The Contribution of Foreign Master's Students to US Start-Ups *

Michel Beine (University of Luxembourg, IZA and CES-Ifo)
Giovanni Peri (UC Davis, NBER and IZA)
Morgan Raux (Aix-Marseille School of Economics, and IZA)

August 1, 2025

Abstract

We estimate the effect of an increase in the share of foreign-born master's students on new start-ups in the US created by master's graduates. To identify causality we use idiosyncratic variation of university-cohort-specific out-of-state fees, generating differential foreign students' enrollment and we corroborate the results using a shift share instrument. For an increase by ten percentage points of foreign student share in a Master's cohort, 0.7 to 0.9 additional start-ups are created by that cohort within 5 years of graduation. Conservatively, 23 to 38% of the effect is due to positive spillover on start-up founded by US-origin Master's graduates.

JEL Classification: F22, I23, I25, M13.
Keywords: Foreign Students, Entrepreneurship, Startups, Co-founding.

*We are grateful to Anna-Maria Mayda, Frederic Docquier, Ian Stuhler, Richard Freeman, Justin Wiltshire and to participants to the Harvard Seminar on Economics of Science, the First European Conference on Macroeconomics of Immigration, to the OECD Migration conference, and to Seminars' participants in UNAM (Mexico City), Victoria University, University of Bari, Georgetown University, Southern Methodist University, University of Linz and University of Ghent, University of Trier for their useful comments. We thank Alessandro Caiumi and Beau Bressler for their help in editing the paper. We do not have any financial interest or conflict of interest to disclose. Michel Beine, email: michel.beine@uni.lu; Giovanni Peri, email: gperi@ucdavis.edu; Morgan Raux, email: morgan.raux@univ-amu.fr.

1 Introduction

The creation of innovative firms that bring to market a new product or service, 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-leading institutions in technology, science, and business[2]. A Forbes article in 2019, entitled "New Fortune 100 CEO Study: The Top Graduate Schools Attended By Fortune 100 CEOs" [3] 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. US Universities are creators of human capital and their graduates translate it into 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]. Specifically, we estimate whether a higher share of foreign students among master graduates increases

[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] 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.

[3] (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)

[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 an excellent sample for our identification strategy and to test our main hypothesis.

the rate at which that group creates start-ups within 5 years of graduation.

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), patented innovation (Kerr and Kerr, 2016; Bernstein et al., 2022) or, more generally, on 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 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 2014[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 Master's students in US universities 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 a Master graduating cohort in a University and year.

Second, we merge these data with information on US start-ups, i.e. new companies that brought to market a new product or service between 1999 and 2019. 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, includes the name, the University, and the year of graduation of the founder(s) of each start-up. They also include the year the firm

[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

was founded, its location and some of the economic characteristics of the firm, such as the amount of capital raised and its sector. This information allows us to construct the count of firms founded within 5 years after graduation by Master's graduates for each university and cohort of graduation. The analysis covers the cohorts who graduated during the period 1999-2014.

Then, still using IPEDS, we construct the average out-of-state and in-state average tuition for Master's students for each school and year. Such tuition is differentiated between in- and out-of-state students in public universities, which is the sample we use for this approach (Master's tuition is the same for in and out of state in private universities) The difference between out-of-state and in-state tuition at the time of enrollment, after controlling for school fixed effects and time varying school characteristics, is used as an instrumental variable for the sample of public universities. 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 share of foreign students in the entry Master's cohort[6].

Additionally, we implement our estimates on the sample of private universities while adopting 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 approach, developed in the immigration literature to capture variation of inflows across local areas, by a series of papers, including Card (2009) and Peri et al. (2015). The property and validity of such IV approach were formalized by Goldsmith-Pinkham et al. (2020). In our case, using data on F-1 (student) visas that provide the origin of the foreign students, we measure the share of students from ten main countries of origin enrolled in each University as of 2001. We then use these shares to distribute the flows of F1-visa students from each of those countries in the 2004-2014 period across schools. The variation in foreign Master's students generated by this IV hinges on the aggregate change in flow size for students from each nationality of origin, interacted with the initial share (network) of those foreigners in each university. As in the case of tuition-based IV we test that these IV are not correlated with pre-period inflow of foreign students and measures of start-up intensity. We find remarkably similar estimated magnitudes across

[6]We used a similar approach 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, we subject this approach to additional testing and validity checks and we complement it with the shift-share approach.

the two alternative IV strategies, consistent with the hypothesis that an increase in foreign student share has a causal impact on the creation of start-up companies by master graduates from that Master's cohort.

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 of international graduates in that same university-cohort cell. We use as instruments, in the public university sample, the "idiosyncratic component of the difference" between out-of-state and in-state tuition, i.e., the residual after controlling for a set of time-varying proxies of universities' quality. In the private University sample, instead, we use the school-specific shift-share predictor of foreign students as IV. The main results from the IV estimates show that an increase by 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 between 0.7 and 0.9 in the number of start-up companies founded by that cohort. This represents a significant increase, equal to approximately two standard deviations of start-up creators per cohort in public universities.[7] This effect is similar in magnitude using either the tuition IV or the shift-share IV.

We subject the IVs to a series of validity checks, including checking the correlation with lead (rather than lagged) foreign enrollment and start-up intensity, the correlation with enrollment of US-born (which should not be affected), and we find that most of them are not significant. The tuition IV passes the test a bit more convincingly and it is a stronger IV. The shift-share IV shows a correlation with the pre-period foreign graduate changes. Nevertheless, the two identification strategies deliver remarkably similar estimates while using different samples and relying on very different sources of exogenous variation. Such evidence 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 capital in the first 3 years than the median US start-up of same age, and those that submitted at least one patent application during the same time window (a

[7]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.

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 foreign Master's graduates contributed to the creation of *successful start-up companies*, not just more of them.

We then analyse the part of this effect which is due to the spillovers of foreign students on start-up creation by native entrepreneurs, as well as the mechanisms driving such effect. As the data from Crunchbase do not identify the place of birth of start-up creators (only their full name and degree, university and year) their country of origin is inferred from their first and last names based on an algorithm developed by Forebears.io and applicable to first and last names all over the world. This algorithm, based on the world distribution of first and last names assesses the probability of being a citizen of each country in the world. Using it we distinguish between US-origin and foreign-origin start-up creators in each master-school-cohort. Hence, we can estimate whether in school-cohorts with a larger share of foreign graduates–measured in the IPEDS data– also people of (likely) US-origin were more likely to create start-ups. This would be consistent with the existence of entrepreneurial spillovers within a Master's cohort. The algorithm could generate some "false attributions" most likely considering second or third generation, US-born offspring of immigrants, as foreign-born. Such a most common misattribution is likely to bias down the estimated spillovers. Hence, in a subsample of start-up founders for which we do an extensive online search to identify their exact place of birth, we estimate the frequency of such "false attributions" and we simulate their consequences on the estimates of the spillover effects.

Based on the inferred US-origin status of start-up founders our analysis reveals that spillover effects play a significant role in explaining the creation of start-ups in the US. Depending on the methodology, and using the probabilistic attribution of Forebears, between 23 and 38% of the positive start-up effects of foreign Master's students is a spillover on US-born students. Accounting for the "false-attribution" error, the spillovers on US-born start-up founders could be as large as 40% to 75% .

Finally, we present evidence suggesting that co-founding between US and foreign-born among former students increases substantially when a large share of foreign students is in the Master-cohort and can contribute to accounting for the positive effect. We also find evidence that the share of foreign students increases the "variety" of start-ups generated by that cohort as measured by a list of keywords describing their products and services. This is consistent with the idea that interaction with foreign students expands the range of products and characteristics of the start-ups.

The rest of the paper proceeds as follows. Section 2 frames the contribution of the present paper within the 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 describes the empirical approach to identification and estimation and performs several 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 probabilities and growth rates higher than those created by natives.[8] Jin et al. (2025) also document that start-ups cofounded by native-migrant teams outperform other start-ups in the US.

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 neg-

[8] 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.

ative 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 emigrants analyzing the decline in firm creation due to a large wave of emigration of potential entrepreneurs 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.[9]

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 full 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

[9]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.

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 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 1999-2019, 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 and validate it comparing it to an alternative IV approach. We are the first to attribute part of this effect to spillovers on start-ups created by US-origin graduates, and to analyze co-founding and increased product/service variety as potential channel. 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 their 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 in each year. 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 variables of interest are the size of the graduation cohort, namely the number of Master's students graduating in each university and year over the 1999-2014 period, and among them, the share of foreign students. 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 2014, our period of analysis). Those fees are the average "on the books" fees for Master's students of the university in each academic year. For public universities, these 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, driven by local pressure to differentiate out-of-state fees, or by accounting decisions. We exploit this idiosyncratic variation, across public universities 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. For private universities, the tuition fees are the same for in-state and out-of-state and we do not pursue this IV approach, but a shift-share IV, described below.

3.2 Founders of Startups

The second key source is a database obtained from Crunchbase, which includes information on start-up firms, their founders, and several of their characteristics. [10] 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.

Crunchbase collects data through its partnership that, as of 2023, involved tens of thousands of 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 back-filled. Hence, information prior to that year includes missing data and becomes sparser as we go back in time. Our analysis focuses on cohorts that graduated between 1999 and 2014 and we consider start-ups created up to 5 years after graduation. Hence we span the period 1999-2019. Information from Crunchbase includes the firm's starting date, its sector and location, the names of its founders, their highest degree (Bachelor's, Master's or PHD), the University and major of graduation, as well as the date of the degree award. The data also includes 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, whether 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.

In this paper the main function of the Crunchbase data is to count start-up founders

[10] Data are available at Crunchbase.com

with a Master's degree across all universities and cohorts and to produce a "start-up entrepreneurship rate" across them.[11] 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.[12] Therefore, this threshold excludes the least successful 40% of them. We then use the information on the amount of venture capital investments raised over time to identify companies that raised more capital within their first three years of operation than the median US start-up of the same age. 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 second part of our analysis, we estimate the spillover effect of foreign students on US-origin start-up creators. To that aim, we need to impute the foreign or US-origin of start-up creators from their first and last names. We follow and refine an approach adopted by several previous papers. 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

[11] 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.

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

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 across 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%.[13] 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), rather than the US, which can generate false positive cases, i.e. US-born individuals identified as foreigners. Of course, there is also the possibility of false negatives, where foreign entrepreneurs are mistakenly classified as native ones by the algorithm. We will use this algorithm to identify US-origin start-up creators, to isolate the spillover effect of foreign-born students on native-born start-up rates.

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 is from each specific country. In our main analysis, 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 a cell by the probability of being of US or foreign origin. As a result, we obtain a probabilistic count of with foreign- or US-origin in each cohort, and this allows us to separate the effect of the foreign share in the Master cohorts on the overall number of start-ups, from the spillover effect on the number of US-origin start-ups in the cohort.[14].

How accurate is the inference procedure based on Forebears? To assess this, we conduct a validation exercise using a sub-sample of about 1500 start-up founders identified in Crunchbase, for whom we collect, through a very intensive process, very accurate information about their foreign-born or US-born status. This process consisted of performing

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

[14]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 measure of US-origin start-ups.

extensive internet searches for each start-up founder in the sample, looking for their personal websites, their LinkedIn profiles, other sites where they are featured or mentioned, YouTube videos in which the person describes their career and start-up, and similar documents. Usually, combining information on the location of their high school, their accent on YouTube, their name and their history and highlights we reached an accurate determination of whether they were US-born. In several cases we found explicit mention of their birthplace, while in other cases we inferred from childhood, K-12 school, accent and other references. Currently we were able to do this research only on 1500 start-up founders due to resources constraints. This procedure remains valuable for evaluating the effectiveness of the Forebears algorithm and can be expanded in the future.

The validation exercise is shown in Table 1 and in it we use the foreign and US-born status based on the most likely country of origin, as predicted by Forebears in a dichotomous way, and we compare those with the US or foreign birth-status of those entrepreneurs, as predicted by the holistic procedure described above. Considering the assignment to US and Foreign-born obtained with the holistic procedure as "correct", the table shows the proportion of false positives and false negatives when applying the Forebears algorithm to this sub-sample. Overall, the algorithm performs reasonably well in identifying foreign-born and US-born status as the probability of correct assignment (on the main diagonal of the table) is 76% of the total sample, with 24% mis-attributed. As anticipated, given the significant role of second-generation immigrants in the U.S. we find more false positives (U.S.-born classified as foreign-origin), equal to about 19% of the total, than false negatives (Foreign-born classified as US-origin), only 5% of the total. This type of error suggests that our estimates of the size of spillovers' effect, which we will present in the later part of the paper, are likely to undercount spillovers on US-born and therefore we can be underestimating the impact of foreign graduation on US-origin entrepreneurship. We will use the percentages of these two type of errors from Table 1 to estimate a confidence interval on the spillover coefficient, in section 7.

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 1999-2014. 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 2 shows summary statistics for different characteristics of the

Crunchbase sample of start-ups. Column 1 includes start-ups whose founders' names are non-missing, while column 2 restricts the sample to start-ups with all the needed information on the name and education of their founders, including degree, school and year of graduation. We see a strong similarity in the average of most characteristics explored, such as average age, average share of co-founded firms, average number of employees, and share of founders by gender. Also, importantly, we observe a similar share of firms (likely) founded by a foreign-born individual. We observe differences in the average size of funding and in the share of firms owning a patent. Nevertheless, the overall 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. Hence, the estimates obtained from the available sub-sample should be applicable to the full Crunchbase sample.

The merged sample examined in column 2 of Table 2 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 ones[15]. For public universities, we use the idiosyncratic variation in the out-of-state tuition net of the 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 distinction in the tuition fees, we use the initial 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 3. 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, albeit slightly higher in the tuition IV sample. 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 (15% 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 (0.46 founders per cohort in private versus 0.17 in public, and 0.51 firms

[15]We exclude for-profit universities from our sample.

created versus 0.19).

In terms of average characteristics of start-ups, graduates of public university Master's programs have a slightly larger probability of founding a company in the same state where they graduated (56% versus 50% 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.

The distributions in terms of funding turn out to be skewed for both samples, with 33% (39%) of the start-ups in the public (private) university sample raising more funds than the median within the first 3 years. Rates of patenting firms within 3 years of their creation are very similar across the two samples, with respectively 3.34% and 3.48% of patenting start-ups. 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 three panels in Figure 1 provide additional characterization of the start-up rates in different subgroups, and they are obtained using jointly the public and private university samples. They show the distribution of start-up rates (per 10,000 graduates) by Master's major, by Carnegie classification category of the university, and by country of origin. The top left Panel shows the start-up rates across Master's majors, including only those with a significant number of graduates. STEM and Business Master's programs are those with the highest rates of entrepreneurship. The top right panel shows the start-up rates differentiating between "R1", "R2" and other institutions according to the Carnegie classification. R2 institutions (high-research universities) are among those with the highest start-up rate.[16] The bottom panel gives the start-up creation rates of foreign students across the most important countries of origin, relying on the F-1 visa data used in the shift-share IV estimations. Rates are particularly high for Indian and Chinese students, confirming the importance of both immigrant communities terms of entrepreneurship and innovative potential in the US (Kerr and Kerr, 2020).

[16]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.

Figure 2 shows as solid lines the average start-up rate separating foreign (black) and US-origin (grey) master students over time. Additionally as Crunchbase also includes foreign firms, the dashed line adds the start-up founded abroad by foreign-origin and US-origin students. Specifically, each point shows the number of start-up founders over the next 5 years, in the cohort graduating in that year, divided by all the graduates of that year. The number is standardized per 10,000 Master's graduates.

The series for US-origin start-up founding rate increased 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 hence it has 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.[17] To avoid these macro-trends affecting our findings, we will not rely on time variation for any of our findings in the empirical analysis. Notice also that including start-ups founded anywhere in the world, does not change much the start-up rates of US-origin Masters, as they mainly create start-ups in the US.

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.[18] 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.

Focusing on the overall start-up rates by graduation cohort, we immediately see a much higher propensity to create start-up companies for foreign origin Masters. [19] In the post-2010 period, foreign-origin graduates had an average start-up rate between 4 and 5 per 10,000 graduates if we account for foreign start-ups, and around 2.5 to 3 per 10,000 graduates considering only US start-ups. Hence they were almost 10 times more likely than US-born to be start-up founders overall, and about 8 times more entrepreneurial if we re-

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

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

[19] To that purpose, as described above, we weight each founder by the probability that their first-last name combination identifies a person with US or foreign origin.

strict the analysis to start-up creation on American soil. This stark difference in start-up rates implies that while being only 10% of the graduating class, foreign-born students were responsible for founding start-ups in numbers comparable to US-origin students. This difference is remarkable and suggests the possibility that the presence of more foreign-born students in a cohort will substantially increase start-ups created by that cohort. Additionally, given this high entrepreneurial propensity of foreign Master's students, it is plausible to think that their presence might generate positive spillover effects on the entrepreneurial rate of US-origin graduates in the same Master's cohorts. This is what we will explore more formally in the empirical analysis.

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. Second, US-origin master students in a cohort with more foreign-born students can show higher or lower propensity to create start-ups if there are spillovers from international students.

We will capture the total effect by estimating how an increase in the share of foreign graduates in each "university-cohort" cell affects the number of founders, or directly the number of start-ups founded by graduates in the same cell within the following 5 years since graduation. 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 the number of U.C. 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 include all founders/firms in a cohort and estimate the total effect. 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 uni-

versities. 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 Figure 2 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's master's cohort cell. When that is the case, the estimated coefficient will capture the effect of replacing one percentage point of US-born graduates with one percentage point of foreign-born graduates in the cohort population. In alternative 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). We then estimate only the part of this impact due to spillover effects of foreign students on US-origin start-up founders. To that aim, we use instead the variable Y_{ut}^{US}, the number of (likely) US-origin founded firms, as our dependent variable.

If the variation of foreign-born graduates across different graduating cohorts of a University was random, the coefficients β in the least squares regression of specification (1) would capture, respectively, the total 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 to be correlated with unobservable variables, we discuss identification and our proposed IV strategies 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 of various types. Local economic conditions, local prices, or local job opportunities are all examples of factors affecting jointly the probability of enrollment of foreign-born students in a university and the business profitability of firms. 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.

The choice of a university/location for any Master's student entails a comparison between the potential benefit from the Master's degree and the total costs of achieving it. International students, who are choosing among several US options from a foreign country, are especially sensitive to factors affecting these calculations. Their decision to attend a specific Master's program 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 for business creation. However, as documented by Beine et al. (2023), enrollment in a university in a specific location also depends negatively on the level of rents and local costs of living, which are strongly 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, and differences in costs play a significant role in their choice.

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, as we will show below, they are hardly correlated with previous trends. Therefore, they provide an exogenous source of cost differences in pursuing a Master's degree between in-state US-born and foreign students, which affects the relative probability of enrollment of international students (and hence their share in a cohort) without being associated with other local variables.

As a second IV strategy, 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 allocation 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, as they are pre-determined, are less correlated with contemporary local characteristics.

Importantly, 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 are consistent with the causal effects of foreign-born students in a master class on the start-up rates of its graduates. We will therefore describe and present both estimates as robustness checks for each other, and of their external validity.

Out-of-state tuition 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 separate these, more systematic components, by first subtracting the in-state tuition and then by controlling for lagged characteristics of the school and location. In the spirit of Borusyak and Hull (2023), we implement a procedure that aims at separating the systematic, non-random variation from the quasi-random one, and test whether the latter component, which will be used as an IV, is uncorrelated with a series of pre-period trends.

This IV approach builds on and expands the method introduced in Beine et al. (2023) and used there to estimate the contribution of foreign-born college graduates to the US high-skilled labor force. Here, we apply it to Master's tuition and fees to predict Master's student enrollment (rather than Bachelors). 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, in US dollars, and we regress this difference on a set of time-varying proxies of universities' quality and financial balances. 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 the 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. This variable still affects the differential cost of attending a Master's program for a cohort of foreign-born (relative to in-state) students, and thus it may affect their enrollment share in that particular university-cohort.

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

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

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, broadly used in the immigration literature (Goldsmith-Pinkham et al., 2020). This approach has been used in the context of foreign education by Machin and Richard (2017), who investigate the effects of foreign education in the UK. 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 2014, interacted with the share of students from each of those origins in each university as of 2001-2002.[20] We include in this sample, which we call the shift-share IV sample, only private universities (not included in the previous IV analysis). Since the data on foreign students by nationality are available starting in 2001, we consider the period of analysis for this specification between 2001 and 2014.

Data on foreign students by nationality of origin are from F-1 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 2001 and 2014, including information on the Master's program, university, year of enrollment, graduation, and country of origin of the student. From the data, we construct the number of foreign Master's students from each country of origin graduating in 2001-2002, the base years, in each US university. For each foreign origin o, we calculate its share of Master's graduates in 2001-2002 in each university

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

u, and then we allocate the post-2003 number of foreign students from the country o in the US in each year t among 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).[21] This predicted share will then be used as an instrument for the observed share of foreign students in equation (1). Further details of this shift-share procedure are provided in the Appendix (Section 1).

Discussion of IV's

An advantage of using the two IV procedures is that they involve different margins of identifying variations. In the "average treatment effect" jargon, each of them involves a different 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 in that school. These types of compliers are more likely to be students from poorer countries and/or from less privileged backgrounds, who are particularly sensitive to the financial cost of a Master's program. The shift-share IV, instead, generates variation based on students from countries of origin that are 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. These two groups have in common their being foreign-born, but they may have different unobservable characteristics. Hence similar results would suggest that those unobserved characteristics and their correlates do not play an important role in the effect we are estimating.

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 somewhat different in the time period due to the shorter number of years of availability of the F-1 data. The sample used for the first IV covers cohorts graduating in the years 1999-2014, while the sample used for the second IV only covers cohorts graduating in the period 2004-2014.

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

Additionally, the two samples differ somewhat in the size of cohorts, average entrepreneurial intensity, but not in the average and variation of foreign born share across cohorts (see Table 3). Estimating a significant and similar coefficient with both approaches would, therefore, be consistent with a robust causal effect of foreign students on start-up rates and would support a significant degree of external validity for one group relative to the other.

5 Strength and Validity of the instruments

Table 4 shows the first-stage coefficients (first and second row) 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 always 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 4 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 specification in column (2) shows that 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. This emphasizes that the change in the differential tuition affects the share of foreign-born rather than the overall size of the cohort. 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, as it means that the total size is barely correlated to the share of foreign-born in the groups. In addition to that, the Kleibergen-Paap F-statistics of the first stage have values around 61, hence much larger than values generating concerns of weak instruments. The Anderson-Rubin test of weak IV rejects the null hypothesis 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. The coefficient implies that an increase of the imputed share of foreign graduates by one percentage point will increase the foreign-born share in the graduating cohort 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. In summary, both instruments are

reasonably strong.

The validity of the tuition IV hinges on the idea that out-of-state tuition fees net of in-state fees affect the enrollment of foreign enrolled students, and is not correlated with other variables. More specifically, changes in out-of-state fee differentials 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 predict the number of US-born students in the cohort. To test these conjectures, Table 5 shows, for each of the two IV strategies, their ability to predict the (hyperbolic sine of the) number of foreign graduates, the (hyperbolic sine of the) number of native graduates, as well as the (hyperbolic sine of the) total number of students[22].

The coefficients of Table 5 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 number of foreign students (column 3), while the number of native students remains unchanged (column 5). 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, increasing their share but leaving unchanged the supply of native students, consistent with what a valid instrument should generate. As for the Shift-share IV, explored in the row below, changes in the shift-share IV predict an increase in the number of foreign students (column 5), and no change in the number of native ones (column 6), leaving the overall size of the graduating cohort unchanged (column 2). This is also consistent with the impact of the shift-share instrument by affecting foreign enrollment and resulting in a significant impact on the share of graduating foreigners. Both IV operate through the hypothesized channel, consistently with their validity.

An additional validity check for our tuition IV consists of testing whether past values in the number of foreign students enrolled and foreign graduates across school cohorts are correlated with current values of the instrument. In particular, correlation with past inflows of foreigners could suggest spurious trends affecting past school-specific conditions and current changes in out-of-state tuition fees. In Table 6 we explore this possibility by showing the correlations of lagged share of foreign students' enrollment (columns 1 and 2), lagged share of foreign students in the graduation cohort (columns 3 and 4) and lagged start-up enterpreneurs rates (columns 5 and 6), with the current idiosyncratic changes in

[22]We take the hyperbolic sine, to proxy an elasticity while still including the 0's. In Appendix Table A3 we show the same regression taking log of the values plus one. The results are extremely similar

out-of-state, net of in-state tuition. We lag the dependent variables 4 or 5 years to capture the past trend of these variables. No coefficient is significant (or close to significance), reassuring us that a persistent correlation of foreign enrollment and entrepreneurship rates across Master cohorts, driven by unobserved variables, does not seem likely. It is only after the occurrence of lower out of state tuition fees, that we observe an increase in foreign-student shares and the potential change in start-up rates.

A lingering concern regarding the use of the tuition-based instrumental variable (IV) is that changes in out-of-state tuition fees may also alter the composition and average quality of US-born students. As out-of-state students may respond to those changes too, if out-of-state students are higher (or lower) quality than in-state, the IV could drive a change in composition and hence quality of national students. We showed in Table 5 that the *number* of US-born students did not change in response to changes in tuition, but did the composition/quality of this group change?

There are only a few measures of the academic performance of US-born (separate from foreign) students in the IPEDS data, and they are only relative to Bachelor's degrees. Hence, we use the graduation rate as such a proxy, looking at wether for bachelor degrees, there is a correlation between the IV computed with tuition charged to Bachelor students and the graduation rates of US-born Bachelor students. Table A9 column (1) in the Online Appendix shows such a coefficient. It is very small and not significant, consistently with the idea that the average quality of US-students did not change in response to changes in out-of state tuition. Additionally, the possibility of a systematic association between the US-student quality and the IV should be much less of a concern for the shift-share IV estimation. We do not see a plausible story why the predicted changes in foreign students, driven by the foreign network size, should be associated with changes in US student quality. Hence, the additional estimates using that IV should be less subject to this potential bias.

The lower panel of Table 6, shows the same pre-trend correlations for the shift-share IV. The difference is that, as we are using the 2001-2002 share composition across schools in terms of foreign-students to construct the post-2003 changes in the IV, we analyze whether these post-2003 changes are correlated to the 1995-2003 variation of total graduates, foreign graduates, US graduates and with the number of 1999-2003 US and foreign start-up entrepreneurs. The estimated coefficients in Table 6 show 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 are consistent with the tuition IV being a

strong and valid instrument: they do not raise concerns of weak instrument bias, show no evidence of pre-trend correlations or correlation with quality and number of US-born students, hence they are consistent with the exclusion restrictions. The shift-share variable seems to be somewhat less strong as an IV, but still mainly uncorrelated with US-born enrollment and with most pre-period trends. Hence, we will use both instruments in the analysis, placing a bit more emphasis on the estimates obtained from the more original and cleaner tuition IV, but considering the shift-share IV estimates as an additional external validity and robustness check.

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 column 1 and IV estimation in all other columns. The upper panel shows the estimates using the idiosyncratic component of tuition fees as an instrument and the 1999-2014 graduating cohorts from the public universities sample. The lower panel shows the estimated coefficients using the shift-share IV as instrument and the 2004-2014 graduating cohorts from the private universities as sample.[23]

Different columns of Table 7 use different outcomes 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 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.

In column (4), we restrict the outcome to 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 Crunchbase data tends to include more successful start-ups. The outcome of column (5) limits the count of start-ups to those that raised more capital within the first 3 years since creation than the median US start-up of the same age. In column (7), we use the count of start-up companies that received at

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

least one patent within the first 3 years of operations as the outcome of interest. This is a rare occurrence as only a few companies file for patents, and the timing of three years is very tight. Only 3% of the companies in our sample are included in such a sample, which therefore identifies a selected group of very innovative companies. Finally, column (7) includes only start-up companies founded in the same state as the university of origin of the cohort. We call this outcome "local" firm creation. By estimating these different specifications, we show the effect of foreign students on the probability of creating not just any start-up but successful companies, selecting them along several metrics of firm outcomes that become more demanding as we move across columns.

Two main and novel findings emerge from the set of estimates in Table 7. First, using either IV strategy, we find statistically and economically significant effects of increasing the share of foreign graduates on the number of start-up founders, on the number of firms, as well as on those surviving after 3 years and raising more capital than the US median start-up of the same age. Second, the estimates exhibit remarkably similar values across the two IV strategies: simple tests of differences of β coefficients fail to reject the null of equal values for all outcome values. This suggests that the share of foreign students, rather than potential omitted variables correlated with the different IVs, is causing the start-up differences.

Considering the estimates starting from column (2) in panel A and B of Table 7, we find that 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 about 0.7 when using the tuition IV, and by about 0.9 when using the shift-share IV. Both estimates are statistically significant. This represents a large effect, equal to the standard deviation of the start-up creators' rate between cohorts, in the tuition IV sample and about one third of the standard deviation for the shift-share IV sample (see Table 3). The coefficient has a causal interpretation: by attracting ten percentage points more foreign-born master students in a cohort, a university generates about 0.7-0.9 more start-up founders among its graduates.

The OLS estimates of the same coefficient, shown in Column (1) of Table 7 are significantly smaller than the IV and not statistically significant. This suggests the existence of a significant downward bias of the OLS estimates, which is related to unobservable variables affecting the cost of living and local start-up rates, over time. We explain and discuss such bias in the section 6.3 below.

The estimates in column (3) show that the effects on extra start-up founders correspond closely to extra start-up companies. The dependent variable, in this case, is the number of start-ups created by graduates in the cohort. The estimates imply that a 10 percentage point increase in the share of foreign students in a master's class leads to 0.7 to 0.9 more start-ups created by its graduates. The estimates mirror the corresponding one for start-up founders in both Panel A and Panel B.

Specifications (4) to (6) consider more selective groups of the start-up created by the master cohort. As several start-ups last only for a short period of time and have a small impact on the economy, simply increasing their number may not be that economically meaningful. As we already described in Section 3 the Crunchbase data already selects the more successful start-ups. Here we only consider the count of firms performing better than the average in some dimensions. In column (4), the count only includes those start-ups still active after 3 years. In column (5), only those who raised more capital within 3 years than the US median start-up of same age. In column (6), those filing for at least a patent in the first three years. The positive and significant coefficients for each of these variables imply that foreign-born graduates increase start-up rates for these high-performing groups of start-ups as well. The coefficients are significant but much smaller for the patenting-firm count, as those are only a very small group (3% of total). For most of the outcome variables, the estimated magnitudes are very similar across the two IV strategies. Finally, in column (7), we show the impact on "local start-ups", namely those founded in the state where the University of the Master cohort is located. Interestingly this coefficient shows that 40 to 60 percent of the additional start-ups are created in the same state (the size of the coefficient is about 40 to 60 percent of those in column (3)), suggesting a strong local agglomeration effect from the University to new local firms. Overall, the effects estimated in Table 7 are quantitatively important, statistically significant, robust across specifications, and consistent between samples and IV choice. A larger share of foreign Master's graduates increases the start-up rate of a Master's cohort, including among those that last for longer and raise more capital.

6.2 Robustness to alternative specifications

We perform some checks to reduce the potential role of omitted variables in the regression. First, the presence of university-specific trends in the entrepreneurial abilities of Master's students, and the share of foreigners, driven by increased (decreased) prestige or popularity,

may represent an unobservable variable causing bias. To mitigate these concerns, in Table 8, using the tuition-based IV only, we estimate specifications similar to those reported in Table 7 above, including more demanding interactions between trends and university-specific fixed effects[24]. In the first check, in panel A, we select for each university, periods of time with stable overall enrollment (i.e., periods in which the number of total students enrolled did not grow more than 5%) and we include a set of fixed effects interacting university-specific period-dummies with university dummies. Each specification is otherwise the same as (2) to (6) in Table 7. These specifications absorb the variations associated with medium-term changes of university enrollment and rely only on idiosyncratic year-to-year variations in foreign share associated to tuition changes. In Panel B of Table 8, we include university-specific time trends, instead, which account for gradual long-term changes in enrollment and outcomes. The results remain largely consistent with those reported in the baseline estimations (Table 7). If anything, the only notable difference is that, after accounting for university-period-specific effects, the estimated impact of foreign graduation on start-up-related outcomes becomes stronger. The same conclusion applies to the robustness of the shift-share IV results, with results largely unaffected with the extended specification.

In Tables 7, 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. 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 an alternative specification that produces the semi-elasticity of the number of start-up firms to percentage (log) changes of foreign graduates, keeping US-born students unchanged. Panels C of Tables 8 and Table A5 report the estimated semi-elasticities.

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. Focusing on the estimations using the tuition IV, we find that on average, an increase of 10% (0.1 log points) in the number of foreign graduates generates about 0.06 more entrepreneurs (i.e. the coefficient of 0.59 times 0.1) in the 5 years following graduation. This is very similar to the point estimates of the effect of an increase in the share of foreign-born reported in Table 7 (equal to a coefficient of 0.07). Similarly, such an increase would produce 0.06 more start-up companies, 0.06 more

[24]We perform the same analysis using the shift-share IV, in the Appendix in Table A5

companies raising more capital than the median US start-up, and 0.01 more start-ups that will obtain at least one patent within the first 5 years of activity. Results are similar with the shift-share IV estimates (panel C of Table A5), with semi-elasticities implying slightly larger effects compared to the baseline estimations. Overall, these alternative specifications confirm the robustness of our results. The findings are consistent with the idea that a larger number of foreign Master's students, attracted by lower tuition fees or network effects, has a significant positive effect on start-up creation, consistently estimated across specifications, and uncorrelated with the changes in native Master's students in a cohort.

The total estimated effect of the foreign share on start-up creation is the sum of a direct effect and a spillover effect. The direct effect is the result of the higher propensity of foreign students to engage in entrepreneurial activities (which was also shown in the stylized facts of section 3.4. The spillover effect, instead, arises if there is a positive impact of the presence of foreign students on start-up creation by native students. We will analyze such a potential spillover effect in section 7.

A relevant question is what the reasons are behind the higher entrepreneurial propensity of foreign students. The limited information on international students' characteristics provided in the IPEDS data does not allow us to explore their differences with natives. The existing literature, however, provides some suggestions. Foreign students may be characterized by a lower risk aversion compared to native students, which can be conducive to higher entrepreneurship. The literature finds evidence that international movers have lower risk aversion than non-movers (Jaeger et al., 2010). Some studies have also found that movers are positively selected on cognitive and other non-cognitive abilities (see Butikofer and Peri, 2017), which can lead to more innovative entrepreneurial attitudes. Another reason may be that a higher propensity to create start-ups derives from the longer and more diverse work experience of foreign students. A significant share of foreign students enroll in master's programs in the US after years of working, and they may have been exposed to ideas and practices in companies outside the US. These two characteristics of international students may also be at the core of their spillover effects: native entrepreneurs might benefit from lower risk aversion and higher work experience brought by foreign students.

An additional potential reason could be the lower opportunity cost of starting a company, relative to starting a career in a company, vis-a-vis a native students. As visas and work permits are hard and rationed for foreign students in the US, starting a company may be perceived as a better alternative and one that may be conducive to remaining in the US. Importantly, however, given the substantial requirements to obtain an investor visa (high

potential growth for the company, significant investment from qualified investors or government in the first 3 years) and given the evidence of foreigners creating highly successful firms, this mechanism does not mean that creating a start-up is a "shortcut" for foreign students to stay in the US. Rather, we should see entrepreneurship as a hard and selective channel to stay in the US, still given the selection of students coming to the US, as one that several of them may consider at the end of their Master program.

6.3 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 share 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. To 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 in the area. 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. Besides housing prices, housing availability may play a role in attracting or diverting foreign students. Although it is difficult to fully control for local housing availability and prices (due to limited data), the IV strategies, by using variables that are likely uncorrelated with local economic conditions, attenuate this issue.

The two correlations shown in Figure 3 strongly suggest the potential for omitted variable bias that we described above. The left panel of Figure 3 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 one-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. The right panel of Figure 3 shows a strong and positive correlation between changes in local rents (x-axis) and de-

meaned 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 in a location are strongly associated with higher housing prices and more start-up activity, and also with fewer foreign students enrolling in local Master's programs. This is consistent with the idea that potential students are sensitive to local costs when deciding about their Master's location, as housing is one of their most relevant costs. The existence of this type of spurious correlation between local economic conditions and Master's enrollment induces a negative bias in the estimation of β by OLS, makes the use of IV especially important. By isolating pull factors for foreign students that are independent from 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.

6.4 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. In this section, we analyze whether the impact of foreign graduates on start-up creation is also heterogeneous across Master's majors or if it is more uniform across most majors. 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. Additionally, we test whether the start-up promoting effect of foreign students in programs located in large metropolitan areas, with many universities, where the local environment may be particularly conducive to entrepreneurship, may be stronger than elsewhere.

Table 9 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, from the rest. Finally, in columns (5) and (6), we separate master programs in areas with a high concentration of universities, or not, to separate potential heterogeneity due to agglomeration of human capital. Panel A reports the estimates for the tuition IV while Panel B gives the results obtained with the shift-share IV.

Two interesting findings emerge. First, the positive effect of foreign-born on start-up

founders is similar between STEM and non-STEM majors, as well as between top majors and other majors. While the Shift-Share IV suggests a somewhat higher effect in the STEM and top 5 majors, the tuition IV does not show such a tendency, and a formal test cannot reject that the two coefficient are equal. So, while we can observe differences in start-up rates across majors, the positive contribution of foreign students seems pervasive to all types of programs. Second, we find that in areas in which the concentration of universities is higher, the positive start-up-creating effect of foreigners is stronger. This can be connected with the higher potential for local collaborations of the foreign graduates in creating start-ups in large areas with multiple universities, (a channel whose relevance we will show below) or because such density of universities provide a "thicker" environment of skills and services to generate start-ups.

7 Spillover effects of foreign graduate share on native start-up creation

Section 6 above has shown the total effects of increasing the foreign share of students in US Master's programs on start-ups created within 5 years. These effects are statistically significant, economically large, robust to several specifications, and pervasive across Majors. In this section we analyze how much of this effect is driven by positive "spillovers" on US-born start-up creators.

This can be analyzed by implementing the same empirical strategy as above, but using as outcome, only the count of start-ups founded by (likely) US nationals in the university-cohort of Master graduates. The presence of foreign-origin students in a Master class may have a "spillover" effect on US-natives, if the exposure to the different skills, higher entrepreneurial inclination, and international connections of foreign-origin students makes US nationals better at or more motivated in launching a start-up.

During a 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. Foreign students can bring different skills and experiences that will later be reflected in different types of firms, leading to more variety in innovation. This could then inspire US-origin students to follow career paths inspired by international students and continue collaborating with them after graduation. Notice that a significant

proportion of foreign students, attending a master's in the US, leave after graduation (as they do not have a legal path to remain) and many of them become entrepreneurs abroad, as suggested by Figure 1. For those, their only impact on the US economy, post-graduation, will be their spillovers on US nationals. Alternatively, it is theoretically 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 rates by natives. In this section we first assess the size of the spillover effect. We then look at possible mechanisms that might explain them.

7.1 Spillover effects of foreign graduate share

In Table 10, we estimate the effect of increasing the share of foreign Master graduates on the count of start-ups by US-origin entrepreneurs, namely what we call the "spillover effect". The dependent variable includes only startups by likely-US origin founders, identified using their first and last names as described in Section 3.3. Specifically, when aggregating start-up creators, for each of them, we calculate the probability of being of US origin and we use those probabilities as weights to obtain the number of US-origin start-up creators in each cohort. Similarly, when the outcome is the number of start-up firms (rather than founders), we attribute a fraction of them to US-origin founders, based on the probability that its (potentially multiple) founders are US-origin[25]. Notice that this method of defining US-origin start-up founders implies that the difference between the total start-up count in each cell of our sample and the count of US-origin start-ups corresponds exactly to the count of foreign-origin start-ups. Hence, by construction of the variable, the estimated total effect (from section 6 above) minus the estimated spillovers (in this section) gives an estimate of the direct effect of foreign-born students on start-ups.

Table 10 presents the estimated spillover effects, using a very similar structure and both instrumentation strategies and similar specifications as done in Table 7 to estimate the total effects. Panel A shows the estimates based on the tuition IV using the sample of public universities, while Panel B displays those derived from the shift-share IV using the sample of private universities. Overall, the results suggest that spillovers are positive, significant, and represent a significant part of the total U.S. start-up creation effect. Focusing on the number of founders (column 1) or on the number of startups, the effect of an extra 10 percent foreign

[25] 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 0.584 US-origin.

students in a cohort is 0.25-0.27 extra startups by US origin founders, representing 23% to 38% of the total effect estimated in Table 7 above. As we have seen that the algorithm's most frequent error is to classify second- or third-generation Americans as being of foreign origin, these estimates likely represent lower bounds of the spillover effects, as some of the firms classified as foreign-origin could have been founded by US-born Americans with foreign-origin names.

Two additional points are worth emphasizing. First, the positive and significant spillover effects hold across all outcomes considered, including when we focus on the better performing startups, i.e. those surviving for three years or more (column 3), those receiving more capital in the first three years than the median US start-up of same ae (column 4), and those filing at least one patent in the first three years (column 5). Second, for most of these spillover effect, we show remarkable similarity between the estimates obtained using the different IV and samples (i.e., between Panel A and B). This remarkable consistency between the estimated spillover effects is consistent with robustness and external validity of our findings, indicating that the size and significance of estimated spillovers do not depend on a specific source of variation or on a specific sample.

An alternative and somewhat coarser (but easier) method to impute the origin of start-up founders is to simply use the top predicted country of origin from Forebears, to classify founders as either U.S. origin or foreign origin. This is what we call dichotomous imputation. While the dichotomous imputation has its advantages and drawbacks compared to the probabilistic method (as discussed in Section 3.3), it is useful to assess how the choice of imputation procedure influences the magnitude of the spillover effects. Therefore, Table 11 report the estimates using the dichotomous imputation approach. The structure of Table 11 is the same as Table 10.

The estimates of Table 11 using the dichotomous imputation approach strongly confirm the statistical and economic significance of the spillover effects of foreign graduates on US startup creation. Their estimated magnitudes are up to two times larger than those of Table 10. Those in Panel A (using tuition IV) imply that a 10 percent larger foreign share in the master class, generates 0.5 more startups by US origin graduates. This represents two-thirds of the total effect magnitude. Panel B estimates, using the shift-share IV estimation, are slightly lower and the spillover effect represents about half of the total startup effects of foreign Master students.

The comparison between the results of Table 10 and Table 11 suggests that the measurement error in the imputation method can be consequential in the estimation of the

spillover effects, as they are somewhat different using the probabilistic and the dichotomous approach. Mis-classifying a second or third generation American as a foreign-origin entrepreneur can affect the estimates of the spillovers. To evaluate the role of such potential imputation errors, more systematically, we implement a simulation exercise that provides a range of estimates that account for such errors. We proceed as follows. First we use the probabilities in Table 1 for false positives (i.e. probability of being assigned foreign status while being native) and false negatives (i.e. probability of being assigned US-status while being foreign born), equal, respectively, to 31.7% and 12.42%, to re-assign a number of positive and negatives (randomly chosen) to their "correct group", matching these probability, in a series of 1000 sample simulations. Then for each of the 1000 simulated samples, we estimate the spillover coefficient in a regression following the same specification As column (1) of Table 11. Figure 4 shows the empirical distributions of the estimates of the spillover effects of foreign graduate share on US start-ups creation estimated using both instrumentation strategies. Vertical bars indicate the mean of each empirical distribution. Inspection of Figure 4 indicates that both the shape and the first moment of the two distributions are strikingly similar across the two IV estimations. In both cases, we find that the average spillover effect is slightly higher than 0.5, which corresponds to between 57% and 75% of the total effect. Moreover, the range of these estimates is very similar. The spillover effect estimated from the tuition IV method ranges from approximately 0.03 to 0.073, while the range for the shift-share IV estimates is more compressed. The key takeaway from this analysis is that, despite the underlying imputation errors of the Forebears algorithm and regardless of the IV estimation method used, the spillover effect of foreign graduates on native start-up creation is substantial, always statistically different from 0, and represents a significant portion of the total effect on U.S. entrepreneurship.

7.2 Explaining the spillover effects of foreign graduate share

7.2.1 Start-up Co-Founding between US and foreign graduates

A natural channel through which foreign-born classmates could encourage the entrepreneurship of Americans is by reducing the barriers to cooperation, helping combine ideas and co-found a startup. We ask the following question: Do US-origin master graduates show a higher propensity to co-found a startup if they have attended a master's cohort with a larger foreign-student share? 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, sev-

eral US-born Master's graduates may become more likely to collaborate with them or other foreign students and, through these collaborations, create a start-up. [26] Complementarity in skills, abilities, and knowledge is crucial in firm creation. 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, increasing the total number of start-ups.

Table 12 shows the estimated coefficients for specifications that are similar to column (3) in Table 7, with the difference that the dependent variable is the count of startups *with at least 2 co-founders*. The outcome of interest, therefore, is the number of co-founded firms. The estimates of column (1) shows coefficients between 0.07 and 0.08 –in Panels A and B. This implies that a large part of the additional start-up creation induced by the presence of more foreign students in a Master cohort is co-founded, as these coefficients are similar to those of Table 7. This finding holds for both IV methods. Column (2) shows that a higher share of foreign-born encourages more co-founded companies with people from the same University and from the same university cohort (column 3). Column (4) shows that being in a cohort with more foreign students encourages more US-foreign co-founding of startups, and part of that extra co-founding is between foreign and US-origin students in the same university (column 5). Overall, the results suggest that a larger presence of foreign students in a Master's cohort enhances the collaborative approach to founding start-ups of natives, encouraging them to team up with foreigners and people from the same university and class. The interactions during master's studies are likely to foster future collaboration in start-up creation, in general, and cooperation with foreigners in particular.

It is important to note that, due to data availability from Crunchbase, our definition of co-founding may not include all the initial collaborations that are important to create a start-up. Specifically, we define collaboration only as co-founding. However, sometimes a set of people, providing initial scientific, financial and technical advice, are crucial to start a firm, and among those, the presence of foreign-born can be even more prevalent. As legal and procedural restrictions can make it harder for immigrants to be legal founders of a firm, they may contribute as initial employees in the firm's team rather than as founders. As a result, the impact of foreign students on various forms of collaboration in start-up creation can be even larger than estimated.

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

7.2.2 Impact of foreign graduate share on diversity and innovation

Like other types of immigrants, foreign students bring different types of skills, experience, and tastes relative to natives. These differences create the potential for larger differentiation in firm creation: foreigners span a range of ideas and preferences that are different from those of natives, expanding therefore the set of possible products and services brought to market by the startups. A large literature in economics has shown the positive impact of immigration on economic performance due to an increase in diversity of skills, ideas, and products (Alesina et al. (2016), Ottaviano and Peri (2005), Peri (2009), Bahar et al. (2022)). Our research adds the dimension of diversity of startups.

To do this, we characterize the product or service differentiation across these startups and we explore to what extent a higher presence of foreign students leads to greater differentiation of the start-ups, relative to those created, on average, by natives.

To capture the degree of differentiation of startup companies created by a university-cohort, we leverage the data descriptions of each startup's activity. Each startup's product/service is characterized by choosing among a set of keywords in a pool of 770 different terms.[27] Using these keywords, we construct an index that measures the similarity or diversity of a start-up relative to others in the dataset. By restricting comparisons to start-ups within a specific category, we can further assess how a given start-up is positioned in relation to the average competitor in the sector.[28] More precisely, we compute the Euclidean distance between two 770-by-1 vectors capturing the presence (1) or absence (0) of a keyword, between each start-up and the average vector of all US-based firms created in that year. Then we also calculate the distance between each firm and the average vector of all foreign-based firms created in that year. This allows us to quantify the distance (overlap) of each start-up from the average of national start-ups in the "keyword" vector space. In each cohort, we count how many start-ups were created with distance above the median distance to the US average vector overall or in the same category and use them as dependent variables in Columns (1) and (2) of Table 13. Similarly, we count how many start-ups were created with a distance below the median distance to the foreign start-up average vector (overall or in the same category) and use them as dependent variables in Column (3) and

[27] The number of keywords selected is not fixed. It is determined by the reporting start-up. For instance, the start-up "Omnidrive" is characterized by the following keywords: E-Commerce, File Sharing, Internet, Social Media, Software, Web Hosting.

[28] A category is the equivalent of a sector of economic activity for traditional firms, and categories are self-reported. For example, a category might be " Commerce and Shopping" or "Internet Services".

(4) of Table 13.

Table 13 presents the estimates of these coefficients for both IV estimations. The coefficients of columns (1) and (2) reflect the effect of foreign students on the creation of startups rather “different" from the average U.S. start-ups. The coefficients in columns (3) and (4) reflect the effect of foreign-born share on the creation of start-ups rather ”similar" to the average foreign start-up. In columns (1) and (3), the average vector is relative to all start-ups, whereas in columns (2) and (4), the average vector is within the same industry category. [29]

The estimates in Table 13 support two key findings. First, foreign students have a significant impact in fostering start-up firms that are farther from the average US-start-up keywords type than the median. Second, higher presence of foreign students encourage creation of start-up that are “closer" than the median to the average start-up keyword type created abroad. If we take the distance from the median US and "foreign" start-up vector as a measure of how differentiated start-ups are relative to those two "reference groups", we can say that more foreign students shift the type of new firms away from what is typically US-created and towards what is more common in the rest of the world. This is likely to expand the range of products and services that they offer relative to a world with only US-born start-up founders.

8 Discussion and Conclusion

In this paper, we consider the connection between two of the most dynamic segments of the US economy: the most innovative segment of business creation and Master's graduates from highly competitive universities. One of the features of this important nexus is that, by stimulating and attracting talent from all over the world, the US economy benefits from highly entrepreneurial foreign graduates. Specifically, in this study, we ask the following question: How much do the international students in US Master's programs contribute to the creation of start-ups in the US? Do they encourage US-born Master's graduates to be entrepreneurs, or do they crowd them out of entrepreneurship?

Our contributions include combining data from different sources that capture the presence of foreign students in all US Master's programs (IPEDS) and data on the creation of US Start-ups (Crunchbase). We use and compare two IV-based identification strategies

[29] Start-ups may report multiple activity categories. In such cases, we average the diversity index across all categories.

that exploit (a) quasi-random variation of out-of-state tuition costs and (b) pre-determined variation in the country of origin networks across universities (shift-share) as drivers of year-to-year changes in the share of foreign Master students in US universities. Then, by matching start-up founders with their university-Master's cohort, we analyze the causal effect of an increase in the share of foreign master students on the number of start-ups created overall and by US-born Master graduates (spillovers).

Our 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 one standard deviation) has an effect of producing about 0.7 to 0.9 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 surviving for 3 years or longer, able to raise more capital than the median and to file for at least one patent by the third year since their founding.

Second, we show that a significant portion of these start-up effects stems from the positive spillover effect on U.S.-origin start-up creators. U.S.-origin graduates become more likely to start a business when they study in a university cohort with a larger share of foreign students. Depending on the instrumentation strategy, we find that between 23% and 38% of the total effect is attributable to the spillover on U.S.-origin students. These estimates of spillovers on U.S.-origin start-up founders are likely underestimates of the spillover effects, as our imputation algorithm, based on names, tends to produce a higher share of false positives—i.e., entrepreneurs classified as foreign-origin who are actually second or third-generation migrants, U.S.-born descendants of immigrants. Simulations estimating the probabilities of classification errors regarding immigration status suggest that the spillover effect could account for an ever more substantial portion of the overall effect of foreign graduates on U.S. start-up creation. Accounting for false Foreign-born attribution, our estimates of the proportion of the spillover effect in the total effect of foreign graduates range from 40 to 75%.

We also provide evidence of two channels explaining the importance of the spillover effects. The first channel is related to the presence of foreign students, increasing the incentives of natives to co-found a start-up. We show that a larger share of foreign students in a given cohort fosters co-founding, within the same university, within the same cohort

and between foreign and native students.

The second channel explores the impact of foreign graduates on the diversity of the start-ups "type" that they help create. By bringing different skills, knowledge, and diverse cultural backgrounds and preferences, foreign students help U.S. start-ups become more diversified, spanning a wider range of products and services. We find evidence that a higher share of foreign students leads to more differentiated U.S. start-ups, as defined by a series of keywords characterizing their services and products. Consistent with the idea that foreign ideas influence U.S. start-ups through foreign students, we also find that these start-up types tend to become more similar to foreign start-ups as defined by the chosen keywords.

Let us notice that 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 after graduation. Many of them (as many as 70% of them according to Beine et al., 2023) leave after graduation due to the lack of available visas. In view of the fact that many of them leave, it is not surprising that a large part of this positive effect post-graduation takes place through the positive spillover on US-born graduates. Potentially, retaining more of these international Master's students could imply a further substantial boost in new firm creation in the US. Policy proposals aiming at providing a path to permanent residence for STEM graduates (as contained, for instance, in the "Dignity Act" proposal of 2023, or proposals to extend the Optional Practical Training (OPT) period for F-1 visas or the AT (Academic Training) period for J-1 visas) could generate a large return in terms of innovative firm creation in the US.

References

Alesina, A., J. Harnoss, and H. Rapoport (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth 21*, 101–138.

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*.

Bahar, D., H. Rapoport, and R. Turati (2022). Birthplace diversity and economic complexity: Cross-country evidence. *Research Policy 51*(8), 103991. Special Issue on Economic Complexity.

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 (2023). Nonrandom exposure to exogenous shocks. *Econometrica 91*(6), 2155–2185.

Butikofer, A. and G. Peri (2017, Sep). Cognitive and noncognitive skills and the selection and sorting of migrants. NBER Working Papers 23877, National Bureau of Economic Research, Inc.

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.

Jin, Z., A. Kermani, and T. McQuade (2025). Native-immigrant entrepreneurial synergies. Technical report, National Bureau of Economic Research.

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*.

Machin, S. and M. Richard (2017). Paying out and crowding out? the globalization of higher education. *Journal of Economic Geography 17*(5), 1075–1110.

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.

Ottaviano, G. I. and G. Peri (2005, 06). The economic value of cultural diversity: evidence from us cities. *Journal of Economic Geography 6*(1), 9–44.

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. (2009, 01). The effect of immigration on productivity: Evidence from us states. *National Bureau of Economic Research, Inc, NBER Working Papers 94*.

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.

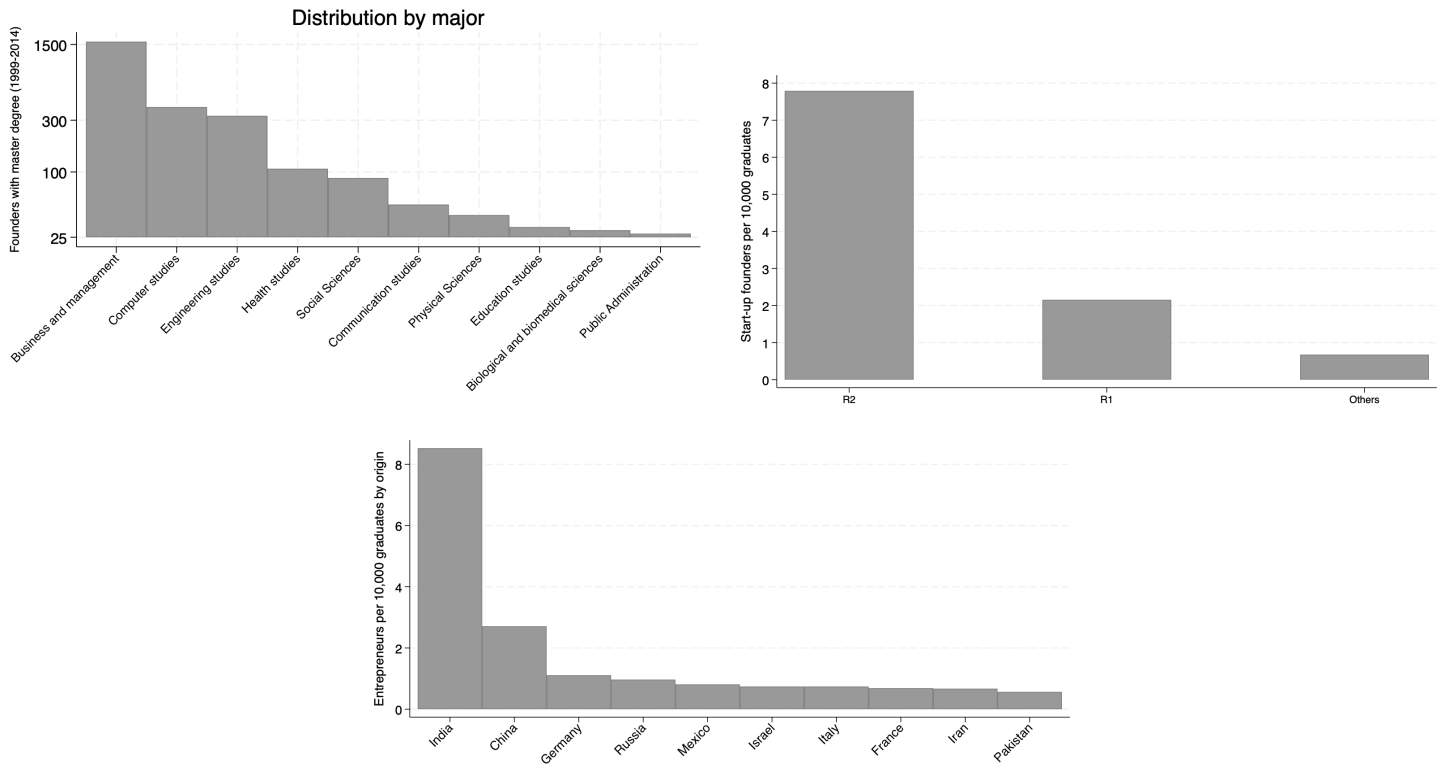
Stuen, E. T., A. M. Mobarak, and K. E. Maskus (2012, 09). Skilled Immigration and Innovation: Evidence from Enrolment Fluctuations in us Doctoral Programmes. *The Economic Journal 122*(565), 1143–1176.

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).

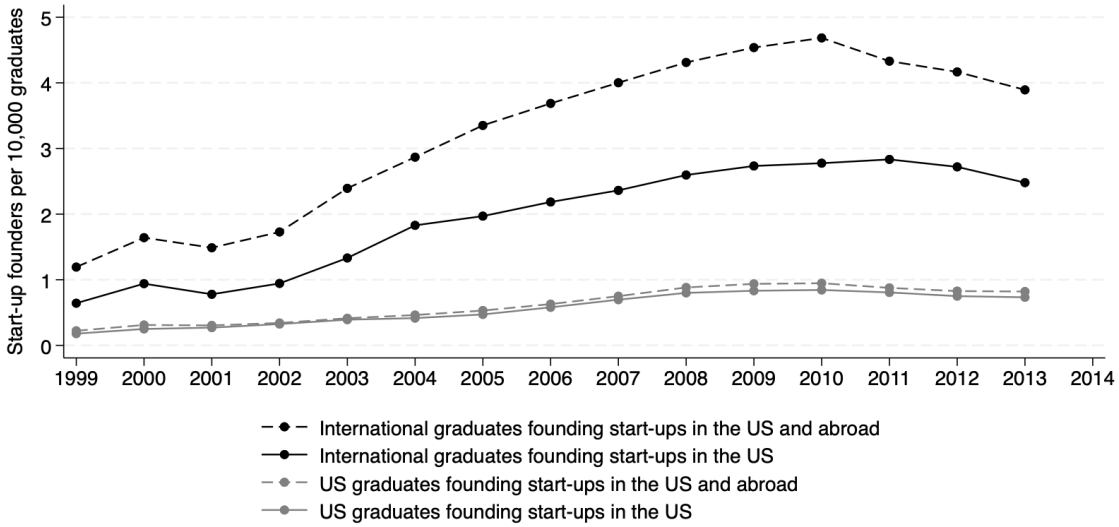
9 Figures

Figure 1: Distributions of foreign Master's graduates start-up founders by study field, Type of university and origin country



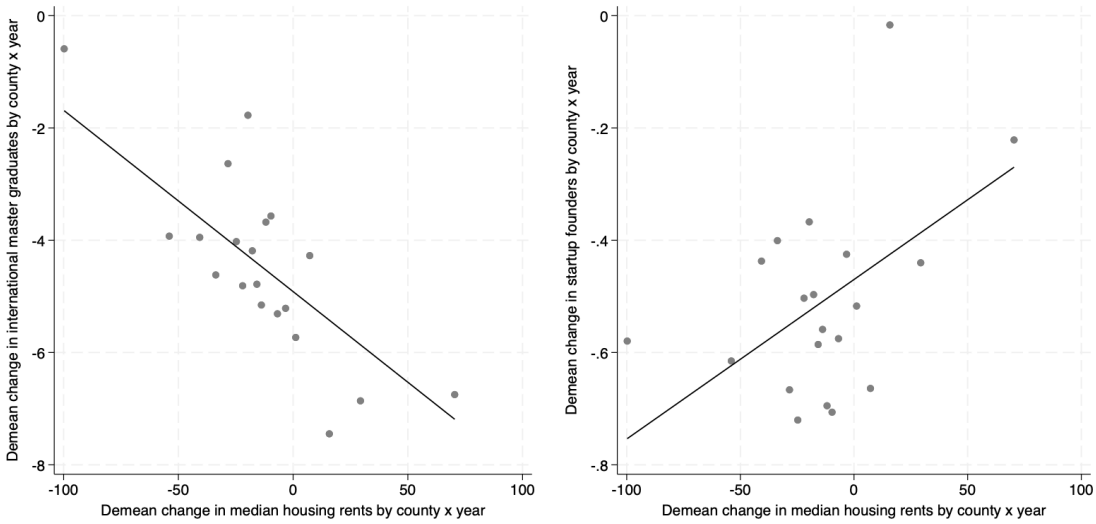
Notes: Panel A: Number of foreign start-up founders by field of thei Masrter's Degree. Panel B: Start-up creation rates of foreign graduates by type of University institutions. Carnegie Classification of Institutions of Higher Education. R1: Doctoral Universities – Very High Research Activity. R2: Doctoral Universities – High Research Activity. Panel C: Start-up creation rates of foreign graduates by country of origin. Country of origin identified using F1 Visa data.

Figure 2: Start-up founder rate of native and foreign master graduates.



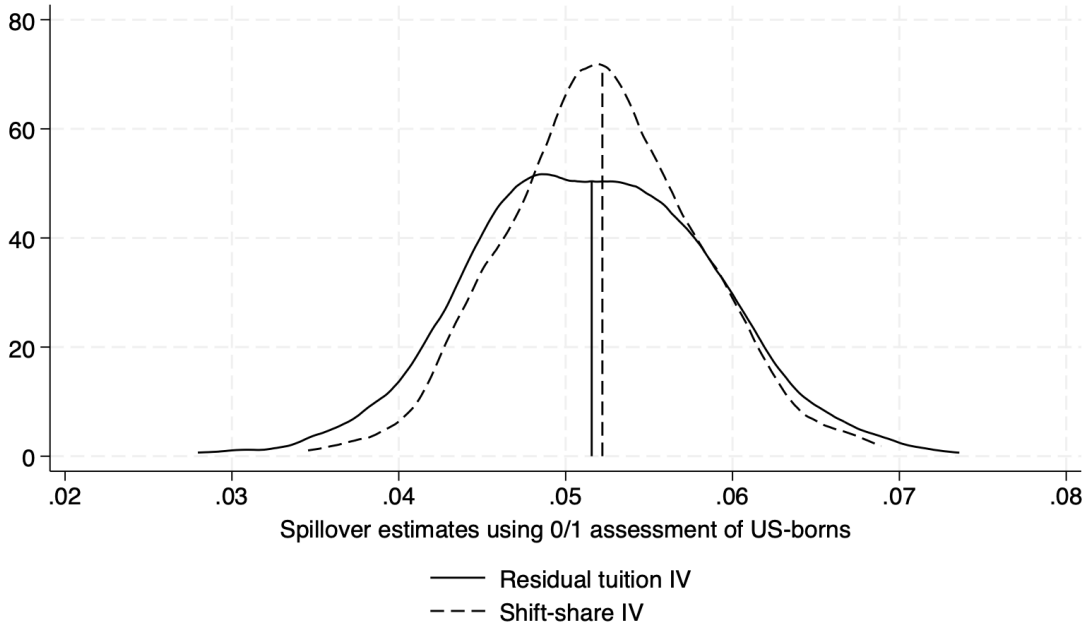
Notes: Evolution of the share of start-up founders within 5 years since graduation per 10,000 Master graduates. Foreign graduates-black line. Native graduates-grey line. The dashed line includes all startup created the continuous line only those created in the US. Decomposition between natives and foreign founders based on probabilistic measures drawn from Forebears.

Figure 3: Relationship between yearly changes in rents, foreign enrollment and start-up creation across US counties, 1999-2014



Notes: Both figures using bin-scatter plots. Left panel: 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. Right panel: 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 4: Distribution of spillover effect estimates



Notes: Empirical distributions of the estimates of the spillover effects of foreign student share on start-ups creation estimated with residual tuition (solid line) and shift-share (dashed line) IVs. Estimates are obtained using 1000 iteration in which each time we re-draw foreign- and US-origin status of the founders following the probabilities of false positive and negatives obtained from Table 1. Smoothed distributions are obtained using a Kernel density estimation, with the selected bandwidth parameter based on Silverman's rule of thumb.

10 Tables

Table 1: Validation of the Forebears imputation.

	Observed US-born	Observed foreign-born	Total
Predicted US-born	616 (68.29 %)	79 (12.42 %) False negatives	695
Predicted foreign-born	286 (31.71 %) False positives	557 (87.58 %)	843
Total	902 (100%)	636 (100%)	1538

Notes: The table compares a sub-sample of 1500 startup founders, manually checked to obtain almost certainty of the place of birth of individuals, with the attribution of U.S.-born and foreign-born using data from Forebears. The manual checks involved gathering information from various sources, including personal websites, LinkedIn profiles, videos, social media, news sites and similar to reach almost certain attribution of the place of birth of the start-up founder.

Table 2: Summary statistics of total Crunchbase Sample and Sample with non-missing education of founders.

	Start-ups with info on founders' name	Start-ups with info on founders' Degree and year	Difference in averages b/w samples
	(1)	(2)	(3)
Number of observations	80329	37596	
Average firm's age	7.47	6.99	-.48
	(5.02)	(4.77)	(.03)
Average number of cofounders	1.44	1.74	.31
	(.72)	(.98)	(.01)
Average capital raised within 3 first years (in millions of US dollars)	12.19	22.58	10.39
	(369.95)	(261.03)	(1.88)
Average firm's imputed number of employees	108.06	117.04	8.98
	(651.86)	(654.26)	(4.17)
Percentage of firm's patenting	1.7	6.16	4.46
	(12.94)	(24.05)	(.13)
Percentage of foreign-born founders	61.29	62.1	.81
	(25.08)	(25.25)	(.18)
Percentage of male founders	85.73	84.82	-.92
	(34.97)	(35.89)	(.22)

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

Table 3: Summary statistics For the Tuition IV and Shift-Share IV samples.

	Tuition IV sample		SSIV sample	
	Mean	Standard Deviation	Mean	Standard Deviation
Total number of graduates	651.11	649.53	482.17	771.57
Number of international graduates	86.13	142.80	63.45	188.56
Number of US graduates	564.98	540.02	418.72	636.78
Share of international graduates	9.63	9.77	7.69	12.30
Share of international graduates in research universities	14.91	9.61	19.58	13.49
Share of international graduates in other universities	6.02	8.08	5.55	10.77
Number of entrepreneurs	0.17	0.76	0.46	2.72
Number of companies	0.19	0.84	0.51	3.01
Share of local start-ups	56.51	45.50	50.64	42.42
Share of start-ups surviving at least 3 years	98.86	8.73	98.59	9.43
Share of start-ups with funding above the median	33.55	41.67	39.53	38.99
Share of patenting companies	3.34	15.24	3.48	13.69

Notes: This table presents summary statistics for each sample used in the two instrumental variable (IV) identification strategies. The tuition IV sample consists of 463 public universities, classified as Doctoral Universities or Master's Colleges according to the Carnegie Classification of Institutions of Higher Education, covering the period from 1999 to 2014. The SSIV sample includes 548 private, not-for-profit universities, also classified as Doctoral Universities or Master's Colleges, covering the period from 2004 to 2014.

Table 4: First stage estimates and power of IV's.

Dependent variable:	Share of international graduates			
	(1)	(2)	(3)	(4)
Sample:	Public universities		Not-for-profit universities	
Standardized predicted residuals from equation (1)	-0.8805*** (0.1121)	-0.8839*** (0.1119)	-	-
Shift-share IV	-	-	0.0782*** (0.0167)	0.0791*** (0.0171)
Log number of master graduates	- -	-0.1964 (0.4731)	- -	0.0355 (0.2482)
Observations	6,019	6,019	5,767	5,736
University fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
First-stage F-statistic	61.67	62.37	21.97	21.46

Notes: Coefficients of the first-stage regressions of the IV estimation procedure. Instruments: Cols (1-2): innovations of non resident tuition fees; estimation period: 1999-2014; sample: public US universities. Cols (3-4) predicted shares of foreign graduates based on shift-shares computed from 10 most important origins of foreign students; estimation period : 2004-2014; sample: private US universities.

Table 5: Impact of instruments on native and foreign graduates.

Dependent variable:	asinh(Number of graduates)		asinh(Number of foreign graduates)		asinh(Number of US graduates)	
	(1)	(2)	(3)	(4)	(5)	(6)
Standardized predicted residuals from equation (1)	-0.0171***		-0.1138***		-0.0063	
	(0.0061)		(0.0200)		(0.0062)	
Shift-share IV		-0.0003		0.0065***		-0.0015
		(0.0011)		(0.0020)		(0.0012)
Observations	6,019	5,736	6,019	5,736	6,019	5,736
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable is the Hyperbolic Sine transformation of the number of Total (1 and 2), foreign (3 and 4) and US (5 and 6) graduates. Cols 1, 3 and 5: instrument: innovation of non resident tuition fees; estimation period: 1999-2014; sample: public US universities. Cols 2, 4 and 6: instrument: shift-share; estimation period: 2004-2014; sample: not-for-profit US universities.

Table 6: Validity test of tuition and shift share IVs: correlation with pre-period trends.

Panel A: Residual tuition IV

Controls :	Share of international graduates (4 years lag) (1)	Share of international graduates (5 years lag) (2)	Share of international enrollees (4 years lag) (3)	Share of international enrollees (5 years lag) (4)	Entrepreneurs per 10,000 graduates (4 years lag) (5)	Entrepreneurs per 10,000 graduates (5 years lag) (6)
Dependent variable:						
Predicted residuals from equation (1)	-6.0862	-3.1070	2.9505	10.1554	-4.7347	-5.5349
	(6.3079)	(6.4109)	(13.8105)	(14.9880)	(6.2046)	(6.6461)
Observations	6,789	6,765	6,819	6,798	6,779	6,754
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Shift-share IV

Controls :	Change in number between 1995 and 2003					
	Graduates (1)	Foreign graduates (2)	US graduates (3)	Entrepreneurs (4)	Foreign entrepreneurs (5)	US entrepreneurs (6)
Dependent variable:						
Change in predicted share of international graduates from SSIV b/w 2004 and 2014	0.1233	1.9070***	-0.0687	-93.0237	-210.2281	-80.3994
	(0.1517)	(0.4835)	(0.1632)	(69.5576)	(165.8605)	(77.8220)
Observations	548	548	548	548	548	548

Notes: Panel A provides the correlation of the tuition based instrument with lagged values of the share of foreign enrolled and graduates 4 and 5 years before (cols 1-4) and lagged values of start-up founder rates. Unit of analysis: university-year. Panel B gives the cross-sectional correlation between the change of the share of international graduates predicted by the shift-share instrument with change in key explanatory variables and outcomes (number of foreign and US graduates and number of US and foreing start-up founders) over the period prior to the IV shares, namely 1995-2003. Col 1-3: previous change in total, foreign and native graduates. Col 4-6: previous change in total number of start-up founders, foreign ones and natives ones. Unit of analysis: university.

Table 7: Impact of foreign-student share on number of startups: main specifications.

Dependent variable:	Number of entrepreneurs		Number of companies	Number of companies survived>3y	Number of companies funded above median	Number of patenting companies	Number of local companies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Residual tuition IV							
Share of international master graduates	0.0005	0.0698***	0.0721***	0.0715***	0.0686***	0.0121***	0.0475**
	(0.0011)	(0.0258)	(0.0277)	(0.0276)	(0.0191)	(0.0042)	(0.0201)
Observations	6,019	6,019	6,019	6,019	6,019	6,019	6,019
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	.	61.67	61.67	61.67	61.67	61.67	61.67
Panel B: Shift share IV							
Share of international master graduates	0.0012	0.0923***	0.0913***	0.0911***	0.0710***	0.0059*	0.0514***
	(0.0018)	(0.0270)	(0.0280)	(0.0279)	(0.0210)	(0.0032)	(0.0174)
Observations	5,767	5,767	5,767	5,767	5,767	5,767	5,767
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-Statistic	.	21.97	21.97	21.97	21.97	21.97	21.97

Notes: OLS and IV estimations of equation (1). Panel A: instrument : innovations to non resident tuition fees (equation 2); estimation period : 1999-2015; sample: public US universities . Panel B : shift-share IV based on 10 most popular origins of foreign graduates; estimation period : 2004-2015; sample: not-for-profit US universities. Col. (1-2): number of start-up founders. Col 3: Number of firms. Col 4: Number of firms still active after 3 years. Col 5. Number of firms with funding above the sample average. Number of companies patenting within 3 years. Number of firms located in same graduation state. Standard errors clustered at university-cohort level.

Table 8: Impact of foreign student share on start-up creation: robustness checks-Tuition-IV.

Dependent variable:	Number of entrepreneurs	Number of companies	Number of companies surviving>3y	Number of companies funded above median	Number of patenting companies
	(1)	(2)	(3)	(4)	(5)
Panel A: Controlling for University × period fixed effects					
Share of international master graduates	0.1117*** (0.0411)	0.1200*** (0.0442)	0.1201*** (0.0440)	0.0992*** (0.0308)	0.0181*** (0.0066)
Observations	5,150	5,150	5,150	5,150	5,150
University x period fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	41.83	41.83	41.83	41.83	41.83
Panel B: Controlling for University × time-trend fixed effects					
Share of international master graduates	0.0695*** (0.0258)	0.0719*** (0.0278)	0.0713*** (0.0276)	0.0686*** (0.0191)	0.0121*** (0.0042)
Observations	6,019	6,019	6,019	6,019	6,019
University x time-trend fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	60.94	60.94	60.94	60.94	60.94
Panel C: Estimating semi-elasticity					
log number of international master graduates	0.5886** (0.2306)	0.6081** (0.2465)	0.6033** (0.2452)	0.5782*** (0.1766)	0.1018*** (0.0373)
Observations	6,019	6,019	6,019	6,019	6,019
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	37.55	37.55	37.55	37.55	37.55

Notes: This Table provides robustness checks of estimations results reported in Table 7, Panel A (tuition IV). In Panel A: equation (1) supplemented with university-period fixed effects. In Panel B: equation (1) supplemented with university specific trends. In Panel C: explanatory variable is (log of) number of foreign graduates rather than share of foreign graduates. Standard errors clustered at university-cohort level.

Table 9: Heterogeneity in the effects, across Majors and locations.

Dependent variable:	Number of entrepreneurs					
	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	STEM majors	non-STEM majors	Top 5 majors	Other majors	CZ with several universities	Other CZs
Panel A: Residual tuition IV						
Share of international	0.0325*	0.0373***	0.0350*	0.0348**	0.1582**	0.0311
master graduates	(0.0176)	(0.0142)	(0.0181)	(0.0144)	(0.0640)	(0.0281)
Observations	6,019	6,019	6,019	6,019	2,633	3,386
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	61.67	61.67	61.67	61.67	12.83	52.36
Panel B: Shift-share IV						
Share of international	0.0542***	0.0381***	0.0530***	0.0393***	0.0950***	0.0351
master graduates	(0.0180)	(0.0129)	(0.0191)	(0.0121)	(0.0301)	(0.0317)
Observations	5,767	5,767	5,767	5,767	4,268	1,499
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	21.97	21.97	21.97	21.97	19.36	5.424

Notes: Estimation of equation (1) with tuition IV. In col (1): Science, Technology, Engeneering and Mathematics (STEM) fields only. Col (2) : non STEM fields only. Col (3): top 5 majors of graduation only. Col (4): majors of graduation outside top 5. Col (5) graduates from universities located in commuting zones with more than 1 university. Col (6): Graduates from universities located in commuting zones with only 1 university. Standard errors clustered at university-cohort level.

Table 10: Spillover effects of foreign student share on US-origin start-ups, identified using probability of US-born from names.

Dependent variable:	Number of entrepreneurs	Number of companies	Number of surviving companies	Number of companies funded above median	Number of patenting companies
	(1)	(2)	(3)	(4)	(5)
Panel A: Residual tuition IV					
Share of international master graduates	0.0275**	0.0275**	0.0275**	0.0319***	0.0039**
	(0.0116)	(0.0125)	(0.0124)	(0.0091)	(0.0016)
Observations	6,019	6,019	6,019	6,019	6,019
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	61.67	61.67	61.67	61.67	61.67
Shift-share IV					
Share of international master graduates	0.0249***	0.0239**	0.0240***	0.0231***	0.0011
	(0.0090)	(0.0094)	(0.0093)	(0.0076)	(0.0012)
Observations	5,767	5,767	5,767	5,767	5,767
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	21.97	21.97	21.97	21.97	21.97

Notes: US and native born status imputed with probabilistic prediction from the Forebears algorithm. See Text for full explanation. Panel A: tuition-based instrument. Panel B: shift-share instrument. See Table 7 for definition of dependent variables used in cols (1-5). Standard errors clustered at university-cohort level.

Table 11: Spillover effects of foreign student share on US-origin start-ups, identified using 0-1 assessment of US-born from names.

Dependent variable:	Number of entrepreneurs	Number of companies	Number of surviving companies	Number of companies funded above median	Number of patenting companies
	(1)	(2)	(3)	(4)	(5)
Panel A: Residual tuition IV					
Share of international master graduates	0.0524**	0.0547**	0.0552**	0.0570***	0.0095**
	(0.0205)	(0.0223)	(0.0223)	(0.0159)	(0.0038)
Observations	6,019	6,019	6,019	6,019	6,019
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	61.67	61.67	61.67	61.67	61.67
Panel B: Shift-share IV					
Share of international master graduates	0.0411**	0.0420**	0.0428**	0.0415***	0.0025
	(0.0161)	(0.0171)	(0.0170)	(0.0145)	(0.0021)
Observations	5,767	5,767	5,767	5,767	5,767
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	21.97	21.97	21.97	21.97	21.97

Notes: US and native born status imputed with dichotomous (0-1) prediction from Forebears, based on the country of origin with largest probability. See main text for full explanation. Panel A: tuition-based instrumentation. Panel B: shift-share instrumentation. See Table 7 for definition of dependent variables used in cols (1-5). Standard errors clustered at university-cohort level.

Table 12: Effect of foreign-born share on number of co-founded startups.

Dependent variable:	Number of cofounded companies				
	All	from same univ.	from same univ-cohort	b/w US/foreign born founders	b/w US/foreign born founders from same univ.
	(1)	(2)	(3)	(4)	(5)
Panel A. Residual tuition IV					
Share of international master graduates	0.0702***	0.0248**	0.0114*	0.0151*	0.0095*
	(0.0194)	(0.0102)	(0.0059)	(0.0088)	(0.0057)
Observations	6,019	6,019	6,019	6,019	6,019
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	61.67	61.67	61.67	61.67	61.67
Panel B. Shift-share IV					
Share of international master graduates	0.0793***	0.0394***	0.0183***	0.0319***	0.0143**
	(0.0214)	(0.0113)	(0.0058)	(0.0110)	(0.0058)
Observations	5,767	5,767	5,767	5,767	5,767
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	21.97	21.97	21.97	21.97	21.97

Notes:Panel A: IV estimation with tuition based instrument. Panel B: IV estimation with shift-share instrument. Dependent variables : col (1): all co-founded firms; col (2) : number of co-founded firms with founders from same university; col (3) : number of co-founded firms with founders from same university and cohort; col (4): number of co-founded firms with founders with different foreign born/native status (predicted by Forebears, dichotomous approach); col (5): number of co-founded firms with founders with different foreign born/native status from same university (predicted by Forebears, dichotomous approach) . Standard errors clustered at university-cohort level.

Table 13: Effect of foreign-born share on count of startups differentiated from US ones and similar to Foreign ones, by keyword type.

Dependent variable:	Number of companies above median median distance with US start-ups	Number of companies above median median distance with US start-ups from same category	Number of companies below median median distance with start-ups abroad	Number of companies below median median distance with start-ups abroad from same category
	(1)	(2)	(3)	(4)
Panel A. Residual tuition IV				
Share of international master graduates	0.0309** (0.0142)	0.0415** (0.0171)	0.0401** (0.0167)	0.0429** (0.0167)
Observations	6,019	6,019	6,019	6,019
University fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
First-stage F-statistic	61.67	61.67	61.67	61.67
Panel B. Shift-share IV				
Share of international master graduates	0.0427*** (0.0129)	0.0464*** (0.0162)	0.0419** (0.0169)	0.0264* (0.0136)
Observations	5,767	5,767	5,767	5,767
University fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
First-stage F-statistic	21.97	21.97	21.97	21.97

Notes: Panel A: IV estimation with tuition based instrument. Panel B: IV estimation with shift-share instrument. Diversity is measured through the Euclidian distance from a set of 770 keywords describing the activity and type of start-ups. Distance of a given start-up is measured from the average frequency of keywords of start-ups. Standard errors clustered at university-cohort level.

Online Appendix

1 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.

1.1 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:

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

1.2 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:

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

1.3 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.

The share of foreign students in a school, obtained from $IPEDS$ wll then be instrumented by the imputed share that uses the shift-share imputation for foreigners as follows:

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

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.

2 Additional Tables

Table A1: Top 10 public and private universities by number of startups.

University	Number of start-ups (1)	Start-ups by (2) 10,000 graduates	Average share of (3) international graduates
Public universities			
University of California-Berkeley	106	33.94	22.77
University of California-Los Angeles	96	24.18	16.64
The University of Texas at Austin	75	16.59	22.24
University of Michigan-Ann Arbor	66	11.85	25.09
University of Washington-Tacoma Campus	36	148.7	1.47
University of Virginia-Main Campus	35	13.66	13.56
University of Central Florida	31	10.61	9.87
University of Colorado Boulder	29	16.47	14.79
University of North Carolina at Chapel Hill	23	7.83	10.19
Georgia Institute of Technology-Main Campus	21	8.84	39.77
Private universities			
Stanford University	483	144.71	33.55
Harvard University	350	62.99	29.03
Massachusetts Institute of Technology	278	111.13	36.25
Columbia University in the City of New York	208	24.85	32.9
University of Pennsylvania	199	42.72	27.38
Northwestern University	135	30.65	22.87
New York University	135	14.35	21.42
University of Chicago	128	32.58	25.53
University of Southern California	102	14.77	32.47
Carnegie Mellon University	94	34.63	44.43

Notes: Sources: Col (1-2): Crunchbase; col (3) : IPEDS.

Table A2: Preliminary regressions of the tuition IV procedure.

Dependent variable: :	Difference between in-state and out-of-state tuitions		
	(1)	(2)	(3)
Share state appropriations	313.4**	1,071***	2,879***
	(133.6)	(126.2)	(130.5)
Number of international graduates (1 year lag)	-	15.46***	3.773***
		(0.227)	(0.286)
University revenues (in dollars)	-	-	3.66e-06***
			(5.38e-08)
Revenues per student (in dollars)	-	-	-0.00111***
			(0.000213)
Instruction expenses per student (in dollars)	-	-	-0.000849*
			(0.000499)
All expenses per student (in dollars)	-	-	0.00112***
			(0.000241)
Constant	4,894***	4,189***	3,324***
	(45.97)	(44.49)	(46.82)
Observations	36,088	36,088	32,337
R-squared	0.115	0.216	0.320
Year FE	Yes	Yes	Yes
Degree FE	Yes	Yes	Yes

Notes: This table presents regressions of equation 2 used to generate $\widehat{\nu_{u,t}}$, the innovations in the difference between out-of-state and in-state tuition fees which are subsequently used as instruments in equations (1) and (3). Sources of all covariates: IPEDS.

Table A3: 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)	-0.0169***		-0.1044***		-0.0062	
	(0.0060)		(0.0170)		(0.0061)	
Shift-share IV		-0.0003		0.0059***		-0.0015
		(0.0011)		(0.0016)		(0.0011)
Observations	6,019	5,736	6,019	5,736	6,019	5,736
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: 1999-2014; sample: public US universities. Cols 2, 4 and 6: instrument: shift-share; estimation period : 2004-2014; sample: not-for-profit US universities. asinh stands for the inverse hyperbolic sine function. It is used because our dependent variables are counts that include some zero values.

Table A4: Validation for Shift-share IV of Indian graduates: correlation with pre-trends.

Controls :	Change in number between 1995 and 2003					
	Graduates	Foreign graduates	US graduates	Entrepreneurs	Foreign entrepreneurs	US entrepreneurs
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:						
Change in predicted share of Indian graduates from SSIV b/w 2004 and 2014	0.0188	0.2631***	-0.0074	-3.8272	-8.5925	-3.3139
	(0.0128)	(0.0399)	(0.0138)	(5.8877)	(14.0376)	(6.5840)
Observations	548	548	548	548	548	548

Notes: Table reports the cross-sectional correlation between the change of the share of Indian graduates predicted by the shift-share instrument with changes in key explanatory variables and outcomes over the period prior to the IV shares. Col 1-3: previous change in resp. total, foreign and native graduates. Col 4-6: previous change in resp. total number of start-up founders, foreign ones and natives ones. Unit of analysis: university.

Table A5: Robustness checks of overall effects estimated with shift-share IV.

Dependent variable:	Number of entrepreneurs	Number of companies	Number of surviving companies	Number of companies funded above median	Number of patenting companies
	(1)	(2)	(3)	(4)	(5)
Panel A: Controlling for University × period fixed effects					
Share of international master graduates	0.0644**	0.0635**	0.0642**	0.0539**	0.0094**
	(0.0319)	(0.0320)	(0.0322)	(0.0254)	(0.0047)
Observations	3,776	3,776	3,776	3,776	3,776
University x period fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	23.93	23.93	23.93	23.93	23.93
Panel B: Controlling for University × time-trend fixed effects					
Share of international master graduates	0.0909***	0.0897***	0.0896***	0.0701***	0.0058*
	(0.0269)	(0.0279)	(0.0277)	(0.0210)	(0.0032)
Observations	5,769	5,769	5,769	5,769	5,769
University x time-trend fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	21.81	21.81	21.81	21.81	21.81
Panel C: Estimating semi-elasticity					
log number of international master graduates	1.2819***	1.2690***	1.2651***	0.9838***	0.0818*
	(0.4636)	(0.4722)	(0.4704)	(0.3593)	(0.0482)
Observations	5,736	5,736	5,736	5,736	5,736
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	13.24	13.24	13.24	13.24	13.24

Notes: This Table provides robustness checks of estimations results reported in Table 7, Panel B (shift-share IV). In Panel A: equation (1) supplemented with university-period fixed effects. In Panel B: equation (1) supplemented with university specific trends. In Panel C: explanatory variable is (log of) number of foreign graduates rather than share of foreign graduates. Standard errors clustered at university-cohort level.

Table A6: Impact of foreign graduation on start-ups: accounting for size of native cohort.

Dependent variable:	Number of entrepreneurs	Number of companies	Number of surviving companies	Number of companies funded above median	Number of patenting companies
	(1)	(2)	(3)	(4)	(5)
Panel A. Tuition IV					
Share of international	0.0695***	0.0718***	0.0712***	0.0687***	0.0120***
master graduates	(0.0257)	(0.0277)	(0.0275)	(0.0190)	(0.0042)
log number of	0.0151	0.0186	0.0181	-0.0085	0.0015
master graduates	(0.0360)	(0.0384)	(0.0381)	(0.0335)	(0.0060)
Observations	6,019	6,019	6,019	6,019	6,019
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	62.37	62.37	62.37	62.37	62.37
Panel B. Shift-share IV					
Share of international	0.0951***	0.0942***	0.0939***	0.0730***	0.0061*
master graduates	(0.0274)	(0.0285)	(0.0283)	(0.0214)	(0.0033)
log number of	-0.0679**	-0.0705**	-0.0689**	-0.0561**	-0.0078***
master graduates	(0.0299)	(0.0311)	(0.0308)	(0.0220)	(0.0030)
Observations	5,736	5,736	5,736	5,736	5,736
University fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	21.46	21.46	21.46	21.46	21.46

Notes: IV estimations of equation (1) supplemented with size of cohort. Panel A: instrument: innovations to non resident tuition fees (equation 2). Estimation period: 1999-2015. Sample: public US universities. Panel B : shift-share IV based on 10 most popular origins of foreign graduates. Estimation period: 2004-2015. Sample: not-for-profit US universities. Definitions of dependent variables in cols (1-5): see Table 7.

Table A7: Robustness: estimates of impact of foreign graduation using origin specific shift-share instruments.

Panel A. Total effect

Dependent variable:	Number of entrepreneurs			
	(1)	(2)	(3)	(4)
	All SS IV	Indian and Chinese SS IV	Indian SS IV	Chinese SS IV
Share of international master graduates	0.0923***	0.1290***	0.0940**	0.1725***
	(0.0270)	(0.0359)	(0.0408)	(0.0638)

Panel B. Direct effect

Dependent variable:	Number of (likely) foreign-entrepreneurs			
	(1)	(2)	(3)	(4)
	All SS IV	Indian and Chinese SS IV	Indian SS IV	Chinese SS IV
Share of international master graduates	0.0511***	0.0737***	0.0548**	0.0961**
	(0.0148)	(0.0219)	(0.0252)	(0.0376)

Panel C. Spillover effect

Dependent variable:	Number of (likely) US-entrepreneurs			
	(1)	(2)	(3)	(4)
	All SS IV	Indian and Chinese SS IV	Indian SS IV	Chinese SS IV
Share of international master graduates	0.0411**	0.0551***	0.0387*	0.0763**
	(0.0161)	(0.0205)	(0.0229)	(0.0326)
Observations	5,767	5,767	5,767	5,767
University fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
First-stage F-statistic	21.97	26.38	15.91	12.18

This table presents some robustness checks of the shift-share IV estimates using specific countries of origin to generate the shift-share instrument. Panel A estimates the total effect using as dependent variable all start-up founders. Panel B estimates the direct effect using as dependent variable inferred foreign born start-up founders using the probabilistic measure from Forebears. Panel C estimates the spillover effect using as dependent variable inferred US born start-up founders using the probabilistic measure from Forebears. In col (1) : IV estimates using shift share based on 10 most important origin countries of foreign graduates. In col (2) IV estimates using shift share based on Indian and Chinese foreign graduates. In col (3) IV estimates using shift share based on Indian graduates only. In col (4) IV estimates using shift share based on Chinese graduates only. Identification of country of origin of graduates based on F1 visa data. Standard errors clustered at university-cohort level.

Table A8: Robustness test clustering standard errors at the cbsa x year level.

Dependent variable:	Number of entrepreneurs		Number of companies	Number of surviving companies	Number of companies funded above median	Number of patenting companies	Number of local companies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Residual tuition IV							
Share of international master graduates	0.0005	0.0698***	0.0721***	0.0715**	0.0686***	0.0121***	0.0475**
	(0.0011)	(0.0259)	(0.0279)	(0.0278)	(0.0193)	(0.0042)	(0.0203)
Observations	6,019	6,019	6,019	6,019	6,019	6,019	6,019
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	.	58.57	58.57	58.57	58.57	58.57	58.57
Panel B: Shift share IV							
Share of international master graduates	0.0012	0.0923***	0.0913***	0.0911***	0.0710***	0.0059*	0.0514***
	(0.0019)	(0.0285)	(0.0292)	(0.0289)	(0.0221)	(0.0033)	(0.0170)
Observations	5,767	5,767	5,767	5,767	5,767	5,767	5,767
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-Statistic	.	22.53	22.53	22.53	22.53	22.53	22.53

Notes: OLS and IV estimations of equation (1). Panel A: instrument : innovations to non resident tuition fees (equation 2); estimation period : 1999-2015; sample: public US universities . Panel B : shift-share IV based on 10 most popular origins of foreign graduates; estimation period : 2004-2015; sample: not-for-profit US universities. Col. (1-2): number of start-up founders. Col 3: Number of firms. Col 4: Number of firms still active after 3 years. Col 5. Number of firms with funding above the sample average. Number of companies patenting within 3 years. Number of firms located in same graduation state. Standard errors clustered at cbsa-year level.

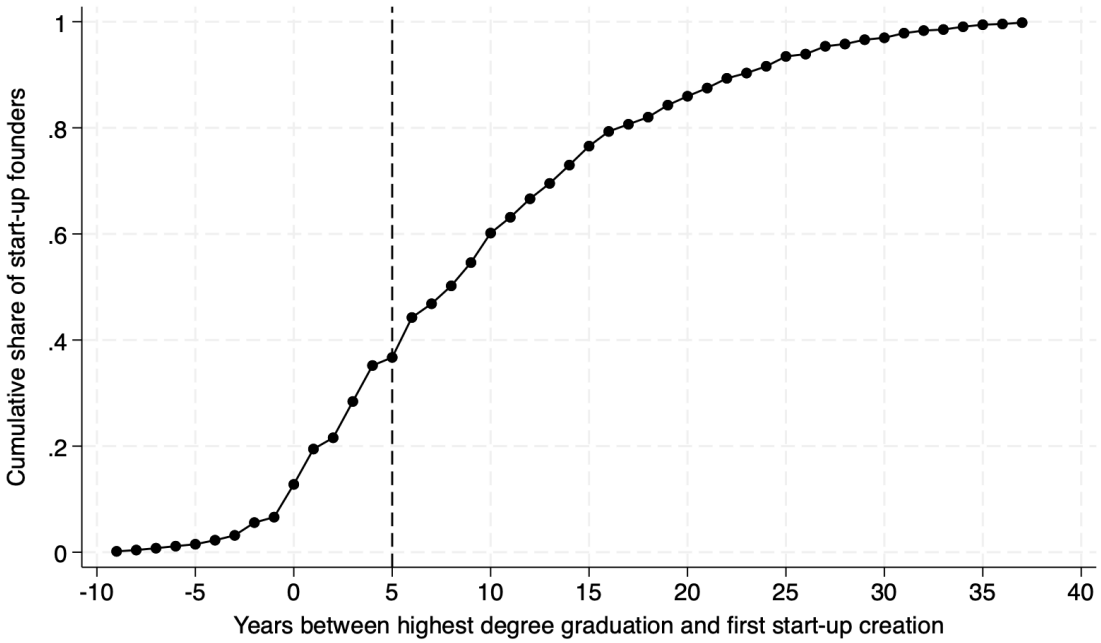
Table A9: Validity test: relationship between residual tuition-IV and graduation rate of bachelor students.

Dependent variable:	Graduation rate US-born bachelor students (1)	Graduation rate International bachelor students (2)	Difference b/w graduation rates of US-born and international bachelor students (3)
Predicted residuals from equation (1)	-0.3787	-0.1552	-0.3350
	(0.2390)	(0.5493)	(0.5818)
Observations	5,561	5,130	5,130
University fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes

Notes: This table regresses the ratio of bachelor graduates per bachelor enrollees by university cohort on the residual tuition IV specifically computed on tuition charged to Bachelor students. Standard errors are clustered at the university-cohort level.

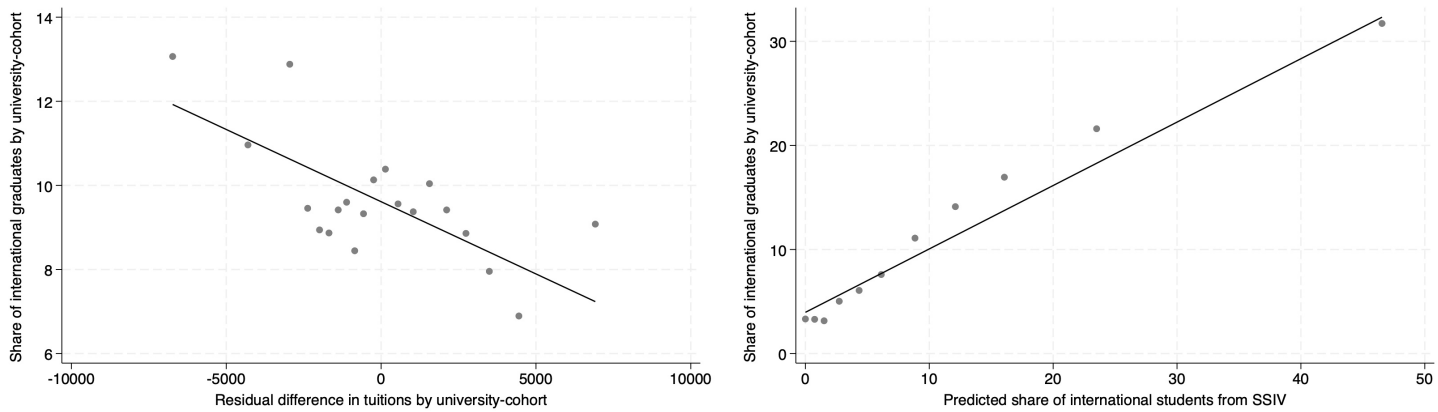
3 Additional Figures

Figure A1: Time between graduation and first start-up creation



Notes: This figure provides the cumulative distribution of the proportion of start-up creation dates compared with timing of graduation. Founders include US-born and foreign born founders. Timing of start-up creation drawn from Crunchbase database. Timing of graduation drawn from IPEDS.

Figure A2: Bin scatter plots of first stages of tuition-based IV and shift-share IV.



Notes: These figures present bin scatter plots of the tuition-based and shift-share instruments with the share of foreign graduates at the university-cohort level. Left panel: tuition-based instrument. Right panel: share of foreign graduates predicted by shift-share instrument.