NBER WORKING PAPER SERIES

THE CONTRIBUTION OF FOREIGN MASTER'S STUDENTS TO US START-UPS

Michel Beine
Giovanni Peri
Morgan Raux

Working Paper 33314
http://www.nber.org/papers/w33314

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2024

We are grateful to participants to the First European Conference on Macroeconomics and of the Immigration OECD Migration conference, and to Seminars participants in UNAM (Mexico City), the University of Bari and the University of Ghent for their useful comments. We thank Alessandro Caiumi for his help in editing the paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w33314

The Contribution of Foreign Master's Students to US Start-Ups
Michel Beine, Giovanni Peri, and Morgan Raux
NBER Working Paper No. 33314
December 2024
JEL No. F22, M13

ABSTRACT

In this paper, we estimate the effect of increasing the share of foreign-born Master graduates on the creation of innovative start-ups in the US. We combine information on international students graduating from Master's programs by university cohort with data on start-ups created in the US between 1999 and 2020 by graduates of those cohorts. To establish a causal link, we use idiosyncratic variation in out-of-state relative to in-state fees charged by universities across Master's cohorts, resulting in differential foreign students' enrollment. We also use changes in the share of foreign students predicted by a shift-share instrument, based on university-level past networks, as an additional identification strategy. For each additional ten percentage points of foreign students graduating in a Master's cohort, we find 0.4 additional start-ups in that cohort. Then, using a name-based attribution of the origin of creators of start-ups, we find that between 30 and 45% of the total start-up creation effect is attributable to a positive spillover of foreign-born on start-up founders of US origin.

Michel Beine
University of Luxembourg
and IZA and CES-Ifo
michel.beine@uni.lu

Morgan Raux
Aix-Marseille School of Economics
morgan.raux@univ-amu.fr

Giovanni Peri
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
gperi@ucdavis.edu

1 Introduction

The creation of innovative firms, known as “start-ups", is a fundamental dimension of the economic and business dynamism of a country. Start-ups are key contributors to employment, real output, and productivity growth (Haltiwanger et al., 2016). They are often more likely to bring high-potential and substantial innovations to the market (Kolev et al., 2022). They hire high-ability workers (Kim, 2018) and frequently pay higher wages for comparable workers, relative to other firms (Babina et al., 2019). The US is consistently ranked among the top countries in the world for entrepreneurial intensity and ease of starting a business.[1] Therefore, analyzing the founders of start-up businesses and their characteristics contributes to our understanding of potential sources of entrepreneurial advantage in the US.

A large share of innovative US entrepreneurs hold a graduate degree from U.S. Universities, many of whom are world leaders in technology, science, and business.[2][3] These institutions are fundamental generators of human capital and innovation, as their graduates translate their talents into creativity, innovation, and new firms generating value-added, productivity growth, and jobs for the US economy. In addition, US universities regularly attract highly talented people from all over the world, especially in many of their graduate programs, channeling them into leadership, managerial, and entrepreneurial roles. In this paper, we quantify the contribution of foreign students with a US Master's degree to start-up creation in the US.[4] We estimate how a higher share of foreign students among master graduates increases the rate at which that group creates start-ups within 5 years of graduation.

[1] See for instance https://www.usnews.com/news/best-countries/rankings/entrepreneurial and https://subnational.doingbusiness.org/en/data/exploretopics/starting-a-business/score

[2] A Forbes article in 2019, entitled "New Fortune 100 CEO Study: The Top Graduate Schools Attended By Fortune 100 CEOs" (available at https://www.forbes.com/sites/kimberlywhitler/2019/09/28/new-fortune-100-ceo-study-the-top-graduate-schools-attended-by-fortune-100-ceos/?sh=6246c2962719), showed that 54% of start-up founders had a graduate degree, and most of them (80%) had a Master degree. Similarly Freeman et al. (2008) found that about 30% of start-up founders are Master's Graduates and 10% of them have a Ph.D.

[3] See one of the most recent rankings (available at https://www.usnews.com/education/best-global-universities/rankings), where US universities take the lion's share of the top 10, 50 and 100 universities in the world. Furthermore, the U.S. has been the country with the largest number of foreign students since the 1950s.

[4] As we explain later, while a large share of start-up founders has a Bachelor's degree as their terminal degree, we focus on Master graduates because the shorter duration of the program, the significance of tuition costs in affecting students' choices, and the large share of foreign students in those programs make them a better sample for our identification strategy and to test our main hypothesis.

Some existing studies suggest that foreigners, in general, and international students in US graduate programs, in particular, are overrepresented among entrepreneurs in the US (Azoulay et al., 2022) and among the founders of US billion-dollar start-ups (Anderson, 2022). Creating innovative companies is an important channel through which immigrants contribute to US innovation and economic growth. However, the academic literature has focused mainly on the contributions of highly skilled immigrants to science (Borjas and Doran, 2012), innovation (Kerr and Kerr, 2016; Bernstein et al., 2022) or the supply of human capital (Peri et al., 2015). In this paper, we focus instead on the impact of high-skilled immigration, specifically foreign Master's students, on start-up entrepreneurship in the US.

To get to the causal relationship between the inflow of foreign Master's students and the creation of new start-ups, we combine different data sources and develop an original identification strategy based on instrumental variables estimation. First, using administrative data from the Integrated Post-secondary Education Data System (IPEDS), we measure the number of foreign-born graduates in the Master's programs of each public and private university by year of graduation from 1999 to 2015[5]. We know from our previous research (Beine et al., 2023) that only about 20% of foreign-born graduates were employed in the US in their first job after graduation. In this paper, we focus on the effect of foreign-born students who graduated with a Master's degree from a US university on start-up creation. The relatively high and variable fees they have to pay for the Master's program, the short duration of their program (1-2 years), and the fact that Master's graduates are the most likely to create start-ups (Freeman et al., 2008) make this group the most appropriate sample for our identification strategy. The unit of analysis is the graduating cohort of the Master programs in each school for each year.

Second, we merge these data with information on US start-ups, which brought to market a new product or service between 1999 and 2020. These data were obtained from Crunchbase.com, the largest crowd-sourcing platform and the most comprehensive source for US start-ups (as discussed in Dimmock et al., 2022). The data, which we obtained directly from the platform, include the name, the University, and year of graduation of the founder(s) of each start-up. They also include the year the firm was founded, its location and some of the economic characteristics of the firm. This information allows us to construct the count of firms founded within 5 years after

[5]We call (interchangeably) foreign-born, foreigners or international students those who attend a Master degree on an F- or J-visa. Although technically there may be other foreign-born students in US masters, including permanent residents or those with irregular status, the preponderance of the visa-holding group and our identification based on fee variation suggest that this is the most relevant group for our analysis

graduation by Master's graduates for each university and cohort of graduation. The analysis therefore covers the cohorts who graduated during the period 1999-2015.

Then, still using IPEDS, we construct the average out-of-state and in-state tuition for Master's students for each school and year. The difference between out-of-state and in-state tuition at the time of enrollment, after controlling for school fixed effects and school characteristics, is used as an instrumental variable. This measure captures idiosyncratic changes specific to the out-of-state tuition faced by different cohorts in a university, which turn out to be a significant determinant of the number and share of foreign students in the entry Master's cohort. A similar approach was used in Beine et al. (2023) to estimate the share of foreign-born students in US colleges working in the US after graduation. In this paper, we specialize the IV to Master's programs only and we subject this approach to additional testing and validity checks.

Additionally, we assess the robustness of our estimates to an alternative and widely used instrumental variable strategy that leverages a different source of variation in the number (and share) of foreign Master's students across school cohorts. This is the shift-share IV, developed in the immigration literature by a series of papers, including Card (2009) and Peri et al. (2015), whose characteristics and validity checks were formalized by Goldsmith-Pinkham et al. (2020). In particular, using data on F-1 (student) visas, we measure the share of students from ten main countries of origin in each school as of 2001, and we use them to distribute the flows of F1-visa students from each of those countries in the 2004-2015 period across schools. The variation in foreign Master's students generated by this IV hinges on the change in flow size for students from each nationality of origin interacted with the initial share (network) of those foreigners in each university.

In our main specifications, we regress the number of start-ups created within 5 years of graduation by Master's graduates from a university-cohort cell on the share (or the log) of the number of international graduates in that same university-cohort cell, instrumented by the idiosyncratic difference between out-of-state and in-state tuition, i.e., the part unexplained by a set of time-varying proxies of universities' quality, or by the shift-share IV. The main results from the IV estimates show that a rise of ten percentage points in the foreign-born share of graduates in a cohort, corresponding to one standard deviation of that measure, results in an increase in the number of start-up founders in that cohort by 0.4 to 0.6. This represents a significant increase, equal to approximately the standard deviation of start-up creators per cohort in public universities.[6] This effect is significant and somewhat larger when

[6]The average size of a public university cohort of Master's graduates in our tuition IV sample is equal to 658 students. The average number of start-up founders in a (public university) cohort is 0.09 and the standard deviation is 0.42.

estimated using the tuition IV or the shift-share IV.

When subject to a series of validity checks, the tuition-based IV passes them more convincingly than the shift-share IV, leading us to consider the estimates obtained using the former as preferred. However, similar findings across the two identification strategies, which span two different samples and different margins of variation, adds to our confidence in the robustness of our quantitative findings and their external validity.

The positive effect of the foreign-born share on start-up creation remains strong and significant when we select companies with increasingly successful features. In particular, they remain economically and statistically significant as we consider start-up companies that survived at least 3 years (which represent most of the start-ups in our sample), those that raised more than $25M capital in the first 3 years, and those that submitted at least one patent application during the same time window (a very selected group that includes only 3% of our start-up sample). These indicators signal survival, the ability to raise capital, and the capacity to generate innovation, suggesting that our results apply to the creation of successful companies. Foreign Master's graduates are associated not only with more start-ups but with more start-ups of high quality.

We then decompose this causal effect into a part driven by foreign-origin and a part driven by US-origin start-up creators. As the data from Crunchbase do not identify the place of birth of start-up creators, their origin is inferred from their first and last names based on an algorithm developed by Forebears.io and applicable to about 30% of the Crunchbase sample. This decomposition assesses whether in master-cohorts with larger share of foreigners, people with likely US-origin (US-likely names) are also more likely to create start-ups, suggesting the existence of entrepreneurial spillovers within a Master cohort. As some of the start-up creators' names attributed to foreign-origin can be second-third generation immigrants, who are US-born, the estimated effect on US-origin names is likely a lower bound of the strength of the pure spillovers from foreign students to US-born start-up creators.

Based on the inferred US and foreign-origin status of entrepreneurs, we find that 30-45% of the effect derives from start-up created by likely US-origin individual, hence most likely a spillover effect while 55-70 % derives from start-ups created by likely foreign-origin graduates (some of which are foreign-born while others may capture additional spillovers on second and third generation immigrants). We also find that the increased probability of creating a start-up for a likely US-origin individual is associated with a higher rate of co-founding with someone from the same university cohort. On top of starting more companies, foreign students seem to encourage US-origin schoolmates to found or co-found more start-ups as well.

The rest of the paper proceeds as follows. Section 2 frames the contribution of the present paper within an extensive literature on the entrepreneurship of immigrants and the economic effect of high-skilled foreign workers. Section 3 describes the data, sources and variable construction. Section 4.1 introduces the empirical approach to the identification and estimation, while providing tests of the power and validity of the instruments. Section 6 presents the main estimates and their extensions, while section 7 is devoted to the estimated strength of spillovers on start-ups created by US-origin graduates. Section 8 provides concluding remarks and discussion.

2 Literature Review

This paper contributes to two complementary areas of research related to immigration, entrepreneurship, and innovation. We first contribute to the literature on the role of foreign-born entrepreneurs. Several studies have pointed out that immigrants have a large propensity to be entrepreneurs and to start companies. This is true not only in the US but also in other countries. Fairlie (2008) and Fairlie and Lofstrom (2015) document the contribution of foreign-born entrepreneurs to the US economy and emphasize the higher rate of entrepreneurship of foreigners relative to natives, as well as their positive impact on the economy. Hart (2011), Hart and Acs (2011), Kerr (2013), and Kerr and Kerr (2020) emphasize the role of foreigners in producing innovative start-ups in the high-tech sectors and in contributing to US patenting and innovation. Azoulay et al. (2022) document the job creation power of immigrants as entrepreneurs, arguing that an inflow of immigrants results in a net increase in labor demand for the US economy due to their role in starting firms. Furthermore, some studies (e.g. Kerr and Kerr, 2016) argue that firms founded by immigrants have survival probability and growth rates higher than those created by natives.[7]

Related to this strand, we also contribute to the literature analyzing spillover effects in entrepreneurship by identifying a "spillover" effect of foreign students on entrepreneurship of US-origin students in the same Master's cohort. Fairlie (2008) estimates a small negative impact of foreign entrepreneurs on the native propensity to become entrepreneurs, analyzing aggregate data from the 1980s. Anelli et al. (2023), instead, show large positive entrepreneurial spillovers from migrants analyzing the decline in firm creation due to a large wave of emigration of potential entrepreneurs

[7]Additional related studies – Kerr and Kerr (2021), Kerr and Mandorff (2023) – focus on the role of co-ethnic networks in promoting foreign-born entrepreneurship and on co-ethnic hiring of foreign-founded firms.

in Italy over the period 2008-2015.

In general, the consensus emerging from this literature is that foreign-born individuals have high entrepreneurship rates and are founders of high-quality and highly innovative firms. Our paper expands this literature by focusing on the role of foreign students and US start-ups who are an important component of high-skilled immigration in developed countries. New to this literature, we use the changes in out-of-state relative to in-state tuition fees as an exogenous factor affecting the number of foreign Master's students which in turn allows us to quantify their contributions to start-ups within 5 years of graduation. Most of the papers on immigrant entrepreneurship are rather correlational and do not use a strong identification strategy. Our paper advances this literature by estimating a causal link between the change in the number of foreign students in a cohort and the increase in the number of start-ups created by that cohort.[8]

The second area of the literature that provides background for this paper relates to the impact of foreign skilled individuals (including students) on the research performance of universities and academic outcomes in the US. The influential paper by Kerr and Lincoln (2010) established the strong role of changes in H-1B visas in fostering "ethnic innovation" by Indian, Chinese, and other foreign-origin groups in the US. That paper is also one of the first to use first and last names of inventors to identify their origin-nationality, a method that we refine in our analysis to identify foreign- and US-origin entrepreneurs. Recent papers (e.g. Bernstein et al., 2022) have shown the important role of high-skilled foreign immigrants in innovation in the US. Closer to our focus, Stuen et al. (2012) shows the positive impact of foreign doctoral students on US innovation (measured by patents). Chellaraj et al. (2008) examine the role of foreign graduate students in promoting innovation in US universities. In one of the few examples considering these effects in a non-US context, Crown et al. (2020) analyze the effect of a visa change in Australia on innovation outcomes. Although some of these studies briefly explore the role of foreign students and graduates in founding firms, most focus on outcomes such as patenting and academic articles, as proxies of scientific and technical innovation.

Additional evidence that the entry of foreigners as students had a significant

[8]Recent papers such as Dimmock et al. (2022), Doran et al. (2022) and Parag Mahajan and Brinatti (2024) have estimated the causal impact of foreign-skilled individuals on US employment, innovation, or firm performance using the H-1B lotteries to produce random variations of foreign-born labor supply across firms. However, due to the very nature of their identification strategy, these papers can only look at existing firm-level outcomes and not at firm creation.

positive impact on US entrepreneurship is provided by papers such as Hunt (2011) and Amornsiripanitch et al. (2023). The former shows that immigrants arriving on student visas are among the most successful in their career achievements, while the latter shows that higher education has been the main entry channel for immigrants who have become entrepreneurs.

Relative to the aforementioned studies, our contribution provides several significant innovations. We introduce a new source of data on innovative start-ups (Crunchbase) scarcely used by economists and we analyze the period 1999-2020, which is the most recent before the Covid-19 crisis. We link this information with data on foreign Master's students by university and cohort and we implement an identification strategy adapted from our previous analysis of foreign-born students (Beine et al., 2023) based on variation of tuition fees. We are the first to attribute part of this effect to new start-ups created by US-origin graduates in the same university cohort and therefore likely to be spillovers. Overall, this paper provides the first credible estimate (which we are aware of) of how much foreign graduate students affected new start-up creation in the US.

3 Data and Summary Statistics

In the empirical analysis, we leverage comprehensive data on the creation of start-ups, their founders, and on some of the firm characteristics. We connect these measures to data on Master's graduates in US universities, which include information on the university, the cohorts, and their composition in terms of native and foreign students as well as the tuition fees charged to each group. One of the contributions of this paper is to create a new database that merges one administrative dataset, the Integrated Postsecondary Education Data System (IPEDS) from the National Center for Education Statistics, and a privately owned database built by Crunchbase, which covers US start-ups and their characteristics. We describe here the datasets, their main features, and how we organized and merged the two.

3.1 International Master's graduates and tuition fees

The IPEDS collects information on the number of domestic and international students enrolled and graduating in each US university, by year and by degree (Bachelor, Master, and Doctorate). International students are defined as the"non-resident alien" students, which include students who are not US citizens and have a temporary visa (usually a F or a J visa) and who, as a consequence, do not have the right to remain in the country indefinitely after graduation. IPEDS counts the number

of degree-seeking students during the fall of each academic year and the number of students graduating each year by degree type and major. “International" and “domestic" students in each category are separately counted. ”International" (or simply "foreign" in our notation) are foreign-born students on an F or J visa. “Domestic” students, include US citizens and (possibly foreign-born) permanent residents. Our main variable of interest is the size of the graduation cohort, namely the number of Master’s students graduating in each university and year over the 1999-2015 period. In our analysis, we include all universities that are classified by the Basic Carnegie Classification as Master’s- or Doctoral Degree- awarding institutions.

Additionally, the IPEDS data include information on the average tuition fees paid by Bachelor’s and Master’s students between 1997 and 2013 (including Master’s students who graduated between 1999 and 2015, our period of analysis). Those fees are the average “on the books" fees for the university in each academic year. They are differentiated across Master’s programs and between in-state and out-of-state/international students. Within public universities, tuition fees are different for in-state and out-of-state students and are subject to a significant amount of annual variation driven by cost, equity, competitiveness, and other considerations. Part of the variation is correlated with financial, cyclical, and economic factors and with the quality of the program and the university. However, even within an institution, there is a large amount of idiosyncratic year-to-year variation, both in the level of tuition and in the difference between in-state and out-of-state tuition fees. We exploit this idiosyncratic variation in our IV approach by using the residual variation over time of out-of-state minus in-state fees as an exogenous determinant of enrollment of international students, after controlling for other factors specific to the university.

3.2 Founders of Startups

The second key source of data is a database obtained from Crunchbase, which includes information on start-up firms, their founders, and several of their characteristics. [9] This is a crowd-sourced dataset that collects information on the birth and history of start-up companies in the US. This database represents the *“premier data asset on the tech/startup world”*, and is used by the capital venture industry to have a complete picture of such a universe (Dalle et al., 2017). Although this database has only been used in a limited way by economists (e.g. Dimmock et al., 2022), several studies in management science have relied on it. Block et al. (2015) and Wang (2018) argue that Crunchbase provides comprehensive coverage for early-stage innovative firms.

[9] Data are available at Crunchbase.com

Crunchbase collects data through its partnership that, as of 2023, involved more than 3,700 investment firms submitting monthly portfolio updates. The data collection process is completed by 500,000 executives, entrepreneurs and investors belonging to the Crunchbase data user community, data analysts who work for Crunchbase, and machine learning algorithms that validate data accuracy and scan for anomalies. The dataset was started in 2005, and information before that date is backfilled. Hence, information prior to that year includes increasingly missing data. Our analysis focuses on cohorts that graduated between 1999 and 2015 and we consider start-ups created up to 5 years after graduation. Hence we span the period 1999-2020. Information from Crunchbase includes the firm's starting date, its sector and location, the names of its founders, their highest degree (Bachelor's or Master's), the University and major of graduation, as well as the date of the degree award. The data also include information on important "events" experienced by the firms, such as "rounds of funding" and "dollar amount raised" by the start-up. Additionally, we retrieved information on the end date of the company, if the company ceased operations. The names of the founders are often linked to their LinkedIn page, which provides additional information about them. We are therefore able to identify the degree, university, and year of graduation of 29,000 founders, which represents about 27% of all founders of US start-ups included in Crunchbase. As this is a large but incomplete sample, in section 3.4, we analyze and compare the main characteristics of start-ups for which we have full information on founders with those for which information is missing to see if the smaller sample is representative of the larger one.

The main purpose of this database is to count start-up founders with a Master's degree across all universities and cohorts to produce a "start-up entrepreneurship rate" across them.[10] Since many start-ups have multiple co-founders, we will calculate the number of new start-up firms per university-cohort pair by assigning each co-founder a fractional amount equal to the inverse of the number of co-founders.

Additionally, we use available information about start-up firms to construct indicators that capture not only the creation of start-ups but also to characterize the creation of especially successful ones. We use the information on companies that ended operations to measure the number of created companies still active 3 years after their creation. According to the Bureau of Labor Statistics Data on the survival of private sector establishments, in the years between year 1993 and 2023 about 60% of firms survive at least 3 years.[11] Therefore, this threshold excludes the least successful 40% of them. We then use the information on the amount of venture cap-

[10]Recognizing that the data may not cover all start-up founders, one can call the outcome "Crunchbase start-up rate" to be specific of the source.

[11]See Table 7 at https://www.bls.gov/bdm/bdmage.htm

ital investments raised over time to identify companies that raised more than $25M within their first three years of operation. This allows us to analyze the probability of founding highly valued firms.

Finally, we supplement this data with information from the US Patent and Trademark Office (USPTO) to measure the number of start-ups that were granted at least one patent within their first three years of operation. With that goal, we combine the list of Crunchbase start-ups with the universe of inventors in the USPTO data by using a fuzzy matching algorithm applied to company names. Being granted a patent in the early years of operation is often a predictor of how innovative a firm will be, an important step to building market power and future profits.

3.3 Inferring US-origin of start-up founders

In the latter part of our analysis, we approximate what part of the effect maybe due to spillovers of foreign students on US-origin start-up creators. To that aim, we impute the foreign or US-origin of start-up creators from their first and last names following an approach adopted by a large literature in social science. Family names have been extensively used in various fields of social science to infer the origin of individuals. For example, this approach has been used to estimate potential genetic connections between populations (see Piazza et al., 1987; Jobling, 2001). In the economic literature, names have been used to connect people to ethnicity or areas of origin. Influential examples are Kerr and Kerr (2018), who use last names to characterize the nature of collaborative patents in innovative public firms in the US, and Clark (2014), who uses family names to estimate rates of social mobility of ethnic groups within countries. Beine et al. (2023) rely on names to infer the ethnic origin of football players and the resulting genetic diversity within national soccer teams in Europe.

To implement our analysis, we rely on the "*forebears.io*" website that provides, for each first and last name in the world, their geographical distribution between countries. In particular, for each first-last-name sequence, the algorithm provides the probability distribution of that person being a citizen of each specific country. The website documentation indicates that the average probability of the algorithm attributing the highest probability to the correct nationality is about 85%.[12] In the US, the probability that the most likely country of birth is correct is only 80%. The most common error is attributing second-generation immigrants to the country of birth of their parents (or ancestors), which can generate a certain amount of false

[12]For details see https://forebears.io/onograph/documentation/appendices/nationality-jurisdiction-probabilities.

positive cases, i.e. US-born individuals identified as foreigners. Our focus will be on using the algorithm to identify US-origin start-up creators, and hence an effect on those attributable to a spillover from foreign students.

By following this procedure and applying this algorithm to each name in the list of founders reported in the Crunchbase dataset, we assign the probability distribution that an individual has origin in specific countries. We then aggregate these probabilities into the probability of being of US- or foreign-origin. As the data in our analysis are aggregated at the "university-graduation-year" cell, we weight each founder belonging to such a cell by the probability of being of foreign descent. As a result, we obtain a probabilistic count of who has foreign or US descent, which will allow us to analyze whether the higher propensity to start-up is due to one or the other group.[13]

3.4 Summary Statistics and Trends

We calculate the probability to found a start-up firm for graduates of each university-cohort pair in the US over the period 1999-2015. This sample includes 13,697 firms and represents about one third of the total Crunchbase sample of new start-ups founded in the period under consideration. Table 1 shows summary statistics for different characteristics of the Crunchbase sample of start-ups. Column 1 includes start-ups whose founder's name is non-missing, while column 2 restricts the sample to start-ups with all the needed information on the name and education of their founders. We see an extreme similarity in the average of most characteristics explored, such as average age, average dollar funding received, the share of firms owning a patent, and, importantly, the share of firms (likely) founded by a foreign-born individual. This evidence suggests that the subsample of firms for which we can fully identify the education details of their founders is a representative (random) subsample of all the firms for which we can have the founder's name in Crunchbase. Only gender distributions of founders slightly differ across samples. Hence, the estimates obtained from the available sub-sample should be externally valid and applicable to the full Crunchbase sample.

The merged sample examined in column 2 of Table 1 covering each university-cohort and the start-ups founded by its graduates is split into two sub-samples in our analysis, depending on the identification strategy used (see Section 4.2). We separate the public universities from the private (not-for-profit) ones. For public universities, we use the idiosyncratic variation in the out-of-state tuition net of the

[13]Note that since we use this procedure only to distinguish between US- and foreign-born entrepreneurs, errors in the algorithm's allocation between foreign countries do not affect our measures.

in-state tuition to assess the causal effect of foreign students. We refer to this as the "tuition IV sample". For private universities, since there is no in-state vs. out of state tuition fees, we use the foreign shares across universities to build our shift-share variation. We refer to this as the "shift-share sample". Therefore, we separately characterize these two different samples and provide their summary statistics in Table 2. Columns 1 and 2 are relative to public universities, while columns 3 and 4 relate to private ones.

The average size of a year-university cohort is similar in the two samples, ranging between 600 and 650 graduates. The average share of foreign graduates is also very similar in the two samples, just below 10%. Within each group, "research" universities (identified as groups R1 and R2 in the Carnegie classification) have a larger share of foreign-born graduates (16% and 20%, respectively), confirming that top research programs tend to attract more students from all over the world. The average number of start-up founders and, therefore, the number of start-ups created by each school-cohort, our main outcome variables, were significantly larger in the private university sample relative to the public university sample (0.27 founders per cohort in private versus 0.09 in public, and 0.1 firms created versus 0.34).

In terms of average characteristics of start-ups, graduates of public university Master's programs have a larger probability of founding a company in the same state where they graduated (59% versus 48% for graduates of private universities). For both samples, we observe a large share of these start-ups (about 98%) surviving for 3 years or more. Compared to a survival rate of about 60% at the national level, this signals that our population of interest is a selected group of high-quality start-ups.

Furthermore, 72% of the start-ups in the public university sample and 65% in the private one raised more than $25 million in capital within the first 3 years, and about 2.5% of the start-ups in both groups developed (at least) one patent within 3 years of their creation. Obtaining a patent early in a firm's life is quite a remarkable achievement even for this selected group of firms, considering that the patenting process takes some time and is very demanding.

The four panels in Figure 1 provide additional characterization of the start-up rates in different subgroups and are obtained jointly using public and private university samples. They show the distribution of start-up rates (per 10,000 graduates) across US states, by Master's major, by Carnegie classification category of the university, and by public/private type. The top left Panel ranks states in terms of start-up rates and shows a very large range of variation across states (from above 2 per 10,000 graduates to less than 0.2 per 10,000).[14] Massachusetts and Califor-

[14]A list of top schools in terms of start-up founders per 10,000 students can be found in the

nia, home to some of the most prestigious universities and Master's programs in the country, have a very high rate of start-up founders.[15] The top right Panel shows the start-up rates across Master's majors, including only those with a significant number of graduates, and one can see that the STEM and Business Master's programs are those with the highest rates. The two bottom panels show the start-up rates differentiating between R1, R2 and other institutions according to the Carnegie classification (bottom-left), and between public and private universities (bottom-right). R2 institutions (high-research universities) are among those with the highest start-up rate, and the rate is higher in private universities than in public universities.

Figure 2 shows the average start-up rate of the graduating cohorts over time. Specifically, each point shows the number of start-up founders in the cohort graduating in that year, as recorded in our Crunchbase data within 5 years of graduation. The number is standardized "per 10,000 Master's graduates". The series shows a start-up founding rate increasing from about 0.3 for the cohort graduating in 1999 to around 1 (per 10,000) for the cohorts graduating from 2009-2014. The fact that this value significantly increases after 2005 may be partly because the database started collecting information in that year, and had more complete coverage of companies founded after 2005. Additionally, as the financial crisis of 2007-2008 prompted more people to pursue a Master's degree, and the following years saw strong economic growth, part of the increase in start-up rate for the post-2006 cohorts can be a consequence of those events.[16] To avoid these macro-trends affecting our findings, we will not rely on time variation for any of our findings in our empirical analysis.

A value of 1 start-up founder per 10,000 graduates in the early 2010s is smaller than the generic entrepreneurship rate in the US, which amounts to 30 new entrepreneurs per 10,000 adults.[17] However, the Crunchbase start-ups considered here are a selected group of firms. They include only companies that bring new products and services to market, and many will grow to become dominant companies in their sectors. Compared to the rate reported in Azoulay et al. (2022) for firms that have more than 50 workers, the rate is similar. They show in Figure 1 (panel B) that about 1 in 10,000 people is a founder of a firm of this type.

Figure 3 shows the same start-up rate by graduation cohort, but separates between graduates of research-oriented universities, defined as universities awarding

Appendix.

[15]Stanford, Harvard, MIT and UC Berkeley are among the schools with the largest rates of graduates founding a start-up. See Table 13 in the appendix.

[16]See Ulvestad and Skjelbred (2023) who show that economic downturns tend to lead more people to continue higher education.

[17]See for instance https://www.statista.com/statistics/693361/rate-of-new-entrepreneurs-us/.

doctorate degrees with high or very high research activity in the Carnegie classification (Categories R2 and R1), and graduates of other institutions. Master's graduates of research-intensive schools are about 6 to 7 times more likely to create a start-up (within 5 years of graduation) than those graduating from other schools. The average rate at which they created a start-up in the more recent years (between 2009 and 2014) was about 3 to 4 in 10,000 graduates. For graduates of other universities, it was only 0.5 out of 10,000 on average. The focus on research and their connections within the world of innovation and business through their faculty and alumni are among the many reasons that make these universities more likely to graduate successful founders of innovative start-ups.

4 Empirical Model and Identification

4.1 Estimating the impact of foreign graduates

An increase in the share of foreign-born graduates in a school-cohort cell might affect the total number of start-ups created by graduates of that cohort in two different ways. First, as we showed that the average probability of founding an innovative start-up firm is significantly larger among foreign-origin graduates than among US-origin ones, an increase in the share of foreigners in the cohort will result in more start-ups. We will approximate such a direct effect by estimating how an increase in the share of foreign graduates in each "university-cohort" cell affects the number of *foreign-origin* founders of start-ups in the same cell in the following 5 years after graduation. Additionally, foreign graduates can inspire US-origin entrepreneurs in the same cohort to found or co-found a start-up. This second would approximate a spillover effect, and it would result in a higher number of start-up founders among natives, too.

To capture the aggregate effect of foreign-born students on the start-up rate of a graduating cohort and to separate the direct and spillover effects, we estimate two different versions of the equation below. The units of observation in these regressions are university(u)-graduation year cohorts (t) for Master's students. For instance, one observation would measure the number of MIT Master's graduates in 2010 or, alternatively, the number of UC Berkeley Master's graduates in 2014. The estimated equation takes the following form:

$$Y_{ut}^{n} = \alpha_u + \alpha_t + \beta FG_{ut} + \varepsilon_{ut} \tag{1}$$

where Y_{ut}^{n}, the outcome variable, is the number of start-up founders (or companies founded) per university u and cohort-year t. Initially, we aggregate founders /

firms in a cohort, and in subsequent specifications we separate them between foreign and US-origin firms/founders, using the subscript n. The terms α_u and α_t capture university-specific and cohort-specific fixed effects. Those fixed effects capture the time-invariant characteristics of different universities. They are likely to be correlated with the selection and characteristics of students, as well as with their probability of founding a start-up. Fixed effects for the year capture differences in the aggregate evolution of start-up creation. These time-fixed effects are particularly important given the evolution over time of start-up creation rates documented in Figures 3 and **??** above. The term FG_{ut} captures the size of foreign graduates in the cohort. In the benchmark regression, this variable is measured as the share of foreign-born graduates in the university masters cohort cell. When that is the case, the estimated coefficient will capture the effect of replacing a one-percentage point of US-born graduates with a one-percentage point of foreign-born graduates in the cohort population. In the additional specifications of equation (1), we measure FG_{ut} as the (log) number of foreign-born graduates and estimate the semi-elasticity of one percent more foreign-born graduates on the number of start-up founders, keeping the number of US-born graduates constant. The coefficient of interest that we report in the tables is denoted by β in Equation (1).

To decompose the effect of foreign-born on start-up creation into an effect by foreign-origin creators and likely spillover by US-origin ones effects, we use either Y_{ut}^F, the number of (likely) foreign-origin founded firms, or Y_{ut}^{US}, the number of (likely) US-origin founded firms as our dependent variables. If the variation of foreign-born graduates across cohorts was random, the coefficients β in the least squares regression of specification (1) would capture, respectively, the direct effect and the spillover effect of an increase in foreign-born graduates on US start-ups. The inclusion of university and year fixed-effects implies that the identification of the coefficient of interest relies on changes in the intensity of foreign-born graduates within a university over different graduating cohorts. Given that such variation is likely still to be correlated with unobservable variables, we discuss identification and our proposed IV strategy in the next section.

4.2 Identification strategy

Potential OLS Bias

A simple estimation of equation (1) using least squares is likely to be subject to omitted variables and endogeneity bias. Some factors affecting the probability of enrollment of foreign-born students in a university and business profitability are the local economic conditions, local prices, and local job opportunities. In general, any

local factor that can produce a correlation between foreign enrollment and start-up rates at the local level can induce a positive or negative bias in the estimated coefficient β, depending on the sign of such correlation.

Crucially, the choice of a university/location for any Master's student is endogenous. International students, who are choosing among several US options from a foreign country, are especially sensitive to these differences. Their location could be positively influenced by local factors related to business opportunities after graduation, such as the presence of other start-ups in the area or the support of local policies and financial institutions to business creation. However, as documented by Beine et al. (2023) for undergraduate students, enrollment in a specific location also depends negatively on the level of rents and local costs of living, both correlated with the area's economic conditions. Compared to US-born students, the location choice of foreign-born students is more sensitive to these factors as they do not have a home bias among US locations.

These confounding factors generate potentially serious omitted variable bias, which we address with an IV strategy. Specifically, we use two alternative instrumental variables that leverage different margins of variation in foreign-born students across University-cohorts. First, we instrument changes in the share of international graduates in public university cohorts with idiosyncratic changes in the difference between the out-of-state and the in-state Master's tuition fees. After controlling for university and cohort characteristics, the tuition fee differences change over time in a quasi-random fashion and are hardly correlated to any previous trend. Therefore, they provide an exogenous source of cost differences in pursuing a Master's degree, which affects the probability of enrollment of international students without being associated with other local variables. Second, as an alternative, we use changes in the share of international graduates in private (not-for-profit) universities using a shift-share instrument. This constructed variable distributes annual flows of new international students by country of origin across university Master's programs, based on their pre-2003 shares. This method leverages pre-existing university-specific networks of foreign-born students, which are very important to attract new enrollment from the same country and are less correlated with other local characteristics.

These two IV strategies provide different sources of identifying variation and enable us to estimate the effects of a change in foreign-born students for different sets of "compliers" with different LATE interpretations and different samples. Significant and similar estimates using these two different approaches and samples would be consistent with a robust causal effect of the greater presence of foreign-born students on start-up rates. The validity tests performed on each IV will indicate which one is more consistent with the identifying assumptions.

Out-of-state tuition fees as an IV

The first IV strategy exploits idiosyncratic changes in tuition and fees charged by US public universities to out-of-state (and foreign) students relative to in-state ones. While part of the tuition charged to Master's students depends on the quality and prestige of the institution and on financial considerations by the university, we isolate its more systematic components by subtracting the in-state tuition and controlling for lagged characteristics of the school and location. In the spirit of Borusyak and Hull (2021), we implement a procedure that allows to separate the systematic, non-random variation from the quasi-random one, and test whether the latter component is uncorrelated to a series of pre-period trends.

This IV approach builds on the method introduced in Beine et al., 2023 to estimate the entry of foreign-born college graduates into the US markets. Here, we apply it to out-of-state Master's tuition fees to predict Master's student enrollment. We begin by running a "stage-0" regression to construct the instrument, specifically considering the difference in the tuition fees charged to out-of-state and in-state students, and regressing it on a set of time-varying proxies of universities' quality and financial situation. These proxies include universities' total revenues, revenues per student, total expenses per student, and instruction-specific expenses per student. We additionally control for the share of State Budget appropriations, the number of past international graduates, and year-fixed effects. Equation (2) shows the corresponding specification:

$$\begin{aligned}(\text{Out-of-state tuition - In-state tuition})_{ut} = \delta_t + \gamma_1' \text{Quality proxies}_{u,t} \\ + \gamma_2 \text{State appropriations}_{ut} + \gamma_3 \text{Int. Graduates}_{u,t-1} + \nu_{ut}, \quad (2)\end{aligned}$$

where δ_t denotes a set of year fixed effects and "Quality proxies" is a vector of variables that capture the other controls listed above.

We then use $\widehat{\nu_{u,t-2}}$, that is, the two-year lagged residual of regression (2), as our instrument in the benchmark analysis. We lag the $\widehat{\nu_{u,t}}$ by two years to account for the tuition charged at time of enrollment to a cohort of graduating Master's students, assuming a typical time needed for graduation of two years. Increases in the non-resident tuition net of in-state tuition, unexplained by the variables of equation (2), capture the more idiosyncratic components of university-specific Master's tuition fees, which affect the cost of attending a Master's program for a given cohort of foreign-born students, and thus their probability of enrollment in that particular university-cohort.

The first-stage equation of our IV procedure takes the following form:

$$FG_{ut} = \delta_u + \delta_t + \gamma(\widehat{\nu_{u,t-2}}) + \zeta_{ut} \quad (3)$$

where δ_u and δ_t are university and time fixed effects and ζ_{ut} is an error term.

Shift-share of foreign-born across university-cohorts as IV

The second IV strategy is a more traditional application of the shift-share approach. In this case, we use the inflow of foreign-born students from the 10 countries with the largest foreign student population in US universities, between 2004 and 2015, interacted with the share of students from each of those origins in each university as of 2003.[18] We include in this sample, which we call the shift-share IV sample, only private universities (not included in the previous IV analysis) and, since the data on foreign students by nationality are available since 2003, we consider the analysis between 2003 and 2015.

Data on foreign students by nationality are from F1 visa data obtained from a FOIA (Freedom of Information Act) request. They are individual records of each student enrolled in a Master's program in the US between 2003 and 2015, including information on the Master's program, university, year of enrollment, graduation, and country of origin of the student. From those data, we construct the number of foreign Master's students from each country of origin graduating in 2003, the base year, in each US university. For each foreign origin o, we calculate its share of Master's graduates in 2003 in each university u, and then we allocate the post-2003 number of foreign students from the country o in the US in each year t between universities, proportionally to that initial share. This allows us to generate an imputed number of foreign students from country o in each university u and year t. We then aggregate across countries of origin and calculate foreign students (of all nationalities as a share of)the university cohort (including US natives).[19] This predicted share will then be used as an instrument of the observed share of foreign students in equation (1). Further details of the shift-share procedure are provided in the Appendix (Section 1).

Discussion of different IV's

An advantage of using the two IV procedures is that they involve different identifying variations. In the "average treatment effect" jargon, each of them involves a different

[18]The countries of origin of the students are the following: China, India, South Korea, Japan, Vietnam, Taiwan, Saudi Arabia, Canada, Mexico, and other countries.

[19]Note that we use the number of native students of the base year so that we abstract from any influence foreign-born students could have on native enrollment in the subsequent years.

group of “compliers". The tuition-fee IV generates variation affecting the margin of foreign students who might be more financially constrained and respond to a variation in the fees of a specific school by enrolling or not. These types of complier are more likely to be students from poorer countries and/or from less privileged backgrounds. The shift-share IV, instead, generates variation based on students from countries of origin experiencing an increase in the number of co-nationals studying abroad and who are sensitive to the pre-existing network in their choice of enrollment. Compliers are those foreign-born students who are more likely to be guided by the network to choose a specific university and those who originate from countries with large networks.

In addition, the two methods are applied to different samples. The first IV strategy can be applied only to public universities that offer different tuition for in-state and out-of-state students, while the second IV is applied to private universities and is limited in time due to the shorter period of availability of the F1 data. The sample used for the first IV covers cohorts graduating in the years 1999-2015, while the sample used for the second IV only covers cohorts graduating in the period 2004-2015. A different estimated value of the coefficient of interest between the two approaches may depend on several reasons, such as the different start-up rates for different groups of compliers or different selection in the samples. As we saw in Table 2, which reports summary statistics for both samples, these differ in the value of the entrepreneurship outcomes. In general, the sample associated with the shift share IV includes observations with a higher average start-up rate for both foreign-born and US-born students, but the share of foreign-born and the characteristics of start-ups are not very different between the two. A significant and stable coefficient estimated with both approaches would be consistent with a robust causal effect of foreign students on start-up rates.

5 Strength and Validity of the instruments

Table 3 shows the first-stage coefficients and the first stage F-statistics (last row) for the tuition IV (columns 1 and 2) and for the shift-share IV (columns 3 and 4). Since the regressions include university- and year-fixed effects, the power of the instrument is based on the within-university variation over time. The estimates in columns (1) and (2) of Table 3 show that the idiosyncratic component of the tuition cost has a strong negative correlation with the share of foreign-born Master's students. An increase of the instrument by one standard deviation produces a decrease in the percent of foreign-born Master's graduates (2 years later) by 0.91 points. This is a significant effect relative to the standard deviation of 10 percentage points for the

endogenous variable. The effect is robust to controlling for the total size of the cohort, captured by the log of the number of graduates in the cohort (column 2), showing that the change in composition (share of foreign-born) is the main driver of the results. Although the inclusion of this variable may represent a "bad control", the fact that it does not change the first stage coefficient, nor the significance of the IV, is reassuring. In addition to that, the Kleibergen-Paap F-statistics of the first stage are around 67, much larger than values generating concerns of weak instruments, and the Anderson-Rubin test of weak IV rejects the null of weak instruments at the standard confidence level.

Columns (3) and (4) show the first stage for the shift-share IV, constructed and implemented for the sample of private universities only. The coefficient implies that an increase of the imputed share of foreign graduates by one percentage point will increase the actual share by 0.26 percentage points. As expected, the intensity of the network that interacts with changes in the total number of foreign students from a country of origin is a strong predictor of changes in foreign-born students. The first-stage Kleibergen-Paap F-Statistic of this IV is close to 30. All in all, both instruments are strong.

The validity of the tuition IV hinges on the idea that out-of-state tuition fees net of in-state fees affect the number of foreign enrolled students, but do not affect other variables. More specifically, changes in out-of-state fees should not affect the enrollment of US-born students (or at least should affect them much less). Similarly, shift-share IV, since it predicts only the change in foreign student inflows, should not be correlated with the number of US-born students in the cohort. To test these conjectures, Table 4 shows, for each of the two IV strategies, their ability to predict the (log) number of foreign graduates, the (log) number of native graduates, as well as the (log of) total number of students.

The coefficients of Table 4 show that an increase in the idiosyncratic component of the tuition fees causes a decrease in the overall size of the graduating cohort (column 1), driven by a decrease in the (log) number of foreign students (column 2), while the number of native students remains unchanged (column 3). The magnitude of the negative and statistically significant coefficient on foreign students is much larger than the non-significant coefficient on natives. Hence, changes in the tuition IV affect the supply of foreign students in a school-cohort, leaving unchanged the supply of native students, consistent with what a valid instrument should generate. With respect to the second instrumentation strategy, explored in the row below, changes in the shift-share IV seem to predict an increase in the (log) number of foreign students, and a corresponding decrease in the (log) number of native ones, leaving the overall size of the graduating cohort unchanged. This may reveal some

crowding-out of native students in response to the shift-share changes. Therefore, the tuition IV seems to pass this exclusion restriction better than the shift-share IV, as it affects enrollment of foreigners but not of natives. Still, both IV significantly affect the number of foreign-born students in the school-cohort, resulting in a significant impact on the share of graduating foreigners.

An additional validity check for our tuition IV consists of testing whether past values in the number of foreign students and foreign graduates across school cohorts are correlated with current values of the instruments. In particular, correlation with past inflows of foreigners would reveal spurious trends or reverse causality on out-of-state tuition fees. In Table 5 we explore this possibility by showing the correlations of lagged foreign students' enrollment and graduation, with the current idiosyncratic changes in out-of-state, net of in-state tuition. The insignificant coefficients at 4, 5 and 6 years lag imply that the past behavior of enrollment does not affect current idiosyncratic fee changes, ruling out a potential reverse causality channel.

In the same vein, for the shift-share IV, we consider how the long-run change in this variable over the 2004-2015 period, which predicts the inflow of foreign students during this period in each university, is correlated to the pre-2003 inflows of foreign students in the corresponding university and in the start-up founders from the corresponding university in the 1995-2003 period, prior to the analysis for the Shift-share sample. Table 6 shows mostly insignificant correlations with these pre-trends, except for the one with changes in foreign graduates between 1995 and 2003.

In summary, the power and validity tests indicate that the tuition IV is a strong and valid IV, with evidence suggesting that the exclusion restriction is reasonably satisfied. The tests show insignificant pre-trends on foreign enrollment and that current fees affect only foreign enrollment without spurious correlations with native enrollment. The shift-share IV, on the other hand, appears to be less strong and somewhat correlated with a measure of pre-trends in the number of foreign students and with native enrollment. Nevertheless, in line with this literature, we will use both instruments in the analysis, placing greater emphasis on the estimates obtained from the newer and more robust tuition IV. The results of the shift-share IV can be used as external validity checks of the tuition-based identification strategy with the caveat of relying on a weaker identification of this approach.

6 Main Results

6.1 The effect of foreign graduate shares on start-up rates

Table 7 shows the main findings of our analysis. It reports estimates of the effect of the share of foreign graduates on start-up creation in a given cohort-year from Equation (1) using Least Squares (in specification 1) and IV (in all other specifications). The upper panel shows the estimates using the idiosyncratic component of tuition fees as an instrument and the 1999-2015 graduating cohorts from the public universities sample, while the lower panel shows the estimated coefficients using the shift-share instrument for the 2004-2015 graduating cohorts from the private universities.[20]

Different columns of Table 7 use different outcome measures as the dependent variable in the analysis. In the first two columns, the outcome is the number of start-up founders in the school cohort within five years after graduation. Column (3) instead counts the number of start-up companies founded (also within 5 years of graduation) by Master's graduates, ensuring that individuals in the same cohort, who co-founded the same company, are not counted twice. This is the most straightforward measure of start-up creation by Master's graduates. Column (4) includes only start-up companies founded by master graduates in the same state as their university. We call this specification "local" firm creation.

In column (5) we only include as outcome start-up companies that survived for at least 3 years. Although the 3-year threshold is reached only by 60% of start-ups on average, in our sample the proportion is much higher, as the database tends to include more successful start-ups. The outcome of column (6) is a count of start-ups that raised at least $25 million in capital in the first 3 years since creation. Finally, in column (7), we use the count of start-up companies that received at least one patent within the first 3 years of operations as the outcome of interest. This is a rare occurrence that affects only 3% of the companies in our sample and identifies very innovative companies. By estimating these different specifications, we show the effect of foreign students on the probability of creating successful companies, using measures of firm outcomes that become more demanding as we move across columns.

Two notable findings emerge from the set of estimates in Table 7. First, using both IV strategies, we find statistically and economically significant effects of increasing the share of foreign graduates on the number of start-up founders in general and when we focus on 3-year survivors and start-ups that raised $25 million. Considering the

[20]Since the sample is different between the two estimation strategies OLS estimates also differ slightly across the two panels.

estimates in column (3), an increase in the share of foreign graduates by 10 percentage points (without any change in the overall size of the cohort), which equals about one standard deviation of that variable across cohorts, results in an increase in start-up founders in the cohort by 0.41 when using the tuition IV, and by 0.73 when using the shift-share IV. Both estimates are statistically significant. This represents a sizeable effect, equal to the standard deviation of the start-up rate between cohorts, in the tuition IV sample and about half of the standard deviation for the shift-share IV sample (see Table 2). The coefficient has a causal interpretation: by attracting ten percentage points more foreign-born master students, a university generates a higher start-up rate of its graduates by about 0.4 new firms in the tuition IV sample. An aspect to note is that the OLS estimate is significantly smaller than the IV and is not statistically significant. This implies a downward bias of the OLS estimates that we will explain and discuss further in the next section.

Second, while the effect of foreign-born graduates is stronger on the simple count of start-up companies, it is still significant when we focus on the creation of high-performing companies (the selected outcomes of columns 5 to 7). The significant coefficients imply that foreign-born graduates increase start-up rates for the high-achieving group of start-ups as well. The IV estimates using the shift-share IV (which showed less reliability in terms of identification) show more variation and less significance for some of these outcomes, while those using the tuition IV are more stable and more significant. Overall, the effects estimated in Table 7 using the IV approach are quantitatively important, statistically significant, robust across specifications, and consistent between samples and IV choice. Our evidence shows that a larger share of foreign Master's graduates increases the start-up rate of a Master's cohort.

Although IV estimates should isolate the causal effect of foreign graduates on start-ups, the presence of university-specific trends in the quality of Master's students, potentially correlated with the share of foreigners, may represent an unobservable variable causing bias. To mitigate these concerns, in Table 8, we estimate specifications similar to those reported in Table 7, but including more demanding interactions between trends and university-specific fixed effects. We start by defining, for each university program, periods of relatively stable enrollment (i.e. periods in which the number of students enrolled did not grow more than 5%). In specifications (1), (3), (5), (7), (9) and (11) of Table 8, we control for a set of fixed effects capturing these university-specific periods interacted with university dummies. These specifications absorb the variations associated with medium-term and long-term changes of university enrollment and rely only on idiosyncratic year-to-year variations in foreign share present in the Master's programs. Furthermore, in the specifications reported

in even columns of Table 8, we control for university-specific time trends, which absorb gradual long-run changes in enrollment and outcomes. As in the previous table, the upper and lower panels of Table 8 show the estimates using tuition IV and shift share IV, respectively. Both the point estimates and the significance of the coefficients are remarkably stable and similar to those of Table 7, suggesting that slow and non-observable trends in tuition and quality of graduates are not driving the estimated effects.

An increase in foreign share by one percentage point leads to a rise in the average number of start-ups in the cohort by 0.04 to 0.07, depending on the IV and sample, in line with previous estimates. In this case, the estimates for the effect on the more selective group of start-ups (those raising $25 million and those with at least one patent) are not quite significant when using the shift-share IV. Nevertheless, the tuition IV shows significant and robust coefficients across specifications.

An interesting question is what are the reasons behind the effect of foreign graduates on start-up creation. The limited information on students' characteristics in IPEDS does not allow us to dig too much into the mechanisms, but we can formulate some conjectures. Foreign students may be characterized by a lower risk aversion compared to native students, which in turn favors entrepreneurship. In general, the literature finds evidence that movers have lower risk aversion than non-movers (Jaeger et al., 2010). Another reason may be that a higher propensity to create start-ups derives from a higher work experience of foreign graduates. A significant share of foreign students enroll in a master program in the US after some years of work at origin and do not enroll directly after graduation. These two mechanisms are also at work in the spillover effects on natives: native entrepreneurs might benefit from lower risk aversion and higher work experience brought by foreign students. Another potential reason is the lower opportunity cost of entrepreneurship compared to native students who have access to a broader set of jobs after graduation. Furthermore, due to the low number of H-1B visas and green cards, foreign students are limited in their transition to the job market as employed workers, which may create an incentive to start a business. However, it is unlikely that they will gain the ability to stay in the US thanks to their business success. At the same time, as we have shown in Beine et al. (2023), the probability of foreign master students staying in the US after graduation is only 20%. Hence, the fact that their presence increases start-up rates even as many of them leave the country may imply a strong spillover on the entrepreneurial rate of US natives. However, these potential reasons require further exploration that is left here for future research.

6.2 OLS bias and importance of the IV

The comparison between the OLS and IV estimates, shown in the first two columns of Table 7, reveals a significant downward bias of the OLS coefficient. This bias could arise because of a negative correlation between omitted factors affecting the creation of start-ups and the probability of foreign enrollment in a university cohort. Possible candidates are omitted variables that capture the cost of living and housing in the area where the university is located. Increases in local prices are typically associated with strong local economic growth (see Moretti, 2004) which, in turn, encourages firm creation. On the contrary, a location experiencing a downturn or slow growth would likely exhibit lower housing prices, making it more attractive to potential students but reducing start-up opportunities. This type of omitted variable would generate a downward bias in the least-squares estimates of the effect of the foreign graduate share on start-up creation. In addition, housing availability and vacancies may play a role. Although it is difficult to fully control for local housing availability and prices (due to limited data), the IV strategy, by using variables that are likely uncorrelated with local economic conditions, can attenuate this issue.

The two correlations shown in Figure 4 strongly suggest the potential omitted variable bias that we described above. The left panel of Figure 4 is a binned scatter plot of the relationship between the demeaned changes in local rents for housing by county and year (x-axis) and the demeaned change in the aggregate number of students enrolled in universities (y-axis), including 5% of the observations in each bin. The data on local rents, from the U.S. Department of Housing and Urban Development, measure changes in the median housing rents for 1 bedroom apartments by county and year. A strong and negative correlation is apparent, consistent with the local housing price reducing the inflow of international students to that location due to increased costs. Similarly, the right panel of Figure 4 instead shows a strong and positive correlation between changes in local rents (x-axis) and demeaned changes in the number of start-up founders by county and year of firm creation (y-axis). Together, the figures suggest that periods of local economic expansion, associated with higher prices and more start-up activity in a local economy also attract fewer students to local Master's programs, possibly due to higher cost of living, consistent with the idea that potential students are sensitive to these local costs when deciding about their Master's location. The existence of this type of spurious correlation between local economic conditions and Master's enrollment makes the use of IV especially important. By isolating pull factors for foreign students that are independent of current economic conditions (idiosyncratic fee differences and interaction of pre-existing networks and push factors), both IV strategies correct for this kind of omitted variable bias.

6.3 Elasticity of start-up firms to the supply of foreign Master's graduates

In Tables 7 and 8 we used as an explanatory variable the foreign share of Master's students in a university cohort. Keeping everything else constant, a change in this variable implies that we are comparing the start-up frequency of a cohort with 1 percentage point more foreign-born and 1 percentage point fewer natives. Such a composition effect will be positive if the propensity to create start-ups is greater for foreign graduates and/or if their presence in the cohort increases the start-up rate for all graduates. In order to quantify more directly the impact of foreign-born on the number of start-ups of a cohort, all else equal –including the number of US-born graduates–, we estimate a specification that produces the semi-elasticity of the number of start-up firms to percentage (log) changes of foreign-graduates. In other words, we use the (log of) number of foreign graduates (rather than their share in the cohort) as the explanatory variable and the number of start-up founders (or firms) at the university-cohort level (as before) as the dependent variable. We rely on tuition IV as our preferred instrumental variable, as we showed in Table 4 of Section 5 that this variable affected the number of foreign students but did not affect the number of US-born students.[21]

Table 9 provides the IV estimates using this tuition-based IV approach. An increase of one percentage point in the share of foreign-born students corresponds to an increase of 10% in the number of foreign graduates. We find that on average, an increase by 10% (0.1 log points) in the number of foreign graduates generates about 0.04 more entrepreneurs (i.e. the coefficient of 0.39 times 0.1) in the 5 years following graduation. This is exactly consistent with the estimates of the effect of an increase in the share of foreign-born reported in Table 7 (estimated coefficient of 0.041). Similarly, such an increase would produce 0.037 more start-up companies, 0.024 more companies raising $25 million or more, and 0.005 more start-ups that will obtain at least one patent within the first 5 years of activity. This specification confirms that the increase in start-ups driven by a larger number of foreign Master's students, attracted by lower tuition fees, is significant, is consistently estimated across specifications, and is uncorrelated with the changes in native Master's students in a cohort.

[21] The shift-share IV, instead, affected the share more than the level of foreign students, and is rather weak in this case.

6.4 Heterogeneity across majors and universities

We observed in the summary statistics of Section 3.4 that Master's programs in some majors such as Computer, Biological, Physical, and Mathematical Sciences are particularly conducive to start-up creation, as the probability of graduates from those programs to be start-up founders is the highest. Additionally, we have shown that research institutions (R1 and R2 in the Carnegie classification) are more likely to graduate Master's students who will create start-up companies.

In this section, we analyze whether the impact of foreign graduates on start-up creation is also heterogeneous across majors and universities. Do foreign-born Master's students stimulate entrepreneurship more or less in majors/universities with a higher propensity for start-ups? Differential effects could be due to differential selection of foreign-born graduates across programs relative to natives, or different levels of complementarity of their skills with those of natives in creating new start-up businesses.

Table 10 shows estimates of the effect of an increase in the share of foreign-born separating majors and universities across three dimensions. First, in columns (1) and (2), we separate STEM and non-STEM majors in each university. Then, in columns (3) and (4), we separate the five Master's majors with the highest start-up rates. Finally, in columns (5) and (6), we separate Master's programs in top research universities (R1 and R2) from those from other universities. In all cases, we consider the tuition-IV sample, which includes public universities only.

Two interesting findings emerge. First, while the positive effect of foreign-born on start-up founders is similar between STEM and non-STEM majors, the foreign-born advantage seems stronger in majors not among the top ones in start-up generating potentials. Foreigners graduating with a Master's degree in health sciences and engineering (not among the top creators of start-ups) are significantly more likely than their native peers to create start-ups. Second, the positive effect of foreign-born is concentrated in research universities. Graduates from these universities have a much larger probability of generating start-ups, in general, and likely foreign students are highly selected in those institutions due to the competitive nature of admission. The heterogeneity analysis confirms that the foreign advantage in start-up creation is broad and far-reaching and particularly strong in firms likely to be at the frontier of innovation and technological progress.

7 Spillovers on US-origin start-ups and Extensions

7.1 Direct and spillover effects of foreign graduates

The results presented in Tables 7 and 9 show robust evidence that an increase in international Master's graduates in a US university cohort increased the number of start-ups founded by that cohort. This is consistent with foreign-origin graduates being more likely, relative to similar US-origin Master's graduates, to be founders of start-ups.

To analyze whether this result is driven by more than the entrepreneurial inclination of foreigners, in this section we identify a likely spillover effect by decomposing foreign-origin and US-origin creators of start-ups. Part of the effect of foreign students on the start-up rate of the master cohort is certainly due to the higher start-up propensity of foreigners. In that case, for a given size, a cohort with a larger share of foreign graduates will automatically have a larger start-up rate because of this composition effect. This composition effect can be considered a "direct" effect.

In addition, a larger presence of foreign-origin students may have a "spillover" effect, on the start-up rate of US-origin students in the same Master's program. The entrepreneurial inclination of foreign-origin students, their complementary skill, their international connections, and their motivations may generate a peer effect on US-origin schoolmates. During the Master's program, there are many opportunities for students to interact and work together as teams, in classes, working groups, executing projects, or collaborating on assignments. Therefore, it is likely that students will form collaborations, combine complementary skills, and develop trust. This could then inspire US-origin students to follow career paths inspired by international students and continue collaborating with them after graduation. Notice that some foreign students might become entrepreneurs outside the US and are therefore and their start-ups not included in our data. Still their spillover effects on US-origin creators of start-ups can still be observed in our data. Alternatively it is possible that the presence of highly entrepreneurial foreign-born may discourage or crowd out natives from being start-up founders. This would produce a negative spillover on the start-up rate by natives.

In Table 11, we test the significance of these spillovers. To do that, we construct a new dependent variable by splitting the number of start-up creators (or of start-ups) in a cohort year into two parts, using the name origin of the founder (inferred from their name as described in Section 3.3) to categorize them as foreign-origin or US-origin. Specifically, when aggregating start-up creators, for each of them, we calculate the probability of being foreign-origin or US-origin, and we use those probabilities

as weights to obtain the number of likely foreign-origin and likely US-origin start-up creators in each cohort. Their sum is, by construction, equal to the total number of start-up founders in the cohort. Similarly, when the outcome is the number of start-up firms (rather than founders), we attribute a fraction of them to foreign- or US-origin founders, based on the aggregate probability of its (potentially multiple) founders.[22]

The upper panel of Table 11 shows the estimated effect of a one percentage point increase in the share of foreign graduates on the number of likely foreign-origin start-up founders (column 1) and start-up firms (column 2), which we referred to as the "direct" effect. The table uses the idiosyncratic tuition component as (our preferred) IV. As in Table 7, the different specifications gradually restrict the dependent variable to count a more selective group of start-ups in terms of success. Column (2) sums all start-ups; in column (3), we only include those start-ups founded in the same state as the university of graduation. Start-ups that survived for at least 3 years are used in column (4), while start-ups that raised at least $25 million capital within 3 years are considered in column (5). Finally, start-up companies that developed at least one patent within 3 years of creation are examined in column (6). The lower panel of Table 11 instead shows the effect of the share of foreign graduates on founders and firms of likely US origin using the same specifications. This captures the "spillover" effect of foreign-born Master's students. As our algorithm is more likely to consider a second- or third-generation American as foreign-origin, this estimate is likely to be a lower bound on the spillover effect on US-born start-up creators (symmetrically, we likely estimate an upper bound of the direct effect). The corresponding results obtained using the shift-share IV are reported in Table 16 in the Appendix.

Several results emerge from this table. The first and most interesting finding is that, in most cases, one can attribute at least 30% of the aggregate effect of the foreign-born share on start-ups to spillovers on US-origin creators. A 1% increase in the foreign share in the Master's leads to a 0.011 increase in start-up creation by US-origin creators, relative to a 0.03 total increase. The remaining part amounts to not more than 70% (coefficient of 0.021 in the lower panel of the table) and is attributed to the direct effect. In most specifications, the spillover effect is significant and about half the size of the direct effect. An increase in the share of foreign graduates by 1 percentage point in a university cohort generates about 0.03 additional start-ups on average. Of those, 0.02 are founded by likely foreign-origin entrepreneurs and 0.01 by likely US-born entrepreneurs.

The second interesting result is that, while the spillover effect is stronger when

[22]As an example, a start-up with 3 founders with the respective probability of being US-origin equal to 0.75, 0.5 and 0 will be considered 0.416 foreign-origin and 0.584 US-origin.

counting total start-ups, it is still significant when considering start-ups surviving for 3 years and start-ups developing at least one patent, which are highly productive and innovative firms. Finally, the weakest spillover effect on natives, interestingly, is on the "local" start-ups. This suggests that foreign-origin master's students may be more likely to create start-ups in the same state where they studied, perhaps because they do not have strong ties with other US states, eventually gravitating around the location where they completed their Master's. Natives, on the other hand, may move out and bring their higher entrepreneurial inclination affected by foreign Master students back to their US location of origin.

As we infer the foreign origin of start-up founders using the name-based algorithm, any mistake in this attribution could result in an over or undercount of the foreign-origin start-up founders. In turn, this would affect the decomposition between direct and spillover effects. Note that this will not affect the total estimated effect. In particular, as we explicitly indicate that the decomposition is between "foreign-origin" and "US-origin", if the algorithm includes among "foreign-origin" entrepreneurs several second- and third-generation immigrants, this attribution will overestimate the direct effect and underestimate the spillovers on US-born (rather than simply US-origin).[23]

To partially address this error, we implement a simulation exercise that serves as a robustness check. We first use information from Forebears (the website that provides the algorithm) on the proportions of false negatives and false positives to infer the place of birth of individuals from their names. The website reports a 20% (46%) probability of incorrectly assigning a non-US (US) native to the opposite group based on the name. We, therefore, re-run the algorithm on all founders and randomly select 20% (46%) of the founders initially predicted as foreigners (natives) and change their classification into native (foreign) founders of start-ups. We repeat this procedure 1000 times and re-estimate in each iteration the decomposition into direct and spillover effects using the tuition-based IV in a specification similar to the one used in column (2) of Table 11.

Figure 6 shows the distribution of the estimated direct and spillover effects.[24] The simulation shows two important features. First, since the probability of misattributing a US-born to foreign origin is greater than vice versa, accounting for this leads us to estimate a higher spillover effect (more natives, including some second-

[23]In other terms, the spillover effect that we estimate here using the Forebears algorithm is the indirect effect of foreign graduates on the start-up creation of native entrepreneurs with American-sounding names. A better identification of native founders with foreign-sounding names will lead to a better estimate of the total spillover effect.

[24]Figure 6 in the appendix presents similar results for shift-share IV estimates.

generation immigrants assigned to foreign-origin, benefit from the foreign student share). The figure shows that, on average, only 55% of the total effect is due to a direct effect and about 45% to a spillover on likely US-origin students. Second, in most realizations, the spillover effect is significantly larger than 0, as its distribution is mostly above the 0.05 value. Spillovers on US-origin students exist and seem very significant. Overall, the spillovers of foreigners on native start-up rates are certainly larger than 0 and likely account for 30-45% of the overall effect of foreign Master's students on start-up creation.

7.2 Start-ups Co-Founding between US and foreign graduates

In the previous section, we documented the economic relevance of spillover effects. Building on this, we analyze whether a channel for these spillovers is start-up co-founding. Namely do US-origin creators of start-up show a higher propensity to have a co-founder due to a larger foreign share in their master class? In our sample, about 25% of start-ups have at least two founders. By developing relationships with foreign schoolmates and learning to work with them, several US-born Master's graduates become more likely to collaborate with them or other students to bring innovative ideas to a start-up. [25] Complementarity in skills, abilities, and knowledge is crucial in firm creation and, if exposure to foreign students with entrepreneurial spirit makes US-born more likely to co-found a company with them, this could be a channel for unlocking the potential for US-born start-up ability and increase the total number of start-ups.

Table 12 shows evidence that co-founding is an important mechanism in explaining the spillover effect. The first column reproduces the overall effect of the share of foreign-born graduates on start-up companies (shown in Table 7, column (3)). An additional 1 percentage point of foreign-born graduates in a cohort increases start-ups founded by that cohort by 0.0397. Column (2) considers the impact of foreign-born in increasing start-ups with more than one founder. The dependent variable is now the count of start-ups with two or more co-founders, and, as we can see, the coefficient is very similar, implying that almost the whole effect originates from start-ups with more than one founder (0.036 extra companies for each additional percent of foreigners). Column (3) uses the number of start-ups with more than one founder from the same school as the dependent variable. It shows a significant increase of 0.01 in companies co-founded by graduates of the same school, which corresponds to

[25] A possible impact of the foreign student presence is to increase trust and awareness of the benefit of collaboration with other entrepreneurs. This could also explain the positive impact on co-founding between native entrepreneurs.

25-30% of the total effect and to about 90% percent of the spillover effect. Columns (4) to (6) maintain the same specifications, but use the shift-share IV. They are qualitatively consistent and somewhat larger than our preferred ones. Overall, a larger share of foreign-born graduates in a university cohort makes US-origin schoolmates more likely to co-found start-ups. This is supportive of the existence of complementarities and peer effects: foreign-born students stimulate US-origin students to team up with them or other entrepreneurs, possibly combining complementary skills to become co-creators of a start-up. Our exercise supports the idea that with a lower percentage of foreigners in a Master's class, US-origin graduates would be less likely to become entrepreneurs.

8 Discussion and Conclusion

In this paper, we considered the connection between two of the most dynamic segments of the US economy: business creation in the area of innovative start-ups and Master's degrees from its highly competitive universities. One of the features of this important nexus is that, by stimulating and attracting talents from all over the world, the US economy is fueled by foreign graduates through enhanced innovative capacity and economic dynamism. We ask the following question: How much do the selected international students in US Master's programs contribute to the creation of start-ups in the US? This is a very relevant channel to assess the contribution of foreign-born to the US economy and to analyze whether this process benefits US-born Master's graduates too or crowds them out of entrepreneurship.

We combined data from different sources and refined an original identification strategy based on a quasi-random variation of out-of-state tuition costs first proposed in Beine et al. (2023). On the one hand, we gathered representative data on Master's students and Master's graduates by university and year of graduation from IPEDS data. On the other hand, we collected data on start-ups in the US, their founders, and some of their characteristics from Crunchbase, an underused database on innovative start-up companies. We match each creator of start-up with his/her university-cohort of Master graduation so that we know the share of foreign-born students in his/her master cohort. This matching allows us to analyze the causal effect of foreign master students on start-up creation. Once we established a causal effect of increasing the share of foreign master students in a cohort on start-up creation by the cohort, we find that a significant part of this effect works through the enhanced entrepreneurial intensity of US-origin (a name-based proxy for US-born) start-up founders.

An important problem limiting all studies of the impact of immigration on economic outcomes across locations is the fact that the number of foreign-born in such

locations is not randomly assigned. In this paper, we addressed this identification issue by constructing two types of instrumental variables that affect the number of foreign-born students in a Master's program, and are less endogenous (and less correlated) to local economic conditions or school-specific characteristics. The first one uses the idiosyncratic component of the non-resident, relative to resident, tuition fees, not predicted by school and local characteristics. This instrument turns out to be a strong predictor of the number of foreign-born students in a cohort, by affecting their costs. The second approach uses a more standard shift-share strategy based on the pre-2003 distribution of students by country of origin across US universities and their subsequent aggregate inflows by nationality. We perform strength and validity checks of the instruments and find that the tuition-based IV passes these tests very convincingly, while the shift-share IV satisfies them only partially. We conducted most estimations using both IV to assess the quantitative robustness of our results, but we consider the tuition IV specification as our preferred one.

The results can be summarized as follows. First, we found that an increase of 10 percentage points in the share of foreign-born students at the cohort-school level (equal to its standard deviation) has an effect of producing about 0.3 additional start-ups created by the average cohort of about 650 students, within the 5 years following graduation. This is a substantial effect, as the average start-up creation in a university cohort is 0.10 for public universities and 0.34 for private ones. Our results also showed that the induced creation of start-ups is to a large extent local, i.e. in the same state as the University of graduation, and involves companies able to raise at least $25 million of capital and achieve at least one patent by the third year since their founding.

Second, we show that a significant part of these start-up effects originates from the spillover effect on US-origin creators of start-up. US-origin graduates become more likely to create a start-up when they study in a university-cohort with a larger share of foreigners. As much as 30-45% of the total effect is due to the spillover on US-origin students, while the remaining 55-70% of the total effect is due to the direct start-up creation by foreign-origin entrepreneurs. This estimate on US-origin start-up founders can be an underestimate of spillovers if some of those classified by name as foreign-origin are second or third generation migrants, born in the US. A large part of this spillover seems connected to the higher propensity of natives who have studied with foreign-born students to co-found a start-up with them or with other foreigners.

The results of this paper are particularly remarkable if one considers that foreign-born Master's students are much less likely than US students to remain in the US. Many of them, due to the lack of available visas, have to leave the US only a few

years after graduation (see Beine et al., 2023). This analysis shows that the benefit of retaining some of these international Master's students in terms of new firms, innovation, and new jobs can be substantial.

References

Amornsiripanitch, N., P. A. Gompers, G. Hu, and K. Vasudevan (2023). Getting Schooled: Universities and VC-backed Immigrant Entrepreneurs. *Research Policy 52*(7), 104782.

Anderson, S. (2022). Immigrant Entrepreneurs and US Billion-Dollar Companies. *NFAP Policy Brief*.

Anelli, M., G. Basso, G. Ippedico, and G. Peri (2023). Emigration and Entrepreneurial Drain. *American Economic Journal: Applied Economics 15*(2), 218–252.

Azoulay, P., B. F. Jones, J. D. Kim, and J. Miranda (2022). Immigration and entrepreneurship in the united states. *American Economic Review: Insights 4*(1), 71–88.

Babina, T., W. Ma, C. Moser, P. Ouimet, and R. Zarutskie (2019). Pay, employment, and dynamics of young firms. *Kenan Institute of Private Enterprise Research Paper*.

Beine, M., S. Peracchi, and S. Zanaj (2023). Ancestral diversity and performance: Evidence from football data. *Journal of Economic Behavior Organization 213*, 193–214.

Beine, M., G. Peri, and M. Raux (2023). International college students' impact on the us skilled labor supply. *Journal of Public Economics 223*, 104917.

Bernstein, S., R. Diamond, A. Jiranaphawiboon, T. McQuade, and B. Pousada (2022, December). The contribution of high-skilled immigrants to innovation in the united states. Working Paper 30797, National Bureau of Economic Research.

Block, J. H., C. O. Fisch, A. Hahn, and P. G. Sandner (2015). Why do smes file trademarks? insights from firms in innovative industries. *Research Policy 44*(10), 1915–1930.

Borjas, G. and K. Doran (2012). The Collapse of the Soviet Union and the Productivity of American Mathematicians. *Quarterly Journal of Economics 127*(3), 1143–1203.

Borusyak, K. and P. Hull (2021). Non-random exposure to exogenous shocks:theory and aplications. Technical Report 27845, NBER Working Papers.

Card, D. (2009, May). Immigration and inequality. *American Economic Review 99*(2), 1–21.

Chellaraj, G., K. E. Maskus, and A. Mattoo (2008). The contribution of international graduate students to us innovation. *Review of International Economics 16*(3), 444–462.

Clark, G. (2014). *The Son Also Rises: Surnames and the History of Social Mobility*. Princeton Universty Press, Princeton.

Crown, D., A. Faggian, and J. Corcoran (2020). Foreign-Born Graduates and Innovation: Evidence from an Australian Skilled Visa program. *Research Policy 49*(9), 103945.

Dalle, J. M., M. den Besten, and C. Menon (2017). Using Crunchbase for Economic and Managerial Research. *OECD Working Papers*.

Dimmock, S. G., J. Huang, and S. J. Weisbenner (2022). Give Me Your Tired, Your Poor, Your High-Skilled Labor: H-1B Lottery Outcomes and Entrepreneurial Success. *Management Science 68*(9), 6950–6970.

Doran, K., A. Gelber, and A. Isen (2022). The Effects of High-Skilled Immigration Policy on Firms: Evidence from Visa Lotteries. *Journal of Political Economy 130*(10), 2501–2533.

Fairlie, R. W. (2008). *Estimating the contribution of immigrant business owners to the US economy*. SBA Office of Advocacy.

Fairlie, R. W. and M. Lofstrom (2015). Chapter 17 - immigration and entrepreneurship. In B. R. Chiswick and P. W. Miller (Eds.), *Handbook of the Economics of International Migration*, Volume 1 of *Handbook of the Economics of International Migration*, pp. 877–911. North-Holland.

Freeman, R., V. Wadhwa, and B. Rissing (2008, 05). Education and tech entrepreneurship. *Innovations: Technology, Governance, Globalization 5*, 141–153.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020, August). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Haltiwanger, J., R. S. Jarmin, R. Kulick, and J. Miranda (2016). High growth young firms: contribution to job, output, and productivity growth. In *Measuring entrepreneurial businesses: Current knowledge and challenges*, pp. 11–62. University of Chicago Press.

Hart, D. M. (2011). Immigrant entrepreneurship and internationalization and globalization of successful u.s. high-tech start-ups. Technical Report 2011-15, GMU School of Public Policy Research Paper.

Hart, D. M. and Z. J. Acs (2011). *Immigration and high-impact, high-tech entrepreneurship*.

Hunt, J. (2011). Which immigrants are most innovative and entrepreneurial? distinctions by entry visa. *Journal of Labor Economics 29*(3), 417–457.

Jaeger, D., T. Dohmen, A. Falk, B. Huffman, U. Sunde, and H. Bonin (2010). Direct Evidence on Risk Attitudes and Migration. *Review of Economics and Statistics 92*(3), 684–689.

Jobling, M. A. (2001). In the name of the father: surnames and genetics. *Trends in Genetics 17*(6), 353–357.

Kerr, S. P. and W. Kerr (2020). Immigrant entrepreneurship in america: Evidence from the survey of business owners 2007 2012. *Research Policy 49*(3), 103918.

Kerr, S. P. and W. R. Kerr (2016). *Immigrant Entrepreneurship*, pp. 187–249. University of Chicago Press.

Kerr, S. P. and W. R. Kerr (2018, 01). Global Collaborative Patents. *The Economic Journal 128*(612), F235–F272.

Kerr, S. P. and W. R. Kerr (2021). Whose job is it anyway? coethnic hiring in new us ventures. *Journal of Human Capital 15*(1), 86–127.

Kerr, W. R. (2013, August). U.s. high-skilled immigration, innovation, and entrepreneurship: Empirical approaches and evidence. Working Paper 19377, National Bureau of Economic Research.

Kerr, W. R. and W. F. Lincoln (2010). The supply side of innovation: H-1B visa reforms and US ethnic invention. *Journal of Labor Economics 28*(3), 473–508.

Kerr, W. R. and M. Mandorff (2023). Social networks, ethnicity, and entrepreneurship. *Journal of Human Resources 58*(1), 183–220.

Kim, J. D. (2018). Is there a startup wage premium? Evidence from MIT graduates. *Research Policy 47*(3), 637–649.

Kolev, J., A. Haughey, F. Murray, and S. Stern (2022). Of academics and creative destruction: Startup advantage in the process of innovation. *NBER Working Papers*.

Moretti, E. (2004). Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data. *Journal of econometrics 121*(1), 174–212.

Parag Mahajan, Nicolas Morales, K. S. M. C. and A. Brinatti (2024). The impact of immigration on firms and workers: Insights from the h-1b lottery,. Technical Report 24-19, Working Papers, Center for Economic Studies, U.S. Census Bureau.

Peri, G., K. Shih, and C. Sparber (2015). Stem workers, h1b visas and productivity in us cities. *Journal of Labor Economics 33(3)*, S225–S255.

Piazza, A., S. Rendine, G. Zei, A. Moroni, and L. L. Cavalli-Sforza (1987). Migration rates of human populations from surname distributions. *Nature 329*, 714–716.

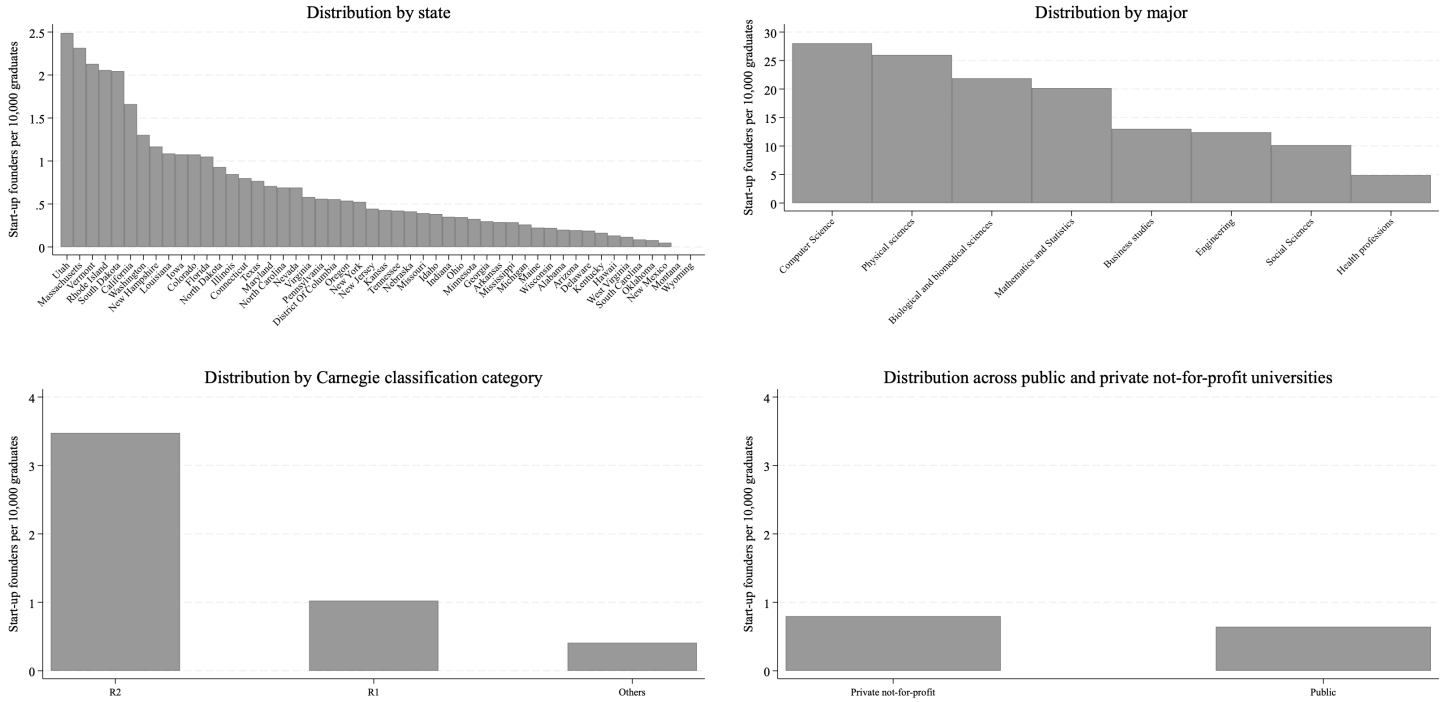
Stuen, E. T., A. M. Mobarak, and K. E. Maskus (2012, 09). Skilled Immigration and Innovation: Evidence from Enrolment Fluctuations in us Doctoral Programmes. *The Economic Journal 122*(565), 1143–1176.

Ulvestad, M. E. and S.-E. Skjelbred (2023). Escaping into a master's degree in times of crisis? master's degree applications and enrolment over the business cycle. *Economics of Education Review 94*, 102404.

Wang, X. (2018). Catering innovation: Entrepreneurship and the acquisition market. *Kenan Institute of Private Enterprise Research Paper* (18-27).

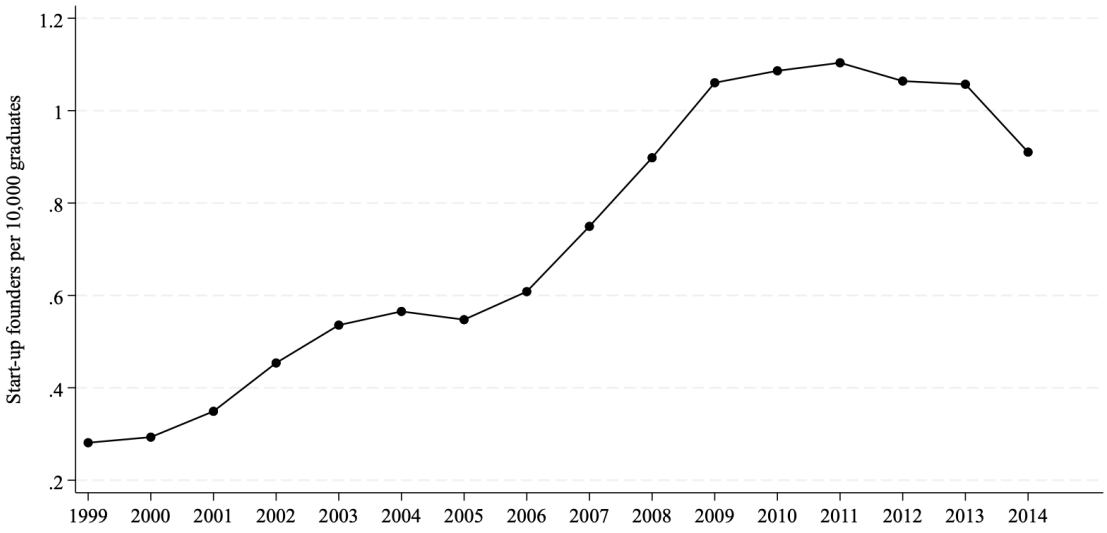
9 Figures

Figure 1: Distributions of start-up founders



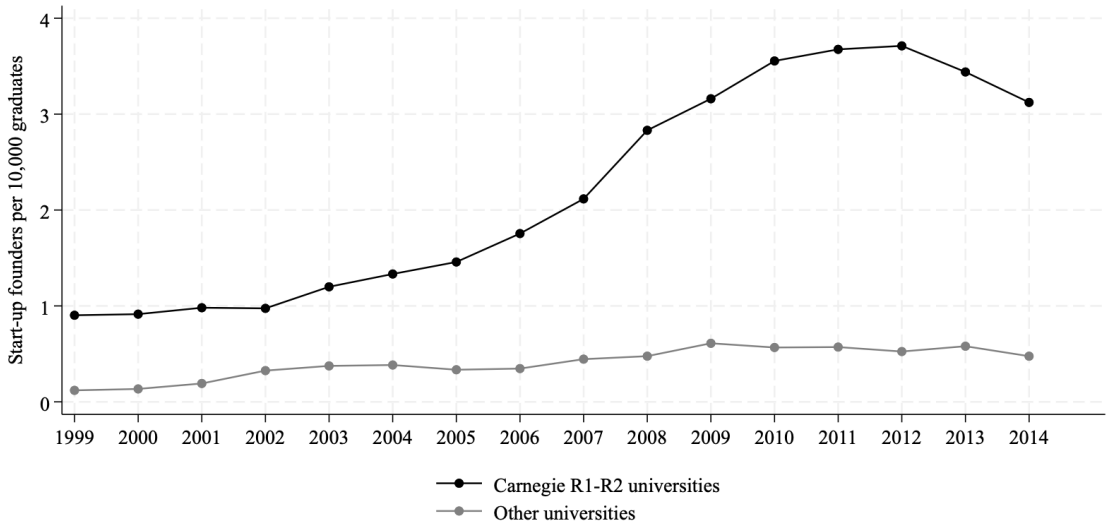
Notes: This graph plots the distribution of start-up founders by state, major, Carnegie classification category, and type of universities.

Figure 2: Master's Graduates who founded a startup within 5 years per 10,000 Master's graduates, 1999-2014



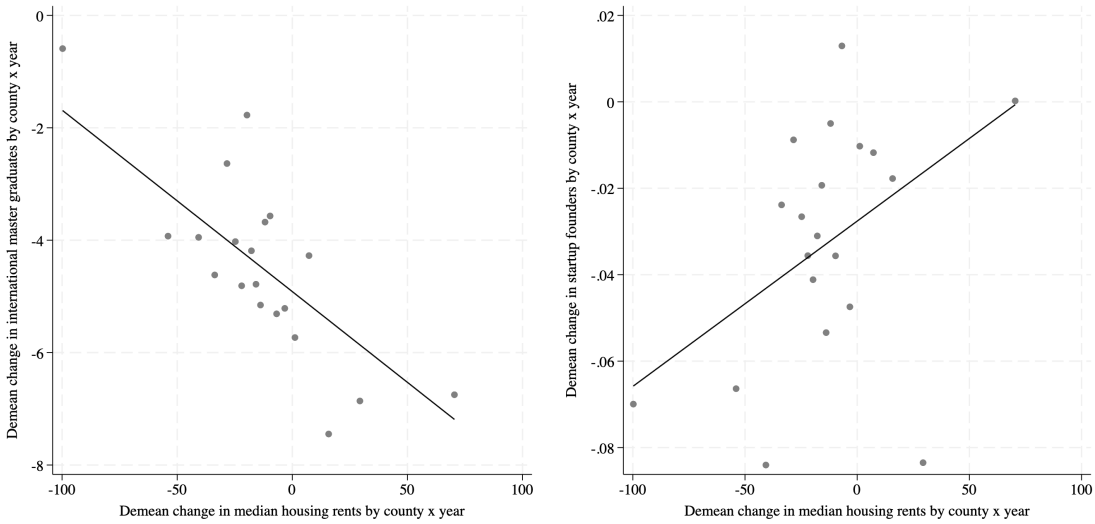
Notes: This graph plots the evolution of the moving average share of entrepreneurs per 10,000 master graduates.

Figure 3: Master's Graduate who founded a startup within 5 years after graduation per 10,000 Master's graduates, by type of university



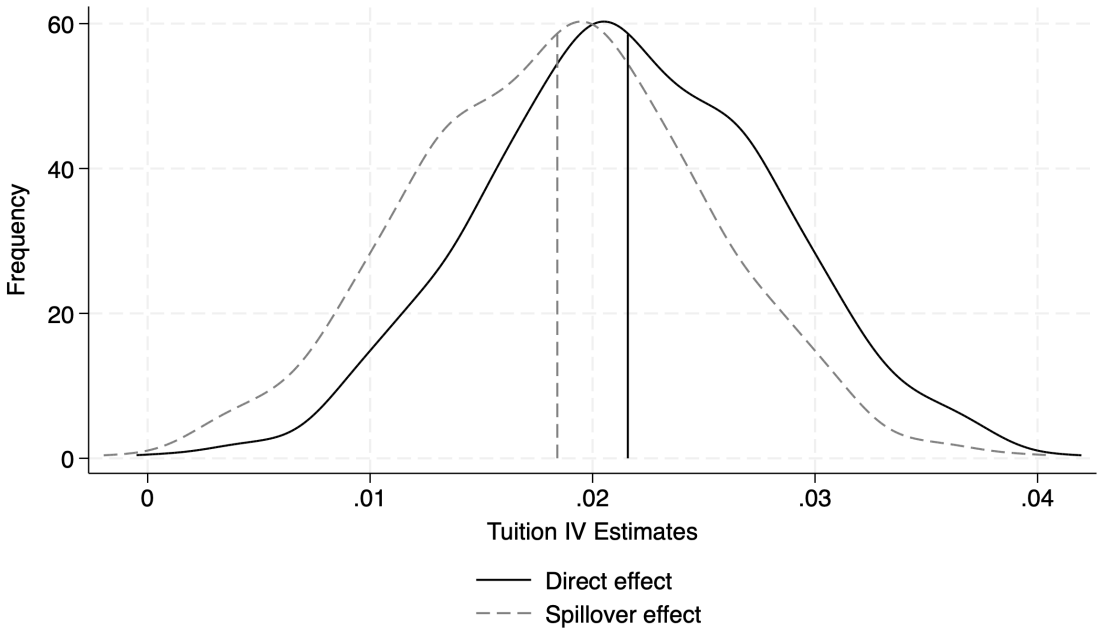
Notes: This graph plots the evolution of the share of start-up founders per 10000 Master graduates in Research universities (R1 and R2 Carnegie categories-black line) and other universities (grey line).

Figure 4: Relationship between housing rents, foreign enrollment and start-up creation



Notes: This graph presents two bin-scatter plots. The left panel plots the relationship between changes in the average number of international master graduates and changes in the median housing rents for 1-bedroom apartments by county and year. The right panel plots the relationship between changes in the average number of start-up founders and changes in the median housing rents for 1-bedroom apartments by county and year. All changes are demeaned by county and year.

Figure 5: Distribution of direct and spillover effects using false negatives and positives



Notes: This Figure plots the empirical distribution of the estimates of the direct (plain line) and spillover (dashed line) effects of foreign graduation on start-ups creation. The estimates are obtained using estimations on 1000 simulated foreign- and US-origin status of the founders. In each simulation, changes in the inferred status are implemented using the probability of false negatives and false positives provided by Forebears for the US.

10 Tables

Table 1: Representativeness of our sample

	Start-ups with info on founder's name (1)	Start-ups with info on founders' education (2)
Number of observations	25,710	13,697
Average firm's age	7.6128	7.2601
	(.0453)	(.042)
Average firm's funding (in millions of US dollars)	31.5543	31.9446
	(2.871)	(2.5194)
Share of firm's patenting	.0999	.1018
	(.0027)	(.0026)
Average firm's size category	1.9274	1.9065
	(.0112)	(.0102)
Share of foreign-origin founders	.5854	.5878
	(.0022)	(.0021)
Share of male founders	.8321	.8034
	(.0034)	(.0034)

Notes: This table compares two samples of start-ups included in Crunchbase data that differ by their missing values regarging founders' information on their educational history. The number of start-ups with information on founders' name is 25710, 13697 when conditioning also on information about university name, highest degree obtained and year of graduation.

Table 2: Summary statistics by university-cohort in both IV samples

	Tuition IV sample		SSIV sample	
	Mean	Standard Deviation	Mean	Standard Deviation
Total number of graduates	658.86	658.77	604.94	887.14
Number of international graduates	89.23	150.56	89.48	234.40
Number of US graduates	569.63	544.17	515.18	720.63
Share of international graduates	9.80	9.92	9.91	13.52
Share of international graduates in research universities	15.13	9.82	20.71	13.57
Share of international graduates in other universities	6.14	8.19	7.23	12.10
Number of entrepreneurs	0.09	0.42	0.27	1.50
Number of companies	0.10	0.51	0.34	1.88
Share of local start-ups	58.77	46.2	48.43	43.38
Share of start-ups surviving at least 3 years	98.73	8.24	98.02	11.21
Share of start-ups that have raised more than $25M	72.07	40.32	65.52	38.54
Share of patenting companies	2.50	13.96	2.31	11.41

Notes: This table presents summary statistics separated for each sample used in each one of the two IV identification strategies. The tuition IV sample includes 463 public universities classified in Doctoral universities or Masters' colleges according to the Carnegie classification of institutions of higher education over the period 1999-2015. The SSIV sample includes 447 private not-for-profit universities classified in Doctoral universities or Masters' colleges over the period 2004-2015.

Table 3: First stage IV estimates: tuition-based and shift-share instruments

Dependent variable:	Share of international graduates			
	(1)	(2)	(3)	(4)
Sample:	Public universities		Not-for-profit universities	
Standardized predicted residuals from equation (1)	-0.9161*** (0.1120)	-0.9155*** (0.1114)		
Shift-share IV			0.2665*** (0.0487)	0.2650*** (0.0485)
Log number of master graduates		0.0328 (0.4654)		0.7323* (0.4174)
Observations	6,443	6,443	4,307	4,306
University fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
First-stage F-statistic	66.93	67.5	29.95	29.8

Notes: Coefficients of the first-stage regressions of the IV estimation procedure. Instruments: Cols(1-2): innovations of non resident tuition fees; estimation period: 1999-2015; sample: public US universities. Cols (3-4) predicted shares of foreign graduates based on shift-shares computed for 10 most important origins of foreign students; estimation period : 2004-2015; sample: not-for-profit US universities

Table 4: Impact of instruments on native and foreign graduates

Dependent variable:	log(Number of graduates + 1)		log(Number of foreign graduates + 1)		log(Number of US graduates + 1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Standardized predicted residuals from equation (1)	-0.0202***		-0.0972***		-0.0087	
	(0.0059)		(0.0170)		(0.0059)	
Shift-share IV		-0.0001		0.0007**		-0.0007***
		(0.0001)		(0.0003)		(0.0002)
Observations	6,443	4,460	6,443	4,460	6,443	4,460
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Cols 1, 3 and 5: instrument innovation of non resident tuition fees; estimation period: 1999-2015; sample: public US universities. Cols 2, 4 and 6: shift-share IV. estimation period : 2004-2015; sample: not-for-profit US universities.

Table 5: Correlation between tuition based IV and previous shares of foreign graduates and enrollees

Dependent variable:	Predicted residuals from equation (1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of international graduates (4 years lag)	-7.0108 (6.7590)	-	-	-	-	-
Share of international graduates (5 years lag)	-	-4.4855 (6.6598)	-	-	-	-
Share of international graduates (6 years lag)	-	-	2.3289 (6.5744)	-	-	-
Share of international enrollees (4 years lag)	-	-	-	3.8075 (13.9184)	-	-
Share of international enrollees (5 years lag)	-	-	-	-	10.3453 (14.8603)	-
Share of international enrollees (6 years lag)	-	-	-	-	-	16.8813 (16.4458)
Observations	7,212	7,193	7,166	7,246	7,227	6,797
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table provides regression coefficient of innovations to non resident tuition fees on a set of covariates. Estimation period: 2003-2015. Sample : public US universities.

Table 6: Validation of the Shift-share: correlation between predicted number of foreign graduates and previous values

Dependent variable:	Change in the predicted number of international graduates from SSIV between 2004 and 2015					
	(1)	(2)	(3)	(4)	(5)	(6)
Change in the total number of graduates between 1995 and 2003	0.0103	-	-	-	-	-
	(0.1699)					
Change in the number of foreign graduates between 1995 and 2003	-	2.9314***	-	-	-	-
		(0.6903)				
Change in the number of US graduates between 1995 and 2003	-	-	-0.1194	-	-	-
			(0.2406)			
Change in the total number of entrepreneurs between 1995 and 2003	-	-	-	167.1870	-	-
				(121.7215)		
Change in the number of foreign-origin entrepreneurs between 1995 and 2003	-	-	-	-	194.7596	-
					(171.7453)	
Change in the number of US-origin entrepreneurs between 1995 and 2003	-	-	-	-	-	514.3149
						(343.9719)
Observations	300	353	353	353	353	353

This table provides estimates of regressions of predicted number of foreign graduates for the year 2003 using the shift-share IV strategy on a set of covariates. Estimation period: 1995-2003. Sample: not-for-profit US universities.

Table 7: Share of Foreign Graduates and Start-ups by University-Cohort: Basic OLS, IV Specifications

Dependent variable:	Number of entrepreneurs		Number of companies	Number of local companies	Number of surviving companies	Number of successful companies	Number of patenting companies
	OLS	IV	IV	IV	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Residual tuition IV							
Share of international master graduates	0.0003	0.0413***	0.0397***	0.0382***	0.0382**	0.0250**	0.0054*
	(0.0007)	(0.0139)	(0.0153)	(0.0124)	(0.0150)	(0.0118)	(0.0028)
Observations	6,443	6,443	6,443	6,443	6,443	6,443	6,443
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep Var	.0856	.0856	.1044	.0632	.1026	.0735	.0029
SD Dep Var	.416	.416	.5131	.3757	.5063	.3812	.0569
First-stage F-statistic		66.93	66.93	66.93	66.93	66.93	66.93
Shift share IV							
Share of international master graduates	0.0018	0.0739***	0.0605***	0.0259**	0.0591***	0.0194*	0.0017
	(0.0014)	(0.0216)	(0.0211)	(0.0121)	(0.0206)	(0.0117)	(0.0018)
Observations	4,307	4,307	4,307	4,307	4,307	4,307	4,307
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep Var	.2666	.2666	.3401	.1571	.3323	.2131	.0092
SD Dep Var	1.5006	1.5006	1.8789	1.1293	1.837	1.168	.1091
First-stage F-statistic		29.95	29.95	29.95	29.95	29.95	29.95

Notes: OLS and IV estimations of equation (1). Upper panel: instrument : innovations to non resident tuition fees (equation 2); estimation period : 1999-2015; sample: public US universities . Lower panel : shift-share IV based on 10 most popular origins of foreign graduates; estimation period : 2004-2015; sample: not-for-profit US universities.

Table 8: Share of Foreign Graduates and Start-up by University-Cohort: Controlling for University-Specific period effects and trends

Dependent variable:	Number of entrepreneurs		Number of companies		Number of local companies		Number of surviving companies		Number of successful companies		Number of patenting companies	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Residual tuition IV												
Share of international master graduates	0.0631*** (0.0215)	0.0412*** (0.0140)	0.0622*** (0.0235)	0.0395** (0.0153)	0.0629*** (0.0197)	0.0381*** (0.0125)	0.0606*** (0.0231)	0.0381** (0.0151)	0.0388** (0.0180)	0.0249** (0.0118)	0.0085* (0.0045)	0.0054* (0.0028)
Observations	5,546	6,444	5,546	6,444	5,546	6,444	5,546	6,444	5,546	6,444	5,546	6,444
University x period fixed effects	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
University x time trend fixed effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	46.68	66.16	46.68	66.16	46.68	66.16	46.68	66.16	46.68	66.16	46.68	66.16
Shift-share IV												
Share of international master graduates	0.0707** (0.0276)	0.0735*** (0.0217)	0.0561** (0.0279)	0.0599*** (0.0212)	0.0285 (0.0178)	0.0256** (0.0121)	0.0516* (0.0271)	0.0584*** (0.0207)	0.0132 (0.0155)	0.0190 (0.0117)	0.0027 (0.0029)	0.0017 (0.0018)
Observations	3,050	4,314	3,050	4,314	3,050	4,314	3,050	4,314	3,050	4,314	3,050	4,314
University x period fixed effects	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
University x time trend fixed effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	29.01	29.34	29.01	29.34	29.01	29.34	29.01	29.34	29.01	29.34	29.01	29.34

Notes: OLS and IV estimations of equation (1). Upper panel: instrument: innovations to non resident tuition fees (equation 2); estimation period: 1999-2015; sample: public US universities. Lower panel : shift-share IV based on 10 most popular origins of foreign graduates; estimation period : 2004-2015; sample: not-for-profit US universities.

Table 9: Elasticity of Start-ups to Foreign Graduates: IV estimates

Dependent variable:	Number of entrepreneurs	Number of companies	Number of local companies	Number of surviving companies	Number of successful companies	Number of patenting companies
	(1)	(2)	(3)	(4)	(5)	(6)
log number of international graduates	0.3896***	0.3740**	0.3597***	0.3604**	0.2356**	0.0512*
	(0.1436)	(0.1545)	(0.1278)	(0.1512)	(0.1158)	(0.0275)
Observations	6,443	6,443	6,443	6,443	6,443	6,443
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep Var	.0858	.1047	.0633	.1029	.0736	.0029
SD Dep Var	.4165	.5137	.3762	.5069	.3816	.057
First-stage F-statistic	32.58	32.58	32.58	32.58	32.58	32.58

Notes: Instrument: innovations of non resident tuition fees (equation 2; estimation period: 1999-2015; sample: public US universities. Local companies: firms located in same state than graduation state. Surviving companies: start-ups still operating 3 years after creation. Successful companies: companies having raised at least 25 Million USD in the first 3 years. Patenting companies: companies having received at least one patent in the first 3 years .

Table 10: Heterogeneity in the effect of foreign graduation

Dependent variable:	Number of entrepreneurs					
	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	STEM majors	non-STEM majors	Top 5 majors	Other majors	Research universities	Other universities
Share of International master graduates	0.0248***	0.0280*	0.0151*	0.0260**	0.1014***	-0.0038
	(0.0096)	(0.0164)	(0.0083)	(0.0111)	(0.0301)	(0.0033)
Observations	6,219	6,384	6,437	6,437	2,621	3,822
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	30.31	22.84	17.49	64.50	38.68	31.22

.

Table 11: Direct and spillover effects of foreign graduation: tuition IV

Direct effects

Dependent variable:	Number of foreign-origin entrepreneurs (1)	Number of foreign-origin companies (2)	Number of local foreign-origin companies (3)	Number of surviving foreign-origin companies (4)	Number of successful foreign-origin companies (5)	Number of patenting foreign-origin companies (6)
Share of international master graduates	0.0265***	0.0218***	0.0225***	0.0205***	0.0150**	0.0029*
	(0.0085)	(0.0082)	(0.0064)	(0.0079)	(0.0061)	(0.0016)
Observations	6,443	6,443	6,443	6,443	6,443	6,443
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	66.93	66.93	66.93	66.93	66.93	66.93

Spillover effects

Dependent variable:	Number of US-origin entrepreneurs (1)	Number of US-origin companies (2)	Number of local US-origin companies (3)	Number of surviving US-origin companies (4)	Number of successful US-origin companies (5)	Number of patenting US-origin companies (6)
Share of international master graduates	0.0157**	0.0110*	0.0063	0.0111*	0.0035	0.0010*
	(0.0062)	(0.0058)	(0.0048)	(0.0058)	(0.0043)	(0.0005)
Observations	6,443	6,443	6,443	6,443	6,443	6,443
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	66.93	66.93	66.93	66.93	66.93	66.93

Notes: Instrument: innovations to non resident tuition fees (equation 2); estimation period: 1999-2015; sample: public US universities.

Table 12: Effect on cofounded start-ups

Dependent variable:	Number of companies	Number of companies with several founders	Number of companies with several founders from same school	Number of companies	Number of companies with several founders	Number of companies with several founders from same school
	(1)	(2)	(3)	(4)	(5)	(6)
	Tuition IV			Shift-share IV		
Share of international master graduates	0.0397***	0.0360***	0.0099**	0.0605***	0.0443***	0.0204**
	(0.0153)	(0.0099)	(0.0049)	(0.0211)	(0.0136)	(0.0083)
Observations	6,443	6,443	6,443	4,307	4,307	4,307
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
widstat	66.93	66.93	66.93	29.95	29.95	29.95

Notes: Dependent variable. Cols (1) and (4): total number of start-ups. Col. (2) and (5): Number of start-ups with at least 2 founders. Col (3) and (6): Number of start-ups with 2 co-founders from the same school. Cols (1-3) : estimation period: 1999-2015; sample: public US universities. Cols (4-6): estimation period : 2004-2015; sample: not-for-profit US universities.

Appendix

11 Shift share IV based on the 10 main origins of foreign students

We use the F1-FOIA data to construct the foreign born students enrolled from each specific country of origin (Denoted as o) as share of the total students enrolled. We also use the F1-FOIA data to measure the foreign students in each origin country in each university in the initial year.

We denote as t_0 the initial year which will be the year 2001. Then the construction of the IV and the analysis will be for years $t = 2002, 2003, ...2013$.

u is the index for individual universities. The universe of universities considered is U

o is the index for countries of origin . The universe of countries of origin is denoted as O.

t denotes years.

The notation $N^{F1}_{u,o,t}$ indicates the number of enrolled students, N, in university u from country of origin o in year t, obtained from source $F1$, where "source" is the dataset of $F1$ obtained using the FOIA.

We describe the procedure to construct the shift-share IV step by step. The procedure should be done separately for "Master" and for "Bachelor" students, obtaining separate cells for each and then we pool them in the regression. Therefore we will not use the indicator "degree", as the same procedure will be done separately.

11.1 Defining and constructing the "shift"

First calculate, the total number of foreign students enrolled (in the US) from country of origin o in year t by aggregating those data across universities as follows:

$$N_{o,t}^{F1} = \sum_{u \in U} N_{u,o,t}^{F1} \tag{4}$$

Then we aggregate further to obtain the total number of foreign students (from any origin, in the US) in year t, based on the $F1$ data, as follows:

$$(NFOR)_{t}^{F1} = \sum_{o \in O} N_{o,t}^{F1} \tag{5}$$

11.2 Constructing the "share"

The number of initial foreign-born students from country o enrolled in university u in year t_0 as share of total students from country o (so that the sum across universities in each year, for each country of origin is 1, and we will use this to distribute the total foreign-born from country o across universities) is as follows:

$$(Share)_{u,o,t_0}^{F1} = \frac{N_{u,o,t_0}^{F1}}{N_{o,t_0}^{F1}} \tag{6}$$

11.3 The "shift-share" IV

The imputed number of foreign-born students from country o enrolled in each university u in Year t using the initial share and the yearly "flow" of enrolled students will be as follows:

$$(\hat{N})_{o,u,t} = (Share)_{u,o,t_0}^{F1} * (N)_{o,t}^{F1} \tag{7}$$

And the total number of imputed foreign students overall for university u and year t will be:

$$(\hat{N})_{u,t} = \sum_{o \in O} (\hat{N})_{o,u,t} \tag{8}$$

This variable, $(\hat{N})_{u,t}$, which varies at the university and year level, is what we will use as Instrument.

If we want to try foreign-students as share of enrolled as explanatory variable (instead of their log number or simple number as we do now), then we will instrument

the share of foreign students in a school, obtained from $IPEDS$ with the following imputed share that uses the shift-share imputation for foreigners:

$$\hat{(Share)}_{u,t} = \frac{(\hat{N})_{u,t}}{(\hat{N})_{u,t} + (NUS)_{u,t_0}} \tag{9}$$

where $(NUS)_{u,t_0}$ is the total number of Native US students enrolled in university u in year t_0 (2000, 2001), so as to keep the number of natives fixed and unaffected by foreign flows in the IV.

12 Additional Tables

Table 13: Universities with largest number of entrepreneurs per 10,000 master graduates

University	Share of entrepreneurs per 10,000 master graduates
Stanford University	37.67
University of Washington-Tacoma Campus	34.92
Massachusetts Institute of Technology	33.59
Babson College	20.27
Harvard University	19.59
Louisiana State University-Shreveport	19.07
Louisiana State University Health Sciences Center-Shreveport	18.18
Kairos University	14.88
Albany Medical College	13.44
University of Pennsylvania	13.36
California College of the Arts	13.24
Maharishi International University	12.87
University of Chicago	11.70
Hult International Business School	11.55
Santa Clara University	10.39
Rhode Island School of Design	9.50
Lasell University	9.26
Vermont Law and Graduate School	8.96
Carnegie Mellon University	8.52
University of Vermont	8.41
Duke University	7.82
University of California-Berkeley	7.71
Rice University	7.67
Vanderbilt University	7.62
University of California-Los Angeles	7.47
Fashion Institute of Technology	6.94
University of San Diego	6.92
Dartmouth College	6.85
Brown University	6.64
Lynn University	6.11

Table 14: Preliminary regressions of the tuition IV procedure

Dependent variable: :	Difference between in-state and out-of-state tuitions		
	(1)	(2)	(3)
Share state appropriations	183.2	1,012***	2,833***
	(133.0)	(125.5)	(129.4)
Number of international graduates (1 year lag)	-	15.33***	3.859***
		(0.216)	(0.272)
University revenues (in dollars)	-	-	3.60e-06***
			(5.18e-08)
Revenues per student (in dollars)	-	-	-0.00110***
			(0.000217)
Instruction expenses per student (in dollars)	-	-	-0.000747
			(0.000508)
All expenses per student (in dollars)	-	-	0.00109***
			(0.000246)
Constant	5,025***	4,289***	3,426***
	(45.40)	(43.88)	(46.09)
Observations	38,174	38,151	34,382
R-squared	0.119	0.222	0.323
Year FE	Yes	Yes	Yes
Degree FE	Yes	Yes	Yes

Table 15: Composition effects of foreign graduates: accounting for size of native cohort

Dependent variable:	Number of entrepreneurs	Number of companies	Number of local companies	Number of surviving companies	Number of succesful companies	Number of patenting companies
	(1)	(2)	(3)	(4)	(5)	(6)
Tuition IV						
Share of international master graduates	0.0415***	0.0399***	0.0383***	0.0385**	0.0249**	0.0055*
	(0.0140)	(0.0154)	(0.0125)	(0.0151)	(0.0118)	(0.0028)
log number of master graduates	-0.0088	-0.0096	-0.0075	-0.0094	0.0040	-0.0010
	(0.0205)	(0.0204)	(0.0187)	(0.0198)	(0.0137)	(0.0027)
Observations	6,443	6,443	6,443	6,443	6,443	6,443
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	67.50	67.50	67.50	67.50	67.50	67.50
Shift-share IV						
Share of international master graduates	0.0744***	0.0609***	0.0261**	0.0594***	0.0196*	0.0017
	(0.0218)	(0.0213)	(0.0122)	(0.0208)	(0.0118)	(0.0018)
log number of master graduates	-0.0795**	-0.0814**	-0.0346*	-0.0760**	-0.0198	-0.0032
	(0.0400)	(0.0374)	(0.0209)	(0.0365)	(0.0178)	(0.0022)
Observations	4,306	4,306	4,306	4,306	4,306	4,306
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	29.80	29.80	29.80	29.80	29.80	29.80

Notes: IV estimations of equation (1) supplemented with size of cohort. Upper panel: instrument: innovations to non resident tuition fees (equation 2). Estimation period: 1999-2015. Sample: public US universities. Lower panel : shift-share IV based on 10 most popular origins of foreign graduates. Estimation period: 2004-2015. Sample: not-for-profit US universities.

Table 16: Direct and spillover effects of foreign graduation: shit-share IV

Direct effects

Dependent variable:	Number of foreign-origin entrepreneurs (1)	Number of foreign-origin companies (2)	Number of local foreign-origin companies (3)	Number of surviving foreign-origin companies (4)	Number of successful foreign-origin companies (5)	Number of patenting foreign-origin companies (6)
Share of international master graduates	0.0535***	0.0536***	0.0213***	0.0507***	0.0238***	0.0014
	(0.0146)	(0.0150)	(0.0073)	(0.0144)	(0.0084)	(0.0011)
Observations	4,307	4,307	4,307	4,307	4,307	4,307
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	29.95	29.95	29.95	29.95	29.95	29.95

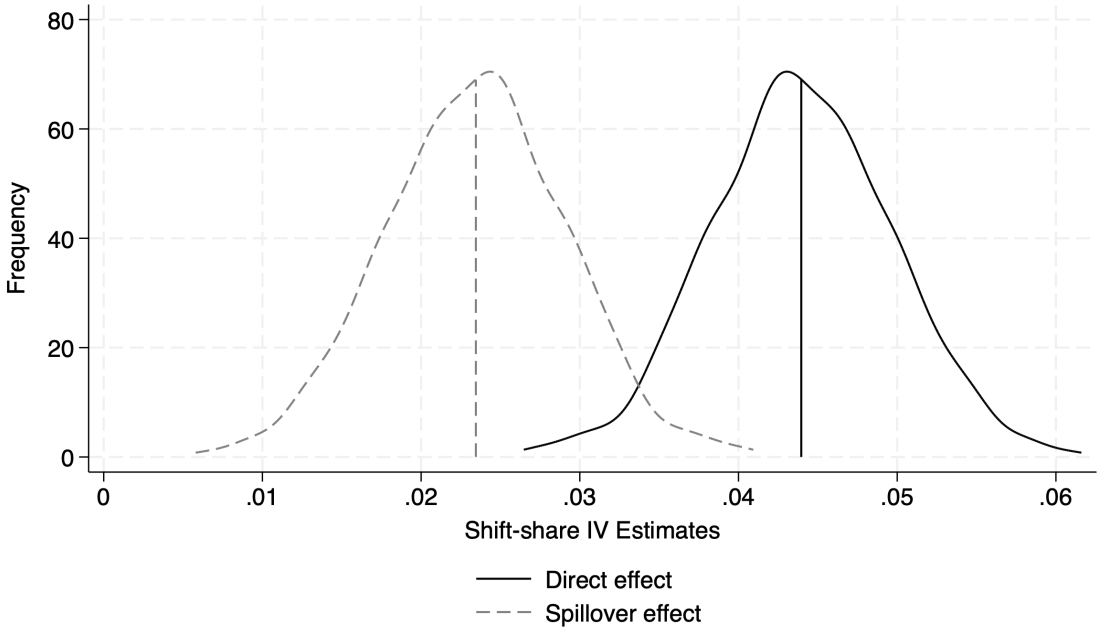
Spillover effects

Dependent variable:	Number of US-origin entrepreneurs (1)	Number of US-origin companies (2)	Number of local US-origin companies (3)	Number of surviving US-origin companies (4)	Number of successful US-origin companies (5)	Number of patenting US-origin companies (6)
Share of international master graduates	0.0191***	0.0125**	0.0065**	0.0119**	0.0041	-0.0000
	(0.0067)	(0.0058)	(0.0032)	(0.0056)	(0.0035)	(0.0006)
Observations	4,307	4,307	4,307	4,307	4,307	4,307
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	29.95	29.95	29.95	29.95	29.95	29.95

Notes: Shift-share IV based on 10 most important origins of foreign students; Estimation period: 2004-2015; sample: not-for-profit US universities.

13 Additional Figures

Figure 6: Distribution of direct and spillover effects estimated with the shift-share IV using false negatives and positives.



Notes: This graph provides the empirical distributions of the estimates of the direct (dashed line) and spillover effects (plain line) using the Shift-share IV. The distribution is obtained using 1000 simulations assigning randomly false negatives and positives of origin assignment using the probabilities provided by Forebears for the US.