. Our main unit of analysis was the technology entrepreneur in each phase of the entrepreneurial process—from conceiving the idea to growing the business. We defined a phase as a milestone reached in the construction and growth of a venture. We employed a phase as a sub-unit because it is a common way for entrepreneurial training programs sponsored by universities, accelerators, investors, and governments to guide entrepreneurs in the development of their businesses. For example, the Duke Entrepreneurship Manual defines the entrepreneurial process as phases that "organize the effort of planning, launching, and building a venture" (Duke, 2021). These phases consist of "idea generation, opportunity evaluation, planning company formation/launch, and growth" (Duke 2021). MIT (Bill Aulet's) Disciplined Entrepreneurship framework includes 24 steps from "understanding your customer" to "scaling your business" (Aulet, 2013). Therefore, phases were a natural way that entrepreneurs categorized their experiences in semi-structured interviews.

Research Context. The high-technology market across international entrepreneurial hubs was the main context for the study. The study specifically sampled entrepreneurs in earlier stages (seed to series A) from entrepreneurial hubs around the world (Table 2.1). The global reach of the study, covering Europe, Latin America, Asia, and North America, enabled us to build more generalizable theory (Yin, 2003). We addressed potential concerns regarding context variation by assessing areas of convergence among these geographies and narrowing our focus to technology ventures in a similar (early) stage of their scaling process within these settings. We focused on the high-technology market because companies in this market are more exposed to international markets than other sectors, given their lack of heavy capital requirements and native adoption of digital tools that enable reaching cross-border customers almost instantaneously.

Figure 2.2: Methods overview

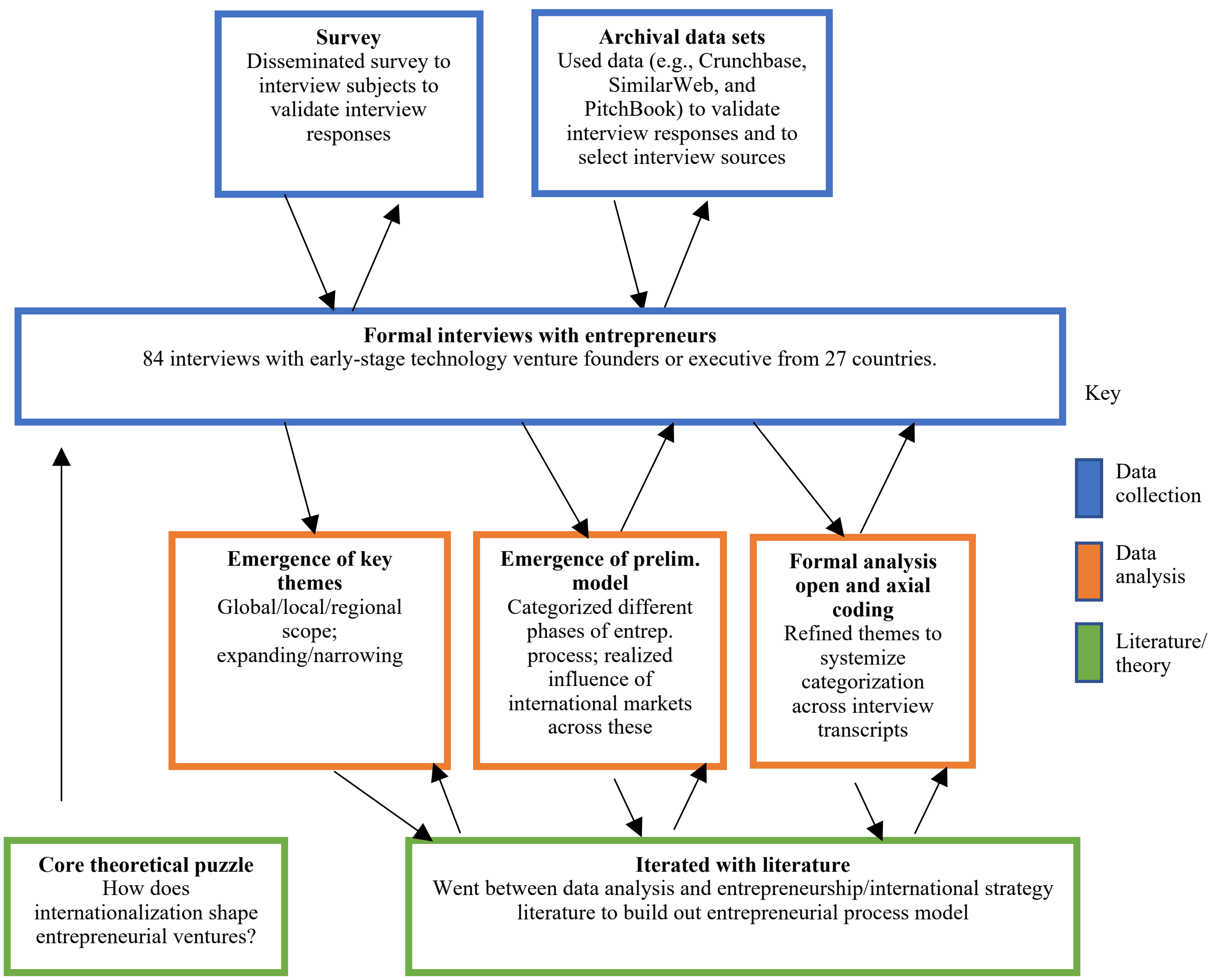


Table 2.1: Interview sample breakdown

HQ Country of Entrepreneurs	Number of Interviewed Entrepreneurs
Argentina	2
Australia	2
Brazil	3
Canada	2
Chile	1
Colombia	1
Denmark	1
Estonia	1
Iceland	1
India	7
Indonesia	2
Ireland	2
Israel	4
Japan	5
Kenya	2
Korea	1
Netherlands	5
Nigeria	1
Pakistan	2
Poland	16
Saudi Arabia	1
Singapore	6
Spain	1
Switzerland	4
UK	3
Ukraine	2
US	6

Data Sources. The main sources of data for this field study were semi-structured interviews conducted with entrepreneurs from 84 technology ventures in the United States, Europe, Latin America, Asia, and Africa. The interviews included questions on the origins of the ventures, the markets in which the ventures sought to reach and actually reached (and why they did so), how they reached those markets, the types of support they received (whether from accelerators, investors, government, etc.), their orientation toward growth, opportunities, and challenges faced in the home entrepreneurial ecosystems, and perceived advantage, vis-à-vis competitors. The interviews lasted generally 30–60 minutes and were done in the office of the venture, in a neutral public location, or virtually (post-COVID). We recruited through Crunchbase and PitchBook, accelerator websites, social media, and emails (where available on company websites), or our existing contacts introduced us to them. We conducted our interviews in phases, where the first 40 interviews helped us realize and develop a process model about how internationalization shapes entrepreneurial experimentation, and the latter 44 interviews allowed us to test and validate this model, as well as reach theoretical saturation.

We corroborated the information we received from the interviews using startup databases, such as Crunchbase and PitchBook to confirm investment information and SimilarWeb to confirm the markets of the ventures through page view data. We further turned to accelerator websites, the startups' own web pages, and a follow-up survey of the interviewed ventures to confirm data on the companies in our sample. The follow-up survey included questions on the sequence of market entry and changes pursued by ventures after entering markets. Through this survey, we were able to capture the exact history of market entry for companies.

Sample. The technology entrepreneur was the main unit of analysis in the field study. Because we were interested in uncovering how such individuals approach internationalization, a theoretical sample (Glaser and Strauss, 1967), which describes entrepreneurs exposed to international markets, was appropriate. We, therefore, targeted entrepreneurs in ventures with at least a minimal viable product, initial clients from international entrepreneurial hubs, and/or other international growth indicators, such as participating in an accelerator program. Consequently, our sample focused on entrepreneurs in ventures that were in the process of international scaling or beginning that process rather than those that were solely locally oriented.

Data Coding. We conducted a combination of iterative coding and memo-writing to analyze the interview data and build a theory of how internationalization influences entrepreneurs (Charmaz, 2014). While our original motivation for field work was open-ended on the internationalization of entrepreneurs, in the spirit of the grounded theory approach, we began open coding around three stages of the entrepreneurial process as we iterated with the literature. We then characterized the relevant geography of the stage—whether global, local, or regional. As we characterized each stage, we noticed that the geography either narrowed or expanded the focus of the entrepreneur, and we proceeded to code these channels. Table 2.2 illustrates the coding structure.

The first stage centered on where entrepreneurs got the ideas for their ventures. We coded the problem ventures were seeking to solve, the geographic location of this problem, and how that geography influenced the ventures' scope of focus. For example, Arg-12 saw that merchants in Latin America faced a problem with getting payments. We coded this problem as a regional problem. When a problem was local to the industry, but not the geography, of the entrepreneur, we coded this as a global problem. For example, Nether-16 drew his idea from a problem that his co-founder observed in the financial industry, where he previously worked. We then characterized whether the geography expanded or narrowed the problem set of entrepreneurs. In the case of Nether-16, we saw that the global geography broadened the choice set of problems that the venture considered to include those of the financial industry, so we coded the mechanism as "expanding."

We also coded the solution that the entrepreneurs proposed and its geographic inspiration. For example, Jap-32 proposed a digital payments solution to address banking challenges in Japan that drew from existing solutions in other markets: "[A US mobile payments provider] hadn't gotten to Japan yet...Not even outside of the US we believe." We coded this solution as a global solution. Because the US mobile provider offered a new idea for a business model to the venture, we coded the global geography as "expanding" the solution set of the entrepreneur.

The second stage focused on where entrepreneurs originally built their minimal viable product in iteration with user and customer feedback. We coded the geographic location of this stage. For example, Nether-15 noted that he was focusing on the Dutch market in order to achieve product-market fit and get feedback. We coded this process as occurring in the local market. The Dutch market enabled the entrepreneur to focus on the core offerings required in his venture's minimal viable product, so we coded geography as "narrowing" the choice set of the entrepreneur.

The third and last stage, while not common in entrepreneurial studies or the experimentation process, was one that originated from our interview data—how entrepreneurs tested beyond their initial early users who gave them feedback in the scoping phase to ultimately grow. We noted this as a process in which entrepreneurs tested the external validity or scope of their products. We coded this phase as global when entrepreneurs tested in country markets across regions, regional when entrepreneurs tested in the country markets within their region, and local when entrepreneurs tested other customers within their country market. For example, we coded Jap-33, who used an online platform (Kickstarter) to launch to customers from Japan, the US, and other parts of the world, as global. The testing made the entrepreneur realize his venture's core value did not resonate in the US and, instead, pushed the entrepreneur to double down on the Japanese market and others with similar regulatory structures that were conducive to such environmental solutions. Because geography focused the entrepreneur, we coded this stage as "narrowing" the choice set of the entrepreneur.

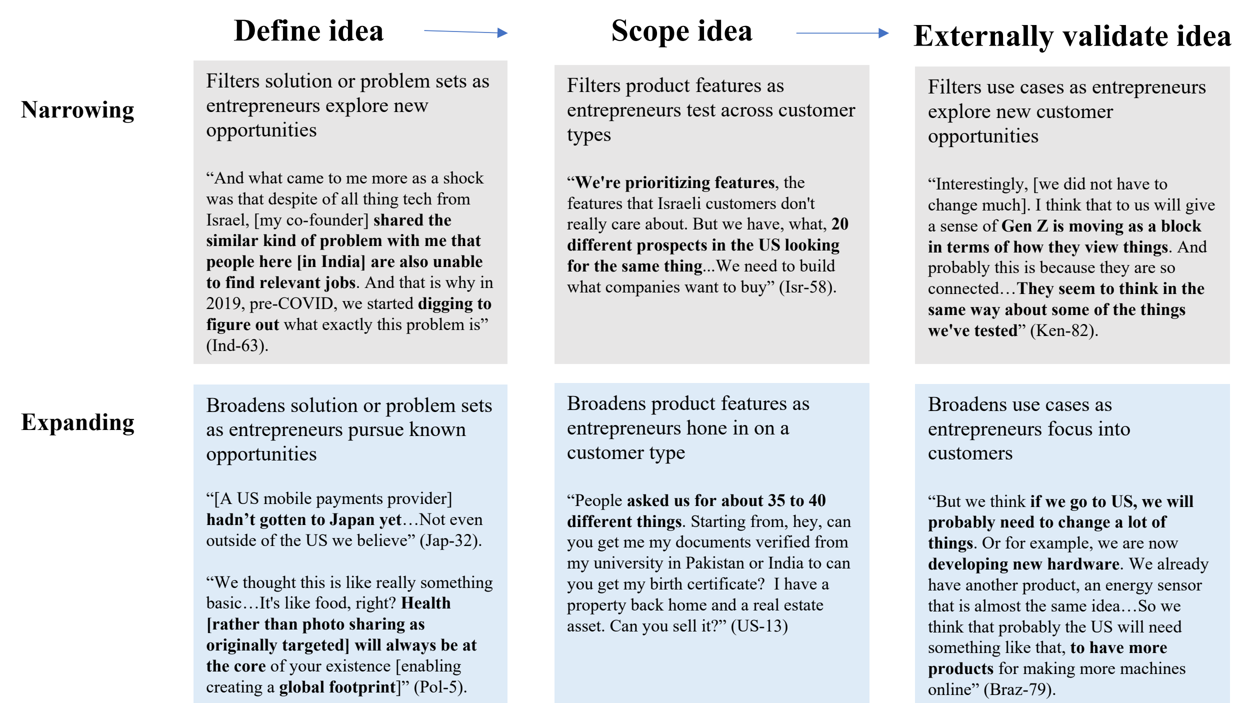
Table 2.2: Interview coding example

Startup ID	**Stage 1 - Problem**		**Stage 1 - Solution**		**Stage 2 - Build**		**Stage 3 - External Validation**	
Swiss-1	Global	Expanding	Global	Narrowing	Local	Narrowing	Regional	Expanding
Ukr-2	Global	Expanding	Global	Narrowing	Local	Narrowing	Global	Narrowing
Pol-4	Local	Expanding	Global	Expanding	Local	Narrowing	Local	Narrowing
Arg-12	Regional	Narrowing	Global	Narrowing	Local	Expanding	Regional	Expanding /Narrowing
US-13	Local	Expanding	Global	Narrowing	Global	Expanding	Global	Narrowing
Nether-16	Global	Expanding	Global	Narrowing	Local	Narrowing	Global	Expanding
Est-26	Local	Expanding	Global	Narrowing	Global	Expanding	Global	Expanding
Jap-32	Local	Narrowing	Global	Expanding	Local	Narrowing	Local	Expanding
Sing-35	Global	Expanding	Global	Expanding	Global	Narrowing	Global	Expanding /Narrowing
Kor-40	Global	Expanding	Global	Expanding	Local	Narrowing	Regional	Narrowing
Irel-41	Global	Narrowing	Global	Expanding	Global	Narrowing	Global	Expanding
Pak-42	Local	Narrowing	Global	Expanding	Local	Narrowing	Global	Expanding
Saud-45	Local	Narrowing	Global	Expanding	Global	Narrowing	Local	Narrowing
Isr-47	Local	Narrowing	Local	Expanding	Local	Expanding	Global	Narrowing
Ind-48	Global	Expanding	Local	Expanding	Global	Narrowing	Global	Expanding
Indo-49	Local	Narrowing	Global	Expanding	Local	Expanding	Global	Narrowing
Col-50	Local	Narrowing	Global	Expanding	Local	Narrowing	Regional	Expanding
Braz-51	Global	Expanding	Global	Narrowing	Local	Narrowing	Global	Expanding
UK-55	Global	Narrowing	Global	Narrowing	Local	Both	Global	Narrowing
Nig-59	Local	Expanding	Global	Expanding	Local	Expanding	Global	Expanding
Den-60	Local	Narrowing	Global	Expanding	Global	Narrowing	Global	Expanding
Spa-64	Local	Narrowing	Global	Expanding	Global	Expanding	Global	Expanding
Ken-65	Local	Expanding	Global	Narrowing	Global	Expanding	Global	Narrowing
Ice-73	Global	Narrowing	Global	Expanding	Local	Expanding	Global	Narrowing
Chil-74	Local	Narrowing	Global	Expanding	Local	Expanding	Global	Expanding
Can-76	Local	Narrowing	Global	Expanding	Local	Expanding	Global	Narrowing
Aus-80	Local	Narrowing	Global	Expanding	Global	Expanding	Global	Expanding

2.4 Findings

In the following sections, we describe how internationalization shaped entrepreneurial experimentation and opportunity identification. We begin by discussing how entrepreneurs initially perceived internationalization as a result of their strategic outcome but then came to realize it could inform and shift their original strategy. We then describe how this internationalization process and the exposure to new geographies allowed them to expand or narrow as they continued to experiment and define, scope, and externally validate their ideas. Figure 2.3 depicts our overall process model.

Figure 2.3: Process model illustration



2.4.1 Internationalization as an output and input of entrepreneurial strategy

Having a global presence and seeking international expansion were often seen to be synonymous with growth and success. One entrepreneur (Pol-5) who had created a medical booking platform shared that he "strived to achieve a billion-dollar unicorn with global market share." He noted: "Everybody says global company...billion-dollar unicorn. But to us, [being a global billion-dollar unicorn] is important. It's like this internal voice saying...we want to have a footprint... [a global footprint]." Similarly, another entrepreneur (Irel-41) shared how his team was "pretty ambitious and obsessed with being a global business...It isn't particularly interesting for us to be an Irish business in the domestic context." This entrepreneur continuously equated ambition and growth with international expansion and global presence. Entering into new international markets, whether it be in Latin America or Europe, was all part of their plans, and having a global "footprint" signified success.

Other entrepreneurs pursued narrower geographies rather than "going global," but did so as a way to achieve internationalization nonetheless. Importantly, they similarly saw these newer international expansion opportunities as a strategy for growth. For example, Arg-12, who started an IT credit solution business, noted that his venture "could be a 1-billion-dollar company only doing what do we need to do in Latin America," tying the venture's local geographic footprint in its region to the "billion dollar" growth objective of the company. Regardless of local geographic expansion or global geographic expansion, entrepreneurs viewed internationalization as an end for accomplishing their strategy for growth (Hymer, 1976; Morck and Yeung, 1992).

Yet, what emerged as entrepreneurs entered international markets in pursuit of distinct strategic objectives, was a sense of new insights about gaps in the markets and a new perspective on their capabilities to address those gaps. The introduction of these insights changed how they viewed the opportunities available to them and represented a new opportunity set. In turn, we discovered that entrepreneurs would subsequently and unexpectedly revise the manner in which they carried out their strategy and engaged in subsequent experimentation. For example, one entrepreneur (Pol-5) described how he sought global markets for his photo-sharing platform solution and thought his path to becoming a global unicorn was predicated on users and downloads from companies across multiple continents, including North America, Europe, and Asia. However, in the process of trying to gain users, he saw how his platform was actually also suited for visual applications in the healthcare space. By expanding through an additional industry vertical (rather than a user base for just one industry), he discovered a better core use case for his product: "We thought this is like really something basic...It's like food, right? Health will always be at the core of your existence." Making this shift is what ultimately allowed this entrepreneur to achieve the global status he envisioned—after first internationalizing and finding ways to improve upon the product and its global reach. Internationalization guided which opportunity the entrepreneur chose.

This notion of geography and internationalization as being a handhold in helping entrepreneurs navigate was shared across nearly all our respondents. As Arg-12 expanded across countries in Latin America, for example to Ecuador, he realized a new customer profile that he should be targeting:

> Our customer has changed. It's not the end user; it is the acquirer, and that's something that we are still learning…We should be charging this massive [company] in Ecuador. So the business models should be…with [that] massive target and not at the same target that we have here in Argentina. We never thought that these types of enterprises [haven't] solved what we built. Like we thought that they should have learned way before we created this technology. Now we are facing a similar learning with…the banks.

Internationalization allowed this entrepreneur to realize that he should not just target small and medium businesses but also large enterprises and banks. He overlooked these latter customer segments in Argentina because he assumed that these types of customers were too technologically advanced to need his IT solution. Moving to a less technologically advanced environment like Ecuador made other types of customers with a need for his IT solution more salient. The entrepreneur subsequently chose to revise his venture's core strategy to focus on enterprises as the main customer because this approach allowed the venture to increase both its financial and social impact returns.

These examples reveal that entrepreneurs targeted international expansion as part of their strategic objective, alongside financial and social returns. However, in doing so, entering new markets also revealed new information about their capabilities and unknowingly changed their perspective on opportunities—whether it be the industry they felt they should be tackling or the customer profile they should be targeting. These realizations led to revisions to their subsequent experimentation and strategy.

2.4.2 Expanding and narrowing mechanisms

We found that international markets shaped entrepreneurs' experimentation by expanding and narrowing their opportunity set as they defined, scoped, and externally validated their ideas. International markets—by making salient differences and similarities relative to entrepreneurs' often locally embedded priors—revealed new information about markets and entrepreneurs' capabilities. This information shifted the number of choices—both expanding and narrowing—that entrepreneurs perceived in their opportunity set with regard to business ideas, technology features for early users, and use cases for growth to mainstream markets. In doing so, it enabled entrepreneurs to both broaden and focus their experimentation efforts. While the former was particularly valuable for entrepreneurs stuck with limited existing opportunities that could get them stuck in local peaks, the latter particularly helped entrepreneurs who were over-exploring opportunities, distracting them from achieving any meaningful objective (March, 1991).

Expanding: Providing increased choice and solution subsets

We found that one way in which international markets changed the opportunity set that entrepreneurs perceived was by revealing information about market gaps and ventures' capabilities—different from entrepreneurs' priors in their previous markets—that expanded the number of choices they perceived. These choices were related to the business idea they were targeting, technology features to include in the minimal viable product for early adopters and use cases for growth in the mainstream market. They helped entrepreneurs realize new possibilities for growing their business in three key corresponding areas that are critical to the success of an early-stage venture: (a) Ideation, (b) Scoping, and (c) External Validation. In the Ideation phase, entrepreneurs evaluated various problems and solutions to address in the market. In the scoping phase, entrepreneurs iterated on their minimal viable product to find product-market fit in an early adopter market. In the external validation phase, entrepreneurs were searching for other (usually more mainstream markets) where they could scale.

Ideation. International markets revealed new problems—gaps or pain points in the market, as well as solutions—a technology, service, or business model—that could address the problems in entrepreneurs' early-stage experimentation. In doing so, they expanded the entrepreneur's choice set of ideas. For example, one entrepreneur with a collaboration technology solution (Pol-11) explained that he noticed in work with US consulting companies that "people or teams are selected...randomly who were available at some specific time...you don't check anything about how the team is set up together...it's super ineffective." This observation revealed a new problem in business team collaboration that he could solve through a technological solution that ended up being the core innovation and application of his venture.

In addition, as mentioned above, Pol-5 evolved his inherent idea—from photo sharing to a medical platform—in pursuit of building a company with a "global footprint." Being exposed to international markets revealed information on problems that the entrepreneur could solve and create value with around the world. Specifically, this information made salient differences in the problems that the entrepreneur was observing locally—photo sharing—to a medical one that customers around the world faced. The entrepreneur could apply his knowledge and past experimentation in platform technologies to address this globally oriented problem.

International markets exposed entrepreneurs to not only problems they could solve in the market, but also solutions that they could apply to pre-identified problems. Jap-32 grew his choice set of ideas to address the inefficient exchange of money in Japan through exposure to mobile payments companies in the US market. He shared that "[A US mobile payments provider] hadn't gotten to Japan yet...Not even outside of the US we believe." In doing so, he saw a new opportunity to create value in Japan by bringing such a mobile payment approach to addressing the deep currency exchange gap that plagued the country. Though the entrepreneur did not physically enter the US or European markets, he observed company examples from afar, which sparked new solution subsets in financial technology—different from those that the company observed in Japan—that could anchor his value proposition to Japanese customers. In this case, international market ideas allowed the entrepreneur to realize a new solution to a previously identified local problem. Whether it be through exposing entrepreneurs to new ideas on problems or solutions, international markets expanded their opportunity choice set that otherwise had a limited number of options that may have constrained their growth.

Scoping. Entrepreneurs in our sample turned to international markets to scope their idea for early adopters to find product-market fit. Sometimes, unintentionally, they revealed to entrepreneurs additional needs that early users or customers had, as well as their own capabilities to address those needs. Entrepreneurs, therefore, realized new ways they could create value for customers. This, in turn, helped entrepreneurs expand the choice set of features they considered in their early technology. The features that international markets exposed led to tangible changes to the product. For example, one entrepreneur (US-13) reached global immigrant audiences through social media channels to sell the value proposition of his immigrant document management platform. Through this process, the entrepreneur learned about different document problems with which immigrant early users wanted help on the platform. The founder describes:

> People asked us for about 35 to 40 different things. Starting from, hey, can you get me my documents verified from my university in Pakistan or India to can you get my birth certificate? I have a property back home and a real estate asset. Can you sell it? I have a special birthday coming up. I don't want just to send flowers. I'm going to make it special.

Learning about these needs allowed the entrepreneur to expand his choice set of features on the platform. The entrepreneur subsequently broadened the functionality of his venture's minimal viable product to address the global needs of its immigrant early users, where he was seeking to find product-market fit. Though these needs differed from what the entrepreneur had observed earlier, they were similar across new international markets. Even across different geographic settings, similar types of features were in demand.

The expanded choice set of features, however, need not be universal for all early users. Pol-11, for example, realized that some needs varied across geographies when expanding his venture's product across early users in different countries. The entrepreneur initially intended to demonstrate the network effects enabled by his platform, such that one user in one country would increase the value of the platform for another user in another country as part of his growth plan. Yet, the founder actually realized that he was dealing with customers who demanded different features across markets:

> We tested different markets, you know, so we did one, we asked one person to do the project in India, and we are checking how the network will spread because the project is a couple of people...So we ended up with...700 users now and from 70 countries, and it's bubbling very nicely, but it's not as effective to conquer the world in one heat. So we've chosen five locations and five different qualities of potential users, and we tried to build small bubbles with the zones and then connect them somehow...We have students in Warsaw. They have a super difficult time showing their projects. We have freelancers in Berlin...

As a result of this observation, the entrepreneur adapted his venture's experimentation efforts to focus on within-geography network effects, as opposed to cross-country ones. The international nature of markets revealed both the universality of demanded features as well as fragmentation in features demanded across customers in different markets, from students in Poland to freelancers in Berlin. In sum, international markets enabled entrepreneurs to update their experimentation with new technology feature options for early users to achieve initial product-market fit.

External Validation. Upon finding product-market fit, entrepreneurs turned to international markets to expand their mainstream customer base and, in doing so, learned about new use cases for growth that informed their subsequent experimentation and strategy. For example, Sing-35 entered several international markets simultaneously with his labor platform solution only to realize that a purely digital approach was not viable in the Asian market. This high-touch approach emerged as a new choice of how to differentiate in the regional market:

> We tried doing it [in Singapore, India, and Indonesia] simultaneously, but we realized it was hard...Operating a digital-only environment. It didn't really work... So that's what we realized is that if clients are open enough and their mindset is in a virtual-only setting, that's great. But most of Asia is not. They want that human touch; they want to see that face-to-face.

The entrepreneur ultimately realized that a one-size-fits-all digital solution was not going to be effective in selling his product in the region. Rather, the founder needed to reorient the venture's strategy and differentiation to a human-facing, relationship-based approach.

Entrepreneurs also uncovered new customer profiles to enable growth in the process of entering new international markets. As referenced above, Arg-12 learned about a new use case to grow demand for his product—

enterprises and financial institutions—by expanding into less economically developed parts of Latin America. As the founder explained: "Now that these technologies are not available in Ecuador, Columbia, the Dominican Republic, Paraguay, and Uruguay, there is a major opportunity to leverage this technology and get all these acquirers and enterprises and merchants integrated today to the API." In learning about this new choice in his growth opportunity set, the founder revised his strategy to focus on larger entities like acquirers and enterprises as the main customer rather than just small and medium companies. The less technologically advanced Latin American markets like Ecuador revealed information on a market gap with regard to the capabilities of enterprises and banks, as well as the capability of the venture to address this need through its existing IT solution.

International markets not only revealed new use cases in terms of the type of customer but also in terms of the context of the technology. Pol-9 created his venture's digital therapy product to address a health disorder common among children in the Polish market. As the entrepreneur sought to expand sales of the venture's products to other markets to grow revenue, he realized new health problems which his product could address. Specifically, when contacted by a US customer, the entrepreneur learned that his product could be addressing a different health disorder that was more common in the American population. The founder recounted:

> Never did I think we could reach a global market...suddenly we heard feedback [from] customers, and we understood we have a fantastic product and we started to check on the market, and we noticed, okay, there is some competition which has a similar hardware solution, but they are [in a different area]...We got a call from Holland. [They said] hey guys, we are working with elderly people, and they have a problem. It's called dementia, Alzheimer's. USA has that problem. Actually, USA, that was interesting. The girl [who was] 12 years old...came with her mom. And after the therapy, the mom gave us a call. Hey guys, I'm here with my daughter. She's doing extremely good. She's smiling, and I want to buy [your product]. You'll need to send [it] to the USA...That's how we started in the USA.

The entrepreneur subsequently revised the venture's strategic plans to focus on addressing this new health disorder, leading to a new avenue for growth. In this case, the US customer supplied information on a new market demand (for addressing Alzheimer's), as well as the venture's capability through its existing digital therapeutics product to address this need. The entrepreneur subsequently reoriented the venture's experimentation and strategy to address this health market gap.

Braz-79 realized that there were large differences between the home market in Brazil and the targeted market in the US, requiring the company to expand its product features:

> But we think if we go to the US, we will probably need to change a lot of things. For example,

> we are now developing new hardware. We already have another product, an energy sensor, that is almost the same idea. It works really fast, easy installing and you'll get the currents and voltage from the machine. And you will know if something is going wrong. But now you can understand electrical failures and mechanical failures to the current and the voltage. And you will see the data in the same platform. So we think that probably the US will need something like that to have more products for making more machines online.

The entrepreneur noticed that the level of maturity and existing solutions in the US market outpaced what was happening in Brazil. As a result, the company would need to adjust its value proposition and subsequent experimentation to be able to succeed in the US and elsewhere globally.

These examples show that international markets exposed entrepreneurs to similarities and differences relative to their priors from previous markets, revealing new market gaps and their ventures' capabilities to address these gaps across their customer base. Internationalization, therefore, revealed new opportunities for growth that entrepreneurs did not consider previously.

Narrowing: Providing focus and fine-tuning

International markets did not always expand entrepreneurs' opportunity choice set. In some cases, internationalization allowed entrepreneurs to focus on a niche. Exposure to international contexts revealed information about the markets themselves and the capabilities of the venture—which may have contrasted with entrepreneurs' prior views in other markets. This information suggested the prioritization of one choice over another or the infeasibility of choices in their initial ideas of problems and solutions, feature sets for early adopters to find an initial product-market fit, and use cases for growth.

Ideation. International markets enabled entrepreneurs to narrow the choice of problems they considered in the early days of their business. Jap-32, for example, narrowed the type of problem he was addressing in his business, focusing on the "constant pressure to be able to turn over hit after hit and...collecting content from songwriters and producers reaching a cap." Realizing this specific but tangible problem in his own local geography and his artificial intelligence (AI) capability to address it allowed the entrepreneur to specialize his AI solution to the music industry. Similarly, Arg-12, by starting in his regional market, focused his problem statement on the financial loan hurdles that local merchants faced. This focus allowed the entrepreneur to address a financial niche that enabled the venture's future experimentation to be more deliberate and efficient in its use of resources.

> [We were] simplifying [the] problem that most merchants in Latin America suffer...Credit cards take between 48 hours up to a year in Brazil to receive payments. And payment schedules are very

> different depending on the type of payment use and the region where they are operating. But also, every single transaction has different discounts, from tax retention to withholdings to charge backs by promotions...And it makes [it] very difficult not only understanding when each payment is gonna hit into the bank account, but also...[the] deficit.

The regional market narrowed the choice set of problems that the entrepreneur perceived—specifically related to credit hurdles faced by small businesses. The founder subsequently revised the venture's experimentation and, ultimately, business model around this market gap.

Ind-63 realized that there was a common job matching problem in his home market in India as there was in Israel, the home of his co-founder. In contrast to his pre-conception that a savvier technological context like Israel would not face such a problem, the entrepreneur realized that a gap remained in the global recruitment market:

> And what came to me more as a shock was that despite...the tech supremacy that Israel has, [my co-founder] shared the similar kind of problem with me that people there are also unable to find relevant jobs. And that is why in 2019, pre-COVID, we started digging to figure out what exactly this problem is.

This realization propelled the entrepreneur to narrow in on this problem space around relevant job matching. The realization of unanticipated similarities across Israeli and Indian markets shed light on the universality of the labor market problem that was valuable to address.

Narrowing the problem scope need not occur in the home market. Nether-15, through the founding team's experience in the architecture industry, observed the inefficiencies in city planning in the Netherlands. The entrepreneur ultimately focused on the city planning problem as a core part of his venture's business idea that then shaped his subsequent experimentation and strategy development.

Together, these examples reveal that international exposure supplied information about market gaps and entrepreneurs' capabilities to address those gaps that helped founders narrow down their choice sets to a single problem that they could address with their technology solution. This narrowing, in turn, allowed entrepreneurs to create a niche that would both help focus their subsequent experimentation and differentiate themselves in the market.

Scoping. As entrepreneurs scoped their ideas to early adopters, international markets helped them narrow which features to focus on to address their early customers' needs better in order to find product-market fit. This process ended up directing their further experimentation efforts. Nether-15 focused on a single country to get feedback and iterate on his venture's minimal viable product in order to find its niche in the market and efficiently use its limited resources. In this case, the entrepreneur intentionally turned to a particular international market to focus his development process, but with the assumption that what it learned from this single market would apply to others in the future.

> Since we need to focus by now more closely on a niche, we want to focus more on this in the Netherlands, but then we want to expand into other markets once we get product-market fit...The main objective here in the Netherlands...is getting product-market fit, getting their feedback. Right now, [it] is about [understanding] what is the data that they need...how should we make this work in order to make the whole process more efficient and help them win the bids. We need to be really strategic [on] how are we going to approach it because resources are small, so [we are] really focusing on where we can actually land something significant that provides us with value.

The intentional focus on a single-country market narrowed the possible features needed to be built because of the homogeneity of the early adopter customer base from which the entrepreneur learned. Moreover, the entrepreneur's focus on the Dutch market for his development effort was a result of his past decision to focus on the city planning problem that emerged from the country in the ideation phase. There was path dependency in the narrowed focus between the first two phases. Thus, the first phase dictated the type of features that the entrepreneur built into the venture's early product and ultimately would test in the mainstream market. In this way, the narrowing effects were persistent.

Other entrepreneurs followed suit in narrowing the scope of features they integrated through their focus on a single country market, but with the assumption that their starting market would not only be applicable but actually have a higher threshold relative to other markets. Nether-16 focused on the Dutch market with a data analytics enterprise solution. The entrepreneur assumed that Dutch enterprise customers were harsher than those from other financial hubs around the world.

> The benefit of the Dutch market is that the Dutch are extremely picky and do not ever want to pay too much...They're also extremely tech-savvy...there's 4G everywhere and 5G in some places...They also are extremely picky on whether they want to do business with you...I think if you can convince Dutch financial institutions to become your clients, then it's way easier than convincing folks in London. And especially in New York...Once you know how to sell to the Dutch, then it is easy also to sell to others.

Thus, the Dutch market narrowed the entrepreneur's choices of features to the ones that would pass the relatively high threshold. In this sense, the founder narrowed his possible lessons to the single market. However, this narrowing was less acute in the Dutch market—with its higher threshold of needs—than in other markets. Thus, geography constrained the scope of experimentation for this entrepreneur, but in a more subtle way than for Nether-15.

Whereas the aforementioned entrepreneurs intentionally focused on the Dutch market to scope features to build their minimal viable product, some entrepreneurs organically came across early adopter customers around the world that ultimately narrowed their product experimentation efforts. Pol-21 produced a generalized video-based artificial intelligence product. However, he only realized that his venture should create a specific feature set geared to the retail market upon being contacted by a Singapore-based foreign corporate client. Pol-21 noted that "we were a small company here in Poland, never in Singapore. We've never seen each other. We only talked over Skype...So I don't know what those people look like. And we just got the deal, and we just started working...and we created that system." The international customer's early feedback and specifications ultimately guided not only the direction of the venture's minimal viable product toward the retail sector, where the entrepreneur sought product-market fit, but also shaped subsequent experimentation and strategy focused on addressing retail enterprises' prediction challenges.

International markets also helped entrepreneurs prioritize features they were already considering. Ukr-3, with a hardware collaboration platform for developers, traveled to international conferences in Europe in order to make its early adopter developers aware of his venture's product. However, through pitching to these international developers, he realized that it was not assembly, but rather version control, that was particularly valuable, revealing a new focus area for the venture's minimal viable product:

> So we've built a prototype, and we went to a web summit conference in Europe. At that time, it was open to the public, and [we] actually did get feedback from hardware developers at that point. We had interactive guides where you could create multiple steps...because we realized that people would just listen to our pitch and then ask to switch back to the prototype for graphical version control.

The entrepreneur subsequently was able to update the venture's experimentation to focus on enabling coordination among hardware developers. The feedback from international early adopter audiences allowed the entrepreneur to focus on a value proposition—his venture's ability to enable coordination through version control—that particularly resonated with audiences he encountered at the international conference.

Isr-58 narrowed in on his venture's product features as it tested with the US market and realized that the demands of early Israeli customers were more advanced than those of US customers—the latter, which made up a larger share of their market.

> So we're prioritizing features, the features that Israeli customers don't really care about. But we have, what, 20 different prospects in the US looking for the same thing. So we're doing something wrong, then we can continue to fight with the market and build what we want to build. We need to build what companies want to buy.

The exposure to the US market narrowed the choice set of product features that the entrepreneur perceived in the build phase. The founder realized that he was getting biased signals on the value of his product from the Israeli market that did not stack up with the demand of US customers, who ultimately made up a larger share of his venture's total market. As a result, the entrepreneur re-prioritized his further experimentation to the features that resonated among US customers.

These examples reveal that international markets, whether intentionally or not, helped entrepreneurs focus the feature set of their minimal viable products for early users to more efficiently find product-market fit and deploy their limited resources into subsequent experiments. This deployment ultimately informed strategic commitments. They did so by revealing information about both similarities and differences across markets that led entrepreneurs to rethink the universe of opportunities in terms of both market needs and the ventures' capabilities to address those needs. Internationalization enabled entrepreneurs to prioritize existing choices.

External Validation. Upon completing their minimal viable products, entrepreneurs turned to international markets to grow the usage of their products beyond early adopters. In this stage, international markets ultimately helped entrepreneurs narrow down their use cases for growth in mainstream markets, whether it be in terms of the type of customer or context of the technology. Arg-12 saw the first. From his past observation of problems and iteration on his minimal viable product in the Latin American market, the founder decided that future growth pathways were also mainly in the region, albeit with an expanded user profile. The entrepreneur noted:

> [We] know how much value we could have in Latin America; I think I have a clear picture of the roadmap. I also have the flexibility to adapt the roadmap. But I really don't have the full picture to add value to the whole world yet. It may happen, but I think we can be a 1-billion-dollar company only doing what we need to do in Latin America.

Thus, past experiences in the Latin American market narrowed the entrepreneur's choices of regions for subsequent growth. Crucially, this example reveals that these past choices to ideate and build in the region shaped his experimentation efforts in the external validation phase, ultimately shaping the venture's strategy to be focused on the regional market. While the entrepreneur found a new business model to scale by noticing that his venture would be able to fulfill the needs of larger enterprises, expanding regionally also constrained his horizontal (geographic) plans for growth.

While selling in the regional market constrained Arg-12's geographic growth plans to that market, it was actually selling in other markets that informed Sing-35 that his venture needed to focus more on the regional, rather than the global, market. The latter entrepreneur tested in Southeast Asia as well as the United States but then realized it was hard to operate multi-regionally because "each country is different... the cultures are different, the laws are different...people operate in different cycles...it didn't work." The entrepreneur decided that his venture was "well-positioned to be part of the growth story [in Asia]" and, therefore, it will not "step into the West, although it seems salivating at times" because it can get an "exit right in [the Asian] market." Testing in international markets revealed information about the disparate needs of customers and his venture's lack of capacity to address all of them. Ultimately, this process made the Singaporean entrepreneur realize that his venture should narrow to the Southeast Asian region despite the monetary incentives of selling to a broader market. Thus, while the Southeast Asian regional market expanded Sing-35's choice set of approaches to differentiate in the market— through a higher-touch, relationship-based approach— it also narrowed the entrepreneur's perceived use cases to expand his venture's product to regional companies and achieve its desired exit.

Though selling in international markets made Sing-35 realize that his venture was trying to do too much and needed to focus its attention on a region given the inter-regional difference in tastes for its product, doing so made Jap-33 realize that his venture's product resonated more with one market over another, allowing him to prioritize among the two.

> We launched the Kickstarter campaign in three days...So then maybe...50% of our customers were from Japan, 30% from the US and 20% from the rest of the world. And I saw that [when] we did the crowdfunding campaign on Kickstarter, and that was basically in English, but still it got the attraction of the Japanese market.

Ultimately, international markets enabled the entrepreneur to narrow his venture's focus on the Japanese, rather than the US, market, given the demand and regulatory structure of the former market that was most conducive to his venture's innovation. This exposure ultimately shaped the entrepreneur's subsequent experimentation and ultimate strategy focused on Japanese consumers first and later European ones, where there was a similar regulatory structure.

In the case of Ken-82, the realization of the—surprisingly—common preferences of students across countries allowed the entrepreneur to narrow into his product's existing feature set, validating the universal nature of his product.

> Interestingly, [we did not have to change much]. I think that...Gen Z is moving as a block in terms of how they view things. And probably this is because they are so connected. They're much more connected than our generation as millennials. So yeah, that's interesting, actually. They seem to think in the same way about some of the things we've tested.

By helping the entrepreneur realize the similarities across country markets—relative to the cultural differences he pre-conceived—exposure to international markets shifted the set of scaling opportunities that the founder observed. The entrepreneur could take a fairly similar use case across country markets with minimal adjustments.

Together, these examples reveal that testing mainstream demand from international markets helped entrepreneurs narrow the scope of their markets to inform their experimentation and ultimate strategy for growth.

2.5 Discussion

Our field study sought to address the question: how does internationalization influence the entrepreneurial process, and how early-stage entrepreneurs identify and exploit opportunities? Relying on interviews and other field data from 84 technology entrepreneurs in 27 countries in Europe, the United States, Asia, Africa, and Latin America, we find that internationalization influences entrepreneurial experimentation and, ultimately, how early-stage entrepreneurs define their businesses. We develop a process model in which internationalization expands and contracts entrepreneurs' opportunity set as they define, scope, and externally validate ideas that can create and capture value. We show how international markets served as inputs into entrepreneurs' subsequent experimentation efforts and strategy, though founders may have intended them to be outcomes of their already-established strategy. International markets became inputs, sometimes unintentionally, by providing information to entrepreneurs about market needs and ventures' capabilities to address them. This information either expanded or narrowed entrepreneurs' choice set of business ideas, technology features for early adopters and use cases for growth to the mainstream market. Notably, we find that international markets acted cumulatively to do so. Changing a product or approach previously by focusing on a particular international market influenced entrepreneurs' subsequent experimentation efforts and strategy.

Our findings also suggest that international markets shape entrepreneurs differently based on their ex-ante state of decision-making. When entrepreneurs had far too many choices that made it difficult to allocate their scarce resources and attention, international markets helped focus them on the ideas, minimal viable product needs, and growth pathways that were most salient. Alternatively, when entrepreneurs had too few choices that constrained them to local peaks relative to their potential, international markets helped expand the choice set to options that could achieve desired growth outcomes. In either case, international markets served as inputs in the experimentation and strategy development process, rather than being only outcomes of those processes.

Our study reveals that the effects of internationalization may be unintentional and profound, contributing to a wide body of scholarship that assesses the often-intended effects of internationalization for companies, including entrepreneurial ventures. Prior work shows that exposure to international markets offers entrepreneurs access to knowledge, investment, talent, and customers, with important implications for performance (e.g., Alvarez-Garrido and Guler, 2018; Balachandran and Hernandez, 2020; Chung and Alcácer, 2002; Cullen and Farronato, 2021; Kulchina and Hernandez, 2016; Lin and Viswanathan, 2016; Stuart and Sorenson, 2003; Wormald et al., 2021). These ventures anticipate these effects and, in fact, target international markets to acquire such benefits, despite the uncertainty, liability of foreignness, and other costs they face to do so (e.g., Zaheer, 1995; Zaheer and Mosakowski, 1997). Our study revealed an unintentional effect that international markets brought to entrepreneurs: a change to their perceived set of choices to grow their firms. In doing so, international exposure shaped the core experimentation and ultimate strategy that entrepreneurs pursued to scale. By revealing unanticipated differences and similarities relative to entrepreneurs' priors, international markets helped entrepreneurs adjust their cognitive lens of opportunities. In doing so, internationalization profoundly influenced the entrepreneurial process and how early-stage entrepreneurs identify and exploit opportunities.

By revealing the unintentional consequences that internationalization has on how entrepreneurs define, scope, and externally validate their ideas, our study relaxes an important assumption behind many works on the internationalization of companies: that the benefits of different international markets are known to companies prior to entry, even if the probability of getting these benefits is unknown. Our study reveals that, at least for early-stage entrepreneurs who are still iterating on their core value proposition and face high uncertainty regarding market demand, these benefits may not be known ahead of time. This imperfect information about not only the probability of success, but also the nature and magnitude of the effects, reveals that international markets may have a profound effect on ventures beyond entrepreneurs' prior expectations. This means that entrepreneurs might face an additional type of uncertainty in international markets—the nature and magnitude of benefits (or costs) they might face. Paradoxically, internationalization also may help resolve some of their core uncertainty that requires experimentation by exposing them to unintentional information—for example, the composition of customer segments and user tastes—in new markets.

2.5.1 Theoretical and practical implications

Our study's findings reveal that internationalization shapes entrepreneurial experimentation and strategy development. While entrepreneurs test to enter and adapt to international markets (Bingham, 2009; Bingham and Davis, 2012; Bingham and Eisenhardt, 2011; Gupta and Khanna, 2019), their iteration does not stop in the particular market. Rather, it continues to shape entrepreneurs' trajectories in the form of which ideas they test. These tests ultimately influence their realization of value and subsequent strategy development. In this way, international markets can actually help ventures reduce some uncertainty on which opportunities are possible to exploit, as opposed to only creating additional sources of uncertainty as typically expected in new contexts requiring experimentation. This appears to be the case because entrepreneurs often observe unanticipated differences in international markets from their priors. Whether it be related to the composition of different customer profiles or the inherent way that business is done by customers more broadly, these differences make salient whether previously conceived opportunities are viable, as well as whether there are new opportunities to exploit. Thus, such unanticipated differences between international markets help entrepreneurs to update their prior beliefs of the nature of opportunities that they can exploit (Kirtley and O'Mahony, 2020) to shape their subsequent testing.

Beyond informing experimentation efforts, we show that internationalization also influences the ultimate strategic commitments that entrepreneurs make. Rather than being part of the opportunity itself as is common in entrepreneurial strategy research (Bingham et al., 2007; Gans et al., 2019; Sørensen and Sorenson, 2003), we show that markets can be an input into the decision-making process to select the opportunity. Inherent differences across international markets—whether it be related to institutional factors, political factors, or cultural tastes—that shape the value of entrepreneurs' innovations may not be anticipated by entrepreneurs. Gaining exposure to these differences re-focuses the attention of entrepreneurs (Ocasio and Joseph, 2018). Context, therefore, matters for entrepreneurial decision-making and is valuable to consider in an assessment of entrepreneurs' choices.

By influencing which ideas entrepreneurs test and ultimately make commitments to, our findings reveal how geography plays an important role in the direction of entrepreneurs' innovations in terms of which markets benefit from these innovations. Entrepreneurs may change their industry focus; for example, from social media to healthcare, as demonstrated by Pol-5. They also may shift the geographic beneficiaries of their innovations; for example, from US to Japanese and European customers, as shown by Jap-33. Through these industry and geographic dimensions, internationalization may ultimately shape who benefits from innovations, as seen in the gender context (Cao et al., 2021; Koning et al., 2020; 2021).

2.5.2 Limitations and avenues for future research

Our theory has two main boundary conditions. First, our study focuses on entrepreneurs in earlier stages of ventures rather than in mature firms. Young companies tend to be particularly responsive to external signals in the environment, given the high uncertainty (Eisenhardt and Bingham, 2017; Kerr et al., 2014b) and lack of organizational inertia that they face, as seen in ventures' responses to customers from nascent industries (McDonald and Gao, 2019) and amid institutional change (Zhang et al., 2016). However, mature firms that are operating under high uncertainty and possess malleable organizational structures may also exhibit internationalization, shaping their experimentation and strategy development. Future research may assess whether such mature companies adapt their core experimentation and strategy as they internationalize.

Second, we focus on entrepreneurs' exposure across countries rather than across geographies within countries because the often larger cross-country differences (Berry et al., 2010; Ghemawat, 2001) provide more scope for learning to guide venture experimentation. However, our theory may extend to geographies within countries where there are substantial differences between them, for example, in terms of culture, regulations, or economic conditions. Future research may assess whether entrepreneurs' entry into new geographies within the same country similarly shapes their core experimentation and strategy.

Our process model also suggests that the narrowing and expansion mechanisms may emerge based on the ex-ante experimentation of entrepreneurs. For example, internationalization may constrain the choice set of entrepreneurs when they are over-exploring or expand them when they are over-exploiting opportunities (March, 1991). Further work may test the variance in these mechanisms relative to entrepreneurs' original experimentation orientation.

2.5.3 Conclusion

Overall, our findings show how internationalization may be a core input into entrepreneurial experimentation and strategy development, rather than just being an outcome of these processes. This appears to be the case because entrepreneurs do not fully anticipate the nature of the information that they confront in these markets. This newly realized information shapes the set of opportunities that entrepreneurs perceive. As a result, international markets may have profound effects on the direction and performance of entrepreneurship around the world.

Our findings contribute to entrepreneurial experimentation research by showing the role which international markets may play in reducing uncertainty and shaping the value realization of startups. They also contribute to internationalization research by showing how not only do startups experiment with entering new markets, but also international markets can shape entrepreneurs' core experimentation. In doing so, we shed light on how the potential effects of international markets may not be fully known ahead of time by founders, revealing how entrepreneurs not only face uncertainty in the probability of success in new markets but also in the nature of the benefits.

We hope our research leads to future work that advances our understanding of the complex relationship between internationalization and experimentation. For example, such work may assess the interplay between experimentation and internationalization among mature firms and within country settings. It also may further investigate when entrepreneurs and managers can forecast the information they confront in different international markets and what might account for informational asymmetries in this regard. Lastly, future work may uncover heterogeneities in how international markets shape entrepreneurial experimentation across different sub-industries, geographies, and the ex-ante experimentation orientations of entrepreneurs.

3

Judging Foreign Startups

3.1 Introduction

Startups, like corporations, are increasingly globalized in terms of their markets, investments, and workforce, partially due to the advent of technology that reduces the cost of expanding internationally (Alcácer et al., 2016; Alvarez-Garrido and Guler, 2018; Brynjolfsson et al., 2019; Ghemawat and Altman, 2019; Kerr, 2016; Lu and Beamish, 2001; Oviatt and McDougall, 2005).[1] As a result, entrepreneurial gatekeepers, particularly accelerators[2] which have diffused around the world (Cohen, et al., 2019b), increasingly evaluate a global pool of startups and must choose the most promising to provide support and funding (Balachandran and Hernandez, 2020). For

[1] Co-authored with Rem Koning and Tarun Khanna

[2] An accelerator is defined as a fixed-term, cohort-based program for startups, including mentorship and/or educational components, that culminates in a graduation event (Cohen et al., 2019b).

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 (X1 Group, 2018; Y Combinator, 2020). At the same time, gatekeepers have missed out on promising international startup opportunities. The same Y Combinator rejected Canada-based online apparel company, Stylekick, which ended up reaching 80 countries, being translated into 14 languages, and ultimately acquired by Shopify (Business Insider, 2018; Mitra, 2018).

Can accelerators choose the most promising startups from this increasingly global pool? Indeed, accelerators are now soliciting applications from across the globe (Cohen, et al., 2019b). However, these organizations may not be able to discern the quality of the startups that apply (Gans et al., 2008; Kerr et al., 2014b; Luo, 2014). Further, they may be particularly inaccurate in discerning the potential of foreign startups because they lack the contextual expertise and information—ranging from knowledge of institutions to differences in consumer tastes—necessary to sort winners from losers. Moreover, judges may carry a bias for or against foreign startups, like the gender, race, and expertise biases documented across a range of entrepreneurial and innovation settings (e.g., Hegde and Tumlinson, 2014; Lee and Huang, 2018; Li, 2017; Niessen-Ruenzi and Ruenzi, 2019).

These concerns are especially acute at the earliest stages of the startup selection process when accelerators make decisions with little more than a quick pitch or text description, often because the number of startups screened makes it too costly to conduct in-depth due diligence on each. In these earliest stages, bias and uninformedness are especially problematic because when judges pass on a startup, they also never get a chance to learn more about the firm and correct any initial mistakes. The sheer number of startups in the earliest selection pool of an accelerator makes "spray and pray" approaches infeasible (Ewens et al., 2018). Offering support to the thousands of startups would be incredibly expensive, leading accelerators to necessarily rely on meaningful filtering and selection (Cohen et al., 2019b).

Thus, understanding whether judges are informed about the quality of local and foreign startups at the earliest screening stage of accelerator decision-making is both necessary if we hope to understand why home bias occurs and how accelerators might address it. Prior research in non-accelerator contexts on foreign discounting[3] shows that trade partners, financial analysts, and investors are more likely to select companies that are nearby, but these studies often conflate crucial differences in the mechanisms underlying the effect (Coval and Moskowitz,

[3]We use the terms foreign discounting, foreign bias, and home bias to refer to the fact that judges are more likely to give lower scores to foreign startups after accounting for startup quality. We do not claim that the presence of bias indicates judges are necessarily xenophobic. As we discussed in our theoretical framework section, there are numerous potential mechanisms that can lead to lower evaluations of foreign startups that are unrelated to startup quality and potential.

1999; 2001; Disdier and Head, 2008; Sorenson and Stuart, 2001). As mentioned above, home bias by accelerators could result from 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 selecting foreign firms. However, such an approach will backfire if the underlying 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 choices of foreign firms are potentially no better than random draws. No matter the threshold, judges will always end up selecting lower-quality foreign ventures than local ones. In this case, remedying the underlying "bias" requires finding judges who can discern winners from losers. This approach might involve assigning judges to only evaluate startups from their home region. Such judges can more quickly and cheaply determine quality without burdensome due diligence. Redesigning how scores are aggregated into decisions may not make a meaningful difference in this scenario. In short, the underlying mechanisms that lead to foreign discounting in startup screening have strong implications for how accelerators should design the first stage of 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, as well as the accelerators who select them, will nearly always confound supply-side forces (the judge's choice of who to pick) and demand-side ones (the founder's choice of where to apply). Further, even when the distribution of potentially selected startups is fully observed (e.g., in venture competitions), startups may selectively choose whether to enter local or foreign competitions, 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, showing that judges discount foreign startups is insufficient to reveal the underlying mechanism. This mechanism ultimately determines how organizations should respond. Specifically, 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 to causally identify if judges exhibit home bias and pinpoint the mechanisms underlying this effect. In the first round of this competition—where judges evaluate text applications—1,040 judges from North America (the United States and Canada), Latin America, Europe, and Israel evaluated 3,780 startups from across the globe. Crucially, in this first round, the accelerator randomly assigned judges to evaluate startups no matter their origin, and no startups could opt out of being evaluated by judges from particular regions. This staged judging process, where judges first evaluate a brief pitch or application before deciding which startups to interview and conduct further due diligence on, is widely used at accelerators, including Y Combinator and Techstars (Cohen et al., 2019b).

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 roughly a third of the effect of a startup going from having no users to some user traction and a tenth of the size of the effect of having raised venture financing. These magnitudes are consistent with prior work documenting home bias in other contexts ranging from financial markets to trade (Coval and Moskowitz, 1999; Disdier and Head, 2008).

Our analysis reveals that 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 relative to foreign firms. Surprisingly, we instead find that judges are equally good at evaluating startup quality whether the startup is from their home region or not. In fact, judges give higher scores to local and foreign startups that go on to raise financing, experience more user growth, as well as have higher employee, valuation, and revenue growth, contrary to prior work showing that judges can struggle to pick startup winners from losers (e.g., Scott et al., 2020). Further, when we conduct back-of-the-envelope calculations, we find that judges passed over 148 promising foreign startups, equating to roughly 1 in 20 startups in our sample. This evidence suggests that simple changes to how accelerators aggregate judges' evaluations may mitigate the impact of home bias on outcomes.

These findings, at first glance, are at odds with prior work from other contexts showing that experts cannot detect the quality of early-stage firms (e.g., Scott et al., 2020) and that when investors can detect quality differences, it is because they have a local information advantage (e.g., Coval and Moskowitz, 2001). Recent work suggests investors increasingly use a "spray and pray" approach to learn about startup quality after making small up-front investments instead of heavily screening which firms to invest in (Ewens et al., 2018). Yet, this work has largely focused on investments and decisions on pre-screened and relatively successful firms. When we restrict our sample to conceptually replicate this prior work, we can actually recover the patterns found in this work. When we only include a more selective range of startups, for example, firms with founders who attended an elite university as in Scott et al. (2020), we find that judges are less capable of evaluating which startups are promising and which are not. Similarly, when we use the application text to restrict our sample to more localized firms as in Coval and Moskowitz (2001), we find that, unlike in our full sample of globally oriented technology startups, judges do possess a local information advantage. These patterns suggest that the quality of a judge depends not only on their innate skills and preferences but also fundamentally on the composition of the pool of startups they are tasked with evaluating.

Our findings make three primary contributions. First, they show that home bias exists in the accelerator setting. Unlike in trade and investment settings (e.g., Coval and Moskowitz, 2001), where home bias has often been studied, we find that judges in the accelerator setting are generally informed but biased against foreign firms when screening early-stage startup ideas. As our conceptual replication of prior work shows, this result does not reflect the innate characteristics of the judges but rather is a combination of judge behavior and the pool of startups being evaluated. Specifically, the pool of startups in the accelerator setting—versus previously studied settings—tend to follow more globally-oriented business models that may be easier to evaluate across countries than firms analyzed in prior research on home bias. This result suggests that future work on evaluation should focus both on who evaluates and, equally importantly, what ends up being evaluated. Indeed, our findings suggest that the widening of the pool of startups (e.g., in terms of educational backgrounds) that judges consider, along with the increasingly standardized business models that these startups adopt, may well imply that accelerators are "better" at screening startups than are investors and mentors studied in prior research (Howell, 2020; Kerr et al., 2014a; Nanda et al., 2020; Scott et al., 2020). Interestingly, as investors increasingly pursue "spray and pray" approaches to learning about startup quality (Ewens et al., 2018), accelerators can serve as complementary sources of early screening to narrow down the "sprayable" pool of startups.

Second, our results suggest that geographic discounting 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). If gatekeepers discount foreign startups, and if most of these gatekeepers still reside in entrepreneurial hubs like in the US, this may potentially result in a gap in startups from non-hub regions. Especially because startups excluded at the first stage undergo no further due diligence, the presence of early bias has the potential to distort the sorts of firms that receive support and succeed. These startups otherwise may not be able to get the same type of support from investors who are increasingly pursuing "spray and pray" models that tradeoff providing support to portfolio startups with investing in more startups to learn about their quality (Ewens et al., 2018). And this discounting does not just impact which startups succeed but also may impact who benefits from their innovations (Koning et al., 2020; 2021). Indeed, if accelerators overlook ideas from these non-hub markets, then there may be too few startups serving customers' needs in these non-hub, often non-western, regions.

Third and finally, we highlight a potential limitation of accelerators when it comes to helping foreign startups gain access to key entrepreneurial ecosystems driven by selection effects. While various studies focus on the treatment effects of accelerator programs, finding positive performance gains for startups (Cohen et al., 2019a; Fehder and Hochberg, 2014; Gonzalez-Uribe and Leatherbee, 2018; Hallen et al., 2020; Howell, 2017; Yin and Luo, 2018; Yu, 2020), our results suggest that the impact of accelerators may be muted for foreign startups because these organizations discount them. This finding shows the value of evaluating selection processes—in addition to treatment effects—in accelerators to fully understand their role in entrepreneurial growth. That said, our results also suggest that relatively minor tweaks to how an accelerator aggregates decisions can address this home bias.

3.2 Theoretical framework

3.2.1 Evaluating Startup Quality

Evaluating early-stage startup quality is especially difficult because of at least three information challenges. First, the success of startup ideas hinges on the interaction of complex factors, including the technology itself, the business model, customer demand, competition, and the founding team (Aggarwal et al., 2015; Gompers et al., 2020; Hoenig and Henkel, 2015; Kaplan et al., 2009; Sørensen, 2007). Second, there are few precedents to anchor startup evaluations. Great startup ideas are inherently novel, and only a subset of those actually succeed in practice (Hall and Woodward, 2010). Third, entrepreneurs may only provide incomplete information about their ideas, as disclosure can eliminate incentives to "pay" for the now "free" to appropriate idea (Arrow, 1962; Gans et al., 2008; Luo, 2014). Consistent with these priors, research shows that entrepreneurial judges often lack the ability to evaluate the quality of startups, and instead, experiment with small investments into startups to learn of their value (Ewens et al., 2018; Kerr et al., 2014a; 2014b; Nanda et al., 2020; Scott et al., 2020).

3.2.2 Contextual Intelligence

Given these challenges in discerning startup quality, when (if at all) can evaluators distinguish winners from losers? Evaluators may be able to do so when they have the expertise (Li, 2017) or intuition (Huang and Pearce, 2015) that compensates for the imperfect information they have on any new venture. Indeed, prior research suggests that expertise is a product of the local region where investors and inventors live and work (Coval and Moskowitz, 2001; Dahl and Sorenson, 2012; Malloy, 2005). However, this locally-developed expertise may not be transferable to foreign contexts because of differences in institutions, culture, language, and markets (Khanna, 2014). Evaluators, therefore, may only be able to use this locally derived expertise to better assess the quality of local, but not of foreign startups. For example, an Israeli judge might be able to use her expertise of Israel's military structure to understand the relative quality of founders of an Israeli company with military experience and not a US company with founders who have military experience. Consistent with this view, prior work has shown that financial analysts are worse at picking foreign stock winners, relative to local stock winners (Coval and Moskowitz, 2001; Malloy, 2005), and information frictions are higher for foreign acquirers (Conti et al., 2020).

3.2.3 Bias in Evaluations

However, reliance on local expertise to evaluate startups may also induce biases. Prior work shows that judges prefer what is more "familiar" (Franke et al., 2006; Huberman, 2001; Lin et al., 2013). In the context of demographics, prior research has found substantial evidence of bias against entrepreneurs from different genders and races (Hegde and Tumlinson, 2014; Lee and Huang, 2018; Niessen-Ruenzi and Ruenzi, 2019). Similarly, in the geographic context, studies in financial and trade markets have detected a home bias for local portfolio stocks or trade partners (Coval and Moskowitz, 1999; 2001; Disdier and Head, 2008).

The literature puts forth at least three reasons why home bias might emerge even if judges are no better at evaluating the quality of local startups. First, judges may cognitively prefer what is more familiar or culturally proximate. For example, a startup from a similar geography as a judge may have a subtle way of framing its pitch that draws on local customs that are especially likely to resonate with the judge (Bell et al., 2012; Chadha et al., 2022; Huberman, 2001). Second, judges may simply be xenophobic against particular nationalities or geographic regions, causing them to give lower scores to startups from foreign places (Arikan and Shenkar, 2013). Inversely, judges may prefer that their own regions benefit from entrepreneurial growth and innovation, leading judges to give higher evaluations to local startups (Bell et al., 2012). No matter which mechanism dominates, in each case, judges give lower scores to foreign startups for reasons unrelated to their ability to detect the startup's quality.

3.2.4 Hypothesis Development

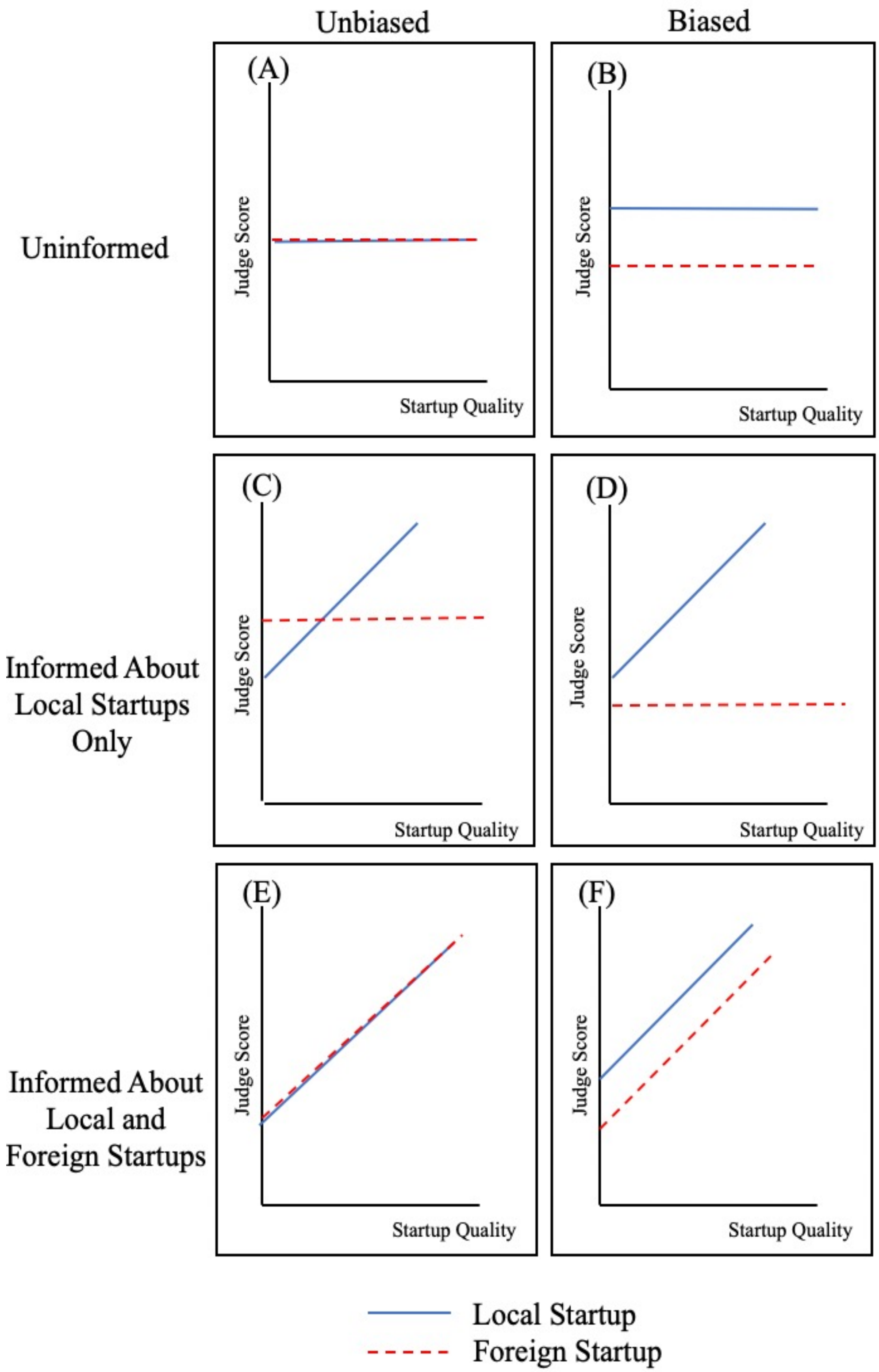
These different mechanisms—evaluation uncertainty, contextual expertise, and bias—generate six scenarios that each call for different strategic responses by accelerators and startups. Figure 3.1 sketches how each of these scenarios reveals a different relationship between startup quality (x-axis) and a judge's evaluation score (y-axis) for startups foreign to the judge (dashed line) and local to the judge (solid line).

In the first row of Figure 3.1, we show the pessimistic cases where judges cannot pick winners from losers. No matter whether judges are biased (cell B)—systematically preferring local or foreign startups—or unbiased (cell A), the selected pool of startups consists of a random share of high- and low-quality firms. In this worst-case scenario, organizations should reduce their attention to screening startups and perhaps re-allocate resources to monitoring selected startups in the hopes of improving firms' future performance (Bernstein et al., 2016).

However, research ranging from work on contextual intelligence to the benefits of investing in and running firms in one's home region (Coval and Moskowitz, 2001; Dahl and Sorenson, 2012; Malloy, 2005) suggests that judges can pick winners from losers locally even if they cannot evaluate the quality of foreign startups. The second row of Figure 3.1 illustrates this scenario. Cell C shows that when judges have a local information advantage and are not biased against foreign startups, they will give higher-quality local startups higher scores. However, they will not necessarily give higher scores to lower-quality local startups. In fact, with better local information, it is likely that judges will give low-quality local startups low scores while erroneously evaluating low-quality foreign startups as better than they actually are. The result is that the lines intersect in cell C. However, if judges are also biased, this shifts the line for local startups upwards, as seen in cell D. While judges still give higher scores to better local startups, all local startups will be judged as better than any given foreign firm. The result is that in cell D and cell B, we see consistent foreign discounting, but each reflects meaningfully different mechanisms. While cell B suggests that organizations would be better off re-allocating attention away from the selection process altogether, cells C and D suggest that organizations would be better off assigning judges to evaluate local but not foreign startups.

Lastly and most optimistically, judges might be able to evaluate the quality of both local and foreign firms, as shown in the third row of Figure 3.1. Startups may follow a similar enough playbook that separating good from bad investments across countries is not significantly harder than within countries. For example, work has shown the benefits of good management appear universal for corporations and startups across the globe (Bloom

Figure 3.1: Predicted relationships between judge scores and startup quality



and Reenen, 2007; Chatterji et al., 2019), as are coding practices (Haefliger et al., 2008). Cell F shows that bias interferes with picking the most promising startups because judges may pass over higher-quality foreign startups for lower-quality local startups. In this case, organizations can simply revise 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) to counter this discount.

The framework presented in Figure 3.1 builds on information-bias tradeoffs discussed in other studies of evaluation (e.g., Boudreau et al., 2016; Li, 2017). Our simple two-by-three reveals that knowing whether judges give lower scores to foreign startups—as is the case in cells B, C, D, and F—is insufficient to understand how an organization might change to address foreign discounting. However, with knowledge of startups' quality, we have sufficient information to separate the different mechanisms.

3.3 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 US, Europe, Israel, and Latin America. There are four rounds in the accelerator program. In the first round (the global round), startups virtually apply to several of the regional locations of the accelerator program. This round is akin to the earliest screening stages of major accelerators that involve the evaluation of inbound text applications. In the latter rounds, the accelerator assigns startups (based on their preferences and judge scores) to one of its regional locations, and judges generally local to that area evaluate the startups. The pool consists of mostly high-tech startups, similar to startups in other top accelerator programs like Y Combinator or Techstars. The startups in the program have collectively raised over $6.2 billion, generated over $3 billion in revenue, and created over 157,000 jobs since the accelerator's inception.

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. Unlike the first round that we analyze, later rounds involve interviews between judges and the startup team, pitches, and further due diligence by judges with expertise in the startup's domain. Startups who make it to the third round (approximately 10 percent of the initial applicant pool) participate in the full in-person 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 tens of thousands of dollars. Across 2013-2019, these four rounds consist of 87,977 startup-judge level observations, including 11,188 unique startups and 3,712 unique judges.

We focus on the global round of the competition, the earliest screening round, where judges—representing executives (60 percent), investors (13 percent), and other professionals (27 percent)—across these international regions initially screen startups from around the world. Judges are well-seasoned. On average, the judges in our sample had already evaluated 56 startups for the accelerator before the evaluation rounds we analyze. Furthermore, 26 of these 56 startups were foreign to the judge. As such, our estimates do not merely reflect how inexperienced evaluators might decide but also capture how more experienced judges screen startups.

Judges evaluate an application that includes self-reported information on the company's background and funding, industry & competitors, and business model & financials. We show the full application template in the appendix. 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. Through a word search analysis of the application text, we find that the home region of the startup is explicitly mentioned in 42 percent of startups' market, traction, and team text. This percent is likely an underestimate because it does not take into consideration implicit mention to the home region, for example, via mention of the past employer and educational institutions of the team. Appendix Table B.8.1 shows robustness checks that startups are not strategically disclosing their location based on their quality or location. Judges review these applications online. Each judge evaluates roughly 20 startups, and each startup receives evaluations from 5 judges on average. Judges recommend whether a startup should move to the next round of the competition, and applicants move on to the next round when at least 50% of the judges recommend the startup should move on.[5] Judges also provide subscores on a scale of 1–10 on the following criteria: startup team, industry & competitors, and business model & financials. The program does not give judges a quota in terms of the number of startups they can recommend. Further, judges must agree to terms that indicate that they "do not expect anything in return," including "future contact" from the startups they evaluate.

To infer judges' location, we use data on the location of the accelerator the judge is affiliated with.[6] As judges

[4]While English applications may mask the 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]Judges provide a 0–5 score on whether they recommend the startup to the next round of the competition; scores above 2 result in startups moving to the next round. While most startups move on to the next round when 50% of the judges recommend them, there are a small number of exceptions to this rule.

[6]The accelerator does not collect data on judges location of residence. It only collects the home accelerator program of each judge.

need to evaluate startups in person during the later rounds of the competition, they tend to be assigned to a physically proximate accelerator. We, therefore, categorize judge locations as corresponding to the accelerator's locations: Europe, Latin America, Northern America (US & Canada), and Israel.

These broad regional categories will lead us to underestimate biases within regions. For example, a UK 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 so potentially discount it. Similarly, measurement error due to some judges being assigned to a home program in which they do not work or reside (e.g., a Chicago-based judge is assigned to the Latin American program) should also bias our estimates towards zero. In Appendix Figures B.1.1–B.1.6, we use additional data on judge locations and Monte Carlo simulations to show that our research design, coupled with the large sample of judge-startup evaluations we observe, allow us to detect foreign bias estimates even in the face of substantial measurement error.

The startups in this global round are of a similar type as those participating in landmark accelerator programs around the world, such as Y Combinator and Techstars. They are largely technology-driven and growth-oriented. Indeed, 39 percent of them are in high tech, 27 percent in general sectors (e.g., retail, consumer products), 17 percent in healthcare/life sciences, 13 percent in social impact, and 4 percent in energy/clean tech. Roughly a fifth of them mention a hub city—such as Silicon Valley, Boston, or London—as identified by the Startup Genome Project (2021), in their market, traction, and team application text (Table B.9.3). The same share also mentions an elite university in their team application text. About 12 percent of the startups mention an MBA, and 9 percent mention a PhD education in their team application text.

3.4 Data

Our data come from the accelerator's 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 the possibility that startups self-select into local programs. Such selection would make it impossible to separate judge from startup effects. Our 2017-18 data consist of 20,579 startup-judge level observations, including 4,420 unique startups and 1,043 unique judges. We remove startups whose headquarters regions do not match any of the judges' home programs to exclude the startups that are foreign to all judges in our sample and therefore lack a local judge score as a basis of comparison.[7] We also remove judges who lack a home program that is part of the main accelerator.[8] This brings our final sample to 17,608 startup-judge level observations, including 3,780 unique startups and 1,040 unique judges.

3.4.1 Measuring Startup Quality

Measuring startup quality is difficult not only for judges, but also for researchers. Early-stage startups rarely have revenue or profits which are common metrics of company performance. Instead, entrepreneurship studies turn to other intermediate milestones to proxy early-stage companies' performance and quality. One common measure is financing from angel investors or venture capitalists (Cao et al., 2021; Howell, 2017; Yu, 2020). This is a common measure because these investors' decisions reflect both selection and treatment effects that should result in startups with financing having higher startup performance. On the selection side, early-stage investors conduct rigorous due diligence on portfolio companies prior to investing, which may enable them to understand the quality of ventures (Gompers et al., 2020). On the treatment side, investors provide added value (Bernstein et al., 2016) and a stamp of approval (Lerner et al., 2018) to startups that enable them to gain subsequent financing and increase their chances of a successful exit, either an acquisition or initial public offering (Catalini et al., 2019). Another increasingly common indicator is user traction, reflecting how much visibility and use a startup is getting from customers and other gatekeepers. Website page visits are becoming a common indicator for the latter in entrepreneurship studies to proxy startup performance (Cao et al., 2021; Hallen et al., 2020; Koning et al., 2022).

We measure both pre-accelerator and post-accelerator measures of financing and website page visits in our analysis. Pre-accelerator measures allow us to assess whether judges can evaluate the quality of startups at the time of evaluation. Post-accelerator measures allow us to evaluate whether judges can evaluate the future potential of startups. Beyond these measures, in Appendix Table B.13.2, we show that the findings hold when we use additional measures of startup quality, including valuation, employee counts, and estimated revenue growth 3–4 years after the accelerator program.

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

[8] Our results are robust to including or excluding judges whose home program is not one of the main accelerator programs.

3.4.2 Dependent Variables

Score – Our first dependent variable is a composite z-score created from the z-scored subscores judges give to startups. These underlying subscores include: customer pain and solution, 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. All but the last range from a scale of 1–10. The latter is on a scale of 0–5. While not all judges complete every subscore evaluation, the vast majority do. Of the 17,608 recommendation evaluations in our data, for 16,339 (93 percent), we have complete subscore information.

Recommend – Our second dependent variable is a binary variable indicating whether a judge recommended the startup to advance to the next round of the competition.[9] Judges separately provide this score on the judging form, so while this measure is correlated with the substantive subscores discussed above, it is not perfectly so. This is the main measure used by the accelerator to determine whether startups move to the next round. However, there are exceptions to this cutoff. In these exceptions, the scores on the numerical dimensions (e.g., customer pain/solution and business model/financials), along with other factors, can play a part in the startup's acceptance into the program.

3.4.3 Independent Variables

Foreign Startup – Our key covariate captures whether the judge and startup are from the same region (e.g., both from Europe, the US/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).

Logged Financing Value (Post) – We use logged financing value six months after the program.[10] This variable indicates the logged amount of USD startups received from investors six months after the program.

Logged Page Visits (Post) – We also use logged monthly page visits after the accelerator program in 2019 (the latest data we have available).

Financing (Pre) – We use logged financing value (in USD) that startups received from investors before the program.

[9] We constructed this as equal to 1 if the judges score was over 2 (on a scale of 0–5) and 0 otherwise, as the accelerator uses this cutoff to determine whether a startup makes it to the next round of the competition.

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

Whether Has Financing – We include a binary variable indicating whether a startup received financing before the program to indicate financing traction.

Logged Page Visits (Pre) – We include logged website page visits 3 months before the initial application review period of the accelerator.

Whether Has User Traction – We use a binary variable on whether a startup reached at least 100 website page visitors on average per month over the last three months before the program to indicate user traction.

In our context, when startups lack page visits or financing data, they generally have so few visits or little financing that corresponding databases like SimilarWeb (that collects companies' page visits) and Crunchbase (that collects startups' funding rounds) do not track them. We, therefore, set missing page visits or financing values to 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.[11]

Accelerator Participation – We also account for whether a startup participated in the accelerator interacted with whether a startup is local or foreign to the judge. 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 3.1 shows summary statistics for our main sample from the global round of the competition, including 17,608 startup-judge level observations, 3,780 unique startups, and 1,040 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 each 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 firms with two exceptions. The first is pre-accelerator user traction: local startups have a higher value on average by 6 percentage points (p=0.000). The second is post-accelerator logged financing: local startups have a higher value on average by 5 percentage points (p=0.002). These exceptions occur because US and Canadian startups, which are more likely to be local to judges since the majority of our data are from US startups and judges, have higher user traction and financing.

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

This difference in traction suggests that controlling for differences in startup quality will be crucial. Table 3.1 also reveals that judges are less likely to recommend foreign startups and rate them as lower quality.

Table 3.1: Summary statistics at the evaluation level

	Local Startup					Foreign Startup					
	Judge-Startup from the Same Region					Judge-Startup from Different Region					Local-Foreign
	No. Obs.	Mean	SD	Min	Max	No. Obs.	Mean	SD	Min	Max	Diff. in Means
Judge Score Measures											
Composite Score	7232	0.01	1.01	-3.31	2.36	9107	-0.12	1.05	-3.31	2.36	0.13***
Overall Raw Score	7706	2.92	1.16	0.00	5.00	9902	2.75	1.13	0.00	5.00	0.16***
Recommend	7706	0.61	0.49	0.00	1.00	9902	0.56	0.50	0.00	1.00	0.05***
Subscore: Customer Needs and Acquisition	7692	6.25	1.85	1.00	10.00	9833	6.06	1.92	1.00	10.00	0.18***
Subscore: Customer Pain and Solution	7694	6.82	1.84	1.00	10.00	9840	6.63	1.95	1.00	10.00	0.18***
Subscore: Financial Business Model	7675	5.72	1.98	1.00	10.00	9787	5.53	2.07	1.00	10.00	0.19***
Subscore: Industry and Competitor	7690	6.11	1.85	1.00	10.00	9827	5.93	1.94	1.00	10.00	0.17***
Subscore: Overall Impact	7686	6.21	1.93	1.00	10.00	9820	6.03	2.00	1.00	10.00	0.18***
Subscore: Regulation and IP	7261	5.91	2.15	1.00	10.00	9175	5.65	2.25	1.00	10.00	0.27***
Subscore: Team and Advisors Investors	7678	6.51	2.01	1.00	10.00	9805	6.31	2.09	1.00	10.00	0.20***
Startup Quality Measures											
Log Pre-Accelerator Total Page Visits	3917	1.37	2.77	0.00	12.50	5816	1.46	2.88	0.00	12.50	-0.09
Log Pre-Accelerator Financing	7706	0.45	1.41	0.00	6.03	9902	0.41	1.33	0.00	6.03	0.04
Log Post-Accelerator Total Page Visits	7706	2.87	3.52	0.00	12.82	9902	2.93	3.61	0.00	12.82	-0.06
Log Post-Accelerator Financing	7706	0.30	1.10	0.00	5.95	9902	0.25	0.98	0.00	5.92	0.05**
Has User Traction	7706	0.59	0.49	0.00	1.00	9902	0.53	0.50	0.00	1.00	0.06***
Has Financing	7706	0.12	0.33	0.00	1.00	9902	0.11	0.32	0.00	1.00	0.01

Notes: The table reports descriptive statistics for the sample of 17,608 startup-judge pairings from the 2017 and 2018 global rounds.
* p<0.05 ** p<0.01 *** p<0.001

3.5 Empirical specification

To assess whether judges systematically give lower or higher scores to foreign startups, we fit the following model (Li, 2017; Malloy, 2005):

$$score_{ijt} = \alpha +_{ij} + judge_{jt} + \mu_{it} + \varepsilon_{ijt} \quad (1)$$

Where $score_{ijt}$ is either a z-scored average or a binary variable on whether judge j recommends startup i to the next round in year t. $foreign_{ij}$ 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.

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

($judge_{jt}$), so that our analysis focuses on judge evaluations of startups within the same year.

We also use several fixed effects to account for differences in startup quality across regions and countries. As with our judge fixed effects, we interact all our fixed effects with the program year to account for the fact that startups can apply in multiple years. In our first specification, μ_{it} in Equation 1 is equal to startup region-year fixed effects. These fixed effects measure startup evaluations within a particular region (e.g., Europe, Latin America, Israel, and Northern America) in each year to account for differences in quality across regions.

We then tighten our specification, with μ_{it} equal to startup country-year fixed effects. These fixed effects focus our analysis on startup evaluations within a particular country in a year to account for differences in quality across countries (within regions). These fixed effects allow us to account for quality differences between, for example, a UK-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), so that μ_{it} is equal to individual startup-year fixed effects. These fixed effects enable us to account for differences in individual startup quality within countries. We cluster robust standard errors at the judge and startup levels. β indicates that the judges discount or boost foreign startups relative to local ones. Returning to the two-by-three in Figure 3.1, this rules out cells A and E, where judges are unbiased and either uninformed or informed. However, a significant β can be consistent with the remaining cells.

To assess whether foreign discounting is driven by judges being better at evaluating local startups or because of bias, we estimate a model similar in spirit 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 3.1, including whether judges are informed and biased (cell F), informed only about local startups and biased (cell D), informed only about local startups and unbiased (cell C), or uninformed about all startups and biased (cell B).

$$score_{ijt} = \alpha + \beta foreign_{ij} + \delta performance_i + \varphi foreign_{ij} * performance_i + judge_{jt} + startupcountry_{it} + \varepsilon_{ijt} \quad (2)$$

Where $performance_i$ indicates logged page visits for the startup one year (for the 2018 cycle) or two years (for the 2017 cycle) after the program. In addition to β, we also are interested in δ and φ. A positive and significant δ indicates that judges are able to discern winners from losers among startups overall. If δ is positive, then future performance correlates with judge scores. A negative and significant φ indicates that judges are less sensitive to the quality of foreign versus local startups. A concern with our approach is that the accelerator itself impacts the post-accelerator performance of startups, which confounds the judges' selection of startups with the treatment effect of the accelerator. Further, this treatment effect might differ for startups from different regions. To account for these possible treatment effects, we control for startups' participation in the accelerator program and this participation interacted with whether the startup is foreign or local to the judge.

3.6 Results

3.6.1 Are foreign judges actually randomly assigned?

Our ability to measure the presence and impact of foreign discounting hinges on the assumption that startups and judges are randomly assigned. To check random assignment, we use chi-squared tests shown in Tables 3.2–3.3. These chi-squared tests allow us to measure whether there is a 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 difference (p=0.809) 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 based on 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 (p=0.006), a fairly small share (0.26 percent) of our sample, representing 25 judge-startup pairings out of 9,733 total in 2018. However, when we take out Israeli judges, we see a similar situation as in 2017 (p=0.256). The distribution is again consistent with random assignment. Our results hold if we include or exclude these Israeli judges from our data. These patterns suggest that the natural experiment that is at the heart of our story is, in fact, randomized.

Table 3.2: Chi-squared table for the 2017 global round showing distribution of judges to startups is no different than what we would expect from random chance

Pearson chi2(4) = 1.5988 Pr = 0.809				
	Judge Subregion			
Startup Subregion	Europe	US & Canada	Israel	Total
Europe	229	791	206	1,226
	239.3	*783.7*	*203*	
US & Canada	1,008	3,322	860	5,190
	1,013.00	*3,317.60*	*859.4*	
Israel	300	921	238	1,459
	284.8	*932.6*	*241.6*	
Total	1,537	5,034	1,304	7,875

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

Table 3.3: Chi-squared table for the 2018 global round showing distribution of judges to startups is no different than what we would expect from random chance when excluding outliers

With Israeli Judges: Pearson chi2(9) = 22.9832 Pr = 0.006 Without Israeli Judges = Pearson chi2(6) = 7.7603 Pr = 0.256				
	Judge Subregion			
Startup Subregion	Europe	Latin America	US & Canada	Israel
Europe	568	153	1,389	25
	595.8	*177.7*	*1,348.20*	*13.4*
Latin America	705	213	1,539	11
	688.7	*205.4*	*1,558.40*	*15.5*
US & Canada	1,406	432	3,134	23
	1,393.90	*415.7*	*3,154.10*	*31.3*
Israel	37	12	84	2
	37.7	*11.2*	*85.2*	*0.8*
Total	2,716	810	6,146	61

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

3.6.2 Is there foreign discounting of startups?

We now turn to whether judges discount foreign startups. In Table 3.1, 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 startup evaluations where the judge and startup do not match geographies versus those that do.

Figure B.2.1 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 different from one another (p=0.000). However, this graph may reflect the fact that most judges in our sample are US-based. Thus, startups that are foreign are more likely to be those that are non-US based, and non-US based startups may be worse quality on average than US-based firms.

We account for these regional quality differences in our regression models. To begin, Column 1 in Table 3.4 shows that when we only control for judge-year fixed effects, judges give 0.2 standard deviation lower scores to foreign vs. local startups (p=0.000). Column 2 adds in startup region-year fixed effects to account for regional variations among startups. Our estimate shrinks to -0.06 (p=0.002). Columns 3–4 add more restrictive startup country-year and startup-year fixed effects, respectively. Our results are virtually identical. These results show that there is little in the way of systematic differences between startups within regions. Overall, Table 3.4 shows that regional differences in startup quality account for about two-thirds of the foreign discounting effect, and judges account for one-third. A potential concern with these estimates is that it could be that only US judges are biased against foreign (i.e., non-US startups). While judge fixed effects will account for differences in harshness among US and other judges, we also show in Figure B.3.1 and Table B.3.1 that US, EU, and Israeli judges are all more likely to recommend local over foreign startups. This suggests that our findings are not idiosyncratic to US judges.

Column 5 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 benchmark the judge bias effect against the effect of key startup milestones. The home bias effect (-0.06, p=0.001) is about 30 percent of the size of a startup having user traction and about 8 percent of the size of the effect of a startup having raised a round of financing at the time of the application. The fact that the whim of a judge matters about one-third as much as having some traction suggests that the foreign bias effect is non-trivial. We confirm that the regression results are not driven by differences in the probability of judges giving incomplete subscores to foreign relative to local startups. Table B.5.1 shows that judges are equally as likely to give foreign and local startups incomplete subscores.

Table 3.4: Regressions showing that judges give lower scores to startups from outside their home region even when we control for judge and startup fixed effects

	(1)	(2)	(3)	(4)	(5)
	Judge's Total Score				
Foreign Startup	-0.204***	-0.061**	-0.061**	-0.061***	-0.058**
	(0.021)	(0.020)	(0.020)	(0.016)	(0.018)
Has Traction					0.201***
					(0.029)
Has Financing					0.712***
					(0.023)
Observations	16,320	16,320	16,320	16,264	16,320
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

Of the 17,608 recommendation evaluations in our data, for 16,339 (93 percent) we have complete subscore information. Standard errors (shown in parentheses) are clustered at the judge and startup levels. Fixed effects shown below observations.
* $p<0.05$ ** $p<0.01$ *** $p<0.001$

Table 3.5 is similar to Table 3.4, but it 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 (p=0.000) before accounting for startup quality differences. This coefficient remains significant and negative, but it falls to 4 percentage points (p=0.000) when accounting for startup region-year fixed effects (Column 2), startup country-year fixed effects (Column 3), and startup-year fixed effects (Column 4), indicating that judge preferences account for about 40 percent of the foreign bias effect.

Further, this foreign discounting result is robust to alternative measures of foreignness, different sub-sample restrictions, and regional quality controls. We show in Appendix Tables B.4.1 – B.4.2 that the foreign discounting

Table 3.5: Regressions showing that judges are less likely to recommend startups from outside their home region even when we control for judge and startup fixed effects

	(1)	(2)	(3)	(4)	(5)
	Judge Recommends Startup?				
Foreign Startup	-0.091***	-0.036***	-0.038***	-0.039***	-0.037***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Has User Traction					0.088***
					(0.015)
Has Financing					0.345***
					(0.010)
Observations	17,593	17,593	17,593	17,590	17,593
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. Fixed effects shown below observations.
* p<0.05 ** p<0.01 *** p<0.001

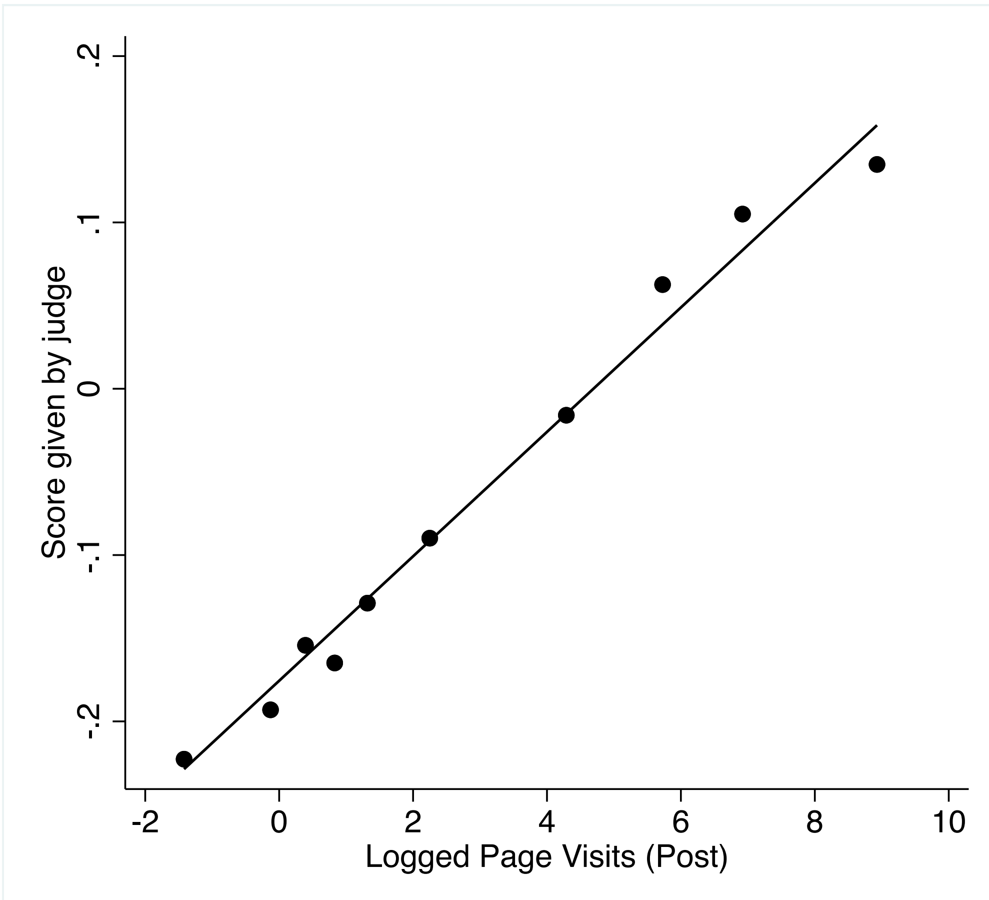
effect holds when we use raw weighted and non-weighted measures of judges' final recommendation score. In Appendix Table B.6.2, we also show that our foreign discounting effect holds when we measure foreignness using (1) geographic distance between the judge's HQ region and the startup's country of operation, (2) whether the region is explicitly mentioned in the startup's application text, and (3) how "regional" a startup appears based on the text in its application. In Appendix Table B.7.1, we further demonstrate that the foreign discounting effect holds when we exclude investor judges who might prefer local startups because they represent a more promising investment opportunity than more distant firms. We also show in this section that our results hold when we exclude Latin American startups, which suggests that differences in English ability and training do not account for our result (Table B.7.2). In Appendix Table B.9.1, we show that the foreign discounting effect holds when we directly control for measures of a country's startup quality, including GDP per capita, patent applications, venture capital availability, and hub status. We also show our results hold when we directly control for founder quality measures, including whether the team has a PhD, MBA, or elite university affiliation (Table B.9.2). In Appendix Table B.10.1, we confirm that the foreign bias result holds when judge-startup industries match or not. We further show that the results hold when we control for whether the startups are headquartered in a hub (Table B.11.1). Finally, while the focus of our paper is on isolating discounting in the first stage of the accelerator evaluation and screening process, we also show that our findings generalize when estimated on a larger sample of accelerator data in which judges are far from randomly assigned. In Appendix Table B.12.1, we show that our findings hold across all rounds and years of the program and that foreign bias occurs even in the later rounds of the program when judges interview and evaluate the startup team in person.

Together, these results reveal that judges consistently give lower evaluation scores to foreign versus local startups.

3.6.3 Is foreign discounting the result of judges being better evaluators of local startups?

We now turn to testing if this foreign bias is the result of differences in judges' expertise or is rooted in a preference for local vs. foreign firms. To begin, we assess whether judges can select winners from losers amongst all startups, no matter their origins. Figure 3.2 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 netting out judge-year and startup country-year fixed effects, as well as startups' participation in the accelerator.[12] The graph shows that better-performing startups are given higher scores. Judges can pick winners from losers in the full sample.

Figure 3.2: Binscatter showing that judges give higher scores to startups with more growth one- to- two- years after the accelerator program



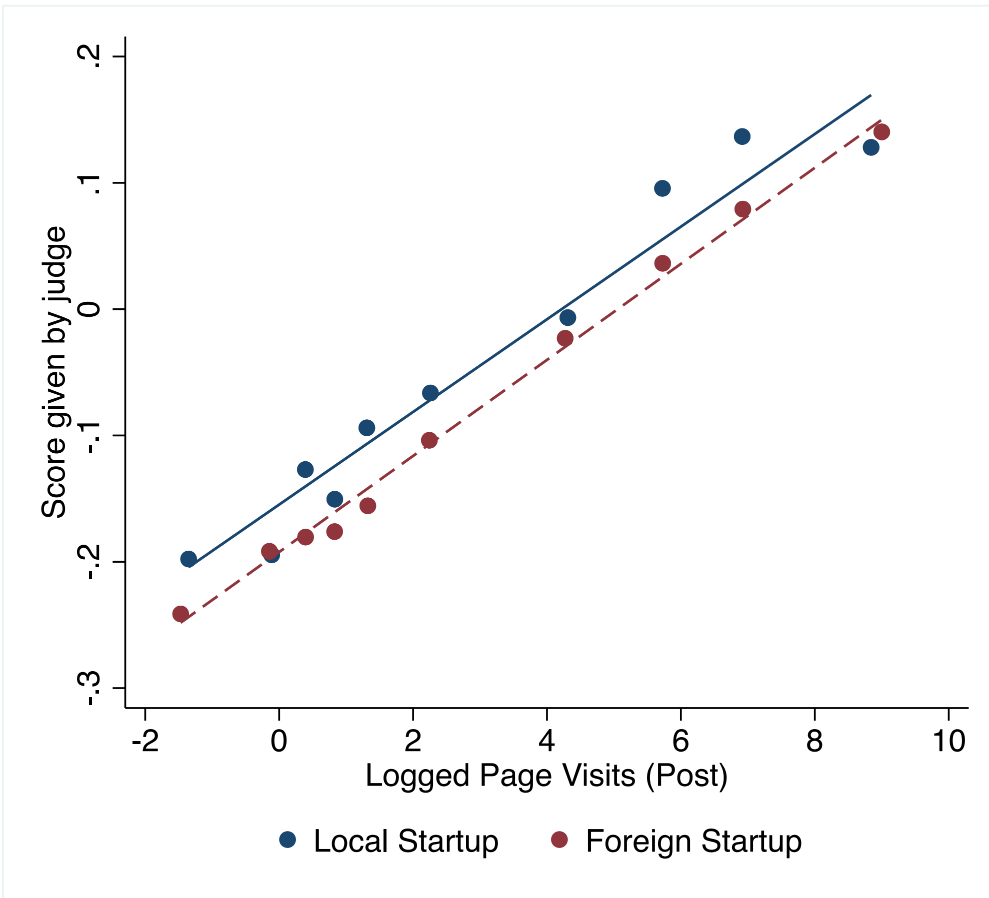
Binscatter after accounting for judge x year and startup country x year fixed effects and accelerator participation. We use 10 bins.

To what extent is this ability to detect the quality of startups driven by evaluations of local startups? To answer this question, in Figure 3.3, we split the evaluations into startups that are foreign to the judge (dotted line) and startups that are local to the judge (solid line). We see that both lines have a positive slope, suggesting that judges can separate high-potential startups from those destined to fail. The fact that the solid line depicting local startup evaluations is above the dashed line across the quality spectrum suggests that judges give an across-the-board

[12]Startups may participate in any of the four regions of the accelerator (US, Israel, Latin America, or Europe) and may not be necessarily foreign to this location, even if they were foreign to a judge's location in the initial screening.

penalty to foreign startups no matter their quality. Further, the solid and dashed lines are similarly sloped. It does not appear that judges are better able to pick winners from losers among local versus among foreign startups. Figure 3.3 matches cell F in Figure 3.1 and so suggests that judges are informed about local and foreign startups, but are simply biased against foreign firms.

Figure 3.3: Binscatter showing that judges give higher scores to startups with more growth one- to- two- years after the program, but they consistently discount foreign startups no matter their eventual success



Binscatter after accounting for judge x year and startup country x year fixed effects and accelerator participation. We use 10 bins.

We next turn to regressions to further confirm that judges are not any better at evaluating local startups. Column 1 in Table 3.6 reveals that there is no difference in the relationship between startup quality and judge scores by local startup origin, as seen in the coefficient on the interaction term between foreign startups and logged post-page visits ($foreign_{ij}xperformance_i$) (p=0.921). Consistent with Figure 3.3, we do indeed find that judge scores correlate with startup quality, shown by the positive coefficient on the main effect for logged post-accelerator page visits. In Column 2, we control for accelerator participation and the possibility that accelerator participation matters more for foreign firms. While accelerator participation has a positive effect on post-accelerator startup page visits, and while this effect is slightly greater for local startups, it does not meaningfully account for the foreign discounting effect nor a judge's ability to evaluate startup potential. We also confirm that the result holds if we exclude startups that participated in the accelerator all-together as shown in Column 3. We get similar results when using logged financing 6 months after the program as our measure of startup quality, as shown in Columns 4–7. There is no difference in the relationship between startup quality and judge scores by local startup origin, no matter if we control for or exclude startups who participated in the accelerator.

As with our foreign bias results, our findings here appear quite robust. Our findings hold no matter the measure of startup quality that we use. In Appendix Table B.13.1, we show that our findings hold when we use pre-accelerator page traction, page visits, and financing as our quality measures. Our findings also hold if we instead use post-accelerator valuation, employee, revenue growth, and a composite index measure of startup success (Table B.13.2). The findings also are consistent if we split our sample by foreignness: the r-squared statistics are similar for foreign and local startup samples when we regress judges' scores on startup quality and quality on score, as shown in Appendix Table B.14.1. In Appendix Table B.17.1, we show that judges can detect the quality of local and foreign startups with similar precision no matter if their region is explicitly stated or not in the application, suggesting that the startup location provides marginal (if any) informational value to judges.

3.6.4 Reconciling Results with Prior Work

These results suggest that judges can detect the quality of all startups with relatively equal precision, though they discount foreign startups, reflecting cell F in Figure 3.1. Yet, prior work either suggests that judges cannot detect the quality of startups at all and instead experiment with small investments, as shown in cells A and B (Ewens et al., 2018; Kerr et al., 2014b; Nanda et al., 2020; Scott et al., 2020) or have a local information advantage as shown in cells C and D (Coval and Moskowitz, 2001; Malloy, 2005). Why do our results contrast with this prior work?

Crucially, our sample differs in two important respects from this past research. First, by focusing on the earliest screening stage of the evaluation process, judges evaluate a much broader range of startups. In contrast to the global and heterogenous sample of startups analyzed by our accelerator's judges, the sample in Scott et al. (2020)'s study are all startups with founders from MIT. Samples evaluated by venture capital research too, tend to comprise pre-selected and high-quality Silicon Valley founders. This suggests that the judges in our sample may well be more informed because they are evaluating startups that vary more in their quality than the already

Table 3.6: Regressions showing judges (1) give higher scores to more successful startups, (2) are equally good at evaluating success for local and foreign startups alike, and (3) still discount foreign startups

	(1)	(2)	(3)	(4)	(5)	(6)
			Judge's Total Score			
Foreign Startup	-0.065**	-0.056*	-0.039	-0.053**	-0.047*	-0.040*
	(0.025)	(0.024)	(0.025)	(0.019)	(0.020)	(0.020)
Log Post-Accelerator Page Visits	0.050***	0.036***	0.043***			
	(0.004)	(0.004)	(0.004)			
Foreign Startup * Log Post-Accelerator Page Visits	0.000	0.003	-0.001			
	(0.005)	(0.005)	(0.005)			
Log Post-Accelerator Financing				0.170***	0.027*	0.178***
				(0.008)	(0.012)	(0.040)
Foreign Startup *Log Post-Accelerator Financing				-0.010	0.009	-0.026
				(0.011)	(0.015)	(0.055)
Accelerator Participation		0.682***			0.701***	
		(0.032)			(0.042)	
Foreign Startup * Accelerator Participation		-0.109**			-0.109*	
		(0.041)			(0.055)	
Observations	16,320	16,320	14,475	16,320	16,320	14,475
Judge x Year	Yes	Yes	Yes	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes	Yes	Yes	Yes
Startup x Year	No	No	No	No	No	No
Accelerator Participation	Yes	Yes	No	Yes	Yes	No

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

pre-selected firms analyzed in prior work. Second, our sample is dominated by globally oriented technology startups. Indeed, every startup in our sample applied to the global round of an online accelerator, suggesting in their choice that they are likely less "localized" than most firms and especially less localized than the non-traded goods-producing, small, or remote firms analyzed in Coval and Moskowitz (1999; 2001).

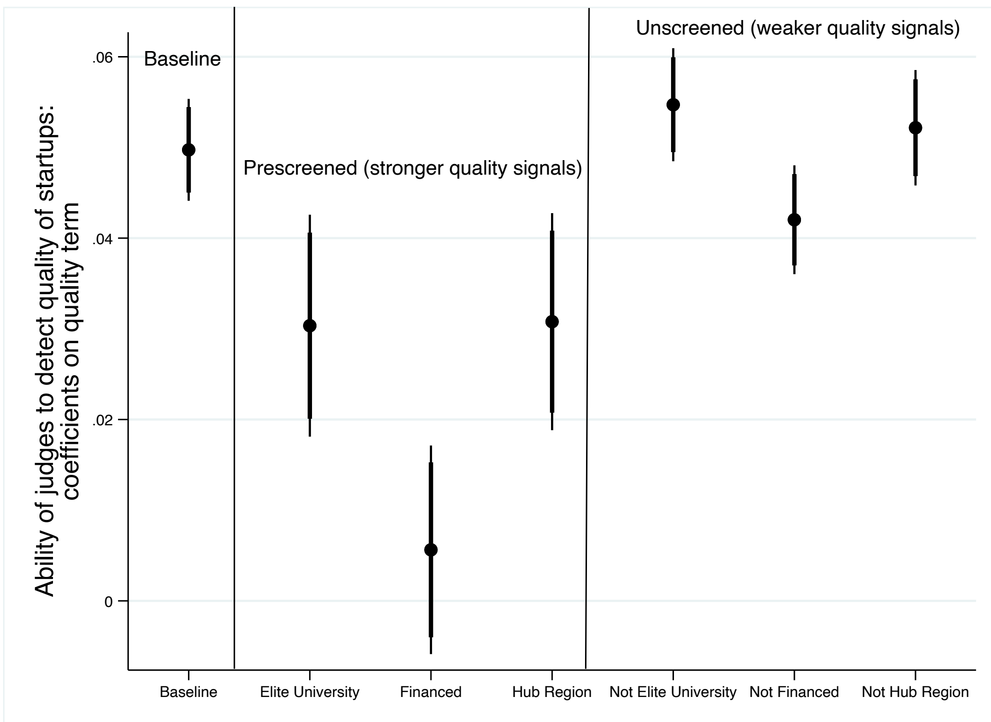
To test our explanation for the first difference, that our pool is much more diverse than prior research, we split our sample into startups with founders affiliated with an elite university (based on the application text), whether the startup is financed at the time of application, and whether the startup mentions being part of a hub city in its application text. These splits let us separate startups that have already been screened (founders affiliated with elite schools, already financed, and startups that have decided to work from a hub) to those that have not. For each sample, in Appendix Table B.15.1, we show regression results similar to Table 3.6.[13] Figure 3.4 shows coefficient plots from these regressions, with the estimates reflecting how much of the judge's score is responsive to differences in startup quality. Consistent with our arguments, we find that judges are worse at picking winners from losers among the pre-screened samples. The coefficients in the pre-screened pools are closer to zero, suggesting scores are less reflective of differences in quality. Thus, our results do not contradict prior works, such as Ewens et al. (2018), Kerr et al. (2014b), Nanda et al. (2020), and Scott et al. (2020). Instead, these results show that accelerators may be better and have an easier time screening good from bad startups because they cast wider nets.

Intriguingly, we also find in Appendix Table B.15.1 that our foreign bias estimate might increase when judges evaluate pre-screened startups, with the foreign discounting coefficient being larger for elite university-affiliated, financed, and hub-affiliated startups than those that are not. This suggests that when judges assess startups that have already met a higher quality threshold, they might rely more on the startup's location. Without easily detectable quality differences, judges may default to picking between startups based on their location.

Further, once the accelerator does narrow down to the approximately 120 startups that it accepts into the final program, judges' ability to detect quality may decline making "spray and pray" or other experimentation techniques employed in venture capital (Ewens et al., 2018) valuable to learn about startup quality. Indeed, in Appendix Table B.16.1, we show that judges' ability to detect the quality of startups declines when evaluating companies accepted into the accelerator program relative to those in the top-of-the-funnel global round.

[13] These regressions do not include an interaction term between foreign startup and the quality term because we are interested here in isolating the ability of judges to detect the quality of startups overall (as opposed to their relative ability to detect the quality of local versus foreign firms, which we later evaluate in Appendix Table B.18.2).

Figure 3.4: Coefficient plot of judge sensitivity to the quality of startups across sub-samples of startups. The bars show 90 percent and 95 percent confidence intervals.



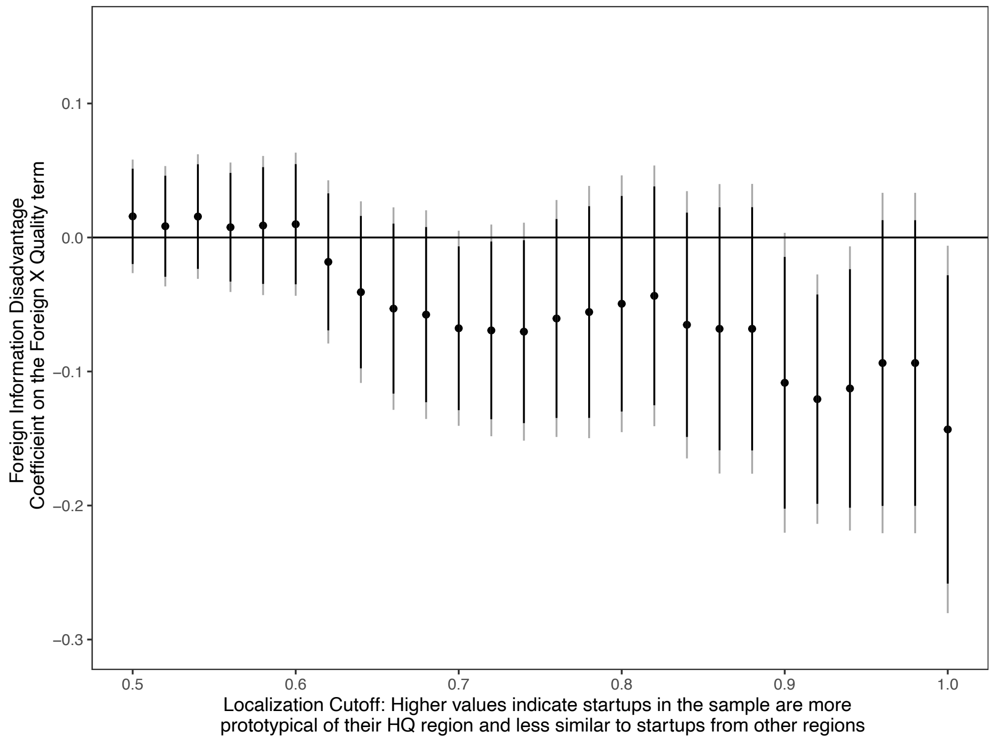
The plot shows coefficients from regressing judges' scores on post-accelerator log page visits, controlling for participation in the program across different sub-samples of startups. It shows 90 and 95 percent confidence intervals.

To test our second discrepancy, why judges lack a local information advantage in our setting, we again split our sample. This time we restrict our sample to startups that are particularly "localized" following Coval and Moskowitz (1999; 2001)'s approach, as it is these firms where local information advantage is likely to matter. To measure a startup's localness, we use the application text and exploit the fact that some words are often used by startups from particular regions. For example, terms like "Jerusalem" and "IDF" are particularly used by Israeli startups and not startups from other regions. Appendix Table B.6.1 provides details. Specifically, for every word in our corpus, we calculate the log-odds ratio that is used in one particular region versus any other region. By aggregating these word-level log-odds ratios, we can calculate a standardized score for how "North American," "Israeli," "Latin American," and "European" each startup application is. To get our final sample of "localized" startups, we restrict our sample to firms where (1) the startup's home region score is greater than X standard deviations and (2) the startup's region score is less than X standard deviations from all other non-home regions. We set X to be 0.5, 0.75, and 1 standard deviations, each reflecting an increasingly localized sample of startups. These two restrictions ensure that the startup is both very localized to its own home region, but also does not happen to read like it is from any other region.

In Appendix Table B.18.2, we replicate our Table 3.6, but only including startups that meet these localization cutoffs. The models include our measures for whether a startup is foreign, our proxy for startup quality, and an interaction term between the two. If judges are worse at evaluating foreign startup quality, the coefficient on quality should be positive and the interaction term negative. Indeed, as Table B.18.2 shows, as we restrict the sample to the most localized startups, we see that judges remain able to detect quality differences, but only for local startups.

To shed further light on this pattern, Figure 3.5 plots the key coefficient, the interaction term between startup quality and whether the startup is foreign, for "localization" cutoffs ranging from 0.5 to 1 standard deviation. If judges are worse at evaluating foreign startups when the sample of firms only includes very localized firms, then the estimates should gradually become more negative. Indeed, the plot shows exactly this, with the interaction term dropping from 0 to a statistically significant negative estimate at about 0.75 standard deviation. Consistent with the idea that most startups are globally focused in our sample, just under 5 percent of startups in our sample are "local enough" to meet the 0.75 cutoff.

Figure 3.5: Coefficient plot of judge local information advantage across different "local" sub-samples of startups. The bars show 90 percent and 95 percent confidence intervals.



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

Our finding thus far show that judges give lower scores to foreign startups on average. However, it is possible that this discounting has little impact on which startups move on to the next round. For example, perhaps judges discount high-quality foreign startups who, though rated somewhat lower, still end up selected for the next round, no matter the discount. 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 decision.

To estimate the number of "missed foreign startups," for whom foreign discounting does make a marginal difference, we estimate what judge decisions would be if we removed their foreign bias. Unfortunately, each judge does not tell us nor the accelerator how foreign-biased they are. Fortunately, the fact that nearly all judges evaluate multiple foreign and multiple local startups lets us estimate a "foreign bias" fixed effect for the vast majority of judges in our sample. This judge-level effect is simply the difference in the average rating a judge gives to foreign versus local startups. Moreover, we can estimate this judge-level bias while simultaneously estimating startup fixed effects. This lets us isolate bias net of any average quality differences between startups that are foreign or local to the judge. These estimates of an individual judge's bias allow us to then "debias" each judge's scoring and so test if different startups would have made it to the second round if selection relied on these debiased scores instead of the judge's actual decisions.

Specifically, we first regress each judge's total score on the fixed effects of each startup application, a fixed effect that absorbs each judge's evaluation of local startups, and a fixed effect for each judge's evaluation of foreign startups. In comparison to when we include judge fixed effects in Equation 1, we are not merely accounting for each judge's overall "harshness," but instead accounting for each judge's individual harshness towards foreign and local firms. We then use these fixed effects, instead of an estimated "foreign discounting" coefficient, to unpack and address judge bias. Consistent with our primary findings that judges discount startups by 0.06 standard deviations, we find the average judge fixed effects for the foreign startups they evaluate is 0.07 standard deviations lower than for the local startups they evaluate. We also find that some judges appear especially biased, with the 25th percentile judge discounting foreign startups -0.35 standard deviations more than local startups and the 5th percentile -0.8 standard deviations.

We then use these individual fixed effects to "debias" each judge's score. For example, imagine a judge is relatively harsh, giving foreign startups they evaluate scores that are -0.5 lower than the average judge and local startups -0.3 lower, even after accounting for startup fixed effects. In this case, we would estimate that this judge has an individual foreign bias of -0.2 (-0.5 minus -0.3). To "debias" this judge, we would add 0.2 to the score for each foreign startup they evaluated. More generally, we repeat this procedure for each judge to account for the distribution of biases in our data. As mentioned above, this offset is net of startup quality since we include fixed effects for each startup when estimating the judge bias fixed effects. To convert these scores into the recommendations the accelerator uses to select startups, we have our "debiased" judges select the same number of recommended startups as we observe in the actual data, but we select those with the highest scores according to our debiased estimates. We use this assumption of the same number of recommendations because judges can recommend as few or as many startups as they would like. There is no numerical score cutoff that leads to a recommendation. Finally, we follow the accelerator's rules and mark a startup as moving on to the next round if 50% or more of the judges recommend the startup.

Using our debiased scores, we find that removing home bias would lead to 148 startups moving from non-recommended to recommended and 86 moving in the other direction. Together home bias appears to lead to mistaken decisions—if the goal is to only select the highest quality startups—for 234 startups, just over 6% of applicants. Moreover, as we show in Appendix Figure B.19.1, these "missed" startups have promise. The estimated quality of these "missed" startups is similar to the majority of actually selected startups. While the highest-quality startups make it to the next round regardless, home bias appears to cause the accelerator to miss out on a non-trivial number of promising ventures.

3.7 Conclusion and implications

We find that judges can equally discern the quality of local and foreign startups with similar ability in the earliest stage of the evaluation process. 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-third of the effect of having some user traction or a tenth of the effect of going from no financing to having some venture financing. Back-of-the-envelope estimates suggest that this bias potentially excludes about 1 in 20 promising entrepreneurial ideas. These results reveal that judges are informed about the quality of both local and foreign startups, but they still discount foreign firms.

However, we also find that observed judge behavior depends on the pool of startups that judges are tasked with evaluating. Judges are worse at evaluating quality when the startups have already been screened and when foreign startups have more localized business models. However, as accelerators increasingly consider a wider pool of global ventures, and startups continue to adopt standardized technology-driven business models, our findings suggest accelerators and similar business plan competitions may increasingly play an important, if biased, role in screening early-stage startups (Chatterji et al., 2019; Haefliger et al., 2008; Howell, 2020). Our findings contrast with past work showing that venture capital firms struggle to screen promising ventures from bad ideas (Kerr et al., 2014a; Nanda et al., 2020; Scott et al., 2020). That said, our findings that observed judge effectiveness and bias depend on the pool of startups evaluated reconcile this difference. Venture capitalists likely struggle to predict success because they evaluate a pre-screened pool of startups, screening that is increasingly first done by accelerators that can separate winners from losers at the very earliest stages of the entrepreneurial process.

Our results also highlight how accelerators increasingly complement venture capitalists' "spray and pray" approaches (Ewens et al., 2018). Specifically, our findings reveal that accelerators effectively screen the thousands upon thousands of startup applications they receive. When they can no longer pick winners from losers, they can then refer the startups they accelerated to venture capitalists who can use "spray and pray" approaches to learn about startup quality through sequential experimentation (Ewens et al., 2018; Hallen et al., 2023; Howell, 2020). However, this linkage also suggests that foreign bias by accelerator judges may lead investors to pass over promising foreign startups since those foreign startups never even make it into the accelerator to begin with. Even if investors are not home-biased (c.f., Lin and Viswanathan, 2016), bias by judges, mentors, and other gatekeepers earlier in the entrepreneurial process may well explain why startups in many parts of the world fail to scale (Wright, 2023).

This logic also suggests that the foreign bias we identify here may impact the direction of innovation. If accelerators pass over startups from remote regions, which are more likely to be foreign to accelerators, they reduce the probability that innovations addressing the needs of those markets will survive and grow. Even if these foreign startups employ globally standardized business models and practices, their innovations and target customers may still disproportionately benefit the home market. This distortion is similar to effects seen in studies of bias in gender and race contexts (e.g., Koning et al., 2020; 2021).

Turning to practice, our results also suggest that accelerators may benefit from opening their initial screening processes to startups more globally, given their ability to discern startup quality at the top-of-the-funnel, no matter the startup's location. Accelerators have the potential to identify firms that might not have received any support otherwise. That said, later rounds of evaluation, where there is likely an opportunity to use local references and networks, may still require localized capabilities to best pick which global startups are most promising. Crucially, however, any global approach depends on accelerators revising their processes to reduce the impact of bias—in this case, foreign bias—that all too often enters the evaluation of diverse and heterogenous samples (Brooks et al., 2014; Cao et al., 2021).

For entrepreneurs, our findings suggest caution when acting on feedback from accelerators. While past studies show that accelerators, by providing signals on the startup's quality to the entrepreneur, are an important source of learning (Cohen et al., 2019a; Howell, 2020; Lyons and Zhang, 2018; Yu, 2020), such signals may lead entrepreneurs astray when they originate from a non-representative sample or from biased actors (Cao et al., 2021). In our case, judges' foreign discounting implies that the signals from accelerators may be distorted for firms from regions under-represented amongst accelerator judges. Given the increasingly recognized importance of entrepreneurial learning in startup performance (Koning et al., 2022), the fact that there is less bias in local signals also provides a novel mechanism to explain why ventures tend to perform better when located in a founder's native region (Dahl and Sorenson, 2012).

Overall, we find that startups face a "liability of foreignness" (Zaheer, 1995). Notably, we do not find that judges face a disadvantage in evaluating foreign startups. Instead, we find that judges can discern the quality of startups across regions in the early screening stage. This may be because technology and business models have standardized into a "playbook" that is comparable across countries, for example, with the proliferation of codified management (Bloom and Reenen, 2007; Chatterji et al., 2019) and technology practices (Haefliger et al., 2008). Further, the existence of such a playbook may reduce the need for private information (Coval and Moskowitz, 2001; Malloy, 2005) or contextual intelligence (Khanna, 2014) to evaluate foreign opportunities. Future work should continue to explore how the changing nature of startups and their strategies impact gatekeepers' ability to screen promising ventures from bad ideas.

4

Open Source Software and Global Entrepreneurship

4.1 Introduction

Open source software (OSS) became mainstream with little fanfare (DiBona et al., 2005; DiBona and Ockman, 1999).[1] Mainstream software development and applications widely employ open source software today. Two decades of experience have routinized resource sharing (Lakhani and von Hippel, 2003; Lakhani and Wolf, 2003) and communications between programmers with different backgrounds (Aksulu and Wade, 2010, Von Krogh et al., 2012). Open source reduces the time to develop innovative software modules, eliminates hassles from negoti-

[1] Co-authored with Frank Nagle and Shane Greenstein

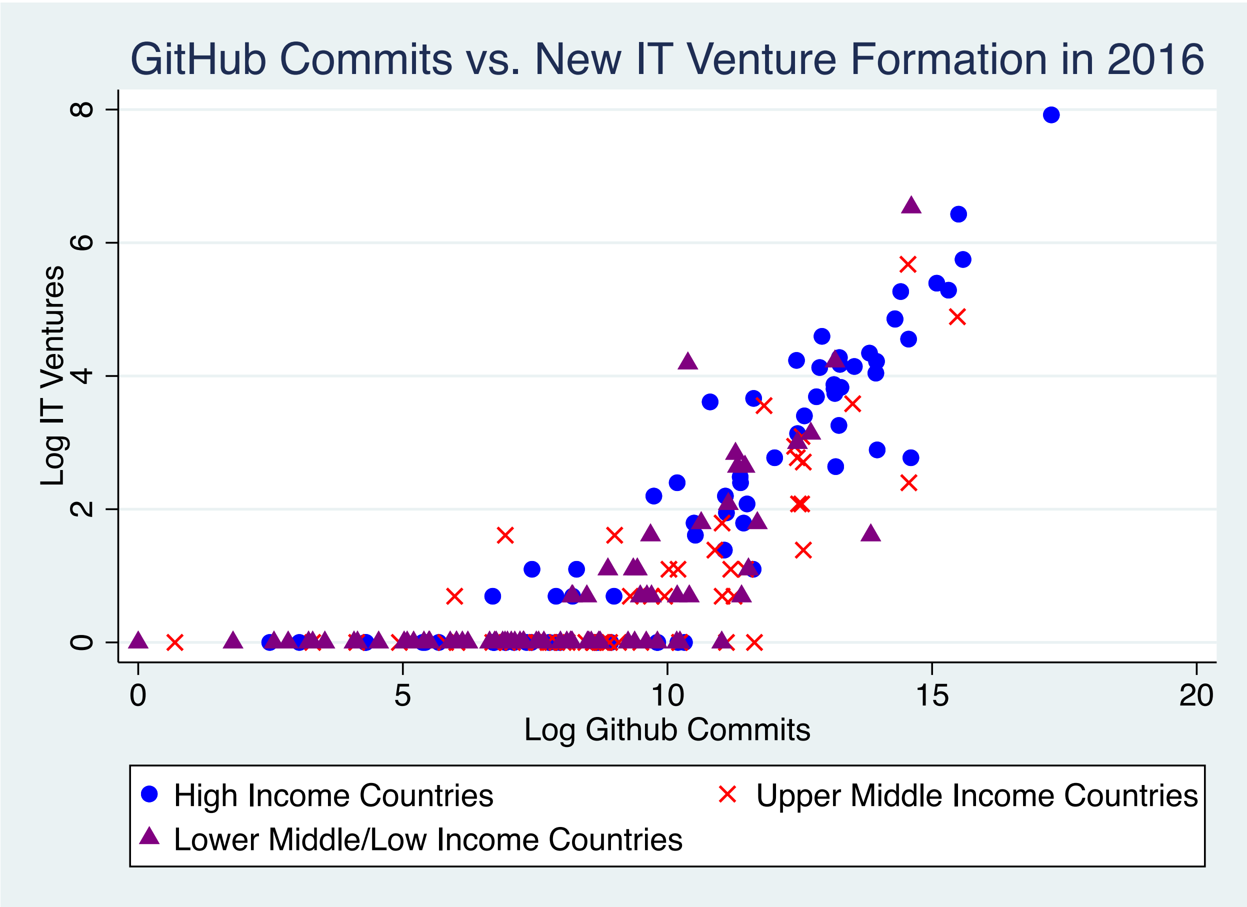
ating intellectual property, and reduces friction associated with raising capital for software development (Nagle, 2019b; Wen et al., 2016). Today, open source is an essential component of artificial intelligence, web-enabled commerce, and most big data software.

While the benefits of participating in open-source communities have been documented within high-income countries (Lerner and Schankerman, 2010; Nagle, 2018), the focus on developed economies limits the observation and ignores the state of global labor markets. For example, programmer workforces have grown in the middle-income countries of Central Europe and Asia, and account for tens of billions of dollars of services a year (Agrawal et al., 2016; Barach et al., 2020; Stanton and Thomas, 2015). Just like their counterparts in developed economies, programmers around the globe employ open-source tools, speak the vocabulary of open-source, and interact with open-source libraries (Nagle et al., 2020). Further, the dynamism and accessibility of open source could represent an opportunity for low- and middle-income countries to reach the technological frontier more quickly than if they needed to develop such software from scratch or obtain it from costly sources, lowering the challenges of "catching up" in areas where knowledge about software and related business processes fosters capabilities in new geographies (Lee and Lim, 2001).

Figure 4.1 can motivate why research on this topic should not be limited to only high-income countries. It shows a plot from 207 countries and illustrates the last year of data from our study, 2016. The figure shows the correlation between broad measures of open source participation and entrepreneurship activities in that country, shown in log-log form. We ask the reader to momentarily defer questions about definitions (which we address below) and focus on the forest and not the trees in the raw data. The figure illustrates that open source and entrepreneurship arise outside of high-income countries. While income clearly plays a visible role in the correlation, many questions remain unaddressed by a simple graph like this.

In this study, we consider four questions related to OSS and entrepreneurship across countries. We consider this relationship on two dimensions, rate and direction, consistent with long-running work on inventive activity. First, we ask whether more participation in open source in a country correlates with more entrepreneurship activity (a higher rate), and, next, whether the evidence is consistent with a causal interpretation. We pioneer the development of instrumental variables for this situation related to the supply and demand for OSS, as well as the networking of relationships in which it arises. We next investigate whether the evidence suggests participation in OSS in a country substitutes for human capital endowments and per capita income levels or complements those endowments (resulting in enhancing or diminishing the rate). We then consider whether OSS has related effects

Figure 4.1: GitHub commits vs. new IT venture formation 2016



on the direction of entrepreneurship, in particular considering whether OSS imparts its more mission-oriented and global-oriented ethos on these new ventures. Relatedly, we also ask whether OSS influences the direction of ventures by leading to high-quality entrepreneurship, as proxied by financial outcomes.

To address these questions, the study pioneers a global approach for measuring open source participation and entrepreneurship activity for 2000-2016. While there are various ways to measure open source participation and entrepreneurship within the US, only some of the existing approaches provide a viable approach to measuring activity outside US borders and over time. We utilize data from GitHub, the largest repository of OSS in the world, which is widely adopted across countries. We match it to a measure of worldwide entrepreneurship sourced from Crunchbase. No other source provides a better-standardized proxy over time and across the globe.

We first establish a correlation between GitHub participation and entrepreneurship. This holds for two definitions of entrepreneurship, one that stresses all information technology entrepreneurship in a country and another that focuses on entrepreneurship-related directly to open-source software. This is unsurprising, but the result is reassuring because the magnitudes are plausible. For example, a one percent increase in GitHub commits (code contributions) in a given country in a year is associated with a 0.2–0.4 percent increase in information technology (IT) ventures and a 0.03–0.1 percent increase in OSS ventures in that country the following year—roughly 5–10 new IT ventures and 0.007–0.02 OSS ventures per year per country on average.[2] While not the focus of this study, multiple mechanisms may explain these effects. These include a reduction of both search costs for human capital and costs for the communication of knowledge, an increase in access to complementary assets, and a standardization of programming practices. Further, although these effects may seem large, it is important to note there is a wide variance when considering the existing endowments of countries.

We next find that the statistical relationship is stronger in countries with higher human capital and income endowments. The evidence supports the view that OSS complements existing endowments, and, importantly, no evidence supports the view that OSS substitutes for endowments. While seemingly straightforward, this finding is novel and informs policies for shaping the spread of OSS to areas other than the richest countries in the world.

Beyond influencing the rate of entrepreneurship, we also find that increases in OSS shape the direction of this activity, which is perhaps more surprising. A one percent increase in GitHub commits leads to a 0.2–0.4 and 0.04–0.1 percent increase in the number of globally- and mission-oriented IT ventures, respectively. Further

[2]The baseline number of new OSS ventures in most countries is quite small when compared to all IT ventures, hence the large difference in impact.

evidence suggests that OSS contributes to high-quality outcomes. A one percent increase in GitHub commits is associated with a 0.4–0.8 percent increase in the value of new venture financing deals, a 0.2–0.5 percent increase in the number of new financing deals, and a 0.1–0.3 percent increase in the number of technology acquisitions in the following year, suggesting $509–1,019 million in new venture financing, 5–13 new financing deals, and 0.4–1.1 acquisitions per country per year.

These results contribute to several research agendas. This is the first study to benchmark variance in OSS participation across the globe, as noted. Our study implies that policy for OSS has larger global consequences than has previously been recognized. This contrasts with prior research investigating OSS in developed economies (Nagle, 2019a, Kogut and Metiu, 2001; Lerner and Schankerman, 2010). We also contribute statistical evidence for how OSS contributes to innovative entrepreneurial development in some countries and, relatedly, why entrepreneurship occurs in some countries more than others. We also add insight into the complementarity between OSS and endowments, which provides evidence about a path for how some low- and middle-income countries may encourage entrepreneurial newcomers in software-intensive activities, allowing them to catch up and, perhaps, eventually overtake established players (Lee and Malerba, 2017).

We further contribute to understanding the rate and direction of innovative activity (Lerner and Stern, 2012; Nelson, 1962). Understanding the incentives to invent new computing has a long history. For example, while the first volume on this topic (Nelson, 1962) does not mention computers, it does analyze the incentives to invent and commercialize transistors. The second volume (Lerner and Stern, 2012) studies innovation in computing at several established firms, highlighting how new entrants can cause innovative activity at leading firms. It also highlights two of the key themes investigated in this study. One chapter considers questions about the role of entrepreneurship in fostering incentives to innovate, and another chapter raises questions about whether programmers prefer a cooperative setting oriented towards multiple contributors.

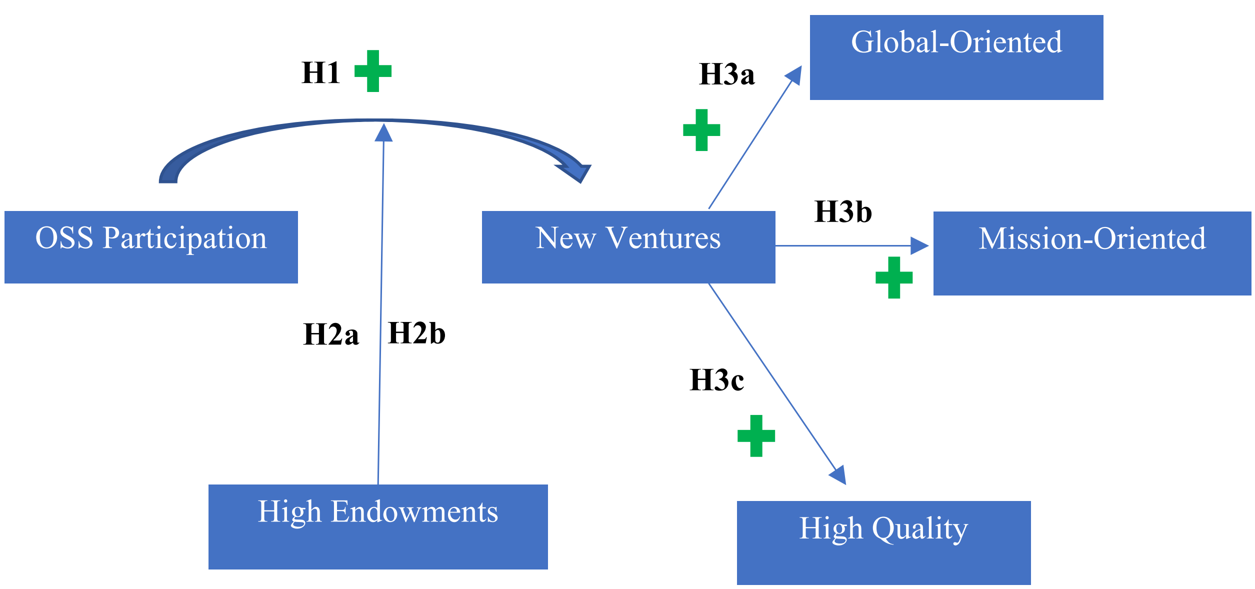
We additionally add to investigations of "digital dark matter," namely, the role that intangible inputs and unpriced digital goods such as OSS play in producing new outputs (Greenstein and Nagle, 2014; Keller et al., 2018; Robbins et al., 2018). In this study, we show how to measure one of the outputs associated with OSS, namely, entrepreneurial activity. This study demonstrates how to measure the impact of OSS by matching large-scale GitHub platform data with commonly-used and publicly available firm-level and country-level data. As a statistical matter, we demonstrate that OSS participation can serve as a valuable predictor variable for quality ventures and entrepreneurial ecosystems around the world. The implementation of our instrumental variable strategy is also novel and may provide guidance to future research that compares OSS across countries.

4.2 Framework and hypotheses

Does open source encourage innovative entrepreneurship? We divide the analysis into two parts—focusing first on the rate of this entrepreneurship and then on the direction. First, we ask whether there is a positive correlation and, if so, whether the evidence supports a causal interpretation. Next, we consider whether open source has the same impact across countries with similar or different endowments. We also divide that into two questions. First, whether open source has a similar impact in countries with high and low human capital and high and low per capita income. Second, what kinds of ventures does open source spur? Does it contribute more to entrepreneurial ventures with a more public mission and global orientation, or to both? We build upon existing frameworks and extend them to craft hypotheses.

We summarize our hypotheses in Figure 4.2 and provide details below. These questions first focus on (H1) whether OSS has a positive effect on venture founding. The questions then focus on whether this relationship is particularly strong in higher or lower-endowed countries (H2a–b). Finally, we assess whether the effect leads to more (H3a) globally-oriented, (H3b) mission-oriented, and (H3c) high-quality ventures.

Figure 4.2: Possible relationships between OSS participation and entrepreneurship



To adapt the hypotheses to our particular empirical context, GitHub, requires some definitions, shown in

Figure 4.3. Although our arguments extend to other OSS repositories, the nomenclature varies from platform to platform, but most have similar attributes, albeit called by different names. Participants are contributors on an OSS platform, who may or may not be employed by a firm to contribute to a particular project. Projects are aggregations of software code around a common goal. Each participant contributes to at least one project, and some individuals contribute more to a particular project while others less. Organizations are groups of projects that share a common goal and may be affiliated with a firm or a shared interest. Participants may be members of an organization or not.

Figure 4.3 illustrates. Participant 1 contributes more to project 1 than does participant 2, and participant 4 contributes to more projects than participant 3. Participants who contribute to OSS as part of their employment are likely to be members of their employers' organization (e.g., participant 1). Other participants may share interests (e.g., participant 2), or be unaffiliated with employers (e.g., participant 3).

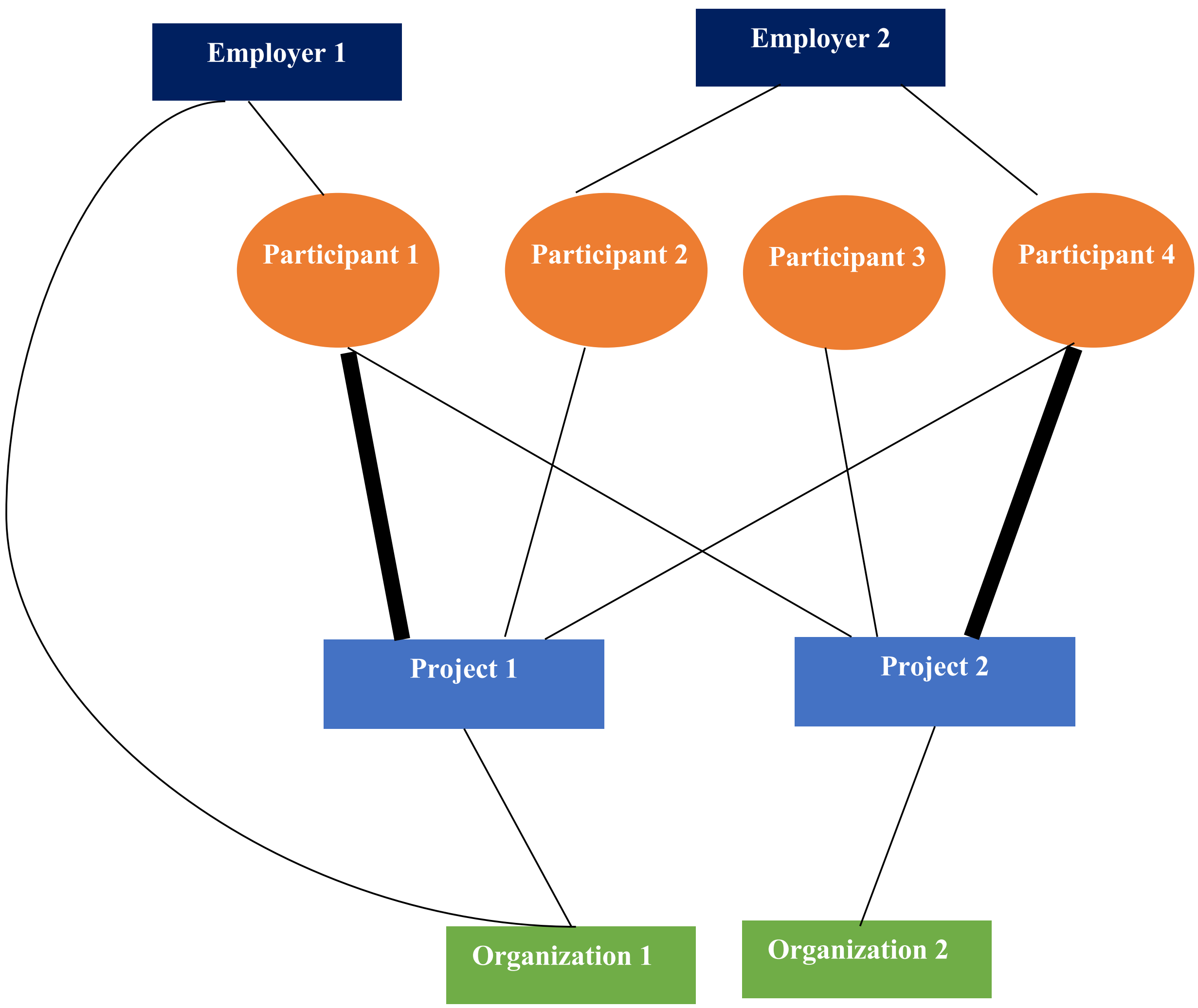
4.2.1 Does OSS increase the rate of entrepreneurship?

Consider Figure 4.2. There are three possible signs of the impact of OSS on entrepreneurship: there is no effect, a positive effect, or a negative effect. No literature points in the negative direction, so we focus on the possibility of no effect or a positive effect.

Does participation in OSS increase the level of entrepreneurial activity? First, OSS might reduce costs in searching for human capital. Talented coders may self-select into participation, and experience on the platform may improve their talent (Nagle, 2018). Second, OSS might increase access to complementary assets, such as community infrastructure and a feedback and recognition system. Such assets are also valuable for the production of commercialized products within a venture (Chatterji, 2009; Elfenbein et al., 2010). Third, OSS could reduce costs to the communication of knowledge, just as in peer networks within company settings (Gompers et al., 2005; Nanda and Sørensen, 2010), within entrepreneurial clusters (Arzaghi and Henderson, 2008), and inside diaspora/ethnic communities (Kerr, 2008; Nanda and Khanna, 2010). Finally, it standardizes coding practices and sharing of programming solutions (Haefliger et al., 2008), and establishes "best practices" (Varian and Shapiro, 2003).

The null is also plausible. There are two arguments. In the first, OSS platforms attract companies like Microsoft and IBM, creating incentives for participants to use the platform for advertising their skills and potentially gain employment. The extrinsic career motivations also may incentivize them to remain employees for incum-

Figure 4.3: OSS context illustration



bent companies (Blatter and Niedermayer, 2009; Hann et al., 2013; Lerner and Tirole, 2002). A second argument focuses on the lack of competitive advantages of participating in OSS. OSS enables access to coordination activities and new ideas, but with few barriers to entry. Moreover, OSS itself is not a source of a rare and non-imitable resource, as competitors may freely use the same code. Summarizing the hypothesis against the null is H1.

H1: An increase in OSS participation in a country leads to an increase in venture founding in that country.

4.2.2 Does open source spur entrepreneurship equally across country contexts?

We next consider how OSS contributes to entrepreneurship in countries with greater or fewer endowments. We face a constraint in measuring endowments because they tend not to change quickly, and there are only a limited number of countries to compare with one another. Our approach focuses on two aspects that vary widely across countries, namely, in terms of income (GDP per capita) and human capital. The first is a broader level of endowment than the latter, although the latter may be more germane to the current study, given the complexities of using OSS. Prior research points to two directions. On the one hand, OSS can serve to level the playing field across more and less endowed countries. By reducing costs to developing talent (Nagle, 2018) and products (Chatterji, 2009; Elfenbein et al., 2010), open source can compensate for income and human capital constraints in countries. In this sense, OSS and country endowments are substitutes.

On the other hand, alternative perspectives suggest that OSS and country endowments are complements. Digital technologies concentrate financial and talent capital in well-endowed nations like the United States (Bloom et al., 2012b). Building upon OSS might require human capital to know how to operate on the platform. Prior work suggests that open-source platforms have increasingly reduced documentation and support, requiring prerequisite knowledge to participate on them (Lerner and Tirole, 2002). Further, knowledge of OSS opportunities might require being part of developer networks that agglomerate in well-endowed nations, especially those with higher human capital.

The theory points in both directions, so it is an empirical question. Thus, we consider the competing hypotheses:

H2a: OSS and country endowments are substitutes such that the impact of OSS participation on venture founding is greater for countries with lower endowments.

H2b: OSS and country endowments are complements such that the impact of OSS participation on venture founding is greater for countries with higher endowments.

4.2.3 What aspects of the entrepreneurship ecosystem does OSS impact?

The following hypotheses consider how OSS impacts the focus and direction of these newly created ventures. Does OSS contribute more to entrepreneurial ventures with a more mission or global orientation, or both? Further, does it contribute to the creation of high-quality ventures? Although the benefits of the latter have been well studied, there is not yet much systematic evidence about the impact of global-oriented and mission-oriented ventures on the local ecosystem, or about their long-term performance. However, possible benefits from such a shift in the direction of these new ventures may accrue in ways similar to that found the context of the gender of inventors, which can influence the type of consumers that benefit from innovations (Koning et al., 2021). Likewise, globally-oriented ventures would be more likely to serve customers with ubiquitous needs rather than niche local ones. We might also expect that mission-oriented ventures would be more likely to serve the needs of underserved communities or tackle large public problems, as these are perceived by the entrepreneurs who founded the venture. In this study, we focus primarily on the origins of this shift in direction.

The global composition of the OSS community may lead OSS contributors to a broader awareness of the global demand for specific new products and services. This international exposure may stimulate the international orientation of their ventures and shape their ability to detect international opportunities. It also may shape their ability to execute on them by understanding risks and leveraging global support/customer networks (Bruneel et al., 2010; Crick and Jones, 2000). This supports the next hypothesis:

H3a: An increase in OSS participation in a country leads to an increase in globally-oriented venture founding.

Next, consider whether factors other than private profit motivate the venture. A "mission-oriented" startup is one that engages in socially-impactful activities, such as promoting gender equality, economic opportunity, environmental sustainability, improved health, education, and broadening access to finance. OSS places importance on the community (Shah, 2006; Von Krogh et al., 2012) and attracts contributors with pro-social motives (Nagle et al., 2020; Von Krogh et al., 2012). Therefore, similar to above, increased exposure to OSS might lead to more awareness of such values and encourage new ventures to take a more mission-oriented approach. This supports the following hypothesis:

H3b: An increase in OSS participation in a country leads to an increase in mission-oriented entrepreneurship in that country.

Why might OSS contributors lead to higher-quality ventures? Both a selection and treatment effect could matter. As for selection, OSS contributors may have higher technical talent than the general population, and that can translate into better products (Nagle, 2019b). As for treatment, the OSS platform aggregates resources—talent, co-founders, and a collaborative coding environment, and that enables coordination. It also enables contributors to observe problems and solutions that may have a global market, such that the solutions can benefit from both a big market on the revenue side and economies of scale on the cost side. Two common proxies of venture quality include (a) the extent of financing (Catalini et al., 2019) and (b) whether they are acquired (Guzman and Stern, 2020). Thus, ventures formed by OSS contributors may receive more financing, as well as have a higher probability of being acquired. Summarizing:

H3c: An increase in OSS participation in a country leads to an increase in the quality of newly founded ventures in that country, as proxied by venture financing and acquisition.

4.3 Measurement

The data sample consists of 3,519 observations, encompassing a panel of 207 countries over 17 years (2000 to 2016). Unfortunately, the panel becomes unbalanced due to missing observations among some of the exogenous variables.

The sample draws from different levels of development. We consider this a good feature, as it retains variance among a novel sample for studies of open source. We begin with virtually all of the 75 high-income countries, which is 36 percent of the sample. We also have good representation from 55 European and Central Asian countries, and this is 27 percent of the sample. We will sometimes reduce the sample size to accommodate the availability of data—principally when using the Human Capital Index. When these are included, 58 high-income countries remain and are 32 percent of the sample. The 49 remaining European and Central Asian countries are 27 percent of the sample.

We also lose observations due to missing variables in some years, especially among Sub-Saharan and upper-middle-income countries and in the initial lagged year (2000). That reduces the number of observations in the final sample for specifications with all control variables included, which consists of 2,747 country-year observations. Once again, it continues to sample from a disparate set of circumstances, and therefore maintains the generalizability of the results.

4.3.1 Measuring entrepreneurship

Country-level data on entrepreneurship comes from Crunchbase, a source that has been used in many studies of entrepreneurship (e.g., Scott et al., 2020; Yu, 2020). The Crunchbase database has grown to become a primary data source for investors as well as in scholarly research. It has been used in over 90 scientific articles (e.g., Dalle et al., 2017; Koning et al., 2019; Scott et al., 2020; Yu, 2020). The variable of interest is the number of new technology ventures founded per year in a given country.

While the VC funding statistics from Crunchbase are similar to alternative sources (Dalle et al., 2017; Kalemli-Ozcan et al., 2015; Kaplan and Lerner, 2016), it comes with a number of challenges. The Crunchbase dataset launched in May 2007, and contributors have backfilled data on companies founded prior to that date, which studies such as Block and Sandner (2009) have used. Crunchbase focuses on younger firms and updates on a daily basis because of the partially crowdsourced nature of the dataset. That necessitates controls for time and motivates a range of robustness tests.

Crunchbase classifies new companies into categories.[3] Next, we identify companies focused on IT and companies focused narrowly on Open Source Software. To examine their orientation, we consider whether they are global or mission-oriented. We construct the global and mission variables through two approaches: word searches of the company descriptions[4] and a supervised logistic regression algorithm.[5] Because of space limitations, we

[3]Examples of sub-categories in our sample include: business information systems, cloud data services, and video chat (information technology); natural language processing, task management, and open source (software); and cloud infrastructure, data center automation, and network hardware (hardware). We explored broad/narrow categories. The broad definition includes information technology, such as cloud data services, network security, and data integration, hardware, and software. The narrow definition is only open-source software companies. As it turns out, this narrow definition is highly correlated (0.6) with the broad definition and, therefore, statistically points in a similar direction.

[4]Global orientation is measured through the use of the words international and worldwide in the company descriptions. Mission orientation is measured through the use of the words empower, gender, women, and climate in the company descriptions.

[5]We manually train the logistic regression algorithm on 1,001 startup descriptions (2 percent of the venture data) by classifying each firm as mission-oriented (1/0) and/or globally-oriented (1/0). We then take 20 percent of these data as test data and see how accurate our algorithm is at classifying these data based on the logistic regression function derived from the other 80 percent, compared to what we actually manually coded this 20 percent. This yields a test accuracy rate of approximately 93 percent accuracy for global orientation and 95 percent accuracy for mission orientation. While there is no standard test error accuracy, generally above 50 percent indicates that classification performs better than random. We then apply this logistic regression function to the rest of our data to get a universal measure of global and mission orientation for our study.

only include the machine learning-created measures in our results.[6]

We also assess their quality, as proxied by entrepreneurial financing and acquisitions. We use the Preqin database to measure financing using 1) the total value of all venture investments in information technology companies that occurred in a given country in a given year; and 2) the total number of venture investments in information technology companies in a given country in a given year.[7] The data have been used by other studies such as Axelson et al. (2013) and Chakraborty and Ewens (2018). For acquisitions, we use data from Crunchbase on the number of acquisitions of IT companies. Crunchbase logs transaction-level data on events in which any of the companies it covers are acquired; we aggregate these events to the country of interest in a given year.

We take log(1+VARIABLE) to account for skewness and the value of zero.

4.3.2 Measuring OSS

Our data on open-source activity in a country comes from GitHub, the most widely used repository for hosting OSS projects. Created in 2008, GitHub became the central repository for most major open-source projects (GitHub, 2019) and became a repository for open-source projects founded before 2008, which moved to the platform to take advantage of its useful tools. Based in San Francisco, it contains 35 million public repositories as of March 2020 (https://github.com/search?q=is:public) and, including private repositories, passed more than 100 million total in 2018.[8] Microsoft purchased GitHub in June 2018 for $7.5 billion.[9]

GitHub provides a consistent and standardized measure of activity in open source in a given country in a given year. However, while frequently used in technical studies of OSS (e.g., Conti et al., 2021; Medappa and Srivastava, 2019), to our knowledge, it has rarely served as a global source of data for business and economic studies (For an exception, see Nagle, 2019a).

GitHub participants must create user profiles with basic information about themselves and their backgrounds. That enables measures of the country-level contributions. Prior research has shown that roughly 50 percent of participants include the country in which they reside in their profile (Nagle, 2019a). No evidence suggests the

[6] Word search results are available upon request from the authors.

[7] Preqin claims to cover 70 percent of all capital raised in the private equity industry, with 85 percent of the data gathered via Freedom of Information Act requests targeting public pension funds (thus helping to reduce self-reporting bias) and the rest coming directly from fund managers.

[8] Private repositories generally contain proprietary code owned by companies and do not meet the definition for open-source software.

[9] At the time of this writing, Nat Friedman, reporting to Scott Guthrie, executive vice president of Microsoft Cloud and AI, leads GitHub.

presence of reporting biases.[10]

The measure of participation is the number of new commits (on average one line of code) made to OSS projects hosted on GitHub yearly by developers that have self-identified as living in a given country from 2000-2016.[11] We also observe whether these commits originate from organization-affiliated or individual user accounts. Commits from organization-affiliated accounts are those that come from users who joined an organization prior to the date of commit. Organization affiliation first emerged in GitHub after 2010 (Neath, 2010). About 30 percent of commits come from organization-affiliated accounts, while the rest come from individual user accounts.[12] The data prior to 2008 reflect projects that were migrated to the platform.[13]

Once again, we take log(1+VARIABLE) to account for skewness.

4.3.3 Empirical specification

The hypotheses concern the statistical relationship between open source and entrepreneurship. After finding robust associations, we utilize a variety of econometric tools to help us add evidence for a more causal interpretation.

Consider estimating Equation (1).

$VENTURE_{it} = \alpha_{it} + \beta_1 GITHUB_{i,t-1} + \beta_2 CONTROLS_{i,t-1} + \gamma_t + \varepsilon_{it}$ (1)

$VENTURE_{it}$ indicates a logged variable in country i in time t which is either the number of new IT or OSS startups founded in that country in that year, as defined above. $GITHUB_{i,t-1}$ indicates a logged variable in country i lagged one year before, in time t-1 that measures the number of commits coming from a country in a given year. Many factors shape entrepreneurship and OSS each year, such as the state of demand for IT, the optimism of investors, and the state of political uncertainty. Such trends are measured with γ_t, which reflects year fixed-effects. The estimation will use robust standard errors clustered at the country level.

$CONTROLS_{i,t-1}$ are lagged control variables. These affect both entrepreneurship and contributions to OSS. We include country-level GDP per capita and the log of the country's population sourced from the World Bank.

[10] In particular, there would be a bias concern if there was evidence that people from some countries over-reported their country, while people from other countries underreported their country in a systematic way.

[11] A commit can be numerous lines of code and a commit can represent the deletion of lines of code. On average, however, commits consist of a change to one line of code (Nagle, 2018).

[12] We obtain these data from the Google BigQuery hosting of the GHTorrent database, which is a mirror of all of the activity on GitHub.

[13] This necessitates testing results with and without the earliest data. The results generally hold if we only use data from 2008 onwards only.

We control for internet connectivity via the log of the number of internet users per capita, sourced from the International Telecommunications Union. We also control for the Human Capital Index (HCI),[14] sourced from the United Nations, which measures the skill level of the workforce.[15]

While Equation (1) will be used to test H1 and H3 (the latter using the alternate definitions of $VENTURE_{it}$ defined above), testing H2 relies on interacting $GITHUB_{i,t-1}$ with a measure of country-level endowments, operationalized as either GDP per capita or HCI as follows:

$$VENTURE_{it} = \alpha_{it} + \beta_1 GITHUB_{i,t-1} + \beta_2 GITHUB_{i,t-1} * ENDOW_{i,t-1} + \beta_3 CONTROLS_{i,t-1} + \gamma_t + \varepsilon_{it} \quad (2)$$

The coefficient of interest is β_2, which indicates whether the relationship between lagged log GitHub commits and venture founding is stronger or weaker in more endowed countries. If this coefficient is negative, then the relationship is weaker in more endowed countries, suggesting that OSS substitutes country endowments, supporting hypothesis 2a. If the coefficient is positive, then OSS complements country endowments, supporting hypothesis 2b.

4.3.4 Endogeneity and reverse causality

OLS estimates of equations 1 and 2 provide estimates of statistical association, but potentially contain reverse causality. If it exists, we would expect reverse causality to impart a positive bias in the OLS estimate. One possible approach to this concern is to include country fixed effects in the estimates to identify "within" estimates. However, that approach will fail due to insubstantial variation in many variables within a country over time. Therefore, as indicated for both Equations 1 and 2, we additionally include many controls that change over time and vary across countries. However, controls alone are not sufficient.

Our approach to causality is to include instrumental variables. These must plausibly shift the likelihood that a country will contribute to OSS, without shifting the likelihood of entrepreneurial startups, thus satisfying the exclusion restriction. As the first paper to explore this relationship, we err on the side of many instruments. We test these instruments individually as well as aggregated in each equation. Both versions yield similar results. For simplicity, we show the aggregated version in the tables. We consider five instrumental variables as candidates. Three are related to the costs of supplying open source and two to the demand for it in a country, and in each

[14]The Human Capital Index is only available in 2003, 2004, 2005, 2008, 2010, 2012, 2014, 2016, and 2018. We impute the values for the unavailable years as averages of the index values of the years before and after.

[15]This index is comprised of the adult literacy rate, combined enrollment ratio of primary, secondary, and tertiary schooling, expected years of schooling, and average years of schooling.

case, we consider only those that do not plausibly cause entrepreneurship. Using multiple instruments is consistent with the literature on strengthening causal identification, especially when no instrument is precisely suited to the context, and preserves the LATE interpretation of the results (Angrist and Krueger, 2001; Mogstad et al., 2021). Although none of these instruments are a silver bullet that fully addresses the endogeneity concerns, in aggregate, their usage helps to take a step towards a causal interpretation of the results.

We begin with three variables related to the supply of open source. These are comprised of variables that reduce costs to open source participation, but do not increase the likelihood of entrepreneurship. Motivated by the literature on contributions to open source during "slack times" (Agrawal et al., 2018), we construct three variables that operationalize an increase in the availability of skilled contributions during "slack times," which we hypothesize increase during economic downturns. The skills are proxied by high human capital, digital skills,[16] and internet users. Each is defined as above-median[17] levels of human capabilities, either broadly reflected in human capital or narrowly reflected in digital skills or internet users, and each is interacted with below-median economic growth. The literature has used such interaction instruments in various scenarios (e.g., Angrist and Krueger, 1991). These three follow a related logic. Each should increase open-source use through the supply side. These interacted instruments satisfy the exclusion restriction because weak economic growth should not increase entrepreneurship, particularly for higher-quality founders (Conti and Roche, 2021). Low demand may also decrease the optimism over near-term increases in the level of final demand for an entrepreneur's product.

Finally, two instruments measure demand-side pulls on open-source use. These pulls come from government-level open-source policies employed by 64 countries worldwide from 2000 to 2009. We use the OSS-related policies, approved and implemented at the country level as captured in an aggregated database constructed by Lewis (2010).[18] Although these policies are implemented for different reasons in different countries, most are to either 1) allow government entities more flexibility in technology procurement and/or 2) reduce costs related to proprietary software. It has long been known that government procurement practices can influence the behavior

[16]The digital skills variable is aggregated from the World Economic Forums Global Competitiveness Index, which is on a 1–7 scale, where 1 indicates that the active population does not at all possess sufficient digital skills (e.g., computer skills, basic coding, digital reading), and 7 indicates it does to a great extent. It is only available for 2017, so we use a single aggregated value for each country. We then calculate above and below median values.

[17]Above median reflects the top two quartiles of the country-year dataset, and below median reflects the bottom two quartiles of the country-year dataset.

[18]These policies are categorized as either advisory, research and development (RD), preference, or mandatory. Of these policies, the mandatory instrument has the strongest relationship with open source contributions, as we would expect as such policies put more stringent demand requirements for open source. The ultimate instrument is aggregated across the four categories of policies as a binary variable.

of firms (Flamm, 1988). Further, prior research has shown that procurement policies related to OSS can lead to increases in OSS contributions from across the country (Nagle, 2019a). We treat countries without policies in this database as 0 for the entire time period. For countries that do have such a policy, the value of the instrument is 0 before the policy comes into place and 1 after (and including) the year when the first policy is employed.[19] We restrict analysis using this instrument from 2000 to 2009 due to unavailable granular data on country-level OSS policies after 2009. The instrument fulfills the exclusion restriction because the only way that the open source policy should affect new venture formation is through participants increasing their engagement with open source.[20] Table C.1.1 shows examples of some of these policies and their motivations. Generally, policymakers pursued these policies to achieve security, efficiency, and local economic development. The latter goal was to decrease government costs (especially those related to foreign software companies like Microsoft and Oracle) and to spur growth in the local IT labor pool due to the reliance on more open technologies. Because none of these goals were to directly spur entrepreneurship (and any increase in entrepreneurship would likely occur through increased OSS contributions), the instruments comply with the exclusion restriction.

Other experiments with demand-side instruments did not bear fruit.[21]

We assemble several estimates using these instruments, and, though different in sample and specification, each points in the same qualitative direction and, thus, does not reject a causal interpretation. Specifically, these in-

[19]For example, Argentina implemented its first open source policy in 2004 (an advisory policy), so the policy instrument was 1 from 2004 onwards and 0 in 2003 and earlier.

[20]For robustness, we also include a form of the open source policy instrument interacted with weak economic growth as the second demand-side instrument. As discussed above, weak economic growth may have a positive push on open source use, but should not have an independent positive impact on entrepreneurship.

[21]One surprisingly weak candidate instrument relates to a language change on the GitHub platform. On July 13, 2010, the GitHub platform announced switching to English-only. Then, on November 18, 2016, the GitHub platform announced support for several other languages: Japanese, French, Serbian, German, Swedish, Croatian, Polish, and Dutch (GitHub, 2010). We also tested continuous versions of the instrument, using data from the UN and an academic study on Twitter on the percentage of country populations speaking/engaging in certain languages (Mocanu et al., 2013). These continuous data are available only for a subset of countries (approximately half). Details are available upon request. Other candidates use the method developed by Bartik (1991). These employ trade relationships between countries as exogenous shifters in the potential supply of open-source contributions. This instrument hypothesizes that large trading partners will have an increased influence on the exposure of individuals in the focal country to external ideas and contributions, shifting the likelihood of individuals engaging in open source. More specifically, we identify each country's top three trading partners from the prior year, and consider the OSS activity in those countries. These instruments surprisingly show a negative relationship with OSS, suggesting that OSS activity may depart from traditional trade patterns. To comply with the monotonicity requirements of the instruments, we do not include them in our main model. However, our results do hold using these variables as part of a group of instruments. Further research may investigate why the trade and OSS patterns diverge.

strumental variables necessitate that we estimate the relationship using two different samples. The first includes all non-policy variables that allow us to cover the full-time span of data (2000-2016). The second consists of all variables, which allows us to improve our identification, albeit with weaker power, as we only cover a subset of data in our sample (2000-2009). Further, in addition to using the government policies as an instrument, we also consider them in an alternate framework that uses a staggered difference-in-differences estimation. Although the results are consistent with the primary analysis, we do not rely solely on this estimation since the availability of data on government OSS policies ends halfway through our dataset.

Separately and in aggregate, the first-stage results of the various instruments are promising. As seen in Table C.2.1, the candidate instruments are predictive of changes in OSS contributions. Further, they all have first-stage F-statistics that are well above the rule-of-thumb threshold of 10 (Staiger and Stock, 1997; Stock and Yogo, 2005). Thus, the usage of these instruments can help add support for a causal interpretation of the main estimates.

To help readers navigate the estimation, we compile all hypotheses, implementations, and results in the final table. It is presented at the end of the estimation in Table 4.9.

4.4 Estimates

Table 4.1 provides summary statistics. As expected, the variables have skewed distributions. For example, the average number of GitHub commits per year is 76,000 and the range is 0–31.2 million. A correlation matrix for these variables is shown in Table C.3.1 in the Appendix.

4.4.1 OSS and new venture founding

We first test hypothesis 1 on whether OSS positively impacts venture founding. Estimates show that the positive relationship arises in a variety of specifications.

Table 4.2 shows OLS and 2SLS estimates that indicate a positive and statistically significant relationship between logged GitHub commits and logged IT ventures (Columns 1–3) and logged OSS ventures (Columns 4–6). All include year fixed-effects. Table 4.2 shows three estimates for each of these dependent variables: the first with an OLS specification, the second with a 2SLS specification using non-policy instruments that use the full span of the data (2000–16), and the third with a 2SLS specification using all instruments that span a subset of the data

Table 4.1: Summary statistics

Variables	# of Countries	N	Mean	SD	Min.	Max.
Log GitHub	207	3519	4.27	4.39	0.00	17.25
GitHub Commits (1000's)	207	3519	76307.28	822612.40	0.00	31200000.00
New Ventures						
Log IT Ventures	207	3519	0.98	1.55	0.00	8.12
IT Ventures	207	3519	24.33	168.06	0.00	3357.00
Log OSS Ventures	207	3519	0.07	0.32	0.00	3.61
OSS Ventures	207	3519	0.22	1.73	0.00	36.00
Controls						
Log Population	207	3514	15.20	2.39	9.15	21.04
Population (Thousands)	207	3514	32600000.00	128000000.00	9420.00	1380000000.00
Log GDP Per Capita	197	3275	8.55	1.54	5.27	12.16
GDP Per Capita (2010 US$)	197	3275	14436.99	21682.95	193.87	191586.60
Log Internet Users	204	3352	2.76	1.35	0.00	4.60
Internet Users (% Population)	204	3352	29.65	28.14	0.00	98.32
Human Capital Index	186	3073	0.72	0.23	0.00	1.42
Financing Deals						
Log Number of Financing Deals	207	3519	0.62	1.34	0.00	8.68
Number of Financing Deals	207	3519	25.03	255.51	0.00	5901.00
Log Value of Financing Deals	207	3519	1.11	2.40	0.00	12.69
Value of Financing Deals (Millions US$)	207	3519	1273.68	14364.63	0.00	324737.20
Log Number of Acquisitions	207	3519	0.25	0.82	0.00	6.98
Number of Acquisitions	207	3519	3.72	37.20	0.00	1073.00
Global and Mission Ventures						
Log Number of Global IT Ventures	207	3519	0.98	1.55	0.00	8.12
Number of Global IT Ventures	207	3519	24.25	167.57	0.00	3347.00
Log Number of Global OSS Ventures	207	3519	0.07	0.32	0.00	3.61
Number of Global OSS Ventures	207	3519	0.22	1.72	0.00	36.00
Log Number of Mission IT Ventures	207	3519	0.10	0.38	0.00	3.83
Number of Mission IT Ventures	207	3519	0.31	2.22	0.00	45.00
Log Number of Mission OSS Ventures	207	3519	0.00	0.04	0.00	1.10
Number of Mission OSS Ventures	207	3519	0.00	0.07	0.00	2.00
Instruments						
Below Median Econ. Growth X Above Median Human Capital Instrum.	200	3290	0.32	0.47	0.00	1.00
Below Median Econ. Growth X Above Median Digital Skills Instrum.	200	3290	0.36	0.48	0.00	1.00
Below Median Econ. Growth X Above Median Internet Users Instrum.	200	3290	0.31	0.46	0.00	1.00
Below Median Econ. Growth X OSS Policy Instrum. (Before 2010)	200	3290	0.13	0.34	0.00	1.00
OSS Policy Instrum. (Before 2010)	207	3519	0.22	0.42	0.00	1.00

The table presents summary statistics for all dependent, independent, control, and instrumental variables used in subsequent regressions. The values cover 2000–2016. The maximum number of countries covered by these variables is 207, and the minimum is 150. In instrumental variables noting "high" and "low" value combinations of variables, "high" reflects the top two quartiles across the country-year dataset, and "low" reflects the bottom two quartiles across the country-year dataset. All variables vary by year, except digital skills, which is aggregated at the country level given the lack of availability of sufficient yearly data.

(2000-09).[22]

The magnitude of the estimates is substantial. A one percent increase in GitHub commits is associated with a 0.2–0.4 percent increase in IT ventures and a 0.03–0.1 percent increase in OSS ventures in the following year. The coefficient on IT ventures exceeds the coefficient on OSS ventures, likely because the latter is a narrower category, and OSS contributions affect both. Furthermore, the sample size of OSS ventures is substantially smaller than that of IT ventures. Finally, the second stage coefficients are higher in magnitude than they were in the OLS.[23] This result may reflect the reduction in measurement error that may have caused attenuation bias in the OLS estimates (as happens in Bloom et al., 2013).[24]

A simple simulation helps ground the estimates. Taking the coefficients from Columns 1–6 in Table 4.2, an increase of roughly 76,000 yearly GitHub commits—i.e., one percent of the average in the sample—is associated with an increase in approximately 5–10 new IT ventures and 0.007–0.02 new open source ventures per year per country on average in the following year.[25]

These results generate a similar qualitative conclusion about hypotheses 1, irrespective of specification, controls, and instruments, with the statistical precision of the estimates becoming weaker with stricter controls for endogeneity. More open-source software in a country predicts more entrepreneurship. The various estimates also suggest an economically important relationship, even when controlling for different econometric implementations. We conclude, therefore, that there is a broad relationship.

4.4.2 Endowments

We next test competing hypotheses 2a–b on whether country endowments and OSS are substitutes (2a) or complements (2b). Specifically, we use equation (2) to measure whether the relationship between OSS and venture founding is stronger in country contexts with higher GDP per capita and human capital endowments. Table 4.3 applies equation (2) interacting lagged log GitHub commits with lagged logged GDP per capita (Columns 1–2) and with the lagged logged human capital index (Columns 3–4). Both interactions are positive and significant for

[22] These results hold when using each instrument individually.

[23] This suggests the OLS coefficient was biased downward, which is inconsistent with the speculation that reverse causality would bias it upward.

[24] In robustness checks, we confirm that the OLS estimates hold when applied to the smaller sample size used in the 2SLS specifications due to data availability.

[25] There are 24.33 new IT ventures and 0.22 new OSS ventures per year per country on average in our sample. Thus, this large difference in the baseline number of new ventures leads to a substantial difference in the interpretation of the coefficients.

Table 4.2: Impact of OSS on new venture founding

	OLS	2SLS (Non-Policy Instruments)	2SLS - All Instruments (Pre-2010)	OLS	2SLS (Non-Policy Instruments)	2SLS - All Instruments (Pre-2010)
	Log IT Ventures	Log IT Ventures	Log IT Ventures	Log OSS Ventures	Log OSS Ventures	Log OSS Ventures
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged Log GitHub	0.217***	0.391***	0.381***	0.0344***	0.0933**	0.101***
	(0.0256)	(0.0672)	(0.0474)	(0.00984)	(0.0305)	(0.0241)
Lagged Log Population	0.252***	0.0928	0.131**	0.0292*	-0.0249	-0.0226
	(0.0395)	(0.0716)	(0.0463)	(0.0125)	(0.0211)	(0.0192)
Lagged Human Capital Index	-0.101	-0.359	-0.482	-0.121*	-0.209*	-0.192*
	(0.282)	(0.309)	(0.271)	(0.0604)	(0.0879)	(0.0780)
Lagged Log GDP Capita	0.311***	0.221**	0.0659	0.0454*	0.0143	-0.00143
	(0.0743)	(0.0723)	(0.0580)	(0.0212)	(0.0187)	(0.0169)
Lagged Log Internet Users	0.0892	-0.116	0.114	-0.0206	-0.0894*	-0.0730*
	(0.0756)	(0.108)	(0.0885)	(0.0165)	(0.0388)	(0.0310)
_cons	-7.477***	-2.658	-2.028*	-0.956**	0.512	0.518
	(0.819)	(1.463)	(0.939)	(0.342)	(0.422)	(0.364)
N (Country x Year)	2747	2741	1526	2747	2741	1526
N (Country)	182	180	180	182	180	180
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

The table presents estimates regressing the founding of IT and OSS ventures on lagged log GitHub commits, reflecting the extent of entrepreneurial activity associated with open source activity. Columns 1 and 4 present OLS results. Columns 2 and 5 present 2SLS results with non-policy instruments for log GitHub commits spanning the full length of data, 2000–2016. Columns 3 and 6 present 2SLS results with all instruments for lagged logged GitHub commits that span a subset of years in the data, 2000–2009. All columns include robust standard errors, clustered by country. Time fixed effects are relative to the year 2000. The regressions are not perfectly balanced by year, due to missing data in the control variable datasets. First-stage estimates corresponding to the 2SLS specifications are shown in Table C.2.1 in the appendix. Robust standard errors, clustered by country. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

logged IT ventures as dependent variables. They are also positive but not always statistically significant for logged OSS ventures, which may be because of the narrow nature of this category. The magnitudes are higher for the interaction with human capital compared to GDP per capita.

Table 4.3: Impact of OSS on new venture founding by GDP per capita and human capital

	Log IT Ventures	Log OSS Ventures	Log IT Ventures	Log OSS Ventures
	(1)	(2)	(3)	(4)
Lagged Log GitHub x Lagged Log GDP Per Capita	0.0546***	0.0119*		
	(0.0109)	(0.00596)		
Lagged Log GitHub x Lagged Human Capital Index			0.340**	0.0599
			(0.112)	(0.0431)
Lagged Log GitHub	-0.296**	-0.0774	-0.0849	-0.0189
	(0.106)	(0.0507)	(0.104)	(0.0348)
Lagged Log Population	0.263***	0.0315*	0.276***	0.0333*
	(0.0377)	(0.0128)	(0.0370)	(0.0136)
Lagged Human Capital Index	-0.140	-0.129*	-0.598*	-0.208*
	(0.271)	(0.0626)	(0.270)	(0.0931)
Lagged Log GDP Capita	0.0190	-0.0182	0.230**	0.0312*
	(0.0690)	(0.0185)	(0.0702)	(0.0149)
Lagged Log Internet Users	0.238***	0.0118	0.146*	-0.0106
	(0.0674)	(0.0146)	(0.0684)	(0.0141)
_cons	-5.391***	-0.501*	-6.414***	-0.768**
	(0.852)	(0.229)	(0.896)	(0.285)
N (Country x Year)	2747	2747	2747	2747
N (Country)	182	182	182	182
Time Fixed Effects	Yes	Yes	Yes	Yes

The table presents estimates regressing the founding of IT and OSS ventures on lagged log GitHub commits interacted with logged GDP per capita and the human capital index. This regression reveals country contexts in which the OSS-venture formation relationship is stronger versus weaker. All columns include robust standard errors, clustered by country. Time fixed effects are relative to the year 2000. The regressions are not perfectly balanced by year, due to missing data in the control variable datasets. Robust standard errors, clustered by country. Standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.4 breaks apart the interaction terms into sub-sample regressions using equation (1) to measure the relationship between lagged log GitHub commits and venture founding for above and below the median of GDP per capita with the same instruments as in Table 4.2. Consistent with the relationship shown in Table 4.3, the relationship between lagged log GitHub commits and logged ventures is larger for above median GDP per capita than for below median GDP per capita for log IT ventures when using the non-policy instruments (Columns 1–2) and for log OSS ventures for both sets of instruments (Columns 5–8).

Table 4.5 repeats this exercise but for above and below median human capital (rather than GDP per capita). In

Table 4.4: Impact of OSS on new venture founding by GDP per capita: 2SLS

	2SLS (Non-Policy Instruments)		2SLS - All Instruments (Pre-2010)		2SLS (Non-Policy Instruments)		2SLS - All Instruments (Pre-2010)	
	Below Median GDP Per Capita	Above Median GDP Per Capita	Below Median GDP Per Capita	Above Median GDP Per Capita	Below Median GDP Per Capita	Above Median GDP Per Capita	Below Median GDP Per Capita	Above Median GDP Per Capita
	Log IT Ventures	Log IT Ventures	Log IT Ventures	Log IT Ventures	Log OSS Ventures	Log OSS Ventures	Log OSS Ventures	Log OSS Ventures
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lagged Log GitHub	-0.161	0.398***	0.380**	0.328***	-0.0130	0.116**	0.0176	0.143**
	(0.174)	(0.1000)	(0.131)	(0.0844)	(0.0206)	(0.0438)	(0.0216)	(0.0451)
Lagged Log Population	0.420**	0.145	0.0935	0.211*	0.0374	-0.0293	0.0166	-0.0534
	(0.156)	(0.113)	(0.0710)	(0.0923)	(0.0212)	(0.0400)	(0.0131)	(0.0486)
Lagged Human Capital Index	0.573	-0.0457	-0.272	-0.462	0.00766	-0.378	-0.0435	-0.447*
	(0.541)	(0.426)	(0.442)	(0.376)	(0.0597)	(0.211)	(0.0646)	(0.212)
Lagged Log GDP Capita	0.160	0.392**	0.00162	0.191	0.00854	0.0515	-0.00107	-0.00442
	(0.131)	(0.146)	(0.0620)	(0.136)	(0.0125)	(0.0526)	(0.00770)	(0.0565)
Lagged Log Internet Users	0.391	-0.107	0.116	0.194	0.0266	-0.177*	0.00546	-0.179*
	(0.207)	(0.213)	(0.112)	(0.176)	(0.0197)	(0.0824)	(0.00998)	(0.0874)
_cons	-7.416**	-5.287	-1.692	-5.299*	-0.626	0.538	-0.248	1.326
	(2.350)	(3.162)	(1.412)	(2.540)	(0.371)	(1.144)	(0.204)	(1.314)
N (Country x Year)	1409	1330	818	708	1409	1330	818	708
N (Country)	105	95	104	87	105	95	104	87
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table presents estimates regressing the founding of IT and OSS ventures on lagged log GitHub commits, split by below and above median logged GDP per capita levels. Columns 1–2 and 5–6 present 2SLS results with non-policy instruments for log GitHub commits spanning the full length of data, 2000-2016. Columns 3–4 and 7–8 present 2SLS results with all instruments for logged GitHub commits that span a subset of years in the data, 2000–2009. All columns include robust standard errors, clustered by country. Time-fixed effects are relative to the year 2000. The regressions are not perfectly balanced by year, due to missing data in the control variable datasets. First-stage estimates corresponding to the 2SLS specifications are shown in Table C.2.1 in the appendix. Robust standard errors, clustered by country. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

all cases across logged IT and OSS ventures, the magnitude of the coefficient on OSS is larger and significant for above-median human capital relative to below-median (Columns 1–8).

Together, these results are consistent with hypothesis 2b: country endowments complement OSS rather than substitute. None of the evidence suggests that endowments substitute for OSS (hypothesis 2a).

4.4.3 Type of ventures

Next, we examine hypotheses 3a–c. This specification uses globally-oriented ventures, mission-oriented ventures, and financing and acquisitions as proxies for high-quality ventures.

Table 4.6 Columns 1–6 show that GitHub commits have a positive and statistically significant association with both globally- and mission-oriented new IT ventures across both OLS and 2SLS frameworks, supporting hypotheses 3a–b. A one percent increase in GitHub commits is associated with a 0.2–0.4 percent increase in globally-oriented IT ventures and 0.04–0.1 percent increase in mission-oriented IT ventures. Columns 7–9 show a similar positive significant relationship between GitHub commits and globally-oriented OSS ventures. A one percent increase in GitHub commits is associated with a 0.03-0.1 percent increase in globally-oriented OSS ventures. However, the relationship is not robust for logged mission-oriented OSS ventures, as shown by the coefficients in Columns 10–12, which, though always positive, are only statistically significant in the OLS specification.

The results estimate a relationship similar in magnitude to the previous findings. They indicate that OSS affects IT-specific entrepreneurial ventures, especially those with global and to a lesser extent, mission orientations.

In Table 4.7, we find that a one percent increase in GitHub commits is associated with a 0.4–0.8 percent increase in the value of financing deals[26] (Columns 1–3), 0.2–0.5 percent increase in the number of new financing deals (Columns 4–7), and 0.1–0.3 percent increase in the number of acquisitions (Columns 7–9). All coefficients carry statistical and economic significance across OLS and 2SLS specifications. A one percent increase in GitHub commits leads to roughly $509–1,019 million in venture financing, 5–13 new financing deals, and 0.4–1.1 acquisitions per year.[27]

[26]This variable indicates the amount of financing provided to ventures in the country in a given year in USD.

[27]These values are calculated by multiplying the coefficients by the average venture value in millions of USD (1,273.68), number of financing deals (25.03), and number of acquisitions (3.72) per year in the sample.

Table 4.5: Impact of OSS on new venture founding by human capital: 2SLS

	2SLS (Non-Policy Instruments)		2SLS - All Instruments (Pre-2010)		2SLS (Non-Policy Instruments)		2SLS - All Instruments (Pre-2010)	
	Below Median Human Capital Index	Above Median Human Capital Index	Below Median Human Capital Index	Above Median Human Capital Index	Below Median Human Capital Index	Above Median Human Capital Index	Below Median Human Capital Index	Above Median Human Capital Index
	Log IT Ventures	Log IT Ventures	Log IT Ventures	Log IT Ventures	Log OSS Ventures	Log OSS Ventures	Log OSS Ventures	Log OSS Ventures
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lagged Log GitHub	0.0333	0.309*	0.240	0.415***	-0.0302	0.142**	0.0507	0.143**
	(0.167)	(0.123)	(0.189)	(0.0932)	(0.0200)	(0.0539)	(0.0483)	(0.0445)
Lagged Log Population	0.297*	0.197	0.212**	0.0827	0.0524	-0.0685	0.0205	-0.0762
	(0.146)	(0.131)	(0.0656)	(0.0968)	(0.0333)	(0.0484)	(0.0161)	(0.0456)
Lagged Human Capital Index	-0.00922	1.219	-0.367	-0.778	0.00685	-0.703	-0.0976	-0.809
	(0.276)	(1.784)	(0.297)	(1.582)	(0.0420)	(0.674)	(0.0801)	(0.653)
Lagged Log GDP Capita	0.141	0.285*	0.0532	0.0572	0.00633	0.00617	-0.00387	-0.0189
	(0.0727)	(0.123)	(0.0565)	(0.112)	(0.0103)	(0.0464)	(0.00789)	(0.0441)
Lagged Log Internet Users	0.281	-0.00528	0.312**	0.0441	0.0505	-0.166*	0.0207	-0.128*
	(0.192)	(0.164)	(0.107)	(0.134)	(0.0313)	(0.0705)	(0.0192)	(0.0554)
_cons	-6.154**	-6.306	-3.541**	-0.719	-0.779	1.771	-0.251	2.093
	(1.925)	(4.080)	(1.162)	(3.237)	(0.508)	(1.550)	(0.242)	(1.460)
N (Country x Year)	1345	1381	640	877	1345	1381	640	877
N (Country)	137	118	89	114	137	118	89	114
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table presents estimates regressing the founding of IT and OSS ventures on lagged log GitHub commits, split by below and above median logged Human Capital Index levels. Columns 1–2 and 5–6 present 2SLS results with non-policy instruments for log GitHub commits spanning the full length of data, 2000-2016. Columns 3–4 and 7–8 present 2SLS results with all instruments for logged GitHub commits that span a subset of years in the data, 2000–2009. All columns include robust standard errors, clustered by country. Time fixed effects are relative to the year 2000. The regressions are not perfectly balanced by year, due to missing data in the control variable datasets. First-stage estimates corresponding to the 2SLS specifications are shown in Table C.2.1 in the appendix. Robust standard errors, clustered by country. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.6: Impact of OSS on global– and mission-oriented venture founding

	OLS	2SLS (Non-Policy Instruments)	2SLS - All Instruments (Pre-2010)	OLS	2SLS (Non-Policy Instruments)	2SLS - All Instruments (Pre-2010)	OLS	2SLS (Non-Policy Instruments)	2SLS - All Instruments (Pre-2010)	OLS	2SLS (Non-Policy Instruments)	2SLS - All Instruments (Pre-2010)
	Log Global IT Ventures	Log Global IT Ventures	Log Global IT Ventures	Log Mission IT Ventures	Log Mission IT Ventures	Log Mission IT Ventures	Log Global OSS Ventures	Log Global OSS Ventures	Log Global OSS Ventures	Log Mission OSS Ventures	Log Mission OSS Ventures	Log Mission OSS Ventures
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Lagged Log GitHub	0.217***	0.390***	0.380***	0.0400***	0.101**	0.0848***	0.0345***	0.0938**	0.101***	0.00123*	0.00348	0.00252
	(0.0255)	(0.0671)	(0.0474)	(0.0108)	(0.0340)	(0.0214)	(0.00982)	(0.0304)	(0.0238)	(0.000576)	(0.00282)	(0.00215)
Lagged Log Population	0.252***	0.0926	0.130**	0.0417**	-0.0147	-0.00194	0.0288*	-0.0256	-0.0230	0.00144	-0.000621	-0.000791
	(0.0395)	(0.0715)	(0.0462)	(0.0141)	(0.0257)	(0.0169)	(0.0125)	(0.0210)	(0.0189)	(0.000865)	(0.00223)	(0.00147)
Lagged Human Capital Index	-0.101	-0.357	-0.479	-0.127	-0.218*	-0.191*	-0.120*	-0.208*	-0.190*	-0.000267	-0.00359	0.000413
	(0.282)	(0.309)	(0.271)	(0.0740)	(0.101)	(0.0760)	(0.0603)	(0.0879)	(0.0778)	(0.00341)	(0.00555)	(0.00481)
Lagged Log GDP Capita	0.310***	0.221**	0.0658	0.0620*	0.0296	0.00565	0.0457*	0.0144	-0.000818	0.00191	0.000723	-0.000284
	(0.0742)	(0.0723)	(0.0579)	(0.0242)	(0.0210)	(0.0153)	(0.0212)	(0.0187)	(0.0168)	(0.00169)	(0.00101)	(0.000596)
Lagged Log Internet Users	0.0892	-0.115	0.114	-0.0258	-0.0976*	-0.0377	-0.0220	-0.0914*	-0.0748*	-0.00140	-0.00403	-0.00177
	(0.0756)	(0.108)	(0.0884)	(0.0199)	(0.0443)	(0.0262)	(0.0166)	(0.0387)	(0.0305)	(0.00165)	(0.00423)	(0.00281)
_cons	-7.466***	-2.658	-2.022*	-1.231***	0.210	0.123	-0.949**	0.522	0.521	-0.0456	0.00638	0.0129
	(0.818)	(1.462)	(0.938)	(0.362)	(0.504)	(0.324)	(0.342)	(0.420)	(0.360)	(0.0240)	(0.0416)	(0.0267)
N (Country x Year)	2747	2741	1526	2747	2741	1526	2747	2741	1526	2747	2741	1526
N (Country)	182	180	180	182	180	180	182	180	180	182	180	180
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table presents estimates regressing the founding of IT and OSS global- and mission-oriented ventures (using a machine learning classification approach) on lagged log GitHub commits. Columns 1, 4, 7, and 10 present OLS results. Columns 2, 5, 8, and 11 present 2SLS results with non-policy instruments for log GitHub commits spanning the full length of data, 2000–2016. Columns 3, 6, 9, and 12 present 2SLS results with all instruments for logged GitHub commits that span a subset of years in the data, 2000–2009. All columns include robust standard errors, clustered by country. Time fixed effects are relative to the year 2000. The regressions are not perfectly balanced by year, due to missing data in the control variable datasets. First stage estimates corresponding to the 2SLS specifications are shown in Table C.2.1 in the appendix. Robust standard errors, clustered by country. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.7: Impact of OSS on quality of ventures

	OLS	2SLS (Non-Policy Instruments)	2SLS - All Instruments (Pre-2010)	OLS	2SLS (Non-Policy Instruments)	2SLS - All Instruments (Pre-2010)	OLS	2SLS (Non-Policy Instruments)	2SLS - All Instruments (Pre-2010)
	Log Venture Value	Log Venture Value	Log Venture Value	Log Num. of Deals	Log Num. of Deals	Log Num. of Deals	Log Acquisitions	Log Acquisitions	Log Acquisitions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lagged Log GitHub	0.362***	0.602***	0.831***	0.206***	0.394***	0.460***	0.115***	0.310***	0.327***
	(0.0438)	(0.147)	(0.107)	(0.0258)	(0.0851)	(0.0640)	(0.0226)	(0.0780)	(0.0615)
Lagged Log Population	0.285***	0.0654	-0.180	0.172***	-0.00106	-0.105	0.0702**	-0.109	-0.112*
	(0.0734)	(0.165)	(0.104)	(0.0413)	(0.0898)	(0.0563)	(0.0263)	(0.0616)	(0.0468)
Lagged Human Capital Index	-0.796*	-1.150**	-1.422**	-0.392*	-0.671**	-0.746**	-0.328*	-0.618**	-0.567**
	(0.355)	(0.420)	(0.432)	(0.191)	(0.249)	(0.237)	(0.147)	(0.233)	(0.193)
Lagged Log GDP Capita	0.662***	0.536***	0.132	0.357***	0.258***	0.0412	0.269***	0.168**	0.0408
	(0.124)	(0.132)	(0.109)	(0.0723)	(0.0734)	(0.0587)	(0.0583)	(0.0525)	(0.0453)
Lagged Log Internet Users	-0.250	-0.533*	-0.504*	-0.141	-0.363**	-0.280*	-0.181***	-0.410***	-0.284**
	(0.129)	(0.220)	(0.199)	(0.0759)	(0.128)	(0.110)	(0.0537)	(0.109)	(0.0889)
_cons	-10.09***	-3.889	2.821	-5.657***	-1.229	1.822	-3.147***	1.229	1.867*
	(1.477)	(3.389)	(2.011)	(0.878)	(1.840)	(1.110)	(0.703)	(1.238)	(0.939)
N (Country x Year)	2747	2741	1526	2747	2741	1526	2747	2741	1526
N (Country)	182	180	180	182	180	180	182	180	180
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table presents estimates regressing the value and number of venture financing deals, as well as acquisitions on lagged log GitHub commits. Columns 1, 4, and 7 present OLS results. Columns 2,3, and 8 present 2SLS results with non-policy instruments for log GitHub commits spanning the full length of data, 2000–2016. Columns 3, 6, and 9 present 2SLS results with all instruments for logged GitHub commits that span a subset of years in the data, 2000–2009. All columns include robust standard errors, clustered by country. Time fixed effects are relative to the year 2000. The regressions are not perfectly balanced by year, due to missing data in the control variable datasets. First-stage estimates corresponding to the 2SLS specifications are shown in Table C.2.1 in the appendix. Robust standard errors, clustered by country. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results suggest that open-source software contributes to more of this narrow implementation of the definition of entrepreneurship.

4.4.4 Robustness

Table 4.9 below shows a summary of the hypotheses and our results. Each result points in the same qualitative direction. These results do not reject the causal relationship in which more OSS activity in a country leads to more venture founding. The evidence suggests OSS causes a wide range of different types of ventures.

We further test the relationship between OSS and global entrepreneurship using a difference-in-difference specification with the policy instrument as the shock (rather than using it as an instrument as done above). Consistent with hypothesis 1 and Table 4.2, Table 4.8 shows that the interaction term between whether countries instituted OSS policies and OSS is positive. Furthermore, OSS spurs new venture founding, particularly after countries implemented policies promoting OSS. Despite data on these global policies only being available through 2009, this alternate specification adds additional evidence to support our main arguments.

If the positive relationship between OSS and entrepreneurship holds in either direction, then we would see a stronger relationship between OSS and the formation of ventures closer in orientation to OSS. Software ventures are closer in orientation to OSS activity than are hardware ventures because OSS inherently involves the creation of software code. Therefore, we should see a stronger relationship between OSS and software ventures than hardware ventures. In the appendix (Table C.4.1), we conduct this test and find results generally consistent with expectations.

Our results also are robust to not lagging all of the controls and instruments. They are also robust to not lagging all of the endogenous variables, along with controls and instruments. Qualitatively similar results emerge.[28]

4.5 Discussion

To summarize, we considered the relationship between OSS and entrepreneurship across countries on the dimensions of both rate and direction. We find that more participation in open source in a country correlates with an increased rate of entrepreneurship activity, as either measured by information technology ventures or open source ventures. Using instrumental variables related to the supply, networking, and demand for open source, we

[28]Results are available upon request.

Table 4.8: Difference-in-difference: Impact of OSS on new venture founding

	Log IT Ventures	Log OSS Ventures
	(1)	(2)
Lagged Log GitHub x Lagged OSS Policy Instrument	0.144***	0.0746***
	(0.0316)	(0.0221)
Lagged Log GitHub	0.143***	0.0130*
	(0.0293)	(0.00592)
Lagged OSS Policy Instrument	-0.393	-0.300*
	(0.211)	(0.125)
Lagged Log Population	0.272***	0.0292*
	(0.0342)	(0.0133)
Lagged Human Capital Index	-0.232	-0.0969
	(0.254)	(0.0558)
Lagged Log GDP Capita	0.0950	0.00340
	(0.0529)	(0.0109)
Lagged Log Internet Users	0.368***	0.0243
	(0.0666)	(0.0152)
_cons	-5.425***	-0.506*
	(0.659)	(0.224)
N (Country x Year)	1530	1530
N (Country)	180	180
Time Fixed Effects	Yes	Yes

The table presents estimates regressing the founding of IT and OSS ventures on lagged log GitHub commits interacted with whether the year is after an OSS policy is instituted in a country. All columns include robust standard errors, clustered by country. Time fixed effects are relative to the year 2000. The regressions are not perfectly balanced by year, due to missing data in the control variable datasets. Robust standard errors, clustered by country. Standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.9: Summary of hypotheses, implementation, and results

Hypotheses	Implementation	Results
H1: An increase in OSS participation in a country leads to an increase in venture founding in that country.	We regress the founding of new ventures on OSS commits from the previous year.	OSS commits positively predict the formation of new ventures. ✓
H2a: OSS and country endowments are substitutes such that the impact of OSS participation on venture founding is greater for countries with lower endowments.	We regress the founding of new ventures on OSS commits in countries with high vs. low GDP per capita and human capital.	GDP per capita and human capital negatively moderate the relationship between OSS commits and the formation of new ventures. ✖
H2b: OSS and country endowments are complements such that the impact of OSS participation on venture founding is greater for countries with higher endowments.	We regress the founding of new ventures on OSS commits in countries with high vs. low GDP per capita and human capital.	GDP per capita and human capital positively moderate the relationship between OSS commits and the formation of new ventures. ✓
H3a: An increase in OSS participation in a country leads to an increase in globally-oriented venture founding in that country.	We regress the founding of new globally-oriented ventures on OSS commits.	OSS commits positively predict formation of new globally-oriented ventures. ✓
H3b: An increase in OSS participation in a country leads to an increase in mission-oriented entrepreneurship in that country.	We regress the founding of new mission-oriented ventures on OSS commits.	OSS commits positively predict formation of new mission-oriented ventures. ✓
H3c: An increase in OSS participation in a country leads to an increase in the quality of newly founded ventures in that country, as proxied by venture financing and acquisition.	We regress the number of venture financing deals, the total value of those deals, and the number of acquisitions on OSS commits.	OSS commits positively predict venture financing deals, the total value of those deals, and the number of acquisitions. ✓

cannot reject a causal interpretation. We further investigate whether the evidence suggests participation in OSS in a country substitutes or complements for endowments, and find evidence consistent with a complementary relationship. We also find that OSS has an additional effect on the direction of entrepreneurship, inducing those oriented toward global audiences or more community-driven missions. OSS also contributes to high-quality entrepreneurship.

Entrepreneurship varies across the globe for many reasons. This study adds open source as an additional cause. The findings are consistent with policies that treat open-source participation as an independent factor shaping the prevalence of innovative entrepreneurship in a country. The evidence that it is complementary to existing endowments should shape policies. Investment in OSS alone, or policies to encourage OSS through government support, will spur little entrepreneurial response in settings with low human capital. However, such activity will spur a greater response with an accompanying program investing in raising human capital. More pessimistically, this evidence suggests OSS will encounter many difficulties spreading to countries that lack programs to raise human capital.

These findings do not suggest that investors or policymakers focus on particular types of entrepreneurial ventures. Investors seeking to invest in emerging entrepreneurial ecosystems with a nurturing endowment set may consider open source as an important factor. Policymakers seeking to build innovative entrepreneurial ecosystems may use open source as a channel of development. Table C.5.1 shows that a one percent increase in 2016 GitHub commits predicts increases in ventures in the subsequent year around the world. While this increase looms large in a high-income economy like the US at 133 ventures, lower-income economies also benefit, for example, with India gaining 24 ventures, Brazil 11, China 5, Indonesia 2, and Mexico 1. Policies that reduce barriers to and/or incentivize participation in OSS may be important stimulants to realize the benefits of OSS for entrepreneurship. These policies need not be tailored to encourage specific outcomes.

This research also highlights a range of unanswered questions about open source's geographic and country-specific features. Many countries, such as India, China, Russia, Korea, and Ukraine, contain large open-source communities. We expect to observe careful studies of the micro-mechanisms that lead communities within those countries to either successfully start entrepreneurial efforts, or fail due to local institutional barriers or resource shortages. How much of entrepreneurship suffers in countries that lack the digital infrastructure to support rapid interaction with international repositories, such as in many African countries, or in countries where repressive governments interfere with internet traffic? Relatedly, how much do countries suffer from a lack of appropriate education or investment in the institutions that enable the development of appropriate human capital? Further, our efforts draw attention to the need for more investigation into the general supply and demand of OSS on a global scale, and how OSS contributions and usage could represent cross-border trade flows that go uncaptured in traditional economic analyses. Future work could explore whether large firms benefit from open source just as much or more than small ones. Questions related to the change in the direction of entrepreneurship induced by OSS can be ripe areas for future research as well. What impact do more global-oriented firms have on a country's entrepreneurial success? What spillovers does a mission-oriented venture have on the success of others?

We expect further research to bring finer geographic evidence to inform inferences. In addition, we hope for further research to explore the creation of OSS. This would involve understanding the variance in the supply and demand for open source across countries in more detail. It also would involve much more evidence about how existing trading and networking relationships contribute to open source. With such evidence in hand, an ideal study may gain insight into the pathways that encourage the growth of an ecosystem in which open source plays a nurturing role. Since many firms participate in open source communities, more participation in open source encourages new ventures, which encourages more participation, and so on. We have seen rapid technological change deepen gaps among countries, and this study shows the potential role of digital goods like OSS in either exacerbating or narrowing these gaps.

References

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

[2] Agrawal, A., Catalini, C., Goldfarb, A., & Luo, H. (2018). Slack time and innovation. *Organization Science*, 29(6), 1056–1073.

[3] Agrawal, A., Lacetera, N., & Lyons, E. (2016). Does standardized information in online markets disproportionately benefit job applicants from less developed countries?. *Journal of international Economics*, 103, 1–12.

[4] Aksulu, A., & Wade, M. R. (2010). A comprehensive review and synthesis of open source research. *Journal of the Association for Information Systems*, 11(11), 6.

[5] Alcácer, J., Cantwell, J., & Piscitello, L. (2016). Internationalization in the information age: A new era for places, firms, and international business networks? *Journal of International Business Studies*, 47(5), 499-512.

[6] Alcácer, J., Chung, W. (2007). Location strategies and knowledge spillovers. *Management Science*, 53(5), 760–776.

[7] 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.

[8] Angrist, J. D., & Keueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings?. *The Quarterly Journal of Economics*, 106(4), 979–1014.

[9] Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, 15(4), 69–85.

[10] Arikan, I., & Shenkar, O. (2013). National animosity and cross-border alliances. *Academy of Management Journal*, 56(6), 1516–1544.

[11] Arrow, KJ. (1962). Economic welfare and the allocation of resources for invention. Nelson R, ed. *The Rate and Direction of Inventive Activity: Economic and Social Factors* (pp. 609–626). Princeton, NJ: Princeton University Press.

[12] Arzaghi, M., Henderson, J. V. (2008). Networking off madison avenue. *The Review of Economic Studies*, 75(4), 1011–1038.

[13] Aulet, B. (2013). *Disciplined entrepreneurship: 24 steps to a successful startup*. John Wiley Sons.

[14] Axelson, U., Jenkinson, T., Stromberg, P., & Weisbach, M. S. (2013). Borrow cheap, buy high? The determinants of leverage and pricing in buyouts. *The Journal of Finance*, 68(6), 2223–2267.

[15] Balachandran, S., & Hernandez, E. (2018). Networks and innovation: Accounting for structural and institutional sources of recombination in brokerage triads. *Organization Science*, 29(1), 80–99.

[16] Balachandran, S., & Hernandez, E. (2020). Mi Casa Es Tu Casa: immigrant entrepreneurs as pathways to foreign venture capital investments. Working Paper. Available at SSRN 3331264.

[17] Barach, M. A., Golden, J. M., & Horton, J. J. (2020). Steering in online markets: the role of platform incentives and credibility. *Management Science*, 66(9), 4047–4070.

[18] Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.

[19] Baron, J. N., & Hannan, M. T. (2002). Organizational blueprints for success in high-tech start-ups: Lessons from the Stanford project on emerging companies. *California Management Review*, 44(3), 8–36.

[20] Bartik, T. J., & Bartik Timothy, J. (1991). Who benefits from state and local economic development policies? Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

[21] Beckman, C. M., & Burton, M. D. (2008). Founding the future: Path dependence in the evolution of top management teams from founding to IPO. *Organization Science*, 19(1), 3–24.

[22] Bell, R. G., Filatotchev, I., & Rasheed, A. A. (2012). The liability of foreignness in capital markets: Sources and remedies. *Journal of International Business Studies*, 43(2), 107–122.

[23] Ben-David, I., Graham, J. R., & Harvey, C. R. (2013). Managerial miscalibration. *The Quarterly Journal of Economics*, 128(4), 1547–1584.

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

[25] Berry, H., Guill én, M. F., & Zhou, N. (2010). An institutional approach to cross-national distance. *Journal of International Business Studies*, 41(9), 1460–1480.

[26] Bhidé, A. (2000). *The origin and evolution of new ventures*. Oxford University Press.

[27] Bingham, C. B. (2009). Oscillating improvisation: How entrepreneurial firms create success in foreign market entries over time. *Strategic Entrepreneurship Journal*, 3(4), 321–345.

[28] Bingham, C. B., & Davis, J. P. (2012). Learning sequences: Their existence, effect, and evolution. *Academy of Management Journal*, 55(3), 611–641.

[29] Bingham, C. B., & Eisenhardt, K. M. (2011). Rational heuristics: The “simple rules” that strategists learn from process experience. *Strategic Management Journal*, 32(13), 1437–1464.

[30] Bingham, C. B., Eisenhardt, K. M., & Furr, N. R. (2007). What makes a process a capability? Heuristics, strategy, and effective capture of opportunities. *Strategic Entrepreneurship Journal*, 1(1-2), 27–47.

[31] Blank, S. (2013). Why the lean start-up changes everything. *Harvard Business Review*, 91(5), 63–72.

[32] Block, J., & Sandner, P. (2009). What is the effect of the financial crisis on venture capital financing? Empirical evidence from US Internet start-ups. *Venture Capital*, 11(4), 295–309.

[33] Blatter, M., & Niedermayer, A. (2009). Career Concerns and Career Choice. Working Paper.

[34] Bloom, N., Genakos, C., Sadun, R., & Van Reenen, J. (2012a). Management practices across firms and countries. *Academy of Management Perspectives*, 26(1), 12—33.

[35] Bloom, N., Sadun, R., & Van Reenen, J. (2012b). Americans do IT better: US multinationals and the productivity miracle. *American Economic Review*, 102(1), 167–201.

[36] Bloom, N., Schankerman, M., & Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4), 1347–1393.

[37] Bloom, N., & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, 122(4), 1351–1408.

[38] 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.

[39] Brooks, A. W., Huang, L., Kearney, S. W., Murray, F. E. (2014). Investors prefer entrepreneurial ventures pitched by attractive men. *Proceedings of the National Academy of Sciences*, 111(12), 4427–4431.

[40] Brown, S. L., & Eisenhardt, K. M. (1997). The art of continuous change: Linking complexity theory and time-paced evolution in relentlessly shifting organizations. *Administrative Science Quarterly*, 1–34.

[41] Bruneel, J., Yli-Renko, H., & Clarysse, B. (2010). Learning from experience and learning from others: how congenital and interorganizational learning substitute for experiential learning in young firm internationalization. *Strategic Entrepreneurship Journal*, 4(2), 164–182.

[42] 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.

[43] Business Insider. (2018). These founders got rejected from Y Combinator four times. Here's how they finally got in. https://www.businessinsider.in/These-founders-got-rejected-from-Y-Combinator-four-times-Heres-how-they-finally-got-in-/articleshow/66644585.cms

[44] Cahn, C., Girotti, M., & Landier, A. (2021). Entrepreneurship and information on past failures: A natural experiment. *Journal of Financial Economics*, 141(1), 102–121.

[45] Camuffo, A., Cordova, A., Gambardella, A., & Spina, C. (2020). A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. *Management Science*, 66(2), 564–586.

[46] Cao, R., Koning, R. M., & Nanda, R. (2021). Biased sampling of early users and the direction of startup innovation. NBER Working Paper, No. w28882.

[47] Carlson, N. A. (2022). Differentiation in microenterprises. *Strategic Management Journal*, 1–27.

[48] Catalini, C., Guzman, J., & Stern, S. (2019). Hidden in plain sight: venture growth with or without venture capital. NBER Working Paper, No. w26521.

[49] Chadha, S., Kleinbaum, A. M., & Wood, A. (2022). Social Networks are Shaped by Culturally Contingent Assessments of Social Competence. Working Paper.

[50] Chakraborty, I., & Ewens, M. (2018). Managing performance signals through delay: Evidence from venture capital. *Management Science*, 64(6), 2875–2900.

[51] Charmaz, K. (2014). *Constructing grounded theory*. Sage.

[52] Chatterji, A. K. (2009). Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry. *Strategic Management Journal*, 30(2), 185–206.

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

[54] Chen, H., Gompers, P., Kovner, A., & Lerner, J. (2010). Buy local? The geography of venture capital. *Journal of Urban Economics*, 67(1), 90–102.

[55] Chung, W., & Alcácer, J. (2002). Knowledge seeking and location choice of foreign direct investment in the United States. *Management Science*, 48(12), 1534–1554.

[56] Clough, D. R., Fang, T. P., Vissa, B., & Wu, A. (2019). Turning lead into gold: How do entrepreneurs mobilize resources to exploit opportunities? *Academy of Management Annals*, 13(1), 240–271.

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

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

[59] Conti, A., & Guzman, J. A. (2021). What is the US comparative advantage in entrepreneurship? Evidence from Israeli migration to the United States. *The Review of Economics and Statistics*, 1–45.

[60] Conti, A., Guzman, J., & Rabi, R. (2020). Herding in the Market for Startup Acquisitions. Working Paper. Available at SSRN 3678676.

[61] Conti, A., Peukert, C., & Roche, M. P. (2021). Beefing IT up for your Investor? Open Sourcing and Startup Funding: Evidence from GitHub. Open Sourcing and Startup Funding: Evidence from GitHub. Working Paper.

[62] Conti, A., & Roche, M. P. (2021). Lowering the bar? External conditions, opportunity costs, and high-tech start-up outcomes. *Organization Science*, 32(4), 965–986.

[63] 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.

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

[65] Crick, D., & Jones, M. V. (2000). Small high-technology firms and international high-technology markets. *Journal of International Marketing*, 8(2), 63–85.

[66] Cui, V., H. Yang, & I. Vertinsky. (2018). "Attacking your partners: Strategic alliances and competition between partners in product markets." *Strategic Management Journal*, 39(12), 3116–3139.

[67] Cullen, Z., & Farronato, C. (2021). Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms. *Management Science*, 67(7), 3985–4003.

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

[69] Dalle, J. M., Den Besten, M., & Menon, C. (2017). Using Crunchbase for economic and managerial research. OECD Science, Technology and Industry Working Paper 2017/08.

[70] Delacroix, J., & Carroll, G. R. (1983). Organizational foundings: An ecological study of the newspaper industries of Argentina and Ireland. *Administrative Science Quarterly*, 28(2), 274–-291.

[71] Delgado, M., Porter, M. E., & Stern, S. (2016). Defining clusters of related industries. *Journal of Economic Geography*, 16(1), 1–38.

[72] Delmar, F., & Shane, S. (2003). Does business planning facilitate the development of new ventures? *Strategic Management Journal*, 24(12), 1165–1185.

[73] Dencker, J. C., Gruber, M., & Shah, S. K. (2009). Pre-entry knowledge, learning, and the survival of new firms. *Organization Science*, 20(3), 516–537.

[74] DeSantola, A., & Gulati, R. (2017). Scaling: Organizing and in entrepreneurial ventures. *Academy of Management Annals*, 11(2), 640–668.

[75] DeSantola, A., Gulati, R., & Zhelyazkov, P. I. (2022). External interfaces or internal processes? Market positioning and divergent professionalization paths in young ventures. *Organization Science*, 0(0).

[76] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

[77] DiBona, C., & Ockman, S. (1999). *Open sources: Voices from the open source revolution*. O'Reilly Media, Inc.

[78] DiBona, C., Stone, M., & Cooper, D. (2005). *Open sources 2.0: The continuing evolution*. O'Reilly Media, Inc.

[79] 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.

[80] Duke Fuqua School of Business (Duke). (2021). The Duke entrepreneurship manual: the entrepreneurial process. https://sites.fuqua.duke.edu/dukeven/selected-topics/the-entrepreneurial-process/.

[81] Dushnitsky, G., & Stroube, B. K. (2021). Low-code entrepreneurship: Shopify and the alternative path to growth. *Journal of Business Venturing Insights*, 16, e00251.

[82] Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550.

[83] Eisenhardt, K. M., & Bingham, C. B. (2017). Superior strategy in entrepreneurial settings: Thinking, doing, and the logic of opportunity. *Strategy Science*, 2(4), 246–257.

[84] Eisenhardt, K. M., & Schoonhoven, C. B. (1990). Organizational growth: Linking founding team, strategy, environment, and among US semiconductor ventures, 1978-1988. *Administrative Science Quarterly*, 35(3), 504–529.

[85] Eisenmann, T., & Wagonfeld, A. (2014). Scaling a startup: People and organizational issues. *Harvard Business School Background Note*, No. 812–100.

[86] Elfenbein, D. W., Hamilton, B. H., & Zenger, T. R. (2010). The small firm effect and the entrepreneurial spawning of scientists and engineers. *Management Science*, 56(4), 659–681.

[87] Ewens, M., Nanda, R., & Rhodes-Kropf, M. (2018). Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics*, 128(3), 422–442.

[88] Fallick, B., Fleischman, C. A., & Rebitzer, J. B. (2006). Job-hopping in Silicon Valley: Some evidence concerning the microfoundations of a high-technology cluster. *The Review of Economics and Statistics*, 88(3), 472–481.

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

[90] Flamm, K. (1988). *Creating the computer: government, industry, and high technology*. Brookings Institution Press.

[91] 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.

[92] 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.

[93] Gans, J. S., Kearney, M., Scott, E. L., & Stern, S. (2021). Choosing technology: An entrepreneurial strategy approach. *Strategy Science*, 6(1), 39–53.

[94] Gans, J. S., Stern, S., & Wu, J. (2019). Foundations of entrepreneurial strategy. *Strategic Management Journal*, 40(5), 736–756.

[95] Gao, C., Zuzul, T., Jones, G., & Khanna, T. (2017). Overcoming institutional voids: A reputation-based view of long-run survival. *Strategic Management Journal*, 38(11), 2147–2167.

[96] Gavetti, G., Levinthal, D. A., & Rivkin, J. W. (2005). Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, 26(8), 691–712.

[97] Gavetti, G., & Rivkin, J. W. (2007). On the origin of strategy: Action and cognition over time. *Organization Science*, 18(3), 420–439.

[98] Gavetti, G., & Porac, J. (2018). On the origin of great strategies. *Strategy Science*, 3(1), 352–365.

[99] Ghemawat, P. (2001). Distance still matters. *Harvard Business Review*, 79(8), 137–147.

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

[101] GitHub Blog. (2010). "GitHub in your language." https://GitHub.blog/20100713GitHub-in-your-language/.

[102] GitHub. (2019). "GitHub Milestones." https://github.com/about/milestones.

[103] Glaeser, E. L., Kerr, S. P., & Kerr, W. R. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *The Review of Economics and Statistics*, 97(2), 498–520.

[104] Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Mill Valley, CA: Sociology Press.

[105] 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.

[106] Gompers, P., Lerner, J., & Scharfstein, D. (2005). Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *The Journal of Finance*, 60(2), 577–614.

[107] 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.

[108] Greenstein, S., & Nagle, F. (2014). Digital dark matter and the economic contribution of Apache. *Research Policy*, 43(4), 623–631.

[109] Gupta, B., & Khanna, T. (2019). A Recombination-Based Internationalization Model: Evidence from Narayana Health's Journey from India to the Cayman Islands. *Organization Science*, 30(2), 405–425.

[110] Guzman, J., & Li, A. (2022). Measuring founding strategy. *Management Science*, 0(0).

[111] Guzman, J., & Stern, S. (2020). The state of American entrepreneurship: New estimates of the quantity and quality of entrepreneurship for 32 US States, 1988–2014. *American Economic Journal: Economic Policy*, 12(4), 212–43.

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

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

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

[115] Hallen, B. L., Cohen, S. L., & Park, S. H. (2023). Are Seed Accelerators Status Springboards for Startups? Or Sand Traps? *Strategic Management Journal*.

[116] Hann, I. H., Roberts, J. A., & Slaughter, S. A. (2013). All are not equal: An examination of the economic returns to different forms of participation in open source software communities. *Information Systems Research*, 24(3), 520–538.

[117] 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.

[118] Hoenig, D., & Henkel, J. (2015). Quality signals? The role of patents, alliances, and team experience in venture capital financing. *Research Policy*, 44(5), 1049–1064.

[119] Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind*. McGraw-Hill.

[120] Howell, S. T. (2017). Financing innovation: Evidence from RD grants. *American Economic Review*, 107(4), 1136–64.

[121] Howell, S. T. (2020). Reducing information frictions in venture capital: The role of new venture competitions. *Journal of Financial Economics*, 136(3), 676–694.

[122] 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.

[123] Huang, L. (2018). The role of investor gut feel in managing complexity and extreme risk. *Academy of Management Journal*, 61(5), 1821–1847.

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

[125] Hymer, S. H. (1976). textitThe international operations of national firms: A study of foreign direct investment. MIT Press.

[126] Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., & Yesiltas, S. (2015). How to construct nationally representative firm level data from the Orbis global database: New facts and aggregate implications. NBER Working Paper, No. w21558.

[127] Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review*, 83(7), 172.

[128] Kaplan, S. N., & Lerner, J. (2016). Venture capital data: Opportunities and challenges. *Measuring entrepreneurial businesses: Current knowledge and challenges*, 413–431.

[129] 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.

[130] Keller, S., Korkmaz, G., Robbins, C., & Shipp, S. (2018). Opportunities to observe and measure intangible inputs to innovation: Definitions, operationalization, and examples. *Proceedings of the National Academy of Sciences*, 115(50), 12638–12645.

[131] Kerr, W. R. (2008). Ethnic scientific communities and international technology diffusion. *The Review of Economics and Statistics*, 90(3), 518–537.

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

[133] Kerr, W. R. (2018). *The gift of global talent: How migration shapes business, economy & society*. Stanford University Press.

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

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

[136] Kerr, W. R., & Robert-Nicoud, F. (2020). Tech clusters. *Journal of Economic Perspectives*, 34(3), 50–76.

[137] Khanna, T. (2014). Contextual intelligence. *Harvard Business Review*, 92(9), 58–68.

[138] Khanna, T., & Palepu, K. (1997). Why focused strategies may be wrong for emerging markets. *Harvard Business Review*, 75(4), 41–51.

[139] Khanna, T., Palepu, K. G., & Sinha, J. (2005). Strategies that fit emerging markets. *Harvard Business Review*, 83(6), 4–19.

[140] Khanna, T., & Palepu, K. G. (2010). Winning in emerging markets: A road map for strategy and execution. Harvard Business Press.

[141] Khanna, T., & Rivkin, J. W. (2001). Estimating the performance effects of business groups in emerging markets. *Strategic Management Journal*, 22(1), 45–74.

[142] Knight, G. (1996). The Born Global Firm: A Challenge to Traditional Internationalization Theory,[in:] *Advances in International Marketing*, eds. S. Cavusgil, T. Madsen.

[143] Knight, G. A., & Cavusgil, S. T. (2004). Innovation, organizational capabilities, and the born global firm. *Journal of International Business Studies*, 35(2), 124–141.

[144] Knight, G. A., & Cavusgil, S. T. (2005). A taxonomy of born-global firms. *Management International Review*, 15–35.

[145] Kirtley, J., & O'Mahony, S. (2020). What is a pivot? Explaining when and how entrepreneurial firms decide to make strategic change and pivot. *Strategic Management Journal*, 44(1), 197–230.

[146] Kogut, B., & Metiu, A. (2001). Open Source Software Development and Distributed Innovation, *Oxford Review of Economic Policy*, 17(2), 248–264.

[147] Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and start-up performance: Evidence from A/B testing. *Management Science*, 0(0).

[148] Koning, R., Samila, S., & Ferguson, J. P. (2020). Inventor Gender and the Direction of Invention. In *AEA Papers and Proceedings*, 110, 250–54.

[149] Koning, R., Samila, S., & Ferguson, J. P. (2021). Who do we invent for? Patents by women focus more on women's health, but few women get to invent. *Science*, 372(6548), 1345–1348.

[150] Kulchina, E., & Hernandez, E. (2016). Immigrants and Firm Performance: Effects on Foreign Subsidiaries versus Foreign Entrepreneurs. *Academy of Management Proceedings*.

[151] Lakhani, K. R., & von Hippel, E. (2003). How Open Source Software Works: "Free" User-to User Assistance. *Research Policy*, 32(6), 923–943.

[152] Lakhani, K. R., & Wolf, R. G. (2003). Why hackers do what they do: Understanding motivation and effort in free/open source software projects. Working Paper.

[153] Landier, A. (2005). Entrepreneurship and the Stigma of Failure. Available at SSRN 850446.

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

[155] Lee, K., & Lim, C. (2001). Technological regimes, catching-up and leapfrogging: findings from the Korean industries. *Research Policy*, 30(3), 459–483.

[156] Lee, K., & Malerba, F. (2017). Catch-up cycles and changes in industrial leadership: Windows of opportunity and responses of firms and countries in the evolution of sectoral systems. *Research Policy*, 46(2), 338–351.

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

[158] Lee, S. (2022). The myth of the flat start-up: Reconsidering the organizational structure of start-ups. *Strategic Management Journal*, 43(1), 58–92.

[159] Lee, S. R., & Kim, J. D. (2022). When do startups scale? Large-scale evidence from job postings. Available at SSRN: https://ssrn.com/abstract=4015530.

[160] Lerner, J. & Schankerman, M. (2010). *The comingled code: Open source and economic development*. MIT Press Books.

[161] 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.

[162] Lerner, J., & Stern, S. (2012). *The Rate and Direction of Inventive Activity Revisited*. University of Chicago Press.

[163] Lerner, J. & Tirole, J. (2002). Some simple economics of open source. *The Journal of Industrial Economics*, 50(2), 197–234.

[164] Lewis, J.A. (2010). *Government Open Source Policies*. Center for Strategic and International Studies.

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

[166] 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.

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

[168] Lu, J. W., & Beamish, P. W. (2001). The internationalization and performance of SMEs. *Strategic Management Journal*, 22(6-7), 565–586.

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

[170] Lyons, E., & Zhang, L. (2018). Who does (not) benefit from entrepreneurship programs?. *Strategic Management Journal*, 39(1), 85–112.

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

[172] March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87.

[173] McDonald, R., & Eisenhardt, K. (2020). Parallel play: Startups, nascent markets, and the search for a viable business model. *Administrative Science Quarterly*, 65(2), , 483–523.

[174] McDonald, R., & Gao, C. (2019). Pivoting isn't enough? Managing strategic reorientation in new ventures. *Organization Science*, 30(6), 1289–1318.

[175] Michelacci, C., & Silva, O. (2007). Why so many local entrepreneurs? *The Review of Economics and Statistics*, 89(4), 615–633.

[176] Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic Management Journal*, 6(3), 257–272.

[177] Mitra, S. (2018). Building a Fast Growth, Cutting-Edge Insurance Brokerage: Karn Saroya, CEO of Cover. One Million by One Million Blog. https://www.sramanamitra.com/2018/10/22/building-a-fast-growth-cutting-edge-insurance-brokerage-karn-saroya-ceo-of-cover-part-1/.

[178] Medappa, P.K. & Srivastava, S.C. (2019). Does superposition influence the success of FLOSS projects? An examination of open-source software development by organizations and individuals. *Information Systems Research*, 30(3), 764–786.

[179] Mocanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q., & Vespignani, A. (2013). The twitter of babel: Mapping world languages through microblogging platforms. *PloS one*, 8(4), e61981.

[180] Mogstad, M., Torgovitsky, A., & Walters, C. R. (2021). The causal interpretation of two-stage least squares with multiple instrumental variables. *American Economic Review*, 111(11), 3663–98.

[181] Monroe, B. L., Colaresi, M. P., & Quinn, K. M. (2008). Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4), 372–403.

[182] Morck, R., & Yeung, B. (1992). Internalization: an event study test. *Journal of International Economics*, 33(1-2), 41–56.

[183] Nagle, F. (2018). Learning by Contributing: Gaining Competitive Advantage Through Contribution to Crowdsourced Public Goods. *Organization Science* 29(4), 569–587.

[184] Nagle, F. (2019a). The Impact of Technology Procurement Regulations on Social Value and National Competitiveness. Harvard Business School Working Paper, No. 19 103.

[185] Nagle, F. (2019b). Open Source Software and Firm Productivity. *Management Science*, 65(3), 1191–1215.

[186] Nagle, F., Wheeler, D. A., Lifshitz-Assaf, H., Ham,H., & Hoffman, J. 2020. *Report on the 2020 FOSS Contributor Survey*. The Linux Foundation Core Infrastructure Initiative.

[187] Nanda, R., & Khanna, T. (2010). Diasporas and domestic entrepreneurs: Evidence from the Indian software industry. *Journal of Economics Management Strategy*, 19(4), 991–1012.

[188] 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.

[189] Nanda, R., & Sørensen, J.B. (2010). Workplace peers and entrepreneurship. *Management Science*, 56(7), 1116–1126.

[190] Neath, K. (2010). Introducing Organizations. GitHub Blog. https://github.blog/2010-06-29-introducing-organizations/.

[191] Nelson, R. (1962). Introduction to The Rate and Direction of Inventive Activity: Economic and Social Factors. In *The Rate and Direction of Inventive Activity: Economic and Social Factors* (pp. 1–16). Princeton University Press.

[192] Nelson, T. (2003). The persistence of founder influence: Management, ownership, and performance effects at initial public offering. *Strategic Management Journal*, 24(8), 707–724.

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

[194] Ocasio, W., & Joseph, J. (2018). The attention-based view of great strategies. *Strategy Science*, 3(1), 289–294.

[195] QS World University Rankings. (2021). QS World University Rankings 2022. Retrieved from https://www.topuniversities.com/university-rankings/world-university-rankings/2021.

[196] Ott, T. E., Eisenhardt, K. M., & Bingham, C. B. (2017). Strategy formation in entrepreneurial settings: Past insights and future directions. *Strategic Entrepreneurship Journal*, 11(3), 306–325.

[197] Oviatt, B. M., & McDougall, P. P. (2005). Toward a theory of international new ventures. *Journal of International Business Studies*, 36(1), 29–41.

[198] Ozcan, P., & Eisenhardt, K. M. (2009). Origin of alliance portfolios: Entrepreneurs, network strategies, and firm performance. *Academy of Management Journal*, 52(2), 246–279.

[199] Pisano, G. P. (1994). Knowledge, integration, and the locus of learning: An empirical analysis of process development. *Strategic Management Journal*, 15(S1), 85–100.

[200] PitchBook (2022). Company profile. Retrieved from Pitchbook database.

[201] Porter, M. (1996). "What is Strategy?" *Harvard Business Review*, 74(6), 61–-78.

[202] Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using siamese BERT-networks. arXiv preprint arXiv:1908.10084.

[203] Rivkin, J. W. (2000). Imitation of complex strategies. *Management Science*, 46(6), 824–844.

[204] Rivkin, J. W., & Siggelkow, N. (2003). Balancing search and stability: Interdependencies among elements of organizational design. *Management Science*, 49(3), 290–311.

[205] Robbins, C. A., Korkmaz, G., Calderón, J. B. S., Chen, D., Kelling, C., Shipp, S., Keller, S. (2018). Open Source Software as Intangible Capital: Measuring the Cost and Impact of Free Digital Tools. In Paper from 6th IMF Statistical Forum on Measuring Economic Welfare in the Digital Age: What and How, 19–20.

[206] Santos, F. M., & Eisenhardt, K. M. (2009). Constructing markets and shaping boundaries: Entrepreneurial power in nascent fields. *Academy of Management Journal*, 52(4), 643–671.

[207] Saxenian, A. (1996). *Regional advantage: Culture and competition in Silicon Valley and Route 128*. Harvard University Press.

[208] Shah, S.K. (2006). Motivation, governance, and the viability of hybrid forms in open source software development. *Management Science*, 52(7), 1000–1014.

[209] Shaheer, N. A., & Li, S. (2020). The CAGE around cyberspace? How digital innovations internationalize in a virtual world. *Journal of Business Venturing*, 35(1), 105892.

[210] Shane, S. (2000). Prior knowledge, and the discovery of entrepreneurial opportunities. *Organization Science*, 11(4), 448–469.

[211] Shane, S., & Cable, D. (2002). Network ties, reputation, and the financing of new ventures. *Management Science*, 48(3), 364–381.

[212] Schilling, M. A. (2018). The cognitive foundations of visionary strategy. *Strategy Science*, 3(1), 335–342.

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

[214] Siggelkow, N. (2001). Change in the presence of fit: The rise, the fall, and the renaissance of Liz Claiborne. *Academy of Management Journal*, 44(4), 838–857.

[215] Sørensen, J. B. & Carroll, G. R. (2021). *Making great strategy: Arguing for organizational advantage*. Columbia University Press.

[216] Sørensen, J. B. & Sorenson, O. (2003). From conception to birth: Opportunity perception and resource mobilization in entrepreneurship. In *Geography and Strategy*. Emerald Group Publishing Limited.

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

[218] Sorenson, O., & Stuart, T. E. (2001). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology*, 106(6), 1546–1588.

[219] Staiger, D., & Stock, J.H. (1997). Instrumental Variables Regression with Weak Instruments, *Econometrica*, 65, 557–586.

[220] Stanton, C., & Thomas, C. (2015). Landing the First Job: The Value of the Intermediaries in Online Hiring. *Review of Economics Studies*, 82(2), 810–854.

[221] Startup Genome Project. (2021). Rankings 2021: Top 30 + Runners-up. Retrieved from https://startupgenome.com/article/rankings-2021-top-30-plus-runners-up.

[222] Stinchcombe, A.L. (1965). Social structure and organizations. In: March, J.P., Ed., *Handbook of Organizations*, Rand McNally, 142–193.

[223] Stock J., & Yogo, M. (2005). *Testing for Weak Instruments in Linear IV Regression. In: Andrews DWK Identification and Inference for Econometric Models*, Cambridge University Press: New York, 80–108.

[224] Stuart, T., & Sorenson, O. (2003). The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research Policy*, 32(2), 229–253.

[225] Symeonidou, N., Bruneel, J., & Autio, E. (2017). Commercialization strategy and internationalization outcomes in technology-based new ventures. *Journal of Business Venturing*, 32(3), 302–317.

[226] Teece, D. J. (1982). Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior & Organization*, 3(1), 39–63.

[227] Teece, D. J. (1985). Multinational enterprise, internal governance, and industrial organization. *The American Economic Review*, 75(2), 233–238.

[228] Teece, D. J. (1986). Transactions cost economics and the multinational enterprise an assessment. *Journal of Economic Behavior & Organization*, 7(1), 21–45.

[229] Tambe, P., & Hitt, L. M. (2014). Job hopping, information technology spillovers, and productivity growth. *Management Science*, 60(2), 338–355.

[230] Van den Steen, E. (2017). A formal theory of strategy. *Management Science*, 63(8), 2616–2636. *Varian 2003* Varian, Hal, R., & Shapiro, C. (2003). Linux adoption in the public sector: An economic analysis. Manuscript. University of California, Berkeley.

[231] Von Krogh, G., Haefliger, S., Spaeth, S., & Wallin, M. W. (2012). Carrots and rainbows: Motivation and social practice in open source software development. *MIS Quarterly*, 36(2), 649–676.

[232] Vissa, B. (2011). A matching theory of entrepreneurs' tie formation intentions and initiation of economic exchange. *Academy of Management Journal*, 54(1), 137–158.

[233] Vissa, B., & Chacar, A. S. (2009). Leveraging ties: The contingent value of entrepreneurial teams' external advice networks on Indian software venture performance. *Strategic Management Journal*, 30(11), 1179–1191.

[234] Wang, D. (2015). Activating cross-border brokerage: Interorganizational knowledge transfer through skilled return migration. *Administrative Science Quarterly*, 60(1), 133–176.

[235] Wang, D. J. (2020). When do return migrants become entrepreneurs? The role of global social networks and institutional distance. *Strategic Entrepreneurship Journal*, 14(2), 125–148.

[236] Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21(1), 1–19.

[237] Wen, W., Ceccagnoli, M., & Forman, C. (2016). Opening Up Intellectual Property Strategy: Implications for Open Source Software Entry by Start-up Firms. *Management Science*, 62(9), 2668–2691.

[238] Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2), 171–180.

[239] Wormald, A., Agarwal, R., Braguinsky, S., & Shah, S. K. (2021). David overshadows Goliath: Specializing in generality for internationalization in the global mobile money industry. *Strategic Management Journal*, 42(8), 1459–1489.

[240] Wright, N. L., Koning, R., & Khanna, T. (2023). Judging foreign startups. *Strategic Management Journal*.

[241] Wright, N. L., Nagle, F., & Greenstein, S. M. (2020). Open source software and global entrepreneurship. Harvard Business School Working Paper, No. 20–139.

[242] Wright, N. L. (2023). Where Strategy Matters: Evidence from a Global Startup Field Study. HBS Working Paper, No. 23–041. Available at SSRN 4335210.

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

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

[245] Yang, M. J., Christensen, M., Bloom, N., Sadun, R., & Rivkin, J. (2020). How do CEOs make strategy? NBER Working Paper, No. w27952.

[246] Yin, R. K. (2003). *Case Study Research–Design and Methods*, Sage, Thousand Oaks, California.

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

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

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

[250] Zaheer, S., & Mosakowski, E. (1997). The dynamics of the liability of foreignness: A global study of survival in financial services. *Strategic Management Journal*, 18(6), 439–463.

[251] Zhang, C., Tan, J., & Tan, D. (2016). Fit by adaptation or fit by founding? A comparative study of existing and new entrepreneurial cohorts in China. *Strategic Management Journal*, 37(5), 911–931.

A
Chapter 1 Supplementary Material

Contents

A.1 Interview protocol and scoring 165
A.2 Coding examples 167
A.3 Sampling checks 180
A.4 Machine learning approach to strategy score 184
A.5 Alternative measures of strategy 185
A.6 Robustness checks using LASSO specifications 186
A.7 Results robust to a country index of the ease of recovering from mistakes 188
A.8 Interviews suggest that mistakes are more costly in non-US contexts. 192
A.9 Strategy can help companies avoid mistakes 193
A.10 Content of strategy in US versus non-US firms 195
A.11 US strategies are less geographically oriented. 199
A.12 Organizational approaches are localized. 201
A.13 Mediation analysis using US investor and executives' US work experience 206
A.14 Robustness checks with other outcomes 207

A.1 Interview protocol and scoring

Measuring strategy relied on structured interviews with executives of software companies that leveraged a standardized interview protocol of open-ended questions asking executives to recount their companies' objectives, market scope, moat, organizational design, and organizational culture choices. Supplementary questions elicited how the executives prioritized their next three action items, how they were using Series A funding, what they perceived as their biggest uncertainties, and why they did not pursue particular alternative approaches; for example, expanding across geographies rather than across industries. These questions are below.

- Objective
  - What is your objective for the company over the next 3–5 years? If you had your way, where would the company be at that point?
- Market scope
  - How are you planning to expand your markets and customer penetration?
- Moat
  - In the years to come, what do you see as your moat relative to competitors? Why won't other players take over this market?
- Organizational design
  - How are you planning to build out your team and organization?
- Organizational culture
  - How would you define the culture you are seeking to build or consider you have already achieved in the organization?
- Supplementary questions
  - At the beginning of the conversation, you mentioned your objective for the company. If you could boil it down to the three actions you need to take to make that objective possible, what would they be?

- How are you using or planning to use your Series A funding?
- What is the biggest source of uncertainty that faces your company in the next 3–5 years? How are you addressing it?
- Why did you not pursue [insert particular alternative] to [expand your market or design your organization]?

Research associates (MBA students and those with similar experience) evaluated the transcribed responses to the above questions and used the following scoring rubric to evaluate them.

- **External alignment**: On a scale of 1–5, how logical is the response? (1= response is illogical—the conclusion DOES NOT at all follow from assumptions; it is internally inconsistent. 3= response is somewhat logical but has some internal inconsistencies. 5= response is logical.) On a scale of 1–5, how does this response change your beliefs in the success of this company? (1= reduces your beliefs; 3= doesn't change your beliefs; 5 = increases your beliefs.)
- **Internal alignment**: On a scale of 1–5, how aligned is this response with the executive's [objective, market scope, organizational design, etc.]? (1=not aligned; 3 = somewhat aligned; 5 = very aligned)?

A.2 Coding examples

The following section shows examples of market scope, moat, organizational design, and organizational culture responses and their strategy scores. For simplicity, the section focuses on scoring the fit of these responses with executives' assumptions (external alignment) and objectives (part of internal alignment).

Market scope examples

Table A.2.2 shows examples of companies' market scope responses and scores. A France-based company developing a video communication tool (third row) has a relatively low market scope score. Its plan to expand its market scope through direct sales and gradual internationalization from Europe to the US are reasonable independently but are misaligned with one another. One might expect that grabbing the biggest market in the US first (rather than in Europe) would be important for solidifying market share and would be more fitting with the direct sales approach to get customers with "bigger revenue...and less churn." This misalignment between two independently reasonable approaches results in an average external alignment score. By pushing off entry into a market with high-value customers and risking its stake in that market, the company also risks its ability to achieve its objective to "build a new category of product" and correspondingly achieve a "1 billion-plus valuation." The misalignment between the market scope approach and objective results in a lower internal alignment score.

As an example of a high market scope score, the US-based company developing a recruiting platform (first row) is expanding its market by partnering with enterprise human resource systems to enable others to build on top. By gaining customer value from incumbents and innovations from third-party entrants in the "quarter-of-a-trillion-dollar market" in HR technologies, the company conveys a plan that fits well with its assumptions of increasing value across markets. Unlike the gradual international expansion approach planned by the France-based company, this company targets the big players in the business first so it can get the first-mover advantage to be the data infrastructure on which others can build. The "API-first" approach enables the company to reach major HR systems seamlessly. This interoperable approach aligns well with its objective to be an "integrated data platform behind many of the HR technologies and big people data systems."

The Brazil-based company (second row) pursuing a document automation platform takes a direct sales approach, like the other two companies above. Unlike the deliberate approach pursued by the US company, the Brazilian company started with enterprise sales because it was familiar with the approach and thought it would be “easy.” However, it soon realized it was mistaken; this was a market with “complicated sales cycles.” The executive admitted that “it [would have been] better to start with the middle or SMB market.” However, by focusing on a shorter-term feasible customer approach first, the company chose an option that ended up being difficult to sustain in Brazil, resulting in an average external alignment score. The company has the potential to achieve substantial financial outcomes by targeting large enterprises in Brazil, where there is a promising market opportunity to “solve a huge bureaucracy problem.” This market focus fits with the company’s objective to get “over 100 million reais in revenue” and an IPO. It would be even more fitting if the company addressed how it would mitigate risks around the “complicated sales cycles.” Together, this plan has a relatively strong fit with the executive’s objective, resulting in a high internal alignment score.

Table A.2.1: Market scope examples

Subscore	Desc.	Objective	Market Scope	HQ
21	Recruiting platform using AI/ML to provide talent to enterprises	In the next three to five years, [the company] will be the integrated data platform behind many of the HR technologies and big people data systems.	We created integrated partnerships with the largest HCMs. And this has been really helpful, because our whole message was, we're going to make your system better. And it caused those systems to want to sell us...And because the whole technology was built as an API-first business, we started to get others saying, hey, can we build on top of these APIs? And then there's about a quarter of a trillion dollars in HR tech that's out there in the world.	US
16	Platform using AI to automate procurement of documents in Brazil for companies.	...For the next three to five years...we will achieve over 100 million reais in revenue...which means, probably, we could see some IPO in the next three or five years.	It's hard because here in Brazil, what we are trying to do is solve a huge bureaucracy problem. So more than understanding our market, we are creating disruptive products. And it's a B2B enterprise market. So the sales cycle [is] more complicated. So we sell for the huge—the most important banks in Brazil, real estate companies, and also agro-companies. So I think the strategy is to focus on outbound sales. And we spend a lot of money to get this whole market, the enterprise market. And after that, for the next round, and to get scale, probably, we will divide our product...to have more standalone products and go to the middle and SMB markets.	Brazil

8	Video communication tool for companies	The main proxy for all the rest is to build a new category of product…KPIs that will probably be attached to a goal of…revenue probably around 100 billion ARR [and]…1 billion plus valuation.	Investing a lot in sales…so reaching out instead of depending on the inbound demand. Second thing is obviously expanding to different markets. We've been really present in…Europe. We'll continue to actually expand in Europe…And also the US, North America in general is actually super interesting, really mature market, a lot of competition that's fairly—we haven't really scratched the surface as a company yet…And so once we have done successfully those two regions that we see, we will probably expand to new ones or most likely APAC and so on.	France

Moat examples

Table A.2.2 shows examples of moat responses and scores from three companies. The Singapore-based company creating a mobile-based alternative credit scoring platform (third row) conceives a moat centered around its data from mobile devices to improve the predictive power of customers' models. Its external alignment is weak. The company does not articulate why other players would be unable to build their mobile data moat. While the company notes that it can coexist with other companies providing telco data because customers use both, it is unclear why other players would not have an incentive to produce mobile-derived data. These players could achieve economies of scope by providing both types of data to customers. Further, the mobile-derived data moat does not directly fit with the objective of being a "global company." Whether the mobile-derived data can ensure international adaptability or relevance is questionable.

In contrast to the Singapore-based company that focused its moat on the "process" of its data collection, the US-based company creating a recruiting platform (first row) pursued a moat focused on the ultimate value that the data would create. The company's data, which is "prohibitively expensive" for even large players to collect, can improve recruitment outcomes by "identifying" and "matching" people in a much better way. The response has a high external alignment because it shows how the company's data creates value and is difficult for others to acquire. Further, this hard-to-get data that creates value enables the company to expand its reach to customers in the recruiting space, aligned with its objective to be behind "many of the HR technologies and big people data systems." The HR-data nature of the moat targets the same HR customers that are core to its objective. This alignment contrasts with the mobile-data nature of the Singapore company's moat, which does not necessarily fit with the global nature of its customers, resulting in the US company achieving a higher internal alignment score.

The Finland-based company creating a product-testing technology (second row) perceived its moat to be a combination of the technical nature and user-friendliness of the product. It is unclear why these capabilities are a "winning combination." Specifically, the executive does not articulate the value the capabilities create for customers and why other competitors would not be able to take over the space, resulting in a weaker external alignment. The "plug and play" nature of the technology that can be applied to international audiences somewhat fits with the company's objective to be a "global player." However, it is unclear how it would create sufficient barriers to entry to allow it to achieve "revenue of 100 million euros." The response, therefore, achieves an average external alignment score.

Table A.2.2: Moat examples

Subscore	Descr.	Objective	Moat	HQ
21	Recruiting platform using AI/ML to provide talent to enterprises	In the next three to five years, [the company] will be the integrated data platform behind many of the HR technologies and big people data systems.	Data...The problem with talent data is it's huge. And those are prohibitively expensive for even a multi-billion-dollar organization to try and structure data that way. So we instead were...going to build one, which was a risk. And once we pulled that off, we really started to see that we could identify people in this much better way...We could also match people much better for organizations based on capability. And the people that we were putting there were so much more diverse.	US
12	Creator of technology to spur creation and testing of products and services	We want to become a global player in our arena...We aim to grow, especially in Europe...But our key goal now with the A round is really to establish a great position in Europe. And after that, the aim is to take the next step and enter the US market...Our five-year plan is to reach a revenue of 100 million euros.	We are a plug-and-play platform. That makes it possible that everyone can really run advanced consumer research and testing and thereby be fast and successful with their innovation process. So combining this super-easiness to use with expert approach and methods, I think that's the winning combination...And of course, then thinking a bit longer-term perspective, since we are focusing on a certain domain, we are also developing our AI capabilities...So, of course, the data combined with our algorithms [will be] a strong asset.	Finland
5	Mobile-based alternative credit-scoring platform to enable financial inclusion and access for underbanked consumers	My ambition [is] to bring [the] company [to the] global level.	[Our moat is] data that would increase the predictive power of [customers'] models...So there are companies who are developing scorecards based on telco data. We don't consider them as a direct competitor because we have clients who are using both, right? Scorecards developed based on device, mobile device data. That is [our company]. So we can coexist together.	Singapore

Organizational design examples

Table A.2.3 shows examples of organizational design responses and scores from different companies. The Brazil-based startup creating a platform to automate procurement of documents (third row) seeks to bring in technology, business, and senior-level talent. The response aligns weakly with the executive's assumptions (external alignment) because the company does not prioritize these different types of talent nor explain the value each will bring to the company. Further, the response does not mention why technology, sales, or senior executives would help the company better reach customers to increase revenue and ultimately get an IPO, as noted in its objective. As a result, the response also earns a relatively low score for fit with the objective (internal alignment).

The US-based startup creating a recruiting platform (first row)—with a similar B2B platform model and type of talent as the Brazil-based startup—pursued an aligned organizational design. The approach of moving from a tech team to an increasingly commercial team with API business experience aligns with the objective of wanting to be the "integrated data platform" underlying "many of the HR technologies and big people data systems." The technology team can create a data platform that initially brings value to the early adopter market that automatically provides feedback. Subsequently, the API-experienced customer success and sales side can bring that value to mainstream companies who are less forthcoming with their feedback. Customer success then becomes important to "anticipate the move of the customer." Therefore, the company can have the space to first develop a great technology with minimal outside pressure and only then make it known to the world (the "Trojan horse" approach). This close fit between customer needs in the market and the organizational design response results in a high external alignment score.

The Qatar-based startup creating a one-stop-shop (second row) focused its organizational design on getting intelligent, young, and experienced people from around the world. This talent mix of global experience and aptitude is compelling. Employees need to bring expertise to excel in their duties but also bring new ideas. Global sourcing is particularly valuable in a smaller country like Qatar, where such experienced and high-potential talent is difficult to find. Therefore, the combination of intelligence and expertise that the company brings to the team fits well with its assumptions about local labor market conditions, resulting in a high external alignment. However, the approach fits weakly with the objective of being "a regional app with multiple services." One might expect that the company would be focusing on the "local" nature of talent in different domains to fit its goal of building a "super regional app" in the Middle East. The discrepancy between the global nature of the organizational design and the local nature of the objective drags down the plan's internal alignment.

The Seychelles-based startup creating a crypto network (fourth row) pursued an organizational design plan that strongly aligns with its objective. Building a network of developers with crypto-specific skills and interests fits with the objective of creating "decentralized unchained governance" that is core to a crypto innovation. The company is pairing a decentralized team structure with a decentralized objective by choosing an open-source network-like organizational structure. The decentralized approach allows the company to get technical contributions relatively quickly and cheaply for its near-term product goals. However, the transactional nature of interaction with developers raises questions about how the company will retain talent and tacit knowledge over time to be "scalable." The discrepancy between the transactional approach and the scalable rationale results in a lower external alignment score.

Table A.2.3: Organizational design examples

Subscore	Desc.	Objective	Organizational Design	HQ
23	Recruiting platform using AI/ML to provide talent to enterprises	In the next three to five years, [the company] will be the integrated data platform behind many of the HR technologies and big people data systems.	In the beginning, tech and products were 80 percent of our team...Most startups start the opposite way. We don't ever market things...We appreciate much more of the Trojan horse approach...We are now going to be much more even keel on a 50-50 percentage. My focus is also...[an] incredibly strong customer success...to anticipate the move of the customer. And we are going to bring in some experts that come from API-first businesses.	US
17	One-stop-shop for online shopping, food, etc.	Our aim, to be a super-regional app with multiple services [in the Middle East].	I had to steal people from a similar business model. So in every specific area, I make sure I have 40 percent been there and done that. Now I have employees coming from Colombia to handle my customer experience department. I was making sure people have been there and done that. However, I bring in young people also...So how can I bring experienced people with know-how and, at the same time, bring the smartest people from the university together? So this is my strategy.	Qatar
16	Platform using AI to automate procurement of documents in Brazil for companies	For the next three to five years...we will achieve over 100 million reais in revenue...which means, probably, we could see some IPO in the next three or five years.	We are in three [areas]. It's the tech team, of course, it's the number one. And we face a challenge here...So we're thinking in the rest of the world, it's the same problem...The sales team we probably will spend more money on. And also, we need to hire more senior executives to our teams. So we develop our senior leadership.	Brazil

16	Crypto network for Web 3.0	Our goal is to…eliminate ourselves as a company in three years. And since we're in this process, we will try to stabilize the protocol, make it very concrete, resistant to the change of environment, both from technical and market condition perspectives. And once we're quite sure the protocol will evolve…the core team will transfer the right privilege to the community to realize what we call decentralized unchained governance.	Once someone [is] convinced by our idea and they believe we're doing the right thing and they have the required capability to help with this…crypto-network building things, they can join…And we will provide incentives in various ways…First, we tried to find some mechanism to incentivize hundreds or even thousands of people to join our network and make contributions relatively—it can be comparable with open source software development…It's much more important and scalable than recruitment.	Seychelles

Organizational culture examples

Table A.2.4 shows examples of organizational culture responses and scores. The India-based company (third row) creating a payments platform for offline businesses says that it does not directly define culture in terms of inputs or outputs. While the company initially asserts that "culture is what you live and breathe every day," which cannot be defined, it later does define culture as allowing "people to create stories." The two points are contradictory. This inconsistency and the vagueness of culture as "story building" reveals a weak external alignment. Further, this view of culture does not quite align with its user-oriented objective of "digitizing three million merchants." One would expect the culture to focus on users instead of employees to fit the objective. Thus, the response earned a relatively low internal alignment score.

In contrast to the poor fit with objectives in the India-based company's response, the US-based startup creating an interviewing platform for engineers (first row) had an organizational culture focused on its target. This culture, oriented toward engineer users, aligned with the company's objective to serve the human resource needs of engineers. The company's focus on satisfying engineers is consistent with the observation that it is "harder to get" this group of employee users relative to employer customers, and this user base is "what will make or break the company." Thus, this culture can help the company attract a scarce and critical asset, which enables it to satisfy not only users but also employer customers. By focusing on its core user base, which the executive believes will "make or break this company," the culture response receives a high external alignment score.

The Switzerland-based company creating a payment management application for banking transactions (second row) discussed a cultural approach that weakly aligns with its objective. Through guidance from an external consultant, the company targeted cultural values like problem-solving, trust, and simplicity. While these cultural values seem reasonable on their own, resulting in a relatively high external alignment, it is not clear whether they reinforce other aspects of the company. For example, the response does not discuss user orientation, which would more directly align with the objective of growing clients. This approach contrasts with the US company's objective-oriented approach of aspiring to satisfy engineer users and then working backward to cultural values oriented toward these users. As a result, the Switzerland-based company's culture response gets a lower internal alignment score.

Table A.2.4: Organizational culture examples

Subscore	Description	Objective	Organizational Culture	HQ
21	Interviewing platform for engineers	In five years, I'd want us to be the main way that software engineers find jobs and the main way that companies find software engineers.	Our one core value as a company is to put our engineering users first. So that means that whether they're paying our bills or whether companies are paying our bills, at the end of the day, everything is about our engineers because they are what will make or break this company... Engineers are generally harder to get. So we want to make sure that we make them happy. And that really bleeds over into everything we do. It means that we have to take a certain tone with our users when they write and ask for help...Flows have to be a certain way because we're trying to make a very particular audience happy.	US
14	Payment management application for banking transactions	Growth...is necessary for us to secure the place that we have—the better you're known, the more clients you will get. Until 2023, or maybe 2024, we aspire to something in the vicinity of 300,000 to 400,000 [clients]...This is what we consider to be the critical size when we don't have to force anymore. So when we can rely only on peer-to-peer recommendations or on organic growth. This would very clearly position us among the top three.	We're actually working on building that culture. So we are now at the point where we take in a professional consultant to essentially guide the company to a culture that remains open and problem-focused, but also to establish a number of common values because there's just a number of rotations and people coming and leaving and eventually, you need to get the tacit knowledge written down...I would say trusting, motivated, able to tackle problems face to face without making your opponent the problem...I think simplicity is also something that we have. I can keep things simple to try to make it as simple as possible. Trust...Being passionate, have positive energy, create something that we're all proud of, and challenge to improve...that is something that gives agility.	Switzerland

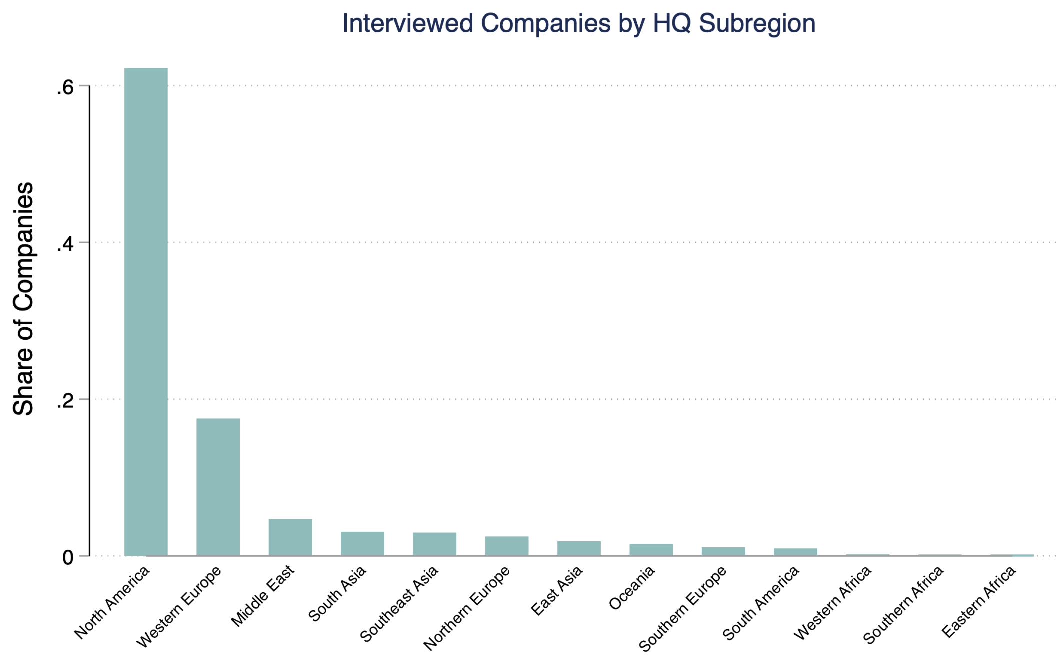
9	Payments platform for offline businesses	We want to digitize 3 million merchants in this country through digital commerce.	There's no definition of culture in the company. Culture is what you live and breathe every day. So we are not defining the culture; it is what is lived and breathed. So what are you? You live for the people. We believe people should create stories in this organization, and they should have a story to tell outside. That's the culture that we want to build. Everybody should go and tell their stories in their life...That's the culture that we want to build.	India

A.3 Sampling checks

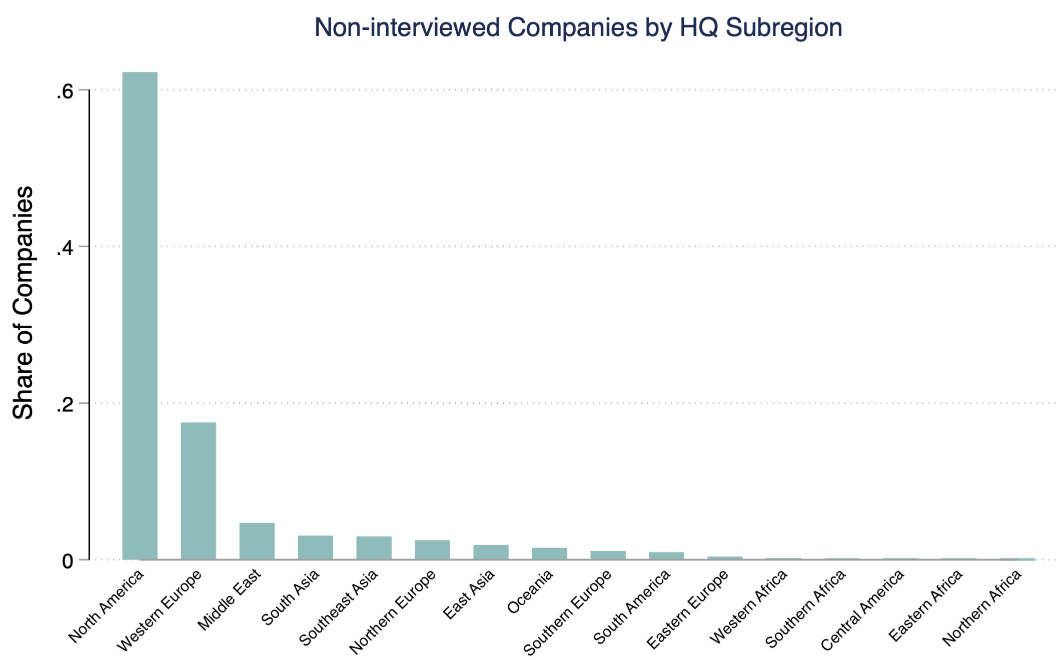
This study interviewed more than 12 percent of the software companies (outside of China) that received Series A funding ($5–20 million) from January 2019–November 2021. There may be a concern that these 253 companies are different than those not interviewed in this sample, biasing the results. For example, perhaps only the worst companies agreed to an interview because they wanted to get advice on how to improve their operations. Alternatively, we could imagine that only the best companies agreed to the interview so as to market their success. To help account for this potential bias, it is valuable to measure differences between the interviewed companies and others in the sampling frame. While Table 1.1 shows that regional identifiers, employee count, initial financing amounts, and valuation do not predict whether a startup was interviewed in the sample, there might still be a concern that the overall distributions vary.

To address this concern, histograms reveal how the overall distributions vary among interviewed and non-interviewed companies for several key variables, where we might be concerned about selection. Figure A.3.1 shows that the regional distribution of the interviewed sample of companies looks similar to that of the non-interviewed companies. Figure A.3.2 shows that the distribution of the number of employees at the time of the interview are similar for interviewed and non-interviewed companies. Figure A.3.3 shows that the distributions of initial financing size among interviewed and non-interviewed companies in the sample are also similar.

Figure A.3.1: Regional distribution of the non-interviewed sample is similar to that of the interviewed sample.

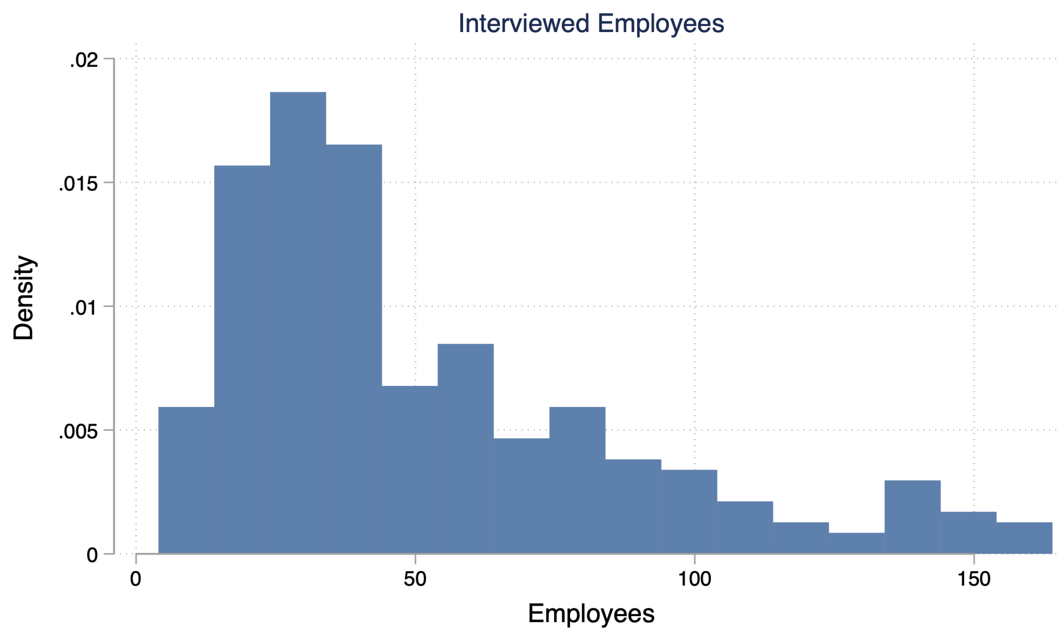


The graph shows 253 companies that have been interviewed.

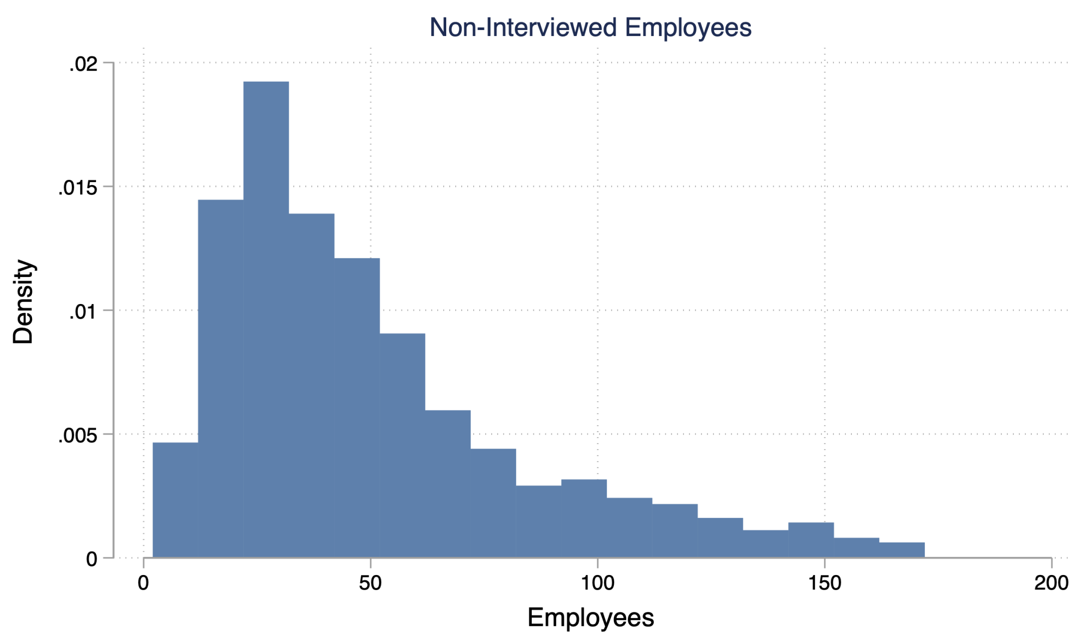


The graph shows 1,576 software companies not interviewed.

Figure A.3.2: Employee distribution of the non-interviewed sample is similar to that of the interviewed sample.

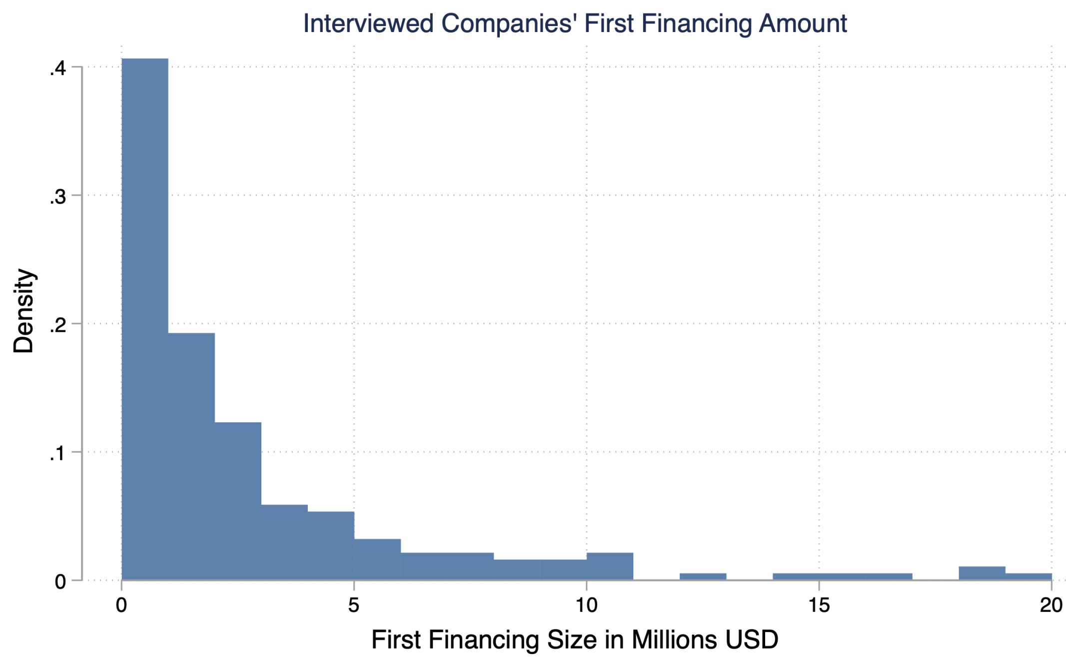


The graph shows companies with no more than 170 employees,
comprising 95% of companies at the time of interview recruitment.

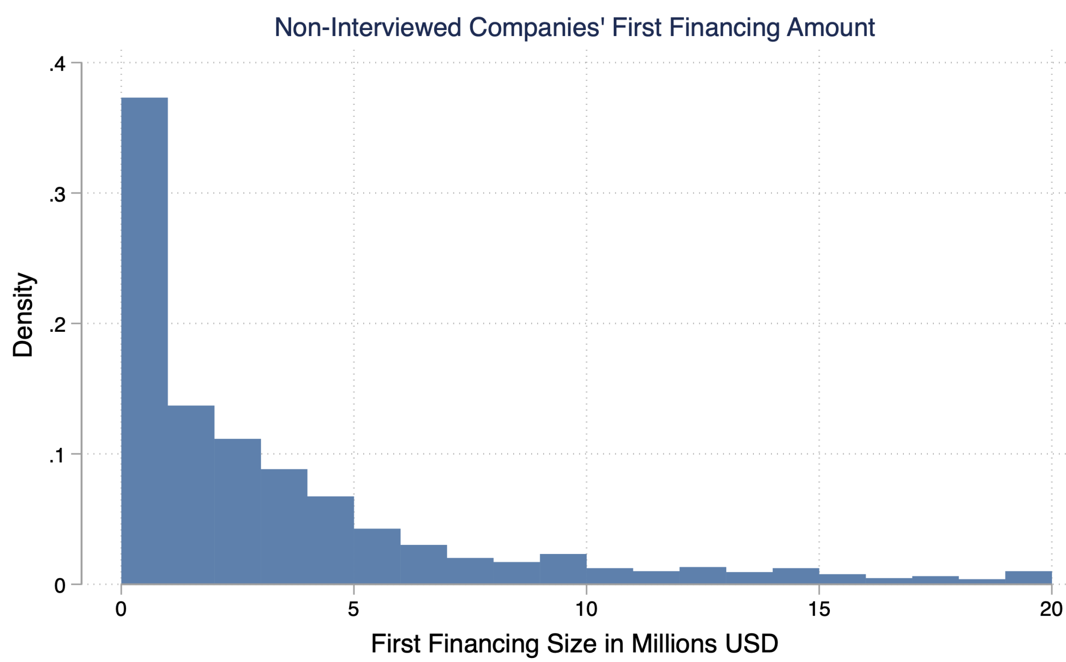


The graph shows companies with no more than 170 employees,
comprising 95% of companies at the time of interview recruitment.

Figure A.3.3: First-financing-size (in USD) distribution of the non-interviewed sample is similar to that of the interviewed sample.



The graph shows financing amounts less than $20 million, comprising over 99% of firms.



The graph shows financing amounts less than $20 million, comprising over 99% of firms.

A.4 Machine learning approach to strategy score

One concern with the strategy measure is that it reflects noisy human coding. A neural-network-based NLP technique makes it possible to create an alternative strategy measure to test the robustness of the human coding. Specifically, the SBERT model measures the similarity of text, taking into account the semantic meaning of sentences by assessing words in the context of their sentences (Devlin et al., 2018; Reimers and Gurevych, 2019). This model allows assessing the similarity of each of the four choices to the objective and to one another as a proxy of fit with the objective and other choices (internal alignment). The model then measures the similarity of the text within each response to proxy fit with assumptions (external alignment). Like the algorithm using the human-coded measures, multiplying the NLP-based similarity scores for each element, summing them together, and normalizing them creates a composite strategy score. This score excludes missing values. Table A.4.1 shows that the final composite score correlates with the primary strategy measure used in the paper (Column 1). The subscores using NLP for each element are also positively associated with the corresponding human-coded subscores (Columns 2–5). The high correlations between the human-coded and computer-generated scores help validate the strategy scores.

Table A.4.1: Human and NLP calculations of strategy are positively correlated.

	(1) Strategy	(2) Market Subscore	(3) Moat Subscore	(4) Org. Design Subscore	(5) Culture Subscore
Strategy NLP	0.133*				
	(0.057)				
Market Subscore NLP		0.125*			
		(0.061)			
Moat Subscore NLP			0.083		
			(0.059)		
Org Subscore NLP				0.124*	
				(0.054)	
Culture Subscore NLP					0.117*
					(0.058)
_cons	0.026	-0.022	0.019	-0.001	0.022
	(0.059)	(0.059)	(0.059)	(0.057)	(0.060)
N	237	253	249	253	238

Robust standard errors (in parentheses) are clustered at the company level.
+ p < 0.10, * p < 0.05, ** p< 0.01, *** p < 0.001

A.5 Alternative measures of strategy

The strategy score in the main tables uses Equation (1), weighting the fit with assumptions of each element of strategy—including market scope, moat, organizational design, and organizational culture—(external alignment) by its fit with the objective and the other responses (internal alignment). There might be a concern that this particular algorithm may produce spurious results.

To address this issue, Table A.5.1 shows several alternative measures of the composite strategy score using a specification similar to that in Table 1.4. Column 1 shows the strategy alternative score created when excluding missing values, which reduces the sample size from 304 evaluations to 269. Column 2 shows an alternative specification that uses only the alignment with the objective rather than with other responses for the internal alignment variable. Column 3 shows a version that is similar to the main score, but double-weights the alignment with the objective in case this fit is more important than the alignment with other choices. Column 4 shows a version that reflects a simple average across the internal and external alignment scores across each of the four elements of strategy (market scope, moat, organizational design, and organizational culture). The results are generally robust across these alternative approaches: US firms have a 0.3 standard deviation higher strategy score.

Table A.5.1: Alternative measures of strategy

	(1) Alt. Score 1	(2) Alt. Score 2	(3) Alt. Score 3	(4) Alt. Score 4
US HQ	0.295^{*}	0.250^{*}	0.271^{*}	0.257^{*}
	(0.124)	(0.117)	(0.126)	(0.126)
_cons	0.464	0.282	0.317	0.390
	(0.518)	(0.499)	(0.552)	(0.552)
N	269	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Filled-In FE	No	No	Yes	Yes
Readability	Yes	Yes	Yes	Yes

Alt. score 1 shows the strategy score without filling in missing values with averages.
Alt. score 2 shows the score weighted by the alignment with the objective.
Alt. score 3 shows the main one but double-weights alignment with the objective.
Alt. score 4 shows a simple average.
All scores have over a 0.93 correlation with one another.
Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.6 Robustness checks using LASSO specifications

Table 1.4 shows that US startups tend to have higher strategy scores than do other companies. One concern with this finding is that it is picking up differences in the content of strategy across companies, which confounds the study's ability to measure differences in whether companies have a strategy. For example, US companies are more likely to have data as a moat: perhaps this form of moat is inherently more aligned with an objective. We might then see that the score reflecting the moat's alignment with the objective is higher for US companies because of the underlying substance—the moat itself.

Additional analyses allow controlling for these confounding content variables. Table 1.4 uses a LASSO model that adds controls for the content of the strategy—comprising all variables shown in the regression tables under objective, market scope, moat, organizational design, and organizational culture—and uses shrinkage to pick the most predictive controls. This model is valuable for "wide" datasets with many coefficients relative to observations.

Table A.6.1 shows results when applying the LASSO model to a specification similar to Equation 2, measuring strategy differences between US and other firms. The results generally hold. US firms tend to have higher strategy scores (Column 1), driven by their moat (Column 3) and organizational design (Column 4) subscores. These differences are actually larger for the latter subscores when controlling for content. The result suggests that the content of the strategy obscures the massive variance in alignment between these two choices. This model suggests that, irrespective of the content of the strategy, US firms tend to have higher strategy scores.

Table A.6.1: The strategy gap between US and non-US firms holds when using a double-selection LASSO model controlling for the strategy's content.

	(1) Strategy Composite	(2) Market Scope Subscore	(3) Moat Subscore	(4) Org. Design Subscore	(5) Org. Culture Subscore
US HQ	0.224*	0.579	0.688^{+}	0.636^{+}	0.293
	(0.108)	(0.353)	(0.411)	(0.334)	(0.384)
N	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Filled-In FE	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes
Strategy Content	Yes	Yes	Yes	Yes	Yes

Double selection LASSO model used.
Standard errors in parentheses.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6.2 shows that strategy predicts performance more for non-US firms, even when controlling for the strategy's content. The coefficient on the interaction term between the strategy score and whether the startup is headquartered in the US is negative when using the performance index (Column 1), logged valuation (Column 2), logged valuation per employee (Column 3), whether the company exited (Column 4), and whether it exited or received future funding (Column 5) as the dependent variables. These results are consistent with Table 1.3.

Table A.6.2: Strategy predicts performance more for non-US firms even when controlling for the strategy's content using a double-selection LASSO model.

	(1) Performance Index	(2) Log Val.	(3) Log Val. Per Employee	(4) Exited	(5) Exited/Raised Future Funding
US HQ	0.303^{+}	0.579^{*}	0.272^{*}	-0.028	0.151^{*}
	(0.155)	(0.226)	(0.107)	(0.045)	(0.073)
Strategy	0.361^{***}	0.454^{*}	0.119^{+}	0.067^{*}	0.016
	(0.105)	(0.177)	(0.069)	(0.029)	(0.048)
US HQ x Strategy	-0.479^{***}	-0.596^{**}	-0.224^{**}	-0.071^{*}	-0.110^{+}
	(0.130)	(0.202)	(0.085)	(0.036)	(0.059)
N	230	184	183	230	230
Evaluator FE	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Filled-In FE	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes
Log GDP Capita	Yes	Yes	Yes	Yes	Yes
Log First Financing	Yes	Yes	Yes	Yes	Yes
Strategy Content	Yes	Yes	Yes	Yes	Yes

Double selection LASSO model used.

Standard errors in parentheses.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.7 Results robust to a country index of the ease of recovering from mistakes

Table 1.3 shows that strategy scores predict the performance of non-US but not of US firms. Qualitative evidence suggests that strategy is more important in non-US contexts to avoid more costly mistakes. Financial, talent, customer, and cultural constraints drive up the price of mistakes.

While the cost of mistakes might particularly vary between US and non-US contexts, there certainly may be more granular sources of variance among countries. Constructing an "ease-of-recovery" continuous index across countries allows for assessing this granularity.

Data from the World Economic Forum's Competitiveness Index serve as the basis for this measure. This database includes variables estimating the financial, talent, customer, and cultural constraints that interview data suggest make mistakes more costly. Specifically, the measure consists of the standardized average of the following standardized variables:

- **Venture capital availability**. Companies with less access to local venture capital have a harder time getting additional funding to replace bad hires or markets.
- **Supply of talent**. Companies with scarce talent nearby have a harder time replacing bad hires.
- **Wage flexibility**. When companies have little ability to negotiate wages with employees, they are likely paying a higher cost to replace bad hires. This variable also reflects broader institutional rigidity in the labor market, making recruitment costly.
- **Domestic market size**. When a company is in a small domestic market, and that market fails, it is far more costly to invest in another international market than it is to invest in another subregion within a sizeable domestic market. So an Australian startup that makes a bad investment in the UK would pay a high cost to then invest in Germany (or even in another part of the UK). It would be less costly for, say, a Texas-based startup that makes a bad investment in Kansas to recover by reallocating investments to Florida.
- **Attitudes toward entrepreneurial risk**. This variable proxies a cultural orientation embracing failure. Startups in contexts with less of this orientation face a higher penalty for any mistake.

Standardizing the average of these standardized variables on a 0–1 scale yields a continuous index. Consistent with the previous results, the US has the highest index value (1); mistakes are the least costly in the US. When creating a binary version of this variable based on the median value, the US stands on its own as the above-median value, consistent with the main US versus non-US variable in this study. Figure A.7.1 shows the index values across countries.

Figure A.7.1: The ease-of-recovery index by country.

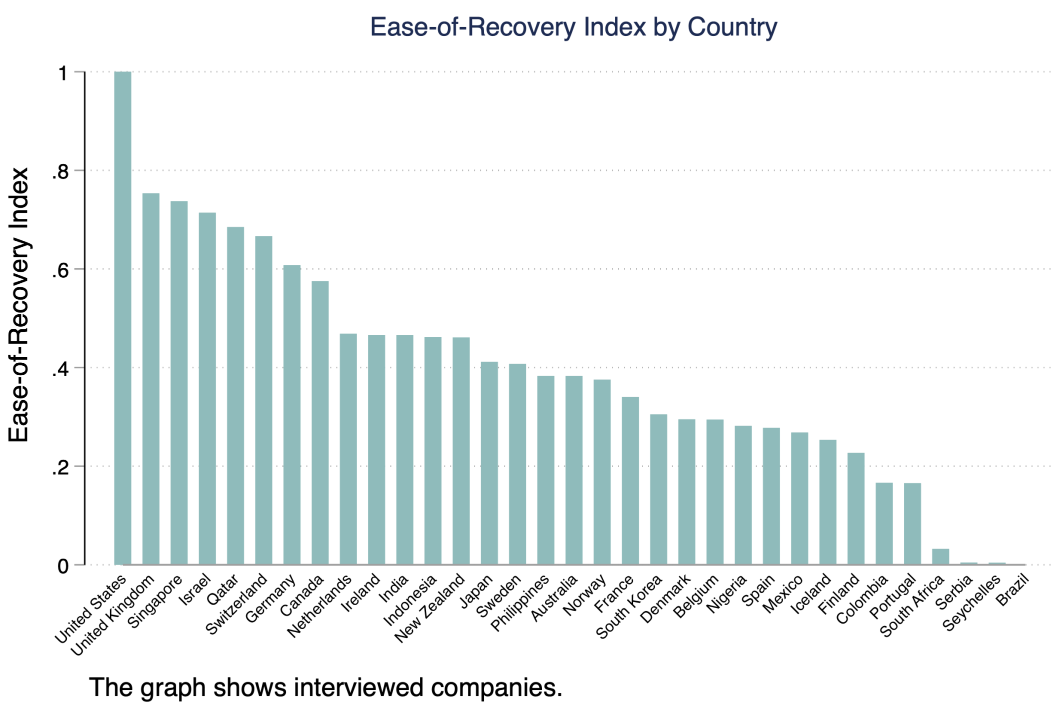


Table A.7.1 applies the same specification as Table 1.3, but replaces the binary US variable with the continuous "ease-of-recovery" index: the results are similar. Strategy scores predict performance—in terms of the aggregate performance index (Column 1), logged valuation (Column 2), logged valuation per employee (Column 3), and whether a company exited (Column 4)—more for firms in countries where it is harder to recover from mistakes.

Furthermore, this same index also predicts strategy scores. Table A.7.2 shows that firms in countries where it is more costly to make mistakes have lower strategy scores by 0.4 standard deviation (Column 1). This result is consistent with the qualitative findings in Section 4.3.2: either directly experiencing mistakes or learning from those of peers, advisors, and investors offers knowledge to entrepreneurs that can inform their strategy. But entrepreneurs in contexts where mistakes are more costly are less likely to access such knowledge and therefore develop a strategy.

Additional robustness checks suggest that the index portrays the cost of mistakes rather than a particular underlying factor or a different construct altogether. The results are weaker when assessing each of the underlying components of the index independently—for example, venture capital availability or the supply of talent—suggesting that the cost of mistakes is not driven by a single resource or cultural factor, but rather by an aggregate of them. Other factors that one might be concerned would confound the ease-of-recovery measure—such as GDP per capita—in fact, do not predict variance in the strategy score, nor its relationship with performance.

Table A.7.1: Strategy predicts performance more where it is harder to recover from mistakes.

	(1) Performance Index	(2) Log Val.	(3) Log Val. Per Employee	(4) Exited	(5) Exited/Raised Future Funding
Ease-of-Recovery Index	0.740*	1.613**	0.856***	-0.015	0.340*
	(0.307)	(0.510)	(0.197)	(0.075)	(0.145)
Strategy	0.596*	1.050*	0.418**	0.132*	0.035
	(0.253)	(0.418)	(0.149)	(0.060)	(0.112)
Ease-of-Recovery Index x Strategy	-0.698*	-1.160*	-0.532**	-0.145*	-0.139
	(0.295)	(0.460)	(0.170)	(0.071)	(0.128)
_cons	4.074**	5.407**	0.208	-0.217	1.196*
	(1.275)	(1.780)	(0.724)	(0.172)	(0.584)
N	230	184	183	230	230
Evaluator FE	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Log GDP Capita	Yes	Yes	Yes	Yes	Yes
Log First Financing	No	Yes	Yes	Yes	Yes
Filled-In FE	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.7.2: Strategy scores are higher where it is harder to recover from mistakes.

	(1) Strategy Composite Score	(2) Market Internal Align.-Obj.	(3) Market Internal Align.-Other	(4) Market External Align.	(5) Moat Internal Align.-Obj.	(6) Moat Internal Align.-Other	(7) Moat External Align.	(8) Org. Internal Align.-Obj.	(9) Org. Internal Align.-Other	(10) Org. Extenral Align.	(11) Culture Internal Align.-Obj.	(12) Culture Internal Align.-Other	(13) Culture External Align.
Ease-of-Recovery Index	0.350^{+}	-0.083	0.009	0.173	0.260	0.132	0.338^{+}	0.140	-0.005	0.323^{+}	0.087	0.141	0.437^{*}
	(0.209)	(0.189)	(0.211)	(0.195)	(0.187)	(0.212)	(0.197)	(0.186)	(0.241)	(0.189)	(0.194)	(0.216)	(0.203)
_cons	0.174	-0.626	0.896	-0.308	0.859^{+}	0.898	0.364	0.008	-0.107	-0.528	0.878	0.120	-0.152
	(0.576)	(0.498)	(0.696)	(0.564)	(0.491)	(0.579)	(0.527)	(0.556)	(0.683)	(0.629)	(0.578)	(0.594)	(0.676)
N	304	304	304	304	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Filled-In	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.8 Interviews suggest that mistakes are more costly in non-US contexts.

Table 1.3 shows that strategy matters more for firms outside the US. The paper posits that this might be because strategy helps companies avoid mistakes, which can be particularly costly in non-US contexts. Section 4.3.1 shows qualitative evidence of how non-US firms seek to avoid mistakes due to resource constraints and a cultural aversion to failure. In contrast, US firms perceive mistakes with a more open mind; many view them as learning opportunities.

Table A.8.1 shows these different orientations to mistakes in the broader data. Among the 108 firms that discuss mistakes in their interviews, US firms are more likely to view them as learning opportunities instead of solely as outcomes to avoid. As an example of a response showing a learning orientation toward failure, one US executive said: "I intentionally hired very smart people. So I let them be smart and execute. And I also let them make mistakes because that's how you learn." Alternatively, a UK executive reveals a non-learning orientation—an aversion to mistakes:

> [We] went through...mistakes like that. And you think, oh, it's easy. We'll fire them. It's not easy. It always creates collateral damage. People can get toxic. Other people get disoriented...Firing the people that don't seem to fit can also be quite disruptive to your organization...I'm a big fan of, like—I really do not like first-principle thinking because I hate making mistakes that other people know the answers to.

The UK company sought to avoid mistakes even if it came at the expense of developing the "first-principle thinking" that might improve subsequent actions. This view contrasted from that of the US company that embraced mistakes in order to learn from them.

Table A.8.1: Perception of mistakes and failure as learning opportunities by whether US HQ

	(1) Mistakes for Learning
US HQ	0.427***
	(0.095)
_cons	0.317***
	(0.066)
N	108
Year Founded FE	Yes
Industry FE	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.9 Strategy can help companies avoid mistakes

By aligning key choices within the company, strategy may help startups avoid mistakes as they scale. Demonstrating this relationship empirically, however, is difficult because data on mistakes are not readily available. This is partly because companies often do not want to disclose their mistakes. The interviews provide retrospective evidence that planning ahead would have helped companies avoid mistakes in their scaling process. For example, one South Korean executive mentioned how it could have avoided a clash in their hiring and culture if they had been "smarter earlier":

> One of the main mistakes that we actually made...right after Series A is that we went on a hiring spree. And the team went from 20 to actually 70 people in less than six months. And that rapid unexpectedly damaged our culture because a lot of new people with different ideas and different working habits came in too fast...We were actually hiring based on just functions—like, hey—we need a finance person...We need someone in compliance... And what was more important was that...someone might not necessarily be the best compliance officer, but...[is] flexible with the system or lack of system that we have right now. And so there was a lot of back progress that we had to go through. And that's just time and money being spent not on the company, but rather fixing mistakes that, if we were smarter earlier, then we probably wouldn't have made in the first place.

To help understand this relationship between strategy and mistakes, I assess how the strategy score correlates with the probability of making a future mistake. As mentioned above, getting data on whether companies committed mistakes is difficult. One proxy of mistakes that is possible to get from existing data (PitchBook) is whether companies closed an office or subsidiary.

Table A.9.1 provides suggestive evidence of a negative relationship between strategy and mistakes. Specifically, it shows that a higher strategy score is associated with a lower probability of closing a regional office or subsidiary—as a proxy of mistakes—since the interview. The negative coefficients on the market (Column 2) and organizational design (Column 4) subscores show this relationship. Together, these results triangulate that strategy can help companies avoid mistakes in scaling.

Table A.9.1: Strategy and office closures

	(1) Closed Office	(2) Closed Office	(3) Closed Office	(4) Closed Office	(5) Closed Office
Strategy	-0.032				
	(0.021)				
Market Subscore		-0.013^{*}			
		(0.006)			
Moat Subscore			0.000		
			(0.005)		
Org. Design Subscore				-0.015^{*}	
				(0.007)	
Culture Subscore					-0.004
					(0.006)
_cons	0.092^{***}	0.285^{**}	0.090	0.318^{**}	0.148^{+}
	(0.019)	(0.097)	(0.078)	(0.108)	(0.086)
N	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Filled-In FE	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.10 Content of strategy in US versus non-US firms

This paper focuses on differences in whether companies have a strategy irrespective of its content. Yet, there are interesting differences to consider in the content of the strategy. This section documents these differences in the content of strategy between US and non-US executives across the core interview questions in the protocol. These tables show a consistent pattern: non-US firms tend to focus more on geography in their objective, moat, organizational design, actions, and funding uses. Surprisingly, many other aspects of strategy, such as the financial nature of objectives, do not systematically vary across US and non-US firms.

Table A.10.1 shows that US and non-US companies tend to have similar financial objectives, such as becoming a unicorn, increasing revenue, getting acquired, and going public over the next 3–5 years. What varies is the geographic nature of their objective and whom they are satisfying: US companies are less likely to focus on being global (Column 5) and more likely to focus on satisfying customers (Column 16).

Table A.10.1: US versus non-US objectives

	(1) Increase Revenue	(2) Get Acquired	(3) Go Public	(4) Grow Users	(5) Become Global	(6) Become Market Leader	(7) Create Market	(8) Become Unicorn (Billion+ USD Val.)	(9) Become Profitable	(10) Raise More Funding	(11) Social Impact Goal	(12) Improve Product	(13) Grow Into Other Industries	(14) Long-Term View	(15) Satisfy Employees	(16) Satisfy Customers	(17) Satisfy Investors
US HQ	0.030	0.039	0.041	-0.063	-0.161**	-0.023	-0.043	-0.016	0.020	0.047	-0.020	0.006	-0.023	-0.005	0.007	0.099**	-0.010
	(0.061)	(0.031)	(0.046)	(0.061)	(0.050)	(0.060)	(0.037)	(0.028)	(0.025)	(0.037)	(0.036)	(0.050)	(0.039)	(0.026)	(0.019)	(0.037)	(0.016)
_cons	0.482	0.123	0.129	0.063	0.029	0.826**	0.029	-0.142	0.080	0.079	0.087	-0.042	0.132	0.253	-0.025	-0.161	-0.062
	(0.341)	(0.165)	(0.256)	(0.390)	(0.261)	(0.317)	(0.209)	(0.133)	(0.145)	(0.320)	(0.181)	(0.268)	(0.229)	(0.169)	(0.080)	(0.171)	(0.044)
N	304	304	304	304	304	304	304	304	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression Table A.10.2 shows how US and non-US companies compare in their approaches to expanding market scope. US firms tend to focus less on geography as a dimension of expanding their markets (Column 8). Instead, they focus on industry as the main dimension of expansion (Column 9). These results are consistent with the overall trend of US companies focusing less on geography in their strategy.

Table A.10.2: US versus non-US market scope approaches

	(1) Expand by Geography	(2) Expand by Industry	(3) Expand by Deepening in Same Market	(4) Expand From Large to Small Customers	(5) Expand from Small to Large Customers	(6) Product-Led Growth	(7) Expand by Acquisitions	(8) Geo is Main Dimension	(9) Industry is Main Dimension	(10) Customer Size is Main Dimension
US HQ	-0.240***	0.157**	0.044	0.072	0.055	0.069	-0.045	-0.243***	0.186***	0.045
	(0.055)	(0.050)	(0.059)	(0.050)	(0.057)	(0.046)	(0.030)	(0.058)	(0.051)	(0.049)
_cons	0.395	0.776**	0.400	-0.404+	-0.255	0.189	0.296*	0.469	0.741**	-0.134
	(0.313)	(0.266)	(0.324)	(0.241)	(0.272)	(0.195)	(0.141)	(0.314)	(0.284)	(0.234)
N	304	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression Table A.10.3 shows how US and non-US companies compare in their approaches to moats against competitors. There is not much difference here, apart from US companies focusing on data as a moat (Column 12).

Table A.10.3: US versus non-US moat approaches

	(1) Technology	(2) Past Learnings	(3) Execution	(4) Internal Organization	(5) Business Model	(6) Switching Costs	(7) Partnerships	(8) Brand	(9) Timing	(10) Local Knowledge	(11) Cost	(12) Data
US HQ	0.039	0.042	0.008	-0.021	0.034	0.061	0.010	0.016	-0.009	-0.034	-0.004	0.155^{**}
	(0.062)	(0.051)	(0.045)	(0.032)	(0.056)	(0.052)	(0.044)	(0.034)	(0.046)	(0.023)	(0.025)	(0.052)
_cons	0.810^{*}	0.123	-0.021	-0.233	0.535^{*}	-0.167	0.203	0.117	-0.334	0.112	0.032	0.362
	(0.316)	(0.305)	(0.221)	(0.146)	(0.267)	(0.250)	(0.175)	(0.199)	(0.287)	(0.182)	(0.087)	(0.291)
N	304	304	304	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression Table A.10.4 shows how US and non-US companies compare in their approach to organizational design. Again, there are few differences here. US companies tend to be less focused on hiring talent for global operations (Column 5), consistent with the overall trend that US companies pursue less geographically oriented strategies. US companies are also more likely to delegate authority from the CEO (Column 8).

Table A.10.4: US versus non-US organizational design approaches

	(1) Hire Technical	(2) Hire Commercial	(3) Hire Overhead	(4) Hire Leadership	(5) Hire by Geo.	(6) Department Structure	(7) Cross-Functional Structure	(8) Delegation from CEO	(9) Outsource Technical	(10) Outsource Commercial	(11) Outsource Overhead
US HQ	0.045	0.049	0.011	0.035	-0.268^{***}	0.047	-0.058	0.065^{+}	-0.032	-0.022	0.019
	(0.050)	(0.046)	(0.062)	(0.062)	(0.061)	(0.047)	(0.054)	(0.039)	(0.034)	(0.020)	(0.020)
_cons	1.182^{***}	0.844^{***}	0.458	0.517	0.593^{+}	0.899^{***}	0.508^{+}	-0.163	0.049	-0.281^{*}	-0.154
	(0.200)	(0.200)	(0.358)	(0.347)	(0.312)	(0.261)	(0.276)	(0.168)	(0.150)	(0.118)	(0.138)
N	304	304	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression Table A.10.5 shows how US and non-US companies compare in their approach to organizational culture. There is not much difference here. US companies tend to be less focused on social impact as a cultural value (Column 9). This result may relate to the geographically focused nature of non-US strategies. For example, Argentinian strategies that focus on, say, financial inclusion across the region inherently foster more of a social impact culture than US strategies that do not have such an economic development angle.

Table A.10.5: US versus non-US organizational culture approaches

	(1) Diversity	(2) Scrappy	(3) Experimentation	(4) Customer-Oriented	(5) Employee-Oriented	(6) Autonomy	(7) Trust	(8) Transparency	(9) Social Impact	(10) Top-Down	(11) Bottom-Up
US HQ	-0.006	0.077	-0.012	0.087	0.038	-0.084	-0.010	-0.034	-0.095^{+}	0.084	-0.071
	(0.050)	(0.057)	(0.055)	(0.057)	(0.044)	(0.059)	(0.044)	(0.052)	(0.054)	(0.051)	(0.048)
_cons	0.392	0.354	0.280	0.715^{*}	-0.007	0.291	0.004	0.021	0.687^{*}	0.629^{*}	-0.002
	(0.250)	(0.345)	(0.288)	(0.278)	(0.235)	(0.325)	(0.179)	(0.276)	(0.296)	(0.268)	(0.276)
N	304	304	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
$+ p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression Table A.10.6 shows how US and non-US companies vary in what they prioritize as their next three actions to reach their objective over the next 3–5 years. US companies are less likely to list expanding globally as an action item (Column 4). This result is consistent with US firms being less focused on geography in their strategy across the board.

Table A.10.6: US versus non-US next three actions

	(1) Have Three Actions	(2) Hiring Talent	(3) Retaining Talent	(4) Expand Globally	(5) Enhance Product	(6) Grow Sales	(7) Ensure Alignment	(8) Hire Leadership	(9) Listen to Market
US HQ	-0.047	0.042	-0.044	-0.103^{*}	-0.003	0.063	-0.044	0.056	-0.009
	(0.044)	(0.066)	(0.040)	(0.041)	(0.062)	(0.060)	(0.042)	(0.035)	(0.047)
_cons	0.920^{***}	0.605	0.143	0.184	0.910^{**}	0.150	0.247	0.140	0.435^{+}
	(0.240)	(0.372)	(0.227)	(0.235)	(0.349)	(0.301)	(0.216)	(0.205)	(0.251)
N	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
$+ p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression Table A.10.7 shows how US and non-US companies compare in how they allocate Series A funding. US companies are less likely to hire talent for global operations (Column 5), consistent with focusing less on geography in their overall strategy. US companies are more focused on hiring commercial talent (Column 3). Consistent with having organizational designs that are more aligned with an objective, commercial talent is essential to grow markets and increase revenue. US companies also are more likely to use their funding to hire technical talent (Column 2).

Table A.10.7: US versus non-US use of Series A funding

	(1) Talent	(2) Technical Talent	(3) Commercial Talent	(4) Overhead Talent	(5) Global Talent	(6) Marketing	(7) Software	(8) Saving for Runway	(9) Intellectual Property	(10) Experimentation	(11) Future Funding	(12) Enhance Product	(13) Physical Infra. Real Estate	(14) Regulatory
US HQ	0.039	0.142*	0.178**	-0.006	-0.108**	0.027	0.002	0.016	-0.013	-0.004	0.024	-0.014	0.011	-0.009
	(0.026)	(0.060)	(0.062)	(0.040)	(0.037)	(0.057)	(0.027)	(0.049)	(0.021)	(0.037)	(0.044)	(0.056)	(0.022)	(0.020)
_cons	0.353**	0.694*	0.518	-0.129	0.177	0.107	0.051	0.047	-0.048	0.210	0.297	0.791*	-0.026	0.066
	(0.135)	(0.343)	(0.317)	(0.209)	(0.184)	(0.283)	(0.168)	(0.244)	(0.101)	(0.185)	(0.201)	(0.312)	(0.077)	(0.071)
N	304	304	304	304	304	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression Table A.10.8 shows how US and non-US companies compare in their perception of their biggest uncertainties over the next 3–5 years. US companies are more likely to list execution (Column 6) and less likely to note expanding markets (Column 7) or economic factors (Column 1). This difference may be because the US external environment is more stable than that of other countries, whether it be because of institutions, funding availability, or quality of talent. As a result, US companies see more of their unknowns in their internal capabilities.

Table A.10.8: US versus non-US biggest uncertainty

	(1) Economy	(2) Political	(3) Market Readiness	(4) Competition	(5) Cyber Threats	(6) Execution	(7) Expand Market	(8) Hire Talent	(9) Retain Talent	(10) Maintain Alignment	(11) Predict Perform.
US HQ	-0.089^{+}	-0.028	0.014	0.021	0.034	0.123*	-0.080*	0.066	-0.013	0.008	-0.002
	(0.045)	(0.046)	(0.045)	(0.056)	(0.026)	(0.056)	(0.034)	(0.054)	(0.033)	(0.034)	(0.026)
_cons	-0.139	0.512	0.548*	-0.070	-0.220	0.401	0.266	-0.033	0.349^{+}	0.080	0.095
	(0.199)	(0.321)	(0.224)	(0.291)	(0.163)	(0.309)	(0.184)	(0.285)	(0.183)	(0.225)	(0.102)
N	304	304	304	304	304	304	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.11 US strategies are less geographically oriented.

Table A.11.1 shows that US startups tend to pursue less-geographically oriented strategies. Additional firm data triangulate this difference. Column 1 of Table A.11.2 shows that US firms are less likely to mention geographic expansion in their pitch decks (for one-third of firms that shared a Series A pitch deck). They are also less likely to include geographically oriented roles in their organizational charts (for one-fourth of firms that shared their organizational charts), though without statistical significance at the five-percent level (Column 2). Figure A.11.1 shows the log-odds words mentioned in the market scope responses across HQ regions. The responses of executives from outside of North America mention geography; for example, startups from the Middle East and Northern Europe mention the "US market" or "North America" (to which they often look to expand). Startups from Western Europe uniquely mention "geographic expansion" and "global business" in their responses. The consistency between what executives state in their interviews and what is in their internal and external documents reduces concerns that the interviews contain inaccurate information or simply "cheap talk."

Table A.11.1: US firms are less likely to focus on geography in their strategy.

	(1) Become Global	(2) Geo-Based Market Scope	(3) Geo-Based Org	(4) Global Expansion Next Action	(5) Fund Talent for Global Ops
US HQ	-0.161**	-0.240***	-0.265***	-0.103*	-0.120**
	(0.051)	(0.056)	(0.062)	(0.041)	(0.039)
_cons	0.038	0.384	0.604+	0.207	0.195
	(0.262)	(0.315)	(0.312)	(0.236)	(0.187)
N	304	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Filled-In FE	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ p < 0.10, * p < 0.05, ** p< 0.01, *** p < 0.001

Table A.11.2: Global strategies in pitch decks

	(1) Pitch Deck - Geographic	(2) Org Chart - Geo.
US HQ	-0.262*	-0.053
	(0.113)	(0.114)
_cons	0.688***	0.237**
	(0.078)	(0.082)
N	77	67
Year Founded FE	Yes	Yes
Industry FE	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ p < 0.10, * p < 0.05, ** p< 0.01, *** p < 0.001

Figure A.11.1: North American interviewed companies are more likely to mention geography in their market scope responses.

A.12 Organizational approaches are localized.

Table A.10.4 showed that the content of US and non-US organizational designs varied. Specifically, non-US firms were more likely to pursue geographically driven organizational designs and less likely to delegate authority from the CEO. These results suggest organizational approaches may be localized to particular geographies. Interview data and outside sources indeed suggest that organizational designs are specific to particular geographies, resulting in substantial variance across geographies.

Executives in interviews noted how it was harder to apply organizational design approaches from other parts of the world to their own companies. For example, a Philippines-based startup reflected that market practices—but not necessarily organizational practices—were transferable to Southeast Asia based on its experience in a Silicon Valley-based accelerator:

> The vast majority of it is very much applicable. There are some nuances that anybody in a foreign market would have to address—and that's cultural stuff. Like in the Philippines, people are generally more risk-averse than in the US. So you have to consider that when you're recruiting, for example. Saying you're a startup in the Philippines is—when you're interviewing somebody, in their mind, it's like, oh, this company could go out of business soon. I don't know if I want to work for them because I want a steady job. So there are things like that.

Consistent with this qualitative data, the interview responses also reveal heterogeneity in the organizational design responses across headquarters regions. Using a log-odds NLP approach, Figure A.12.1 shows that North American companies tend to use objective-oriented terms like "build product" and "revenue growth" more than companies from other regions. One test to identify the extent of these differences relative to other choice categories of interview responses is to see how well a machine learning algorithm can detect the headquarters region of a company based on only an organizational design response relative to its ability to do so based on other responses. A higher accuracy would suggest that the organizational design response is more localized to the region relative to the others. A supervised machine learning model trained on 60 percent of the responses and tested on the remaining 40 percent assesses this localization. Table A.12.2 shows that the organizational design and culture responses have among the highest accuracy rates across logistic, neural network, and random forest machine learning models, suggesting that they are more localized to the headquarters region than are other responses.

Figure A.12.1: North American interviewed companies' organizational responses tend to mention an objective more than those of companies from other regions.

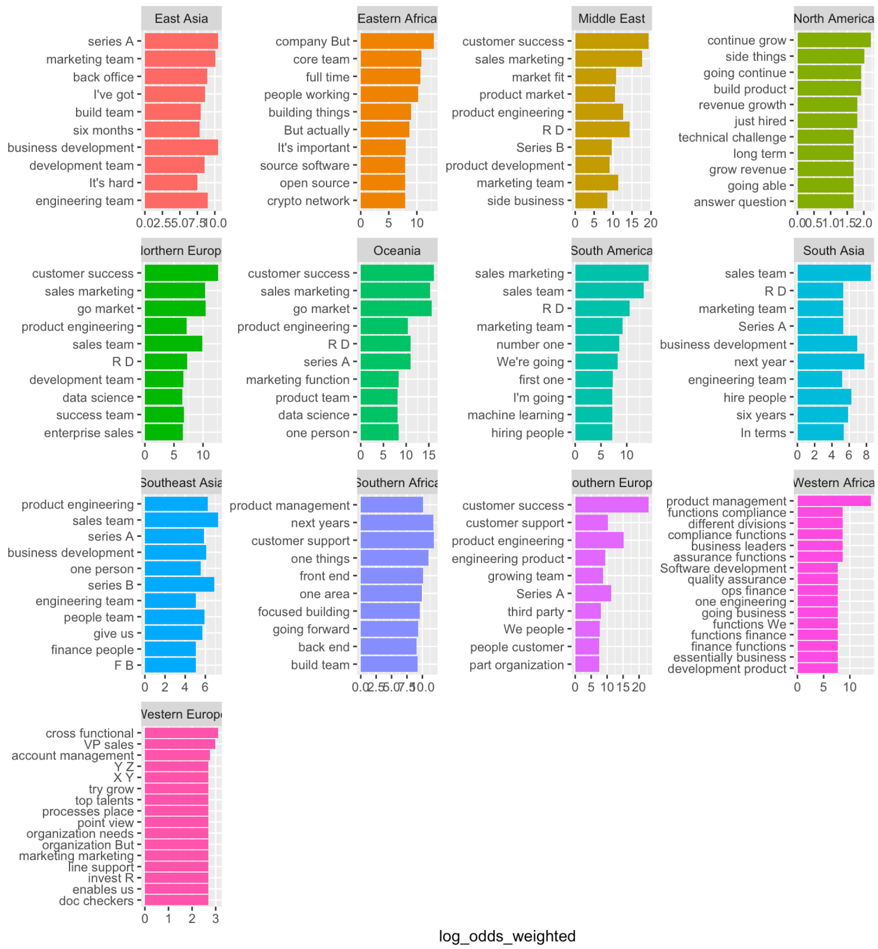
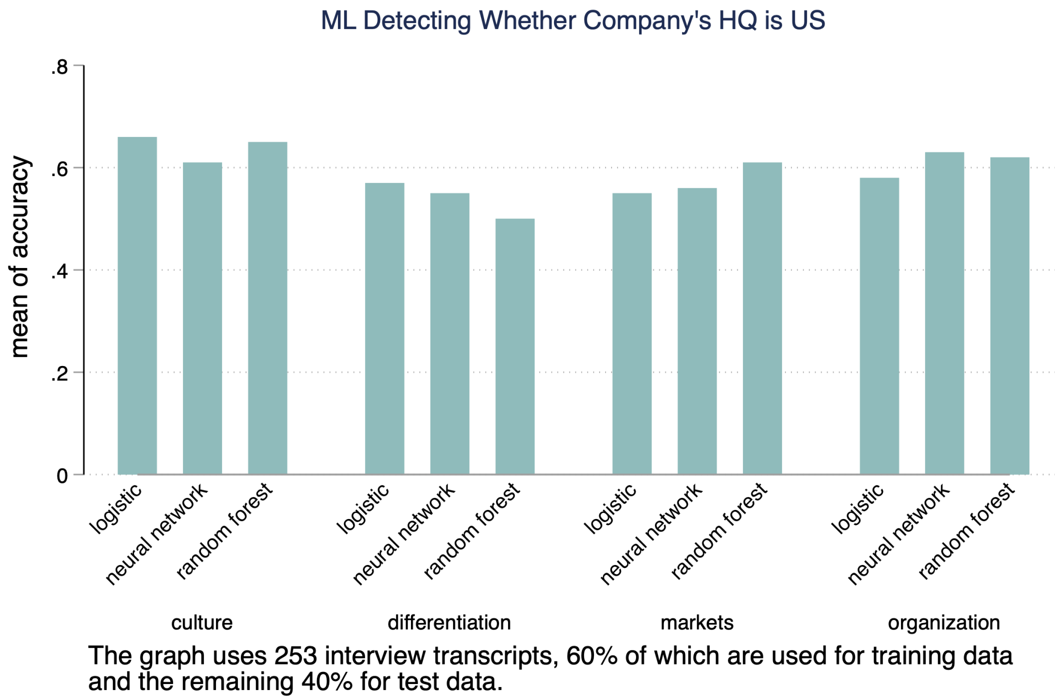
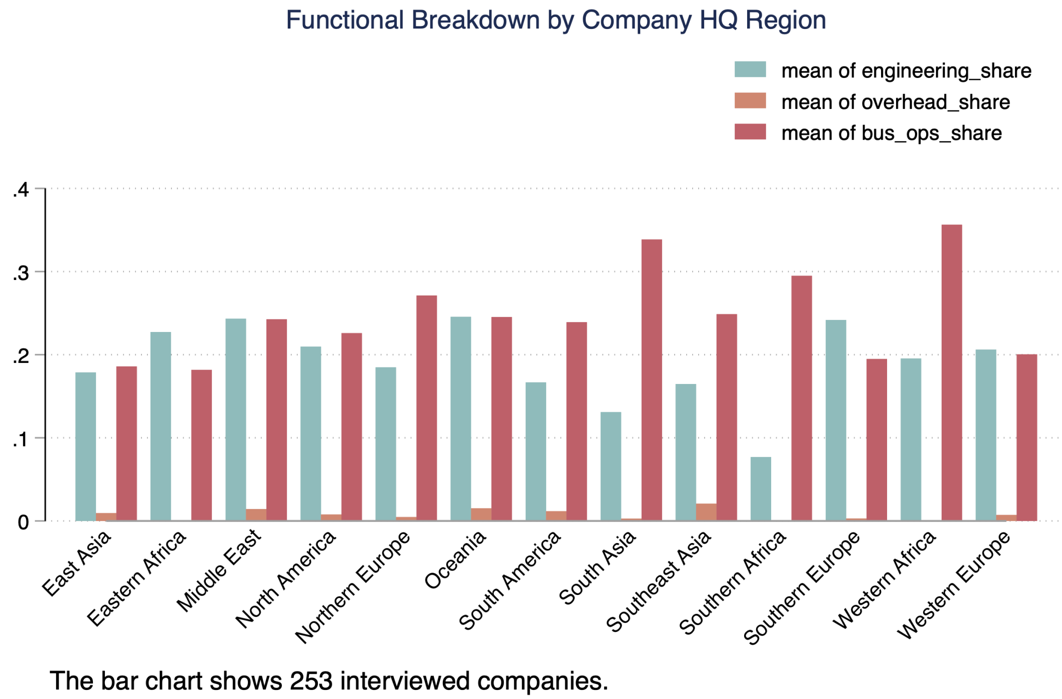


Figure A.12.2: Machine learning models are more accurate at detecting the region of organizational design and culture responses than of other responses of interviewed companies.



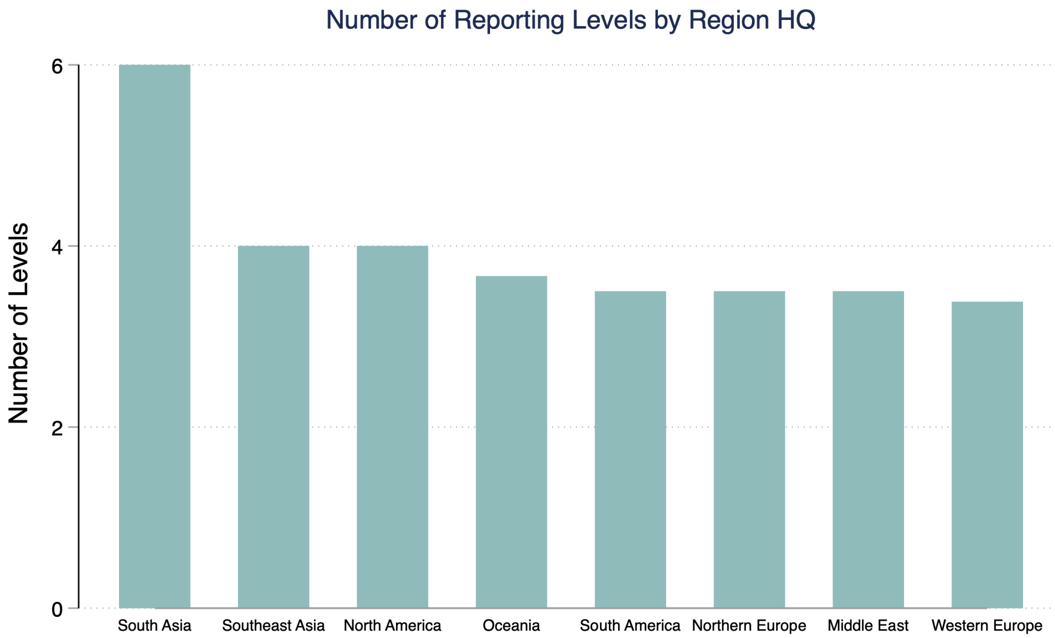
These differences in thinking about organizational design across geographic contexts translate into differences in commitments. Figure A.12.3 shows the distribution of key functions in the organization of the 253 interviewed companies, using online resume data by the region of companies' headquarters. The table shows that this distribution varies widely across regions. Specifically, companies from North America and Western Europe tend to have fairly even shares of engineering and commercial/business talent. Companies from South Asia, Southeast Asia, and South America tend to have a far higher share of commercial/business talent than of engineering talent. This result suggests that organizational structures—in terms of the relative size of departments—vary across geographies.

Figure A.12.3: Organizational structure in terms of breakdown by roles varies across regions.



The organizational design's hierarchical nature is also fairly localized to the headquarters geography. Figure A.12.4 shows the average number of organizational levels from one-fourth of the interviewed companies that provided organizational charts. The number of levels is highest for South Asian companies and lowest for European companies. These results are consistent with the Hofstede "Power Distance" index measuring the extent of hierarchy across national cultures (Hofstede, 2010). Despite the seemingly homogeneous and global "startup" culture that we may often hear about, these descriptive results show that cross-country differences remain.

Figure A.12.4: Organizational "flatness" varies across regions.



The graph shows the average max number of reporting levels in a subset of interviewed firms that shared organizational charts.

A.13 Mediation analysis using US investor and executives' US work experience

Table 1.6 shows that non-US firms are less likely to rely on direct experience, company examples, advisors, and investors to inform their strategy, suggesting that the variance in knowledge inputs may shape why they have lower strategy scores. If knowledge frictions help explain why US companies have higher strategy scores, then we would expect the difference to disappear when accounting for firm executives' exposure to the US, where there has been a concentration of scaling successes and, therefore, access to scaling knowledge (Conti and Guzman, 2021; Kerr and Robert-Nicoud, 2020; PitchBook, 2022).

To test this prediction, Table A.13.1 shows a mediation analysis. It assesses whether accounting for executives having US work experience or US investors reduces the US strategy premium. To do so, it adds US executive work experience and US investor controls in Equation 2. If the coefficient on whether a company is headquartered in the US declines or loses significance at the five-percent level relative to the same coefficient in Equation 2, this would suggest that US direct or investor experience is a mechanism behind the US strategy premium.

The analysis shows that the US's overall strategy premium declines by over half when incorporating controls for whether companies had US investors (Column 2). It declines similarly when accounting for whether executives had US work experience (Column 3). The US strategy premium disappears when incorporating both US investors and executive work experience in the US (Column 4). These regression results suggest that US work experience and advisors help explain why non-US firms have lower strategy scores.

Table A.13.1: Gap between US and non-US strategy scores declines when accounting for US investors and executive work experience in the US.

	(1) Strategy	(2) Strategy	(3) Strategy	(4) Strategy
US HQ	0.269*	0.118	0.130	0.001
	(0.128)	(0.130)	(0.142)	(0.138)
Has US Investor(s)		0.433*		0.403*
		(0.187)		(0.188)
Exec Has US Work Experience			0.378*	0.345*
			(0.159)	(0.160)
_cons	0.322	-0.178	0.209	-0.245
	(0.569)	(0.592)	(0.566)	(0.581)
N	304	304	304	304
Evaluator FE	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Filled-In FE	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.14 Robustness checks with other outcomes

Table 1.3 shows that non-US firms' performance is more sensitive to strategy than that of non-US firms. We might be concerned that the trend only holds for the handful of measures shown in that table, such as logged valuation and exit. To account for this concern, Table A.14.1 uses a similar specification as Table 1.3 to assess this relationship across other measures that are part of the performance index. It shows that a similar trend holds for other performance measures, including logged funding (Column 1), logged number of employees (Column 2), whether companies reached 150 employees (Column 3), whether companies reached 200 employees (Column 4), logged page visits (Column 5), and logged technology tool count (Column 6). While not significant at the five-percent level, the general trend is that the coefficient on the interaction term (US HQ x Strategy) is negative, suggesting that non-US firms' performance is more sensitive to strategy. The coefficient on the strategy score is positive, indicating that, on average, a higher strategy score predicts higher growth.

Table A.14.1: Strategy predicts performance more for non-US firms

	(1) Log Total Raised	(2) Log Employees	(3) Reached 150 Employees	(4) Reached 200 Employees	(5) Log Page Visits	(6) Log Tool Count
US HQ	0.176	0.058	-0.009	0.035	0.031	0.001
	(0.156)	(0.138)	(0.050)	(0.038)	(0.319)	(0.085)
Strategy	0.191	0.129	0.043	0.035	0.161	0.038
	(0.123)	(0.084)	(0.041)	(0.032)	(0.211)	(0.047)
US HQ x Strategy	-0.261^{+}	-0.108	-0.069	-0.062^{+}	0.116	-0.079
	(0.149)	(0.104)	(0.044)	(0.035)	(0.312)	(0.067)
_cons	4.878^{***}	8.815^{***}	2.307^{***}	1.910^{***}	8.896^{***}	4.424^{***}
	(1.274)	(1.048)	(0.492)	(0.468)	(1.979)	(0.980)
N	230	228	230	230	230	230
Evaluator FE	Yes	Yes	Yes	Yes	Yes	Yes
Year Founded FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Log GDP Capita	Yes	Yes	Yes	Yes	Yes	Yes
Log First Financing	No	No	No	No	No	No
Filled-In FE	Yes	Yes	Yes	Yes	Yes	Yes
Readability	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors (in parentheses) are clustered at the company level.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B

Chapter 3 Supplementary Material

Contents

B.1 Monte Carlo simulations showing that 0–40% judge location measurement error still allows us to detect foreign discounting and local information advantage effects. . . . 211
B.2 Kernel density plots showing that judges give lower scores to foreign startups 216
B.3 Judges across home regions are biased against foreign startups. 217
B.4 Foreign bias results are robust to raw weighted and non-weighted measures of judges final recommendation score. 220
B.5 Judges are equally likely to give incomplete subscores to foreign and local startups. 222
B.6 Alternative foreignness measures produce similar results. 223
B.7 Results hold when excluding investor judges and Latin American startups. 226
B.8 Robustness to check that startups are not strategically disclosing their region in their applications 228
B.9 Robustness to country and founder quality measures 230
B.10 Foreign bias results hold whether the judge-startup industry matches or not. 233
B.11 Hub location does not explain the foreign discounting effect. 236
B.12 The findings generalize to all years and rounds of the accelerator program. 238
B.13 Robustness using alternative measures of startup quality 239
B.14 Judges are equally responsive to quality when we split our sample by foreignness. . 244
B.15 Judges are worse at detecting the quality of startups that have raised financing, are from hubs, or have founders affiliated with elite universities. 246

B.16 Judges are worse at selecting winners from losers among startups in later stages of the competition. 248
B.17 Judges can equally detect the quality of local vs. foreign startups no matter whether startups regions are mentioned in their applications. 249
B.18 Judges have a local information advantage when assessing very localized startups. . 251
B.19 Additional analyses estimating missed startups. 254
B.20 Application questions that startups answer and judges evaluate in the accelerator program 256
B.21 Robustness when including elite university and hub controls 258

B.1 Monte Carlo simulations showing that 0–40% judge location measurement error still allows us to detect foreign discounting and local information advantage effects.

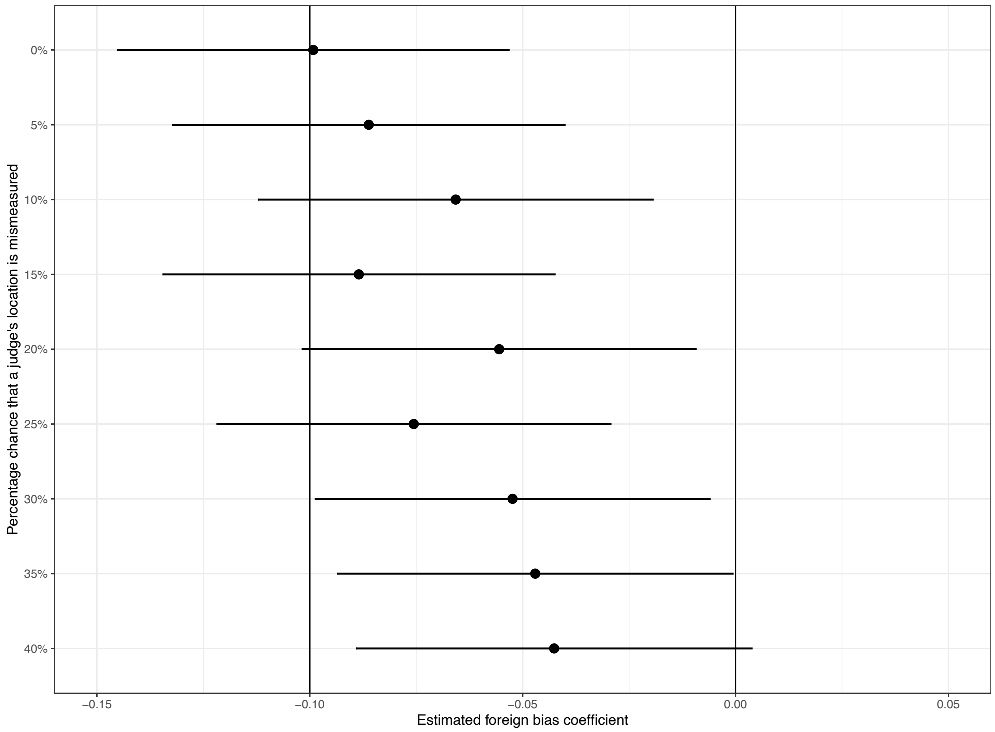
Our data includes information on judges' home regions in the program. There may be a concern that some judges may not actually be from the home regions, which they are associated with in the accelerator, resulting in measurement error that confounds our results. To address this concern, we conduct Monte Carlo simulations to assess how measurement error in judge location impacts our coefficient estimates of both the amount of foreign discounting and local information advantage. In our simulations, we assume that there are 1,000 judges (roughly what is present in our actual analysis). Each judge evaluates 15 startups (again similar to our actual sample). This gives us 15,000 judge-startup pairs, comparable to the 16,320 observations we analyze in Table 3.4. For simplicity, we assume there are two types of judges: US judges and EU judges.

We then evaluate the impact of measurement error on two different models of how judges evaluate startups. In the first model, we assume that startups vary in their quality, judges vary in their harshness, and judges simply discount startups from outside their own region by -0.1 standard deviations, no matter the startup's quality. This effect is slightly larger than our estimate of -0.06, but if there is measurement error in our data, then this figure—as the simulations below show—is likely an underestimate of the true effect. Given this model, we then randomly shuffle judge's locations to vary the mismeasurement rate from 0% (all judge locations are perfectly measured) to 40% (2 in 5 judges are recorded as being from the wrong region). We then run regressions exactly like Equation 1 in the paper, where we regress the score on whether the judge is measured to be foreign or not while including judge and region fixed effects. With this simulation effort, our goal is to test if our coefficient estimates match the true foreign discount rate of -0.1 standard deviations.

Figure B.1.1, displayed below, shows the estimated foreign bias coefficient estimates given data with different measurement error rates. As the error rate goes up, our estimate is increasingly biased towards zero, consistent with intuition and arguments concerning classical measurement error. With an error rate of 30% the estimated effect size is -0.05 but still significant at the 5% level; at 40%, we find a positive but statistically insignificant effect.

Beyond the impact on the main foreign bias coefficient, we also evaluate how measurement error impacts a model where judges are assumed not just to be biased, but also more informed about the quality of local as against foreign startups. In this model, which conceptually mirrors Equation 2 in the paper, we assume that all

Figure B.1.1: Classical measurement error: Foreign bias coefficients



judges, on average, rate a startup one standard deviation higher when the startup's quality goes up by one standard deviation. That said, we also assume judges from a different region are worse at evaluating the quality of these foreign startups (e.g., US judges are worse at evaluating the quality of EU startups and vice versa). Again, following equation 2, we assume that the interaction between "foreign" and "quality" has a value of -0.1, indicating that foreign judges in our simulated data are a tenth of a standard deviation less responsive to differences in quality for foreign as against local firms. We refer to this interaction term as the "ignorance" coefficient. This value reflects how much less sensitive judges are to foreign startup quality relative to local startup quality. Figures B.1.2– B.1.3 below show the estimated "bias" and "ignorance" coefficients as we increase the amount of measurement error. Again, we see estimates on both coefficients shrink toward zero.

Overall, these findings suggest that uniform measurement error should not cause us to find a result when there isn't one. However, we might worry that the rate of measurement error varies across geographies. Perhaps many of the judges from non-US home programs are actually from the US, but judges from the US home program are nearly always from the US. To evaluate how differential rates of measurement error impact our results, we also present results showing what happens if we increase measurement error in only one region. Specifically, we take US judges and assign a given percentage of them to be labeled as "EU" judges. For EU judges, we introduce

Figure B.1.2: Classical measurement error: Foreign bias coefficients

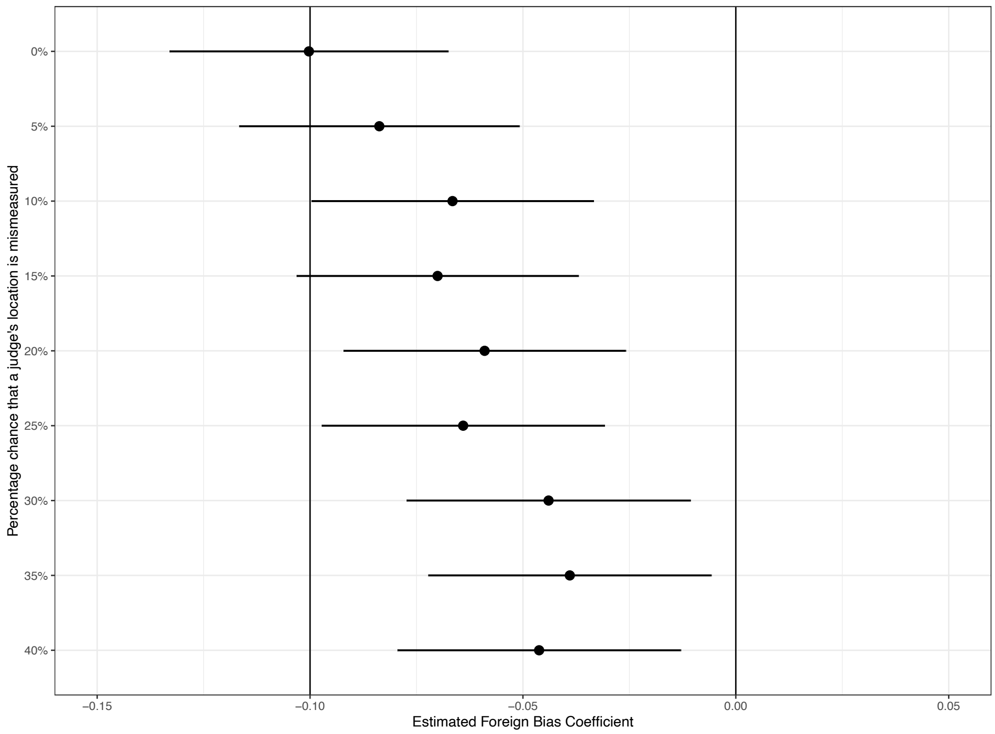
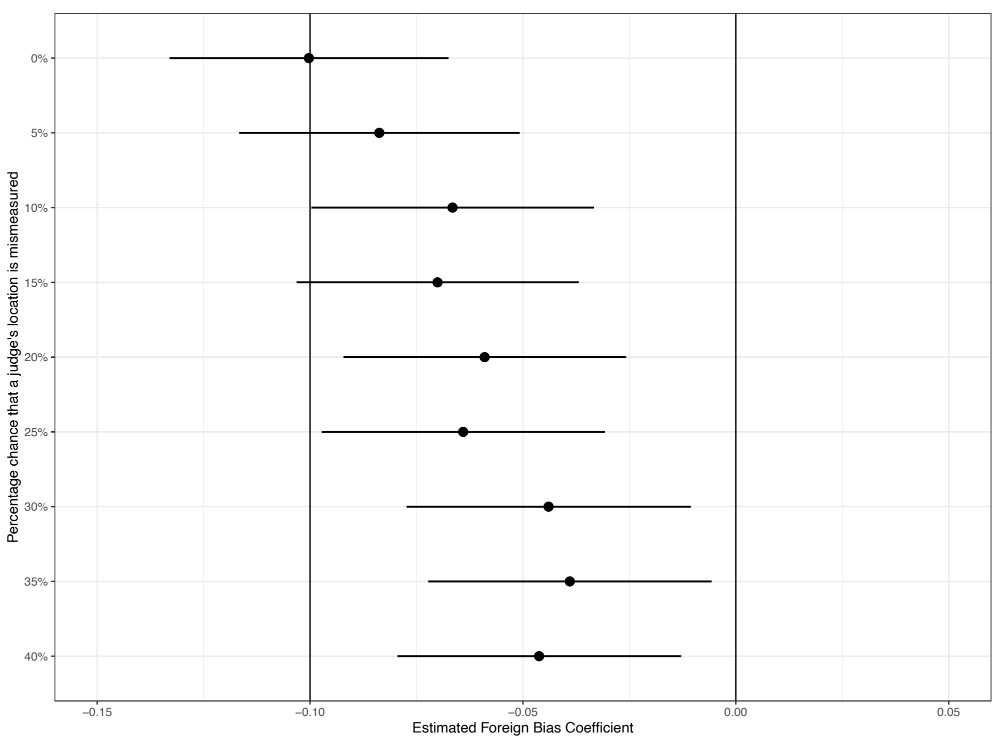
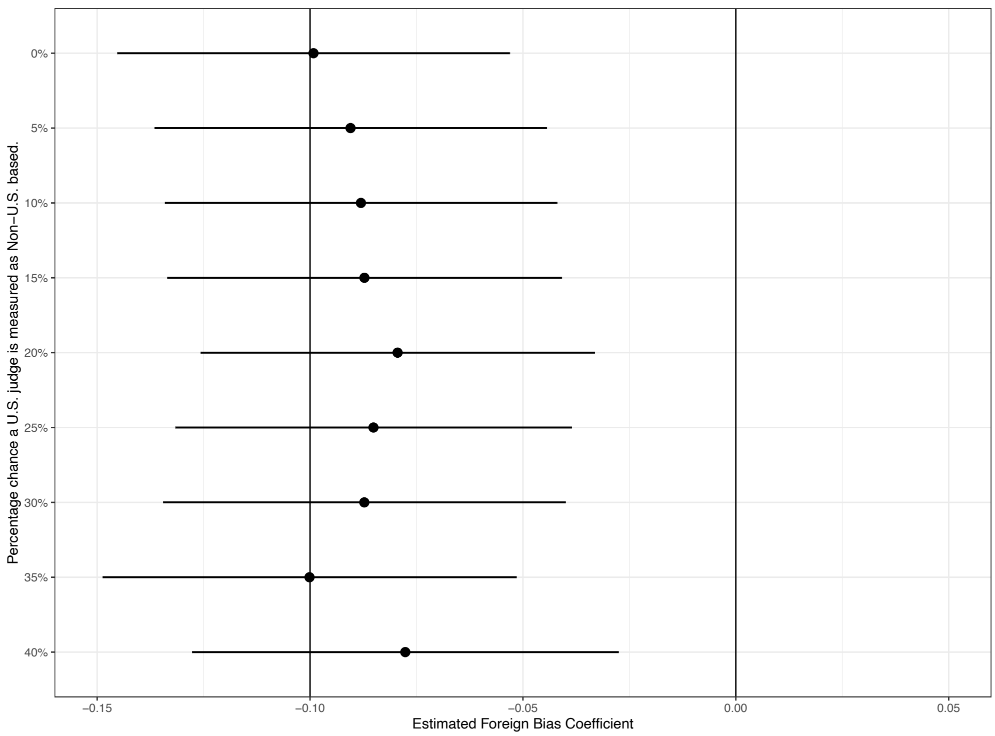


Figure B.1.3: Classical measurement error incorporating information asymmetries: Foreign information coefficients



no measurement error. Figure B.1.4 shows the estimated coefficients for the "foreign bias only" model as we increase the mislabeling rate for US-based judges from 0% to 40%. While the coefficients shrink, the decline is less pronounced than before because only half of our sample is mismeasured. The figure suggests our estimates are likely a lower bound for the foreign bias effect.

Figure B.1.4: Classical measurement error: Foreign bias coefficients with asymmetrical error



Turning to the impact on a model with both foreign bias and foreign ignorance, we find similar patterns. Figures B.1.5 – B.1.6 again show that as the error rate increases, our estimates tend towards zero. It is not the case that when mismeasurement only impacts one region that we overstate the size of the ignorance or bias effects. Instead, effect sizes tend towards zero, as is the case with symmetric measurement error. We do see in these graphs that the bias in the ignorance coefficient towards zero grows faster than the bias in the foreign bias coefficient. However, neither coefficient ends up overestimated.

Overall, while we wish we had richer data on judge location, our simulation results show that even if there is relatively severe measurement error in judge location, we should still be able to detect effects, though the estimates are likely to be lower bounds on the true effect. This robustness is largely due to our large sample size, which lets us recover effects even in the face of relatively large measurement error rates. Again, the simulations suggest that, if anything, our findings are underestimates of the true foreign discounting effects.

Figure B.1.5: Classical measurement error incorporating information asymmetries: Foreign bias coefficients with asymmetrical error

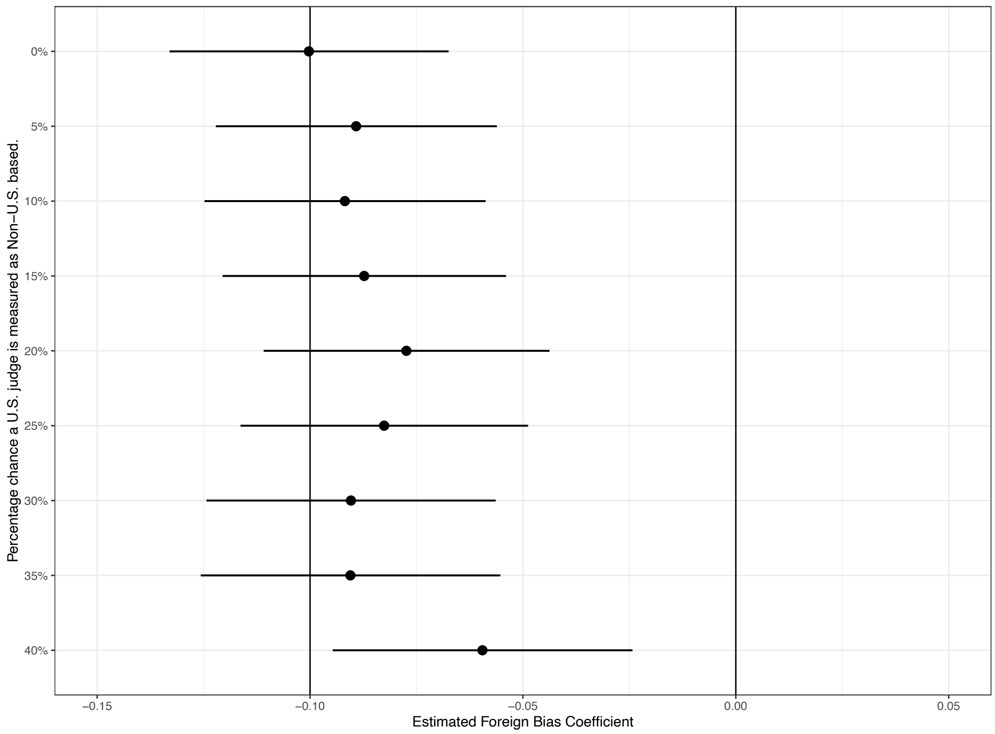
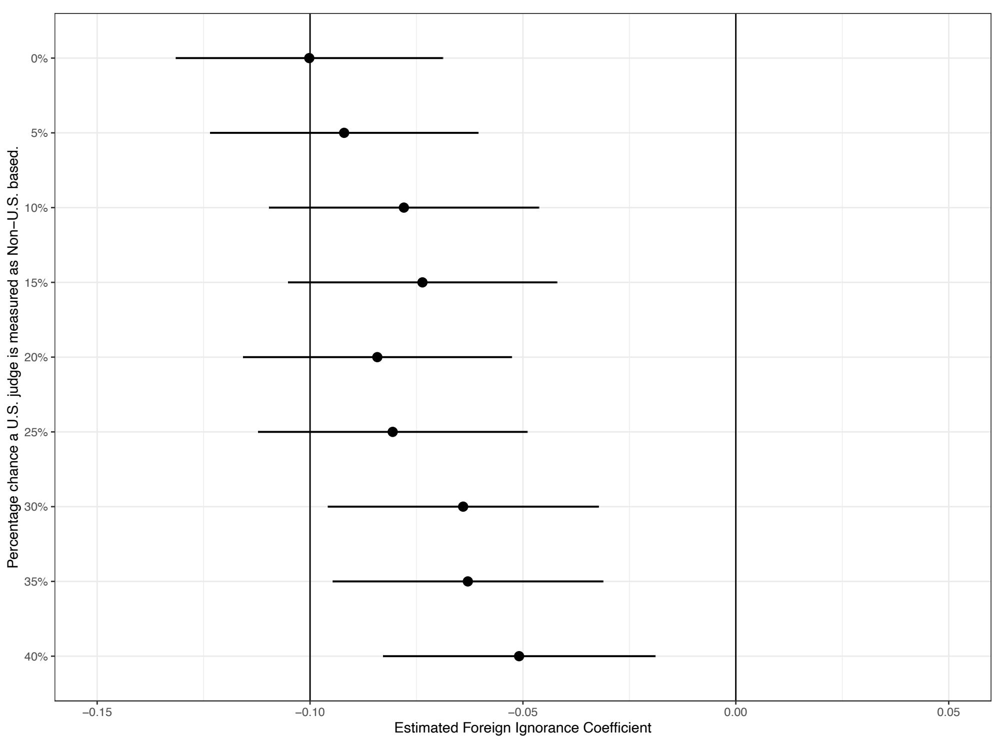


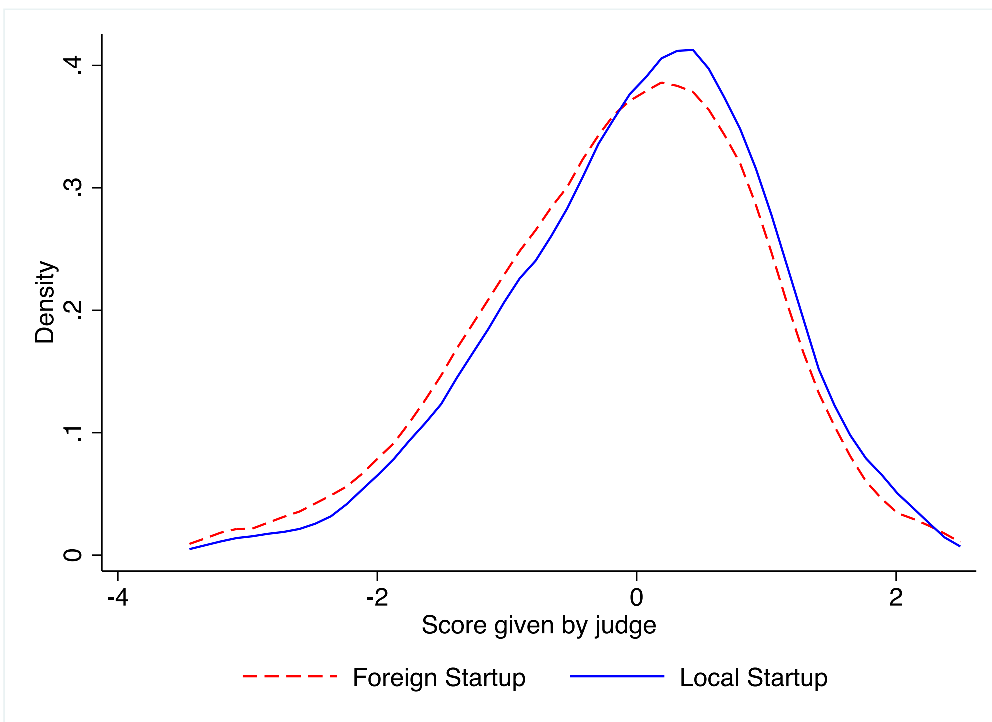
Figure B.1.6: Classical measurement error incorporating information asymmetries: Information coefficients with asymmetrical error



B.2 Kernel density plots showing that judges give lower scores to foreign startups

Tables 3.4– 3.5 and Figure 3.3 show that judges discount foreign startups on average. However, this effect might only apply at the left tail of the scoring distribution, which would suggest that the bias only matters for startups that would never be selected anyway, limiting the economic significance of the foreign bias effect. To assess whether this is the case, we plot the distribution of scores by whether the startup is local or foreign to the judge. Figure B.2.1 presents a kernel density plot of these distributions. The scores for local startups first-order stochastically dominate (i.e., are always to the right) of the scores given to foreign startups. Foreign discounting matters for startups with both bad and good scores.

Figure B.2.1: Kernel density plot of scores by whether the judge and startup are from the same region (local startup) or from different regions (foreign startup)



B.3 Judges across home regions are biased against foreign startups.

Is our foreign bias effect simply the result of judges from a particular geography being biased? For example, perhaps only US judges dislike startups from other countries. Or is the foreign bias broadly based? To adjudicate between these alternatives, in Figure B.3.1, we show the average local and foreign recommendations for North American, European, Latin American, and Israeli judges. In Panel A, we see that US and European judges are less likely to recommend foreign startups relative to local ones, though this does not hold for Latin American and Israeli judges. This likely is because of the lower quality of Latin American startups[1] as well as the relatively small share of both Israeli and Latin American firms in our sample.[2] When we control for a minimum threshold of quality by limiting our sample to financed startups, shown in Panel B, we find that judges from all regions, but Latin America, discount foreign startups. The effect is not merely the result of a single country being particularly harsh towards foreign firms.

To further test whether non-US judges differ from US judges, in Table B.3.1 we run regressions similar to our primary Table 3.4 in the paper, but we interact our "foreign startup" dummy with whether the judge is from Latin America, Europe, or Israel. Even though the judge's home region is fixed, since the "foreign startup" dummy varies within a judge, these interaction terms are still identified when we include judge fixed effects. Since we have fixed effects for judges, the main "judge region" dummies that one would normally include drop in the regression. In this regression, the coefficient on "foreign startup" corresponds to the level of bias exhibited by US judges. Each interaction term then reflects if the bias is different for judges from Latin America, Europe, or Israel.

Column 1 in Table B.3.1 shows the results of regressing the judge's score on whether the startup is foreign along with our judge region interaction terms. This model includes region fixed effects. Consistent with our results in Table 3.4, we find a significant and negative effect on the foreign startup dummy variable. None of the interaction terms are significant at the 5 percent level, and the coefficients for Latin American and Israeli judges are close to zero. While the coefficient for European judges is positive and large, the confidence intervals overlap with zero. Further, in models 2 and 3 when we include country and startup fixed effects, respectively, we see the size of the coefficients shrink toward zero. Overall, these results show that our findings are not driven by judges

[1]Latin American startups have the lowest probability of being recommended to the next round relative to startups from other regions. These startups are over 20 percent less likely to be recommended to the next round compared to other startups in our main sample.

[2]Latin American startups comprise less than 15 percent of all startups in our main sample, and Israeli startups comprise 9 percent of all startups in our main sample.

from one particular region.

Figure B.3.1: Bar graph showing that US, EU, and Israeli judges are more likely to recommend local over foreign startups to the next round of the competition. Panel A is for all startups and Panel B is only for startups that have raised financing.

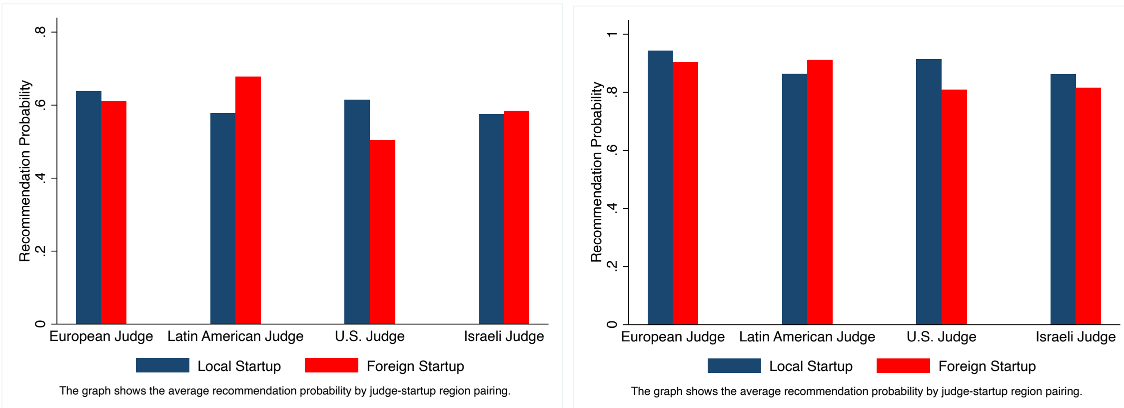


Table B.3.1: Israeli, European, and Latin American judges are similarly discounting foreign startups as are other judges.

	(1)	(2)	(3)
	Judge's Total Score		
Foreign Startup	-0.109**	-0.108**	-0.087**
	(0.038)	(0.039)	(0.033)
Foreign Startup x Latin American Judge	0.099	0.101	0.052
	(0.106)	(0.105)	(0.084)
Foreign Startup x European Judge	0.132	0.130	0.085
	(0.076)	(0.075)	(0.063)
Foreign Startup x Israeli Judge	0.017	0.013	-0.029
	(0.099)	(0.099)	(0.089)
Observations	16,320	16,320	16,264
Judge x Year	Yes	Yes	Yes
Startup Region x Year	Yes	No	No
Startup Country x Year	No	Yes	No
Startup x Year	No	No	Yes

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

B.4 Foreign bias results are robust to raw weighted and non-weighted measures of judges' final recommendation score.

Tables 3.4- 3.5 show that judges discount foreign startups when using a composite measure of judge scores (across all sub-categories discussed in Section III) and a binary recommend measure. A concern with these tables is that they show noisy measures of judges' evaluation of startups because they are not using the ultimate continuous measure of final recommendation that judges give to startups. To address this concern, we show in Tables B.4.1–B.4.2 the same regressions as in Table 3.5, but now using two continuous versions of the judges' final recommendation score as the dependent variables. The first is a weighted continuous measure of judges' final recommendation score to startups (on a 0, 2, 4, 12, 16, and 20 scale) with scores 12 and above indicating a positive recommendation (Table B.4.1). The second is a non-weighted version (on a scale of 0, 1, 2, 3, 4, 5) with scores 3 and above indicating a positive recommendation (Table B.4.2). The results are consistent with those in Table 3.5. Judges discount foreign startups even when we control for judge and startup fixed effects.

Table B.4.1: Regressions showing that judges score lower startups from outside their home region even when we control for judge and startup fixed effects: Weighted raw recommendation score

	(1)	(2)	(3)	(4)
	Judge's Total Score			
Foreign Startup	-1.347***	-0.636***	-0.631***	-0.622***
	(0.120)	(0.114)	(0.102)	(0.107)
Has Traction				1.223***
				(0.181)
Has Financing				4.936***
				(0.138)
Observations	17,593	17,593	17,590	17,593
Judge x Year	Yes	Yes	Yes	Yes
Startup Region x Year	No	Yes	No	No
Startup Country x Year	No	No	Yes	No
Startup x Year	No	No	No	Yes

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

Table B.4.2: Regressions showing that judges score lower startups from outside their home region even when we control for judge and startup fixed effects: Non-weighted raw recommendation score

	(1)	(2)	(3)	(4)
		Judge's Total Score		
Foreign Startup	-0.277***	-0.130***	-0.131***	-0.127***
	(0.025)	(0.023)	(0.021)	(0.022)
Has Traction				0.258***
				(0.037)
Has Financing				0.969***
				(0.030)
Observations	17,595	17,595	17,592	17,595
Judge x Year	Yes	Yes	Yes	Yes
Startup Region x Year	No	Yes	No	No
Startup Country x Year	No	No	Yes	No
Startup x Year	No	No	No	Yes

Standard errors (in parentheses) are clustered at the judge and startup levels.
Fixed effects shown below observations. The table includes two additional data points (0.01% of the total sample) due to labeling changes.
* p<0.05 ** p<0.01 *** p<0.001

B.5 Judges are equally likely to give incomplete subscores to foreign and local startups.

Table 3.4 shows that judges score foreign startups lower than local ones. This table excludes evaluations that received missing subscores from judges. We might be concerned that judges are more likely to give missing subscores to foreign startups. Perhaps this may be because they have a harder time understanding them and, therefore, "punt" the decision by leaving the score blank. If this were the case, then there would be missing foreign startups in the analysis that may confound our main result. Since foreign startups are given lower scores on average than are local startups, these missing foreign values would likely bias our foreign discounting result upward. In Table B.5.1, we regress whether a judge gives an incomplete score on whether the startup is foreign to the judge. We find that judges are equally likely to give incomplete scores to local and foreign startups. This suggests that incomplete subscores are not driven by whether startups are foreign to judges. Therefore, incomplete subscores are unlikely to confound our results.

Table B.5.1: Probability that judges give incomplete subscores

	(1)
	Incomplete Subscore
Foreign Startup	0.002
	(0.005)
Observations	17,590
Judge x Year	Yes
Startup x Year	Yes

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

B.6 Alternative foreignness measures produce similar results.

Our primary measure of whether a startup is foreign is based on the headquarters region of the startup. However, given that the startup application does not list the HQ location of the startup, it remains possible that the discounting could be related to factors unrelated to the startup's location. Further, even if we find evidence that judges are biased against foreign locations, we know that "foreignness" to the judge is likely non-binary but continuous. There likely is a spectrum of how foreign a startup is to a judge. Finally, judges might be picking up on the fact that ideas and founder experiences from different countries might be different. For example, an Israeli judge is likely to understand the pros and cons of backing a team that was part of the IDF's Unit 8200, whereas a Latin American judge might not.

We address each of these three concerns here. First, we used a simple dictionary-matching procedure to check if the application text explicitly mentioned the country name when describing the market the startup was targeting, the description of the team, and the startup's coming traction. Admittedly, this approach will miss indirect location indicators, for example, if the application mentions the team is based in the city of Ghent, indicating operations in Belgium. However, it will allow us to identify startups that are obviously and explicitly local or foreign to the judge. In Column 1 of Table B.6.2, we show that using this much more stringent measure still results in judges discounting foreign startups. In Table B.6.3 Column 1, we further show that the foreign bias effect that we pick up with the headquarters of a startup is stronger when the home region is explicitly mentioned in the text according to this measure. In fact, when we limit the sample to startups that do not reveal their home region in their application text, we see that the foreign discounting effect weakens substantially (Column 2).

To address the second concern that distance is not binary but continuous, we generate a simple measure using the geographic distance between the startup's HQ location and the center of the judge's home region, using a measure developed by Berry et al. (2010). This measure serves two functions. First, for a US judge, it would classify startups from Germany as more distant than a startup from London. Second, given the European office of the accelerator is based in Switzerland, for a European judge, startups from Germany would be classified as less foreign than from London. To account for the skew in distances, we log geographic distance as is common in studies of international trade and strategy. Table B.6.2 Column 2 shows the results using our logged distance measure. Consistent with our results in Table 3.4, we find that judges give lower scores to startups more distant in geographic space even when using a more continuous measure.

To address the third and final concern—that startups from different regions are different in their ideas and approaches—we use natural language processing (NLP) tools to classify the text of startups from different regions as more or less "of that" region. Specifically, we take the application text for each startup and convert it into a "bag of words." We follow standard practice and first remove all punctuation and capitalization before converting the text to a bag of words. We also remove standard stop words. To remove idiosyncratic words, we only retain words that appear in at least 10 startup descriptions and are used at least 20 times. For example, this approach avoids us from picking up on startup names that might be unique to a firm. Using these words, we then estimate for each word in our corpus of startup text whether the word is more or less likely to occur in a given region or not. We do so using a weighted log-odds ratio procedure as described in Monroe et al. (2008). The end result is that for each word in our corpus, we know how much more likely (and unlikely) it is to appear in North American, Latin American, European, and Israeli startup applications. Table B.6.1 shows N example words that are most and least likely to appear from each region, along with the estimated log-odds. These log-odds give us a quantitative estimate of how regional or localized the startup's application is. To generate measures at the startup level, we simply sum these scores across the application text and divide by the total number of words to account for differences in text length. We then standardize each of these variables to get measures for how North American, Latin American, European, and Israeli each startup is.

Table B.6.1: Example "localized" words by startup home region

North American words	Raw log-odds ratio	Israeli words	Raw log-odds ratio
opioid	3.61	jerusalem	6.77
sbir	3.48	technion	4.98
nih	3.46	idf	4.37
northeast	3.17	wix	2.79
dartmouth	3.08	jewish	2.91

European words	Raw log-odds ratio	Latin American words	Raw log-odds ratio
epfl	5.59	pesos	6.73
gmbh	3.57	mercado	5.58
organisation	3.3	cdmx	5.6
hec	3.76	colombia	3.7
chf	5.58	argentina	3.16

Table B.6.2: Foreign discounting using other measures of foreign startup

	(1)	(2)	(3)
	Judge's Total Score	Judge's Total Score	Judge's Total Score
Foreign Startup (Application Word Search)	-0.039**		
	(0.013)		
Log Geographic Distance		-0.008***	
		(0.002)	
Foreign Startup (Application NLP)			-0.0265
			(0.0074)
Observations	17,590	16,257	17,590
Judge x Year	Yes	Yes	Yes
Startup x Year	Yes	Yes	Yes

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

Table B.6.3: Foreign discounting weakens when the region is not explicit in the application text

	(1)	(2)	(3)
	Judge's Total Score		
Foreign Startup	-0.067**	-0.039	-0.039*
	(0.026)	(0.021)	(0.020)
Foreign Startup * Home Region Explicit in Application			-0.054*
			(0.027)
Observations	6,764	9,356	16,264
Judge x Year	Yes	Yes	Yes
Startup x Year	No	No	No
Region Explicit in App	Yes	No	Yes
Region Not Explicit in App	No	Yes	Yes

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

B.7 Results hold when excluding investor judges and Latin American startups.

Our data includes startups and judges from heterogeneous backgrounds. On the one hand, this improves the generalizability of our analysis. On the other, it opens up the possibility that particularly "weird" subgroups drive our findings. Specifically, while most judges are not investors, some are. For these investor judges, we might worry that they give lower scores to foreign startups because they would rather select a local startup that will be easier for them to invest in. On the startup side, we might worry that judges are simply responding to differences in writing quality. Since applications are in English, startups who are from international contexts where English training and education are weaker may be at a disadvantage. This is likely the case for startups from Latin America as English is both used and taught much more often in North America, Europe, and even Israel.

In Table B.7.1, we show that the foreign bias effect holds when we exclude judges who are investors from our sample. This is not all together surprising since the accelerator guidelines explicitly tell the judges that they should not expect to personally gain by participating.

In Table B.7.2, we exclude Latin American startups from our sample and again find that the foreign bias effect holds. This largely rules out the idea that poor writing coming from a single region is responsible for our results.

Table B.7.1: Foreign discounting effect excluding investor judges

	(1)	(2)
	Judge's Total Score	Judge Recommends Startup?
Foreign Startup	-0.063***	-0.044***
	(0.018)	(0.010)
Observations	13,930	15,013
Judge x Year	Yes	Yes
Startup x Year	Yes	Yes

Standard errors (in parentheses) are clustered at the judge and startup level. Fixed effects shown below observations. The table excludes judges who are investors.
* p<0.05 ** p<0.01 *** p<0.001

Table B.7.2: Foreign discounting effect excluding Latin American startups

	(1)	(2)
	Judge's Total Score	Judge Recommends Startup?
Foreign Startup	-0.058**	-0.035***
	(0.018)	(0.010)
Observations	13,418	14,505
Judge x Year	Yes	Yes
Startup x Year	Yes	Yes

Standard errors (in parentheses) are clustered at the judge and startup level. Fixed effects shown below observations. The table excludes Latin American startups.
* p<0.05 ** p<0.01 *** p<0.001

B.8 Robustness to check that startups are not strategically disclosing their region in their applications

One concern about the foreign bias results in Tables 3.4– 3.5 is that startups are strategically mentioning their region in their applications. This would mean that the foreign bias results are partly driven by startups' decisions to disclose their locations, confounding our ability to evaluate judge decision-making. Thankfully, the design of the program mitigates this problem since startups do not know the location of the judges who will be judging them in the rounds we analyze. Thus, a US startup has no control over whether an EU-based judge will or will not judge them. This makes it challenging to strategically game disclosure.

That said, we might be concerned that lower-quality startups are less aware of potential foreign bias and so are more likely to disclose their region explicitly. This would suggest that quality measures may bias the region identifiers. We check this in Table B.8.1 where we regress whether a startup mentions their region explicitly in their application on startup quality measures. Column 1 shows that different metrics of quality—whether startups have user traction, financing, or are located in a hub—do not predict whether startups explicitly mention their region in their application.

We also might be concerned that non-US startups think there is a US premium (given that the accelerator is US-headquartered) and thus are more likely to NOT disclose their locations. Column 2 shows that non-US startups are actually less likely to disclose their region by about 4 percentage points.

Together, these results suggest that neither quality nor being US-based predict whether startups explicitly mention their regions in their applications, reducing concerns about startups strategically disclosing their location.

Table B.8.1: Startup quality does not predict whether region is explicitly mentioned in applications

	(1)	(2)
	Whether Region Explicit in App	
Log Pre-Accelerator Page Visits	0.009	
	(0.012)	
Log Pre-Accelerator Financing	0.010	
	(0.015)	
Has Traction	-0.025	
	(0.066)	
Has Financing	-0.073	
	(0.088)	
Whether Hub Region	0.047	
	(0.027)	
Whether US HQ		-0.037*
		(0.016)
Observations	9,724	17,594
Judge x Year	Yes	Yes
Startup Region x Year	No	No
Startup Country x Year	Yes	No
Startup x Year	No	No

Standard errors (shown in parentheses) are clustered at the judge and startup levels. Fixed effects shown below observations. The table includes one additional data point (0.01% of total sample) due to labeling changes.
* p<0.05 ** p<0.01 *** p<0.001

B.9 Robustness to country and founder quality measures

To rule out quality differences across countries and startups in Table 3.4, we include region, country, and startup fixed effects. While these fixed effects account for any time-invariant quality differences between startups or countries, they provide little insight into what aspects of a country or a startup judges rate higher. Are startups from wealthier countries rated better? From more innovative countries? From places with more VC funding? In hubs like Silicon Valley?

To address these questions, in Table B.9.1, we directly control for these differences. We include GDP Per Capita (World Bank)[3], patent applications to the USPTO (OECD)[4], venture capital availability (the World Economic Forum's Global Competitiveness Index)[5], and whether the startup affiliates with a hub in its application text (using the Startup Genome (2021) classification shown in Table B.9.3 where we also show the distribution of startups in our sample). As expected, judges are more likely to rate startups from wealthier, move innovative, VC-rich, and hub regions more highly. That said, even when we control for these variables directly, we still find a foreign discounting effect.

Similarly, while startup fixed effects account for all time-invariant differences in founder quality, they provide little insight into what aspects of a founder are valued by judges. To directly test how differences in founder background impact judging, we generate measures for whether a founder has a PhD, MBA, or an affiliation with an elite university. We use the text describing the founding team to generate these measures. Thus, if a founder has a PhD but does not mention it in the text, we would mark her as not having a PhD. To generate our measure of whether the founder affiliates with an elite university, we check if the text contains the name of one of the top 10 elite universities in each of the regions in our sample (Europe, North America, and Latin America), according to QS World University Rankings (2021). Table B.9.2 shows models including these controls. Unsurprisingly, founders with MBAs, PhDs, or elite university affiliations are given higher scores. As above, we still find a foreign discounting effect.

[3]Data may be accessed here: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD.

[4]Data may be accessed here: https://stats.oecd.org/Index.aspx?DataSetCode=PATS$_C$*OOP*.

[5]Data may be accessed here: https://tcdata360.worldbank.org/indicators/h8a7ea3d1?indicator=529&viz= line$_c$*hart*&*years* = 2007, 2017.

Table B.9.1: Foreign discounting effect controlling for country HQ quality

	(1)	(2)	(3)	(4)
	Judge's Total Score			
Foreign Startup	-0.133***	-0.067***	-0.104***	-0.195***
	(0.021)	(0.020)	(0.020)	(0.021)
Log Startup HQ GDP Per Capita	0.180***			
	(0.020)			
Log Startup HQ Patent Apps		0.053***		
		(0.005)		
Startup HQ VC Availability			0.166***	
			(0.017)	
Startup Hub				0.114***
				(0.025)
Observations	16,304	16,308	16,306	16,320
Judge x Year	Yes	Yes	Yes	Yes
Startup x Year	No	No	No	No

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

Table B.9.2: Foreign discounting effect controlling for founder quality

	(1)	(2)	(3)
	Judge's Total Score		
Foreign Startup	-0.059**	-0.061**	-0.062**
	(0.020)	(0.020)	(0.019)
Founder(s) have PhD	0.359***		
	(0.032)		
Founder(s) have MBA		0.193***	
		(0.028)	
Founder(s) Attended Elite university			0.301***
			(0.024)
Observations	16,320	16,320	16,320
Judge x Year	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes

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

Table B.9.3: Distribution of startups in our sample by hub

Hub Name	Number of Startups	Share of Startups	Hub Rank in Startup Genome Project (2021)
Amsterdam-Delta	4	0.11	13
Atlanta	4	0.11	26
Austin	8	0.21	20
Bangalore	1	0.03	23
Berlin	15	0.4	22
Bern-Geneva	52	1.38	36
Boston	339	8.97	5
Chicago	32	0.85	14
Dallas	1	0.03	31
Denver-Boulder	6	0.16	27
Dublin	3	0.08	36
Hong Kong	5	0.13	31
London	50	1.32	3
Los Angeles	5	0.13	6
Melbourne	3	0.08	36
Montreal	3	0.08	31
Munich	13	0.34	31
New York City	38	1.01	2
Paris	30	0.79	12
Philadelphia	6	0.16	28
San Diego	2	0.05	21
Seattle	5	0.13	10
Seoul	4	0.11	16
Shanghai	7	0.19	8
Shenzhen	1	0.03	19
Silicon Valley	16	0.42	1
Singapore	9	0.24	17
Stockholm	3	0.08	17
Sydney	1	0.03	24
Tel Aviv	12	0.32	7
Tokyo	6	0.16	9
Toronto-Waterloo	5	0.13	14
Washington, D.C.	5	0.13	11
Other	3,086	81.64	

Startups are identified into hubs based on whether they explicitly mention the hub in the market, team, or traction application fields.
Source: Startup Genome Project (2021)

B.10 Foreign bias results hold whether the judge-startup industry matches or not.

One concern with our foreign bias results in Tables 3.4– 3.5 is that only less or more informed judges are driving foreign bias results. One way we can measure the informedness of judges relative to one another is by whether their industries match with those of the startups. If this were the case, then we might expect that when judges match industries with startups, they would be less biased against foreign startups because they would have other ways to discern quality of the startups through their industry expertise. They also would be better able to detect the quality of startups overall.

As shown in Table B.10.1, when we adapt the specification from Tables 3.4– 3.5, judges are no less likely to be biased against foreign startups when their industries match, as shown by the interaction term between whether a startup is foreign and whether judge-startup industries match in Column 3. This coefficient is not significant at the 5 percent level.

Further, in Table B.10.2, when we adapt the specification from Table 3.6, we find that judges are no better at detecting the quality of startups when their industries match (Column 2), versus when they do not match (Column 3). Of course, these results might be partly capturing selection on industry matches (unlike geographic matches). The accelerator partly allocates judges to startups on the basis of industry matches. Further, industry categories are broadly construed. For example, both consumer application and cybersecurity technology startups are classified as "high technology" (the largest industry category comprising nearly 40 percent of judges and startups). This means that even within an industry category, there is variation in expertise areas.

There may also be a concern that the foreign bias measure is actually reflecting industry bias. This might be the case because industries are often concentrated in certain geographies. To test for this confounding effect, we assess whether judges are biased against startups from different industries. To do so, we apply a similar specification as in equation 1, but replace whether a startup is foreign to the judge, with whether the startup is from a different region as the judge. In Table B.10.3, we show that judges do not discount startups from a different industry. If anything, the coefficient is actually positive (though not significant at the 5 percent level). This result may reflect that startups are following standardized enough business models that judges from across industries can understand them.

Together, these results suggest that more informed judges, as proxied by industry, are not driving our results.

Table B.10.1: Judges are similarly biased against foreign startups no matter whether their industry matches that of the startups.

	(1)	(2)	(3)
	Judge's Total Score		
Foreign Startup	-0.018	-0.070**	-0.037
	(0.034)	(0.023)	(0.022)
Foreign Startup * Different Industry			-0.039
			(0.027)
Observations	5,179	8,847	16,264
Judge x Year	Yes	Yes	Yes
Startup x Year	Yes	Yes	Yes
Same Industry	Yes	No	Yes
Different Industry	No	Yes	Yes

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

Table B.10.2: Judges can tell the quality of foreign vs. local startups with similar precision no matter whether their industry matches that of the startups.

	(1)	(2)	(3)
	Judge's Total Score		
	Baseline	Same Industry	Different Industry
Foreign Startup	-0.065**	-0.068	-0.057
	(0.025)	(0.039)	(0.032)
Log Post-Accelerator Page Visits	0.050***	0.051***	0.048***
	(0.004)	(0.005)	(0.005)
Foreign Startup * Log Post-Accelerator Page Visits	0.000	0.005	0.000
	(0.005)	(0.007)	(0.006)
Observations	16,320	6,471	9,625
Judge x Year	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes
Startup x Year	No	No	No
Same Industry	Yes	Yes	No
Different Industry	Yes	No	Yes

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

Table B.10.3: Judges do not discount startups from outside of their industry.

	(1)	(2)
	Judge's Total Score	Recommend
Different Industry	0.022	0.010
	(0.015)	(0.009)
Observations	16,264	17,591
Judge x Year	Yes	Yes
Startup x Year	No	No

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

B.11 Hub location does not explain the foreign discounting effect.

Figure B.3.1 shows that judges discount foreign startups when located in hubs like the US, Europe, and Israel, but not necessarily elsewhere, like in Latin America. Indeed, Latin American startups differ from US, European, and Israeli startups because, even if they are located in a major city, that city will be a less developed tech hub than Silicon Valley, London, or Tel Aviv is. This raises the concern that if our effects are driven by Latin American startups, then our findings might reflect differences in how judges rate startups from hubs vs. non-hubs rather than foreign discounting. To test this possibility, in Table B.7.2, we show that our foreign discounting effect remains even when we exclude Latin American startups from our sample.

To address the broader concern that we are picking up discounting of non-hub startups, we first classify every startup in our sample as operating in a hub or not. As described in Appendix 9, we used data from PitchBook and the Startup Genome Project (2021) to map the city a startup is in to whether it is classified as a startup hub. In Table B.11.1 (Column 2), we show that controlling for whether a startup is in a hub or not does not alter our foreign discounting effect.

Even if hub location does not explain the foreign discounting effect, perhaps it is the case that foreign discounting occurs for startups only inside or outside of hubs. To understand whether hub location has such a moderating effect, in Table B.11.1 (Column 3), we add an interaction term between whether a startup is in a hub and whether it is foreign. This interaction term in the third row turns out to not be meaningful, suggesting that hub location does not moderate foreign discounting.

Together, these analyses suggest that hub location is neither explaining the foreign discounting effect nor moderating it.

Table B.11.1: Whether a startup is in a hub does not remove or moderate the foreign discounting effect.

	(1)	(2)	(3)
	Judge's Total Score		
Foreign Startup	-0.060***	-0.060**	-0.051*
	(0.016)	(0.020)	(0.021)
Hub		0.066**	0.089**
		(0.025)	(0.031)
Foreign x Hub			-0.047
			(0.039)
Observations	16,264	16,321	16,321
Judge x Year	Yes	Yes	Yes
Startup Region x Year	Yes	No	No
Startup Country x Year	Yes	Yes	Yes
Startup x Year	Yes	No	No

Standard errors (shown in parentheses) are clustered at the judge and startup levels. Fixed effects shown below observations. The table includes one additional data point (0.01% of the total sample) due to labeling changes.

B.12 The findings generalize to all years and rounds of the accelerator program.

Our main results show that judges discount foreign startups when provided only textual information about the startups during the first round of judging in 2017 and 2018. The benefits of focusing on this restricted sample are meaningful. Judges were randomly assigned to startups on the basis of the home region, allowing us to estimate a foreign discount effect without worry that the estimate is merely the result of selection bias. Further, since judges only evaluated the text application, we know the full set of information that the judges based their decision on.

However, there is no such thing as a free lunch. The focus on this restricted sample may raise concern that our results might be particular to this first round of judging in 2017 and 2018. Fortunately, we have access to a broader sample of judge-startup evaluations spanning all years of the accelerator (2013-2019) and from all four rounds of the program. These latter rounds include in-person interaction between judges and startups. While judges were far from randomly assigned in this sample, the larger dataset allows us to check if our results generalize beyond our unique dataset.

Column 1 in Table B.12.1 shows that judges still discount foreign startups in this broader sample that includes all rounds from 2013-2019. Furthermore, in Column 2, we show that the foreign discounting effect holds even if we only focus on the later rounds of the competition. The effect remains significant and negative. While beyond the scope of our study, this suggests that interviewing or interacting with founders—as is common in the VC due diligence process—will not eliminate foreign bias.

Table B.12.1: Foreign discounting effect across all years and rounds of the accelerator program

	(1)	(2)
	Judge's Total Score	
	All rounds (1-4)	Excluding first round
Foreign Startup	-0.064***	-0.118*
	(0.012)	(0.053)
Observations	69,639	24,883
Judge x Year	Yes	Yes
Startup Country x Year	Yes	Yes
Program x Year	Yes	Yes

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

B.13 Robustness using alternative measures of startup quality

An important concern with the results in Table 3.6, which show that judges are no more informed about local as against foreign startups, is that the findings might be specific to the particular quality proxies that we use. If our proxy for startup quality is too noisy, the lack of a difference in how informed the judges are might simply be the result of our noisy measure, not the fact that the judges are actually equally informed.

To address this concern, here we replicate our result in Table 3.6 using a variety of different quality proxies. Further, we use quality measures from both before and after the program. This lets us rule out the possibility that the post-program measures are somehow influenced by the judges who evaluated the startup and thus are biased. Our measures of quality from before the program include a binary measure of whether a startup has financing at the time of the application, a binary measure indicating whether a startup has user traction (as measured by having at least 100 page visits on average per month over the last three months before the program), logged financing value prior to the program, and logged page visits prior to the program.

However, our true focus is not necessarily observable quality at the time of application and is instead whether the startup will be a success, a much harder quantity to measure before the program. Fortunately, once we control for whether the startup is admitted into the program, it seems reasonable to assume judge scores have essentially no influence on future startup performance. Thus, we can use the realized outcomes for the startup in our sample to measure their quality. Specifically, using data from Pitchbook, we collect measures as of 3–4 years after the program on the logged number of employees, revenue growth, and logged valuation. We also construct a quality index by taking the three variables previously mentioned, as well our two post-accelerator quality measures from the main paper (logged page visits and funding after the program), and then standard normalize each variable. We then add these variables together and again normalize to construct our final index.

Table B.13.1 replicates Table 3.6 but uses our four measures of pre-program startup quality. In all cases, we find that judges give higher scores to higher-quality startups, this estimate is the same for local and foreign startups, and judges discount foreign startups. Figures B.13.1 and B.13.2 present graphical evidence for our two continuous measures of startup quality. The binscatters show that judges are biased against foreign startups as the red dotted line is above the blue solid line across the quality distribution of startups. They are also able to detect the quality of local and foreign startups with equal precision, as shown by the similarly positively sloped red dotted and blue solid lines.

Table B.13.2 replicates Table B.13.1 but uses our four post-program realized success measures of startup quality. We find similar patterns as with our pre-program measures. Column 4, which uses our index measure, shows that a one standard deviation in quality leads to a 0.1 standard deviation increase in the score a judge will give to a startup. This suggests that while judges are informed, they are far from oracles.

Table B.13.1: The ability of judges to evaluate startup quality: Pre-accelerator quality measures

	(1)	(2)	(3)	(4)
		Score		
Foreign Startup	-0.045*	-0.038	-0.053**	-0.039
	(0.019)	(0.027)	(0.020)	(0.027)
Has Financing	0.773***			
	(0.029)			
Foreign Startup * Has Financing	-0.097**			
	(0.038)			
Has User Page Traction		0.264***		
		(0.036)		
Foreign Startup * Has User Traction		-0.042		
		(0.036)		
Log Pre-Accelerator Financing			0.154***	
			(0.007)	
Foreign Startup *Log Pre-Accelerator Financing			-0.012	
			(0.009)	
Log Pre-Accelerator Page Visits				0.036***
				(0.006)
Foreign Startup * Log Pre-Accelerator Page Visits				-0.006
				(0.007)
Observations	16,320	16,320	16,320	9,063
Judge x Year	Yes	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes	Yes

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

Table B.13.2: The ability of judges to evaluate startup quality: Post-accelerator quality measures

	(1)	(2)	(3)	(4)
		Score		
Foreign Startup	-0.052*	-0.048*	-0.047*	-0.044*
	(0.020)	(0.020)	(0.020)	(0.020)
Log Post-Accelerator Employees	0.115***			
	(0.014)			
Foreign Startup * Log Post-Accelerator Employees	0.013			
	(0.017)			
Post-Accelerator Revenue Growth		0.027		
		(0.043)		
Foreign Startup * Post-Accelerator Revenue Growth		0.031		
		(0.054)		
Log Post-Accelerator Valuation			0.123***	
			(0.030)	
Foreign Startup * Log Post-Accelerator Valuation			-0.003	
			(0.038)	
Post-Accelerator Performance Index				0.101***
				(0.016)
Foreign Startup * Post-Accelerator Performance Index				0.027
				(0.020)
Accelerator Participation	0.694***	0.768***	0.743***	0.625***
	(0.032)	(0.030)	(0.031)	(0.037)
Foreign Startup * Accelerator Participation	-0.091*	-0.096*	-0.096*	-0.117*
	(0.040)	(0.039)	(0.040)	(0.048)
Observations	16,320	16,320	16,320	16,320
Judge x Year	Yes	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes	Yes

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

Figure B.13.1: Binscatter showing that judges can give higher scores to startups with more growth before entering the accelerator program, but consistently discount foreign startups no matter their growth.

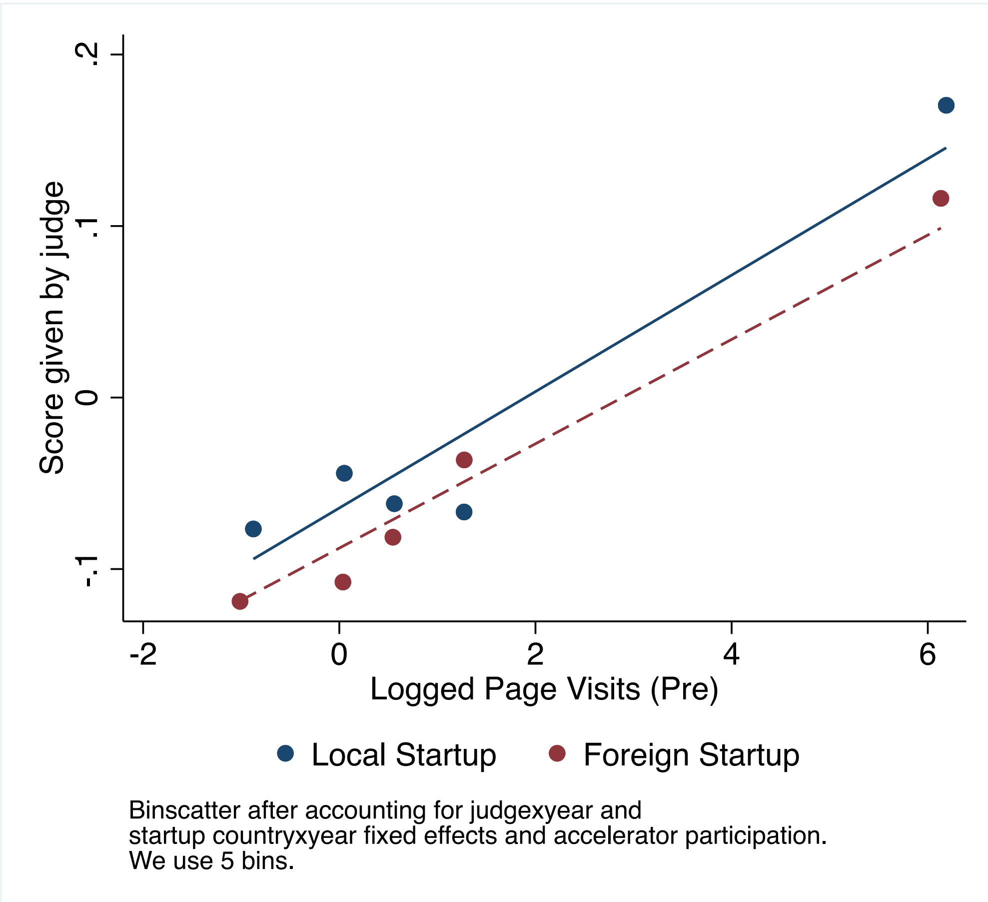
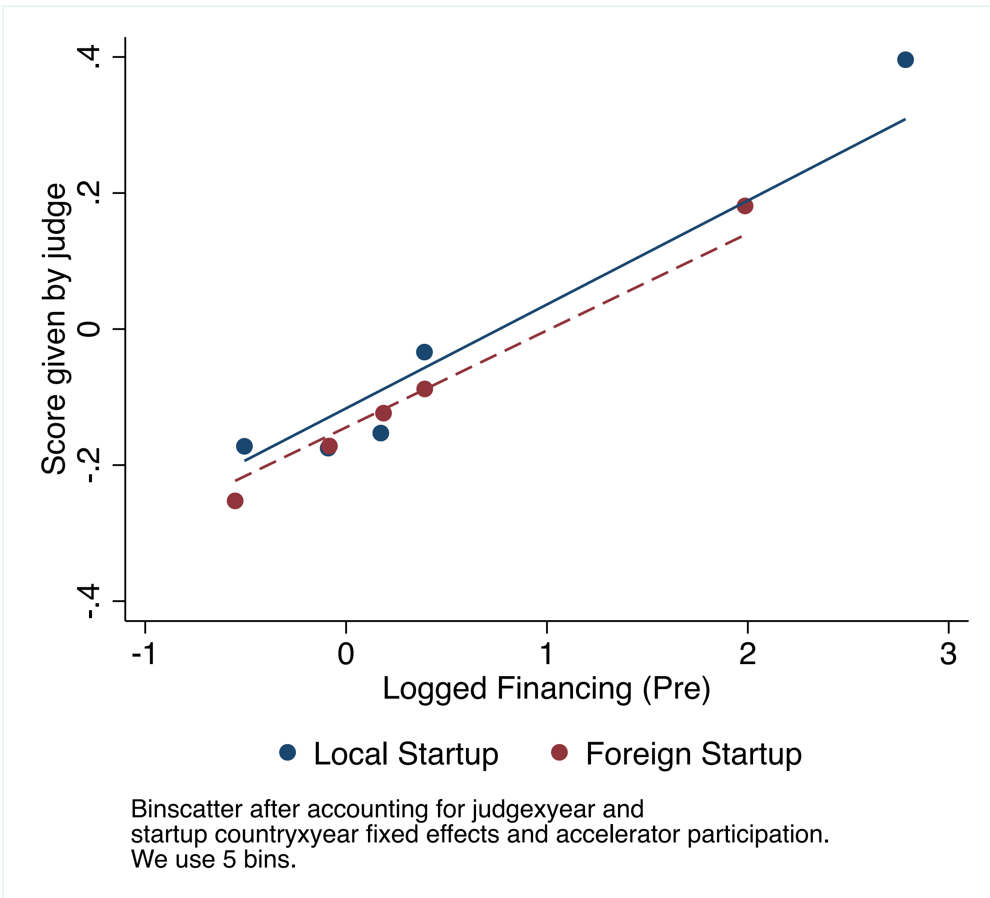


Figure B.13.2: Binscatter showing that judges can give higher scores to startups with more financing before entering the accelerator program, but consistently discount foreign startups no matter their financing.



B.14 Judges are equally responsive to quality when we split our sample by foreignness.

Table 3.6 shows that judges are able to tell winners from losers with equal precision among local and foreign startups. One concern with this estimate is that the extent to which the judge's score accounts for startup quality might vastly differ among foreign vs. local startups. Conversely, our measures of startup quality may account for the judge's score to different extents among local and foreign startups. For example, a European judge may be able to account for much more of a European startup's quality than for a US startup's quality because of their contextual knowledge of the European market. To test this, we compare the R-squared statistics of regressions assessing the relationship between judge scores and startup quality for local versus foreign startups. Specifically, we show the R-squared statistics for models regressing the judge's score on quality (Table B.14.1) and quality on the score (Table B.14.2) are similar when we split the sample by foreignness.

In Table B.14.1, Models 1–2 include no additional fixed effects and show judges give higher quality scores to both local and foreign startups. Moreover, the R-Squared values are nearly identical. In Models 3–4, we include judges-year fixed effects. Again, the within-R2s are nearly identical. Models 5–6 phase in country-year fixed effects. We find similar patterns.

Table B.14.2 mirrors Table B.14.1 but swaps the dependent and independent variables. Since judges are randomly assigned to evaluate startups, there is no need to include judge fixed effects. Further, in this model, judge fixed effects account for outcome differences in the startups the judge evaluates and not differences in judge harshness. That said, the results below are essentially unchanged when including judge fixed effects. Again, we find that judges can crudely predict which startups will succeed and which will not.

At least in our full sample, it does not appear that judges are any better at evaluating local startups than foreign ones. Overall, we find R-squared estimates that, depending on the model, range from 3-to-6 percent.

Table B.14.1: The R-squared values for the extent to which judges are informed are similar when we split our sample into startups foreign to the judge and startups local to the judge.

	(1)	(2)	(3)	(4)	(5)	(6)
	Judge's Total Score					
Log Post-Accelerator Page Visits	0.367***	0.4104***	0.3862***	0.4082***	0.3876***	0.3775***
	(0.0299)	(0.0287)	(0.0267)	(0.0251)	(0.0268)	(0.0255)
Observations	7,232	9,107	7,232	9,107	7,232	9,107
Sample	Local	Foreign	Local	Foreign	Local	Foreign
Judge x Year	No	No	Yes	Yes	Yes	Yes
Startup Country x Year	No	No	No	No	Yes	Yes
R-Squared	0.03279	0.03757	0.44205	0.47552	0.44633	0.49148
Within R-Squared			0.05086	0.05701	0.05087	0.04856

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

Table B.14.2: Regressions showing that judges' scores predict future startup performance.

	(1)	(2)	(3)	(4)
	Log Post-Accelerator Page Visits			
Judge's Total Score	0.0893***	0.0916***	0.0922***	0.082***
	(0.0068)	(0.0065)	(0.0066)	(0.0065)
Observations	7,232	9,107	7,232	9,107
Sample	Local	Foreign	Local	Foreign
Startup Country x Year	No	No	Yes	Yes
R-Squared	0.03279	0.03757	0.05712	0.07963
Within R-Squared			0.03516	0.03016

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

B.15 Judges are worse at detecting the quality of startups that have raised financing, are from hubs, or have founders affiliated with elite universities.

Table 3.6 shows that judges can detect startup quality. This finding contradicts other studies that show that judges often cannot detect the quality of startups, especially startups in consumer and enterprise high-technology sectors (Scott et al., 2020). However, in this past work, judges often evaluate startups from a highly pre-selected pool. For example, the startups in Scott et al. (2020)'s study are all from an elite university (MIT), and so reflect a much higher quality pool than is present in our study, which includes both MIT founders and those with less prestigious educational credentials. Perhaps differences in the variability of the pool of startups being evaluated explain when judges can separate the best startups from the worst.

To test this hypothesis, we restrict our sample to startups that look more like the firms analyzed in Scott et al. (2020). We do so in three ways. First, we measure affiliation with an elite university by whether the application text team portion mentions the name of one of the top 10 elite universities in each of the regions in our sample (Europe, North America, and Latin America), according to QS World University Rankings (2021). Second, we restrict our sample to only startups that had raised funding at the time of their application. Third and finally, we measure if a startup is connected to a hub region like Silicon Valley by whether the application mentions the name of a hub region as defined by the Startup Genome Project (2021).

Table B.15.1 presents regression results with samples split by each of these three measures. We include our primary measure of startup quality, logged post-accelerator page visits, along with our foreign dummy variable. We do not include the interaction term for two reasons. Following Scott et al. (2020), we do not include the interaction as our focus here is on whether judges can detect quality in these sub-samples, not on whether they are more or less informed about local startups.

We find that judges' scores are less responsive to quality differences between startups in the "high quality" sub-samples, as evidenced by the lower magnitude of the coefficient on "log post-accelerator page visits." The coefficient plot in Figure 3.4 reveals that these estimates are statistically different from our baseline estimate. Unlike the pronounced drop we see in the ability of judges to detect quality differences, we find that the foreign discounting effect remains relatively stable across the different sub-samples. While the effect loses significance for startups with founders from elite universities, it is actually larger for both financed and startups connected to hub regions. Though beyond the scope of our paper, this difference could be because judges turn to geography to discern one startup from another when it is otherwise hard to choose among a sample of relatively high-quality startups.

Table B.15.1: Foreign discounting effect across different quality sub-samples of startups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Judge's Total Score						
	Baseline	Founders from Elite Universities	Founders Not from Elite Universities	Financed Startups	Non-Financed Startups	Startups Targeting Hubs	Startups Not Targeting Hubs
Foreign Startup	-0.064**	-0.054	-0.069**	-0.112*	-0.040*	-0.107*	-0.054*
	(0.020)	(0.050)	(0.021)	(0.046)	(0.020)	(0.049)	(0.022)
Log Post-Accelerator Page Visits	0.050***	0.031***	0.055***	0.006	0.042***	0.031***	0.052***
	(0.003)	(0.006)	(0.003)	(0.006)	(0.003)	(0.006)	(0.003)
Observations	16,320	2,899	13,124	1,596	14,323	2,734	13,265
Judge x Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

B.16 Judges are worse at selecting winners from losers among startups in later stages of the competition.

Table 3.6 shows that judges can pick winners from losers among the startups. Can they still do so when they get to later rounds of the competition with a higher quality pool of startups? In Table B.16.1, we apply the same specification as for Table 3.6, but for the latter two rounds of the accelerator. Column 2 shows that judges' ability to detect startup quality declines relative to the baseline results in Table 3.6 for post-accelerator page visits by 0.03 standard deviation (Columns 1–2) and by 0.135 standard deviation for post-accelerator financing (Columns 3–4). These results suggest that judges might be able to screen in early rounds, but as the sample of startups becomes higher quality, their ability declines, consistent with results from Table B.15.1. As a result, they may need to turn to alternative experimentation (e.g., "spray-and-pray") measures—as found in the venture capital context (Ewens et al., 2018)—to detect startup quality.

Table B.16.1: Judges' ability to detect startup quality among startups accepted into the accelerator program declines.

	(1)	(2)	(3)	(4)
	Judge's Total Score			
Log Post-Accelerator Page Visits	0.050***	0.020**		
	(0.003)	(0.006)		
Log Post-Accelerator Financing			0.165***	0.030***
			(0.007)	(0.007)
Observations	16,320	3,898	16,320	9,753
Judge x Year	Yes	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes	Yes
Startup x Year	No	No	No	No
Initial Round	Yes	No	Yes	No

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

B.17 Judges can equally detect the quality of local vs. foreign startups no matter whether startups' regions are mentioned in their applications.

While Table 3.6 shows that judges can detect startup quality of local and foreign startups with similar precision, it may be the case that foreign discounting is absorbing a local information advantage. Thus, judges might be able to detect the quality of both local and foreign startups only if their region is made explicit in their applications, suggesting that the geography of startups is already "priced" into the judgment. In Table B.17.1, we show that judges can detect the quality of startups with even better precision when the region is not explicitly stated in the application of startups. Indeed, the coefficient on logged post-accelerator page visits increases by 0.01 standard deviation when the region is not made explicit in the application (Column 3) relative to when it is explicit (Column 2). Meanwhile, there remains no local information advantage when the region is not made explicit. This suggests that location does not provide an informational value that is already incorporated into the judging. If anything, it seems to worsen judges' sensitivity to the quality of startups, perhaps because it leads them to rely on their location preferences.

Table B.17.1: Judges can equally detect the quality of local vs. foreign startups no matter whether startups' regions are mentioned in their applications.

	(1)	(2)	(3)
	Judge's Total Score		
	Baseline	Region Explicit in Startup App	Region Not Explicit in Startup App
Foreign Startup	-0.065**	-0.082*	-0.054
	(0.025)	(0.036)	(0.034)
Log Post-Accelerator Page Visits	0.050***	0.042***	0.053***
	(0.004)	(0.005)	(0.005)
Foreign Startup * Log Post-Accelerator Page Visits	0.000	0.006	-0.003
	(0.005)	(0.007)	(0.006)
Observations	16,320	6,786	9,395
Judge x Year	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes
Startup x Year	No	No	No
Region Explicit in App	Yes	Yes	No
Region Not Explicit in App	Yes	No	Yes

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

B.18 Judges have a local information advantage when assessing very localized startups.

In Table 3.6, we show that judges detect the quality of foreign and local startups with equal precision. This finding contrasts with past studies (e.g., Coval and Moskowitz, 1999; 2001; Malloy, 2005) that show that judges have a local information advantage. One possible reason for this difference is that the industries and business models (mainly high technology) that represent the majority of our startups are geographically agnostic. They follow standardized models, for example, originating from Silicon Valley, such as software-as-a-service. Indeed, Table B.18.1 shows that the top words appearing in applications of startups across the four regions in our sample are similar. The presence of these common words—such as "market," "sales," and "platform"—suggests that startups may be pursuing increasingly standardized approaches across geographies. In contrast, the companies in older studies finding a local information advantage tend to be more localized in terms of producing non-traded goods, being smaller size, or being located in remote regions (Coval and Moskowitz, 1999; 2001).

To confirm if the difference in sample composition of localized companies may account for the difference in results in our study relative to past studies on local information advantage, we split our sample into startups that include terms in their application text that are more specific to their home region versus not (using the NLP approach described in Appendix B.6.1) and the rest. In Table B.18.2, we show that judges have more of a local information advantage when evaluating startups in such geographically sensitive sectors relative to other startups. However, this localized group of startups is an extremely small share of the sample (less than 5 percent). As a result, our overall result shows the lack of a local information advantage among startups.

Table B.18.1: Top words in application by startup home region

Europe			Israel		
Word	**Count**	**Log Odds Ratio**	**Word**	**Count**	**Log Odds Ratio**
market	2166	0.151	**market**	1008	0.179
business	1414	0.193	**product**	571	0.21
companies	1084	0.0338	**companies**	567	0.206
sales	1027	0.0683	**business**	538	-0.00536
product	1026	-0.0374	**experience**	515	0.312
data	1025	0.0986	**users**	478	0.169
marketing	984	0.0559	**people**	473	0.172
platform	923	-0.0653	**marketing**	456	0.0915
revenue	919	-0.0607	**platform**	452	0.044
people	916	0.0122	based	384	0.337
customers	888	0.0953	**time**	384	0.054
experience	831	-0.0578	**customers**	381	0.0419
time	818	-0.0027	**data**	373	-0.134
food	806	0.776	**social**	373	0.000594
development	805	0.199	**company**	354	0.114
team	783	0.103	**online**	351	0.151
users	782	-0.195	solution	350	0.509
technology	769	0.0173	ip	343	0.203
online	725	0.0679	**technology**	343	0.016
social	715	-0.19	israel	331	4.64

Latin America			North America		
Word	**Count**	**Log Odds Ratio**	**Word**	**Count**	**Log Odds Ratio**
de	1329	3.93	**market**	5777	-0.14
market	1227	-0.044	**business**	3495	-0.236
mexico	1119	4.43	revenue	3455	0.317
companies	1009	0.43	**product**	3439	0.0485
business	979	0.227	**data**	3235	0.182
people	977	0.577	**sales**	3060	-0.0326
platform	772	0.205	**platform**	3015	-0.0729
sales	768	0.201	**users**	2953	0.018
experience	737	0.282	**companies**	2932	-0.315
social	719	0.301	**marketing**	2857	-0.119
marketing	690	0.115	**social**	2627	-0.0383
users	660	0.0901	**time**	2604	-0.011
company	621	0.312	**customers**	2554	-0.0664
product	572	-0.231	**experience**	2536	-0.216
online	569	0.258	**technology**	2491	0.0749
products	549	0.377	**people**	2443	-0.38
time	541	-0.00681	team	2274	-0.0273
customers	538	-0.0168	ip	2178	0.171
usd	530	2.32	**company**	2145	-0.151
services	523	0.481	cost	2133	0.0826

Table B.18.2: As we restrict the sample to more localized startups, we see that judges become better at evaluating startup quality and worse at evaluating foreign startup quality.

	(1)	(2)	(3)
	Judge's Total Score		
Foreign Startup	-0.2835*	-0.2151	-0.3726
	(0.1400)	(0.1998)	(0.2328)
Log Post-Accelerator Page Visits	0.0401*	0.0608*	0.2001**
	(0.0180)	(0.0304)	(0.0576)
Foreign Startup *Log Post-Accelerator Page Visits	0.0158	-0.0645	-0.1432*
	(0.0216)	(0.0455)	(0.0699)
Observations	1,225	485	157
Localness cutoff	0.5 S.D.	0.75 S.D.	1 S.D.
Regional Score Controls	Yes	Yes	Yes
Judge x Year	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes

The first column includes all startups that have a score of 0.5 standard deviations for their home region and are below 0.5 standard deviations for all other regions. The second column raises this cutoff to 0.75. The third column to 1. Standard errors (in parentheses) are clustered at the judge and startup level. Fixed effects shown below observations.
* p<0.05 ** p<0.01 *** p<0.001

B.19 Additional analyses estimating "missed startups."

In the primary manuscript, we discuss how we use rich fixed effects models to estimate that judges reject about 5 percent of promising startups due to foreign bias that, in the absence of such bias, would likely have been selected. These fixed effect models include both startup fixed effects and "foreign discounting" fixed effects for each judge.

Here, we use the startup fixed effects to show that the "missed" startups are not especially low quality but come from the core of the accepted startup quality distribution. To do so, in Figure B.19.1, we plot the kernel density of estimated startup quality, as measured using the startup fixed effects, in blue. The x-axis shows the "quality" in terms of the judge's standardized z-score and so is easily interpretable in terms of standard deviations. The dotted red line shows the estimated quality for the startups missed due to foreign bias. These are the startups we think should have been selected for the next round of the accelerator when judges' scores are "debiased." The figure reveals that missed startups are not just the "bad" startups that would have been rejected regardless. Instead, as shown by the larger overlap between the two distributions, these missed startups are of similar quality to those that make it to the next round. While the very best startups make it to the next round, no matter, this plot points to the idea that judges appear to be discounting promising companies.

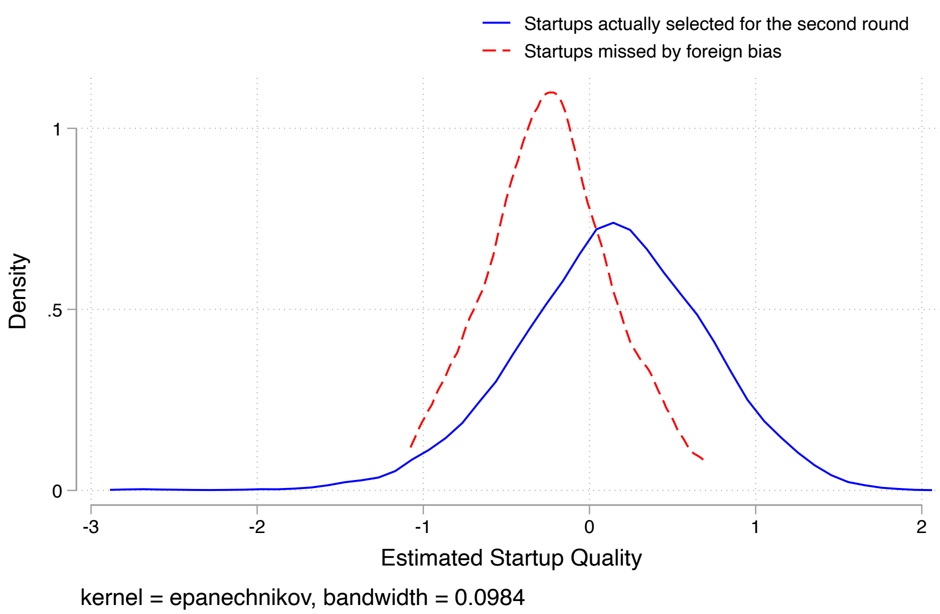
To triangulate our back-of-the-envelope results, we conduct two additional approaches to the calculation estimating missed-out startups. These calculations compare the startups judges would have selected if they only relied on quality-dependent measures and not on the startup's foreign status to the startups they selected when considering foreign status. To isolate the quality-dependent portion of the judges' scores, we regressed judge decisions on our startup quality measures. While crude, this model allows us to recover the judges' weights on different measures of startup quality—both pre-accelerator and post-accelerator—and so construct counterfactual rankings as if judges are unbiased but still selected the same number of startups.[6] We then compare this ranking to two alternatives. The first is the actual recommendation of the judge. The second is the "biased" counterfactual ranking that uses the quality measures and whether the startup is foreign to generate deliberately foreign-biased recommendations. The first alternative tells us how much relying only on quality measures would increase the number of foreign startups. The second reveals how many foreign startups are missed when we introduce foreign bias on top of "unbiased" quality-based evaluations.

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

In these additional back-of-the-envelope counterfactuals, we find that foreign bias impacts the number of foreign startups that are recommended to the next round of the competition. We find that moving to evaluations only based on quality leads to 512 more foreign startups being recommended, accounting for 14 percent of the startups in our sample. When we introduce foreign bias onto the quality-based recommendations, 324 fewer foreign startups are recommended. When we use only the criteria-based score, 312 fewer startups are recommended. These differences suggest that foreign bias leads judges to overlook 9-to-14 percent of startups that, at least based on our quality measures, should have been recommended to the next round.

These calculations show that a higher share of startups—9-to-14 percent—are passed over, suggesting our main estimate of 5 percent is conservative.

Figure B.19.1: Distribution of startups actually selected versus estimated to be missed with foreign discounting.



B.20 Application questions that startups answer and judges evaluate in the accelerator program

We show the application questions that startups fill out and judges evaluate in the venture competition below.

Company Background

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

Customer Pain and Solution

- Problem – please describe what problem (customer pain point) you are trying to solve.
- Solution – what is your solution?

Overall Impact

- 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

- How would you define your potential market, and what is the addressable market size?
- What traction have you made to date with market validation?
- Marketing – what will be your messaging to users/customers, and how do you plan to spread it?
- Sales and distribution – how will you reach your customers?

Industry and Competitors

- Which organizations compete with your value offering now, and who might do so in the future?
- Which organizations complement your offering in the market?
- What are the primary advantages relative to existing or potential competitors?

Business Model/Financials

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

Regulation and IP

- What intellectual property or regulatory requirements exist for your business or in your industry?

Founding Team and Advisors/Investors

- Please share some background information on your team members.

- Please tell us about current or anticipated advisors and investors.

B.21 Robustness when including elite university and hub controls

Table 3.6 shows that judges can equally detect startups that are foreign and local to them. While the regressions control for page visits, financing, and headquarters country as proxies of startup quality, we might be concerned that other measures of startup quality, like whether executives affiliate with an elite university or whether the startup is headquartered in a hub, might confound this result. To address this concern, we show in Table B.21.1 the same results as Table 3.6, but with controls for whether a startup executive is affiliated with an elite university and whether the startup is headquartered in a hub. The results are similar. Judges are no better at detecting the quality of local startups than they are of foreign ones.

Table B.21.1: Regressions showing judges (1) give higher scores to more successful startups, (2) are equally good at evaluating success for local and foreign startups alike, and (3) still discount foreign startups when controlling for elite university affiliation and hub location

	(1)	(2)	(3)	(4)	(5)	(6)
	Judge's Total Score					
Foreign Startup	-0.068**	-0.058*	-0.044	-0.052**	-0.046*	-0.040*
	(0.024)	(0.024)	(0.025)	(0.019)	(0.019)	(0.020)
Log Post Accelerator Page Visits	0.049***	0.036***	0.042***			
	(0.004)	(0.004)	(0.004)			
Foreign Startup * Log Post Accelerator Page Visits	0.002	0.004	0.000			
	(0.005)	(0.005)	(0.005)			
Log Post Accelerator Financing				0.163***	0.024*	0.158***
				(0.008)	(0.012)	(0.038)
Foreign Startup *Log Post Accelerator Financing				-0.015	0.008	-0.022
				(0.011)	(0.015)	(0.060)
Accelerator Participation		0.654***			0.681***	
		(0.032)			(0.041)	
Foreign Startup * Accelerator Participation		-0.124**			-0.119*	
		(0.041)			(0.054)	
Elite University	0.297***	0.252***	0.266***	0.260***	0.242***	0.257***
	(0.023)	(0.022)	(0.025)	(0.023)	(0.022)	(0.025)
Hub	0.043	0.036	0.056*	0.046	0.038	0.058*
	(0.024)	(0.023)	(0.026)	(0.024)	(0.024)	(0.026)
Observations	16,321	16,321	14,476	16,321	16,321	14,476
Judge x Year	Yes	Yes	Yes	Yes	Yes	Yes
Startup Country x Year	Yes	Yes	Yes	Yes	Yes	Yes
Startup x Year	No	No	No	No	No	No
Accelerator Participation	Yes	Yes	No	Yes	Yes	No

Standard errors (in parentheses) are clustered at the judge and startup level. Fixed effects shown below observations. The table includes one additional data point (0.01% of the total sample) due to labeling changes.
* p<0.05 ** p<0.01 *** p<0.001

C

Chapter 4 Supplementary Material

Contents

C.1 Example of OSS policies 261

C.2 Impact of OSS on new venture founding: First stage estimates 262

C.3 Correlations 263

C.4 Impact of OSS on hardware vs. software ventures 264

C.5 Predicted increase in number of ventures with 1% increase in 2016 GitHub commits 265

C.1 Example of OSS policies

Table C.1.1: Example of OSS policies

Country	Type	Data	Details
Argentina	Advisory	2004	The two institutions, which coordinate IT policy and implementation, announced that they promote Linux in all applications in public administration. The rationale for this decision is lower costs, creating local employment, and security.
Australia	Advisory	2004	The Australian Tax Office will consider OSS alongside proprietary solutions.
Bahrain	Preferred	2006	Bahrain's Ministry of Social Development (MOSD) is to become the first ministry in the Kingdom and in the Middle East to base its entire IT infrastructure on open-source technology. Reasons for migration include lower cost, simplified IT management, the ability for modular scalability, and improved security and space efficiency.
Brazil	Preference	2004	Twenty percent of all computers used by the Brazilian ministries are running Linux and other open-source software. In a few months this number should grow to 100 percent. Through its Digital Inclusion Program, Brazil wishes to democratize the use of computers.
Cambodia	Preference	2001	All laws, regulations and policies in the IT sector will reflect the following guiding spirit and philosophy: to uphold the interests of the consumers and general public, to guarantee security of information, while facilitating the broadest possible access to public information to respect individual rights, and to avoid dependency on proprietary systems, instead promoting open systems and interoperability. This is done to avoid dependency of proprietary systems, help reduce poverty, and efficiently develop human resources.
Cuba	Preference	2007	The Cuban government is migrating thousands of its computers to Linux to counter Microsoft and the US.
France	Advisory	2008	The French Ministry of Education is increasing the number of open-source software licenses to France's educational institutions to offer more choices and make users less dependent on software vendors.
Japan	R&D	2007	The central government of Japan says to make Linux and open source a priority for all IT procurements, starting this July. The central government of Japan says it plans to spend around $1.25 trillion yen, or $10.4 billion, on IT over the next year. The government has said explicitly it wants to decrease its reliance on Microsoft as a server operating system platform.
South Africa	R&D	2007	The South African Cabinet today announced that it had approved a free and open-source strategy and that government would migrate its current software to free and open-source software...This strategy will, among other things, lower administration costs and enhance local IT skills.
Russia	Mandatory	2007	The Russian government plans to reduce its dependence on foreign commercial software by installing domestically-developed GNU/Linux open-source software on all of its schools' computers by the end of 2009.

CSIS (2010) categorizes these policies as either advisory, research and development (R&D), preference, or mandatory. The most stringent are mandatory policies, which require the purchase and/or use of OSS. Preference policies encourage the purchase and/or use of OSS, but do not mandate it. Advisory policies permit the purchase and/or use of OSS.
Source: CSIS (2010)

C.2 Impact of OSS on new venture founding: First stage estimates

Table C.2.1: Impact of OSS on new venture founding: First stage estimates

	Instruments Individually	Non-Policy Instruments Together	All Instruments (2000-09)
	Lagged Log GitHub	Lagged Log GitHub	Lagged Log GitHub
	(1)	(2)	(3)
Lagged Below Median Econ. Growth X Above Median Human Capital Instrum.	1.236***	0.746**	0.335
	(0.200)	(0.281)	(0.457)
	F = 244.76		
Lagged Below Median Econ. Growth X Above Median Digital Skills Instrum.	0.797***	0.0951	0.124
	(0.165)	(0.217)	(0.281)
	F = 241.11		
Lagged Below Median Econ. Growth X Above Median Internet Users Instrum.	1.323***	0.786***	1.428***
	(0.194)	(0.234)	(0.403)
	F=248.51		
Lagged Below Median Econ. Growth X OSS Policy Instrum. (Before 2010)	2.312***		0.293
	(0.337)		(0.391)
	F=73.66		
Lagged OSS Policy Instrum. (Before 2010)	1.632***		1.413***
	(0.387)		(0.408)
	F=43.57		
N (Country x Year)		2741	1526
N (Country)		180	180
F		238.31	75.99
Controls	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes

The table presents the first stage results when regressing instruments on lagged log GitHub commits. Column 1 presents the coefficients when regressing each instrument individually. Column 2 presents the coefficients when regressing all non-policy instruments together, spanning the entire length of the data (2000–2016). Column 3 presents the coefficients when regressing all instruments together, spanning a subset of years of the data (2000–2009). All columns include robust standard errors, clustered by country. Time fixed effects are relative to the year 2000. The regressions are not perfectly balanced by year, due to missing data in the control variable datasets. These first-stage estimates correspond to the second stage estimates in Tables 4.2, 4.3–4.5, 4.6, 4.7, and C.4.1. Robust standard errors, clustered by country. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C.3 Correlations

Table C.3.1: Correlations

	Log GitHub Commits	Log IT Ventures	Log OSS Ventures	Log Population	Log GDP Capita	Log Internet Users	Human Capital Index	Log Financing Deals	Log Venture Value	Log Number of Acquisitions	Log Number of Global IT Ventures	Log Number of Mission IT Ventures	Log Number of Mission IT Ventures	Log Number of Mission OSS Ventures	Below Median Econ. Growth X Above Median Human Capital Instrum.	Below Median Econ. Growth X Above Median Digital Skills Instrum.	Below Median Econ. Growth X Above Median Internet Users Instrum.	Below Median Econ. Growth X OSS Policy Instrum. (Before 2010)	OSS Policy Instrum. (Before 2010)
Log GitHub Commits	1.00																		
Log IT Ventures	0.68	1.00																	
Log OSS Ventures	0.34	0.57	1.00																
Log Population	0.35	0.35	0.29	1.00															
Log GDP Capita	0.49	0.62	0.30	-0.20	1.00														
Log Internet Users	0.67	0.56	0.23	-0.15	0.84	1.00													
Human Capital Index	0.29	0.51	0.24	-0.17	0.77	0.64	1.00												
Log Financing Deals	0.70	0.86	0.59	0.42	0.56	0.52	0.39	1.00											
Log Venture Value	0.68	0.83	0.54	0.37	0.58	0.51	0.41	0.96	1.00										
Log Number of Acquisitions	0.51	0.73	0.67	0.31	0.49	0.37	0.33	0.82	0.78	1.00									
Log Number of Global IT Ventures	0.68	1.00	0.57	0.35	0.62	0.56	0.51	0.86	0.83	0.73	1.00								
Log Number of Global OSS Ventures	0.34	0.57	1.00	0.29	0.30	0.23	0.24	0.59	0.54	0.67	0.57	1.00							
Log Number of Mission IT Ventures	0.42	0.67	0.69	0.34	0.30	0.26	0.22	0.66	0.62	0.72	0.67	0.69	1.00						
Log Number of Mission OSS Ventures	0.09	0.18	0.36	0.11	0.07	0.05	0.06	0.19	0.17	0.22	0.18	0.36	0.30	1.00					
Below Median Econ. Growth X Above Median Human Capital Instrum.	0.37	0.51	0.30	-0.05	0.59	0.50	0.59	0.44	0.43	0.40	0.51	0.30	0.26	0.09	1.00				
Below Median Econ. Growth X Above Median Digital Skills Instrum.	0.33	0.38	0.24	-0.11	0.53	0.45	0.42	0.36	0.36	0.36	0.38	0.24	0.21	0.05	0.68	1.00			
Below Median Econ. Growth X Above Median Internet Users Instrum.	0.45	0.40	0.23	-0.10	0.57	0.60	0.44	0.38	0.37	0.32	0.40	0.23	0.21	0.05	0.73	0.71	1.00		
Below Median Econ. Growth X OSS Policy Instrum. (Before 2010)	0.47	0.56	0.36	0.18	0.51	0.43	0.42	0.55	0.53	0.51	0.56	0.36	0.33	0.09	0.65	0.54	0.62	1.00	
OSS Policy Instrum. (Before 2010)	0.49	0.59	0.30	0.31	0.44	0.39	0.39	0.59	0.56	0.46	0.59	0.30	0.34	0.08	0.36	0.26	0.28	0.72	1.00

The table presents correlations between the main dependent, independent, control, and instrumental variables used in subsequent regressions.

C.4 Impact of OSS on hardware vs. software ventures

Table C.4.1: Impact of OSS on hardware vs. software ventures

	OLS	2SLS (Non-Policy Instruments)	2SLS-All Instruments (Pre-2010)	OLS	2SLS (Non-Policy Instruments)	2SLS-All Instruments (Pre-2010)	OLS	2SLS (Non-Policy Instruments)	2SLS-All Instruments (Pre-2010)
	Log IT Ventures	Log IT Ventures	Log IT Ventures	Log Hardware Ventures	Log Hardware Ventures	Log Hardware Ventures	Log Software Ventures	Log Software Ventures	Log Software Ventures
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lagged Log GitHub	0.217***	0.391***	0.381***	0.0778**	0.269*	0.248***	0.177***	0.456***	0.406***
	(0.0256)	(0.0672)	(0.0474)	(0.0231)	(0.103)	(0.0670)	(0.0357)	(0.108)	(0.0936)
Lagged Log Population	0.252***	0.0928	0.131**	0.241***	0.0381	0.00474	0.361***	0.0644	0.0737
	(0.0395)	(0.0716)	(0.0463)	(0.0523)	(0.120)	(0.0982)	(0.0627)	(0.125)	(0.130)
Lagged Human Capital Index	-0.101	-0.359	-0.482	0.0298	-1.665	-2.440*	0.504	-1.972	-3.106*
	(0.282)	(0.309)	(0.271)	(0.606)	(1.129)	(1.070)	(1.250)	(1.586)	(1.544)
Lagged Log GDP Capita	0.311***	0.221**	0.0659	0.303***	0.213*	0.0688	0.407**	0.276*	0.0851
	(0.0743)	(0.0723)	(0.0580)	(0.0775)	(0.0854)	(0.0940)	(0.122)	(0.129)	(0.142)
Lagged Log Internet Users	0.0892	-0.116	0.114	-0.0728	-0.230	-0.0441	-0.0269	-0.257	0.0410
	(0.0756)	(0.108)	(0.0885)	(0.0761)	(0.121)	(0.122)	(0.116)	(0.167)	(0.199)
_cons	-7.477***	-2.658	-2.028*	-6.414***	-1.017	1.203	-8.747***	-0.862	1.242
	(0.819)	(1.463)	(0.939)	(1.300)	(3.041)	(2.652)	(1.595)	(3.181)	(3.401)
N (Country x Year)	2747	2741	1526	1274	1274	694	1274	1274	694
N (Country)	182	180	180	121	121	106	121	121	106
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table presents estimates regressing log IT ventures, hardware ventures, and software ventures on lagged log GitHub commits. Columns 1, 4, and 7 present OLS results. Columns 2, 5, and 8 present 2SLS results with non-policy instruments for log GitHub commits spanning the full length of data, 2000–2016. Columns 3, 6, and 9 present 2SLS results with all instruments for logged GitHub commits that span a subset of years in the data, 2000–2009. All columns include robust standard errors, clustered by country. Time fixed effects are relative to the year 2000. The regressions are not perfectly balanced by year, due to missing data in the control variable datasets. First-stage estimates corresponding to the 2SLS specifications are shown in Table C.2.1 in the appendix. Robust standard errors, clustered by country. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C.5 Predicted increase in number of ventures with 1% increase in 2016 GitHub commits

Table C.5.1: Predicted increase in number of ventures with 1% increase in 2016 GitHub commits

Country	Income Level	Predicted Increase in Ventures in 2016
United States	High income	132.805
United Kingdom	High income	24.852
India	Lower middle income	23.938
Germany	High income	12.837
Brazil	Upper middle income	10.581
France	High income	8.572
Canada	High income	7.498
Australia	High income	6.927
China	Upper middle income	5.422
Spain	High income	4.516
Netherlands	High income	4.426
Japan	High income	3.800
Singapore	High income	2.827
Italy	High income	2.681
Sweden	High income	2.206
Finland	High income	2.072
Indonesia	Lower middle income	1.997
Belgium	High income	1.964
New Zealand	High income	1.784
Poland	High income	1.769
Ireland	High income	1.766
Korea, Rep.	High income	1.422
Denmark	High income	1.413
Norway	High income	1.408
Austria	High income	1.242
Mexico	Upper middle income	1.120
Portugal	High income	1.066
Georgia	Lower middle income	0.909
United Arab Emirates	High income	0.900
Malaysia	Upper middle income	0.879
Israel	High income	0.806
Estonia	High income	0.737
Czech Republic	High income	0.711
Switzerland	High income	0.568
Hungary	High income	0.556
Romania	Upper middle income	0.540
Vietnam	Lower middle income	0.529
Philippines	Lower middle income	0.490
Colombia	Upper middle income	0.481
Greece	High income	0.431
Thailand	Upper middle income	0.400
Chile	High income	0.390
Argentina	High income	0.382
Russian Federation	Upper middle income	0.356
Pakistan	Lower middle income	0.338
Bulgaria	Upper middle income	0.316
Nigeria	Lower middle income	0.305
Bangladesh	Lower middle income	0.267
Uruguay	High income	0.230
Turkey	Upper middle income	0.210
Lithuania	High income	0.201
South Africa	Upper middle income	0.192
Saudi Arabia	High income	0.191
Luxembourg	High income	0.172
Cyprus	High income	0.168
Slovenia	High income	0.147
Sri Lanka	Lower middle income	0.137
Slovak Republic	High income	0.112
Peru	Upper middle income	0.110
Latvia	High income	0.107
Ukraine	Lower middle income	0.105
Kenya	Lower middle income	0.088
Malta	High income	0.072
Nepal	Low income	0.067
Belarus	Upper middle income	0.067
Iceland	High income	0.066
Lebanon	Upper middle income	0.061
Ecuador	Upper middle income	0.058
Ghana	Lower middle income	0.058
Iran, Islamic Rep.	Upper middle income	0.049
Egypt, Arab Rep.	Lower middle income	0.046
Panama	High income	0.044
Croatia	High income	0.042
Costa Rica	Upper middle income	0.039
Myanmar	Lower middle income	0.035
Qatar	High income	0.032
Guatemala	Upper middle income	0.032
Uganda	Low income	0.023
Kazakhstan	Upper middle income	0.022
Cameroon	Lower middle income	0.021
Morocco	Lower middle income	0.019
Kuwait	High income	0.018
Tunisia	Lower middle income	0.016
Cambodia	Lower middle income	0.015
Paraguay	Upper middle income	0.015
Honduras	Lower middle income	0.014
Ethiopia	Low income	0.014
Oman	High income	0.014
Jordan	Upper middle income	0.013
Armenia	Upper middle income	0.013
Albania	Upper middle income	0.011
Zambia	Lower middle income	0.010
Bahrain	High income	0.009
Senegal	Low income	0.008
Kyrgyz Republic	Lower middle income	0.008
Mongolia	Lower middle income	0.007
Belize	Upper middle income	0.002

The table shows countries for all available and non-zero 2016 GitHub commits. We construct these estimates by calculating the estimated percent change in new ventures for each country in 2016 using the full specification OLS model from equation 1, which we then multiply by the average number of ventures in 2016 in each country. The predicted values differ within income groups because of variations in 2016 GitHub commits, number of IT ventures, internet users, human capital, population, and GDP per capita.