Internet Appendix to:

The Role of High-Skilled Foreign Labor in Startup Performance: Evidence from Two Natural Experiments

Jun Chen[a], Shenje Hshieh[b], and Feng Zhang[c,*]

[a] ***Department of Finance, School of Business, Renmin University of China, 59 Zhongguancun Street, Haidian District, Beijing 100872, China.***
[b] ***Department of Economics and Finance, College of Business, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong, China SAR.***
[c] ***Department of Finance, David Eccles School of Business, University of Utah, 1655 Campus Center Drive, Salt Lake City, UT 84112, USA.***

First draft: April 2019
This draft: September 2020

[*] Corresponding author.
Email addresses: chen.jun@rmbs.ruc.edu.cn (Jun Chen), shshieh@cityu.edu.hk (Shenje Hshieh), feng.zhang@eccles.utah.edu (Feng Zhang)

Internet Appendix A:

Additional Tables and Figures

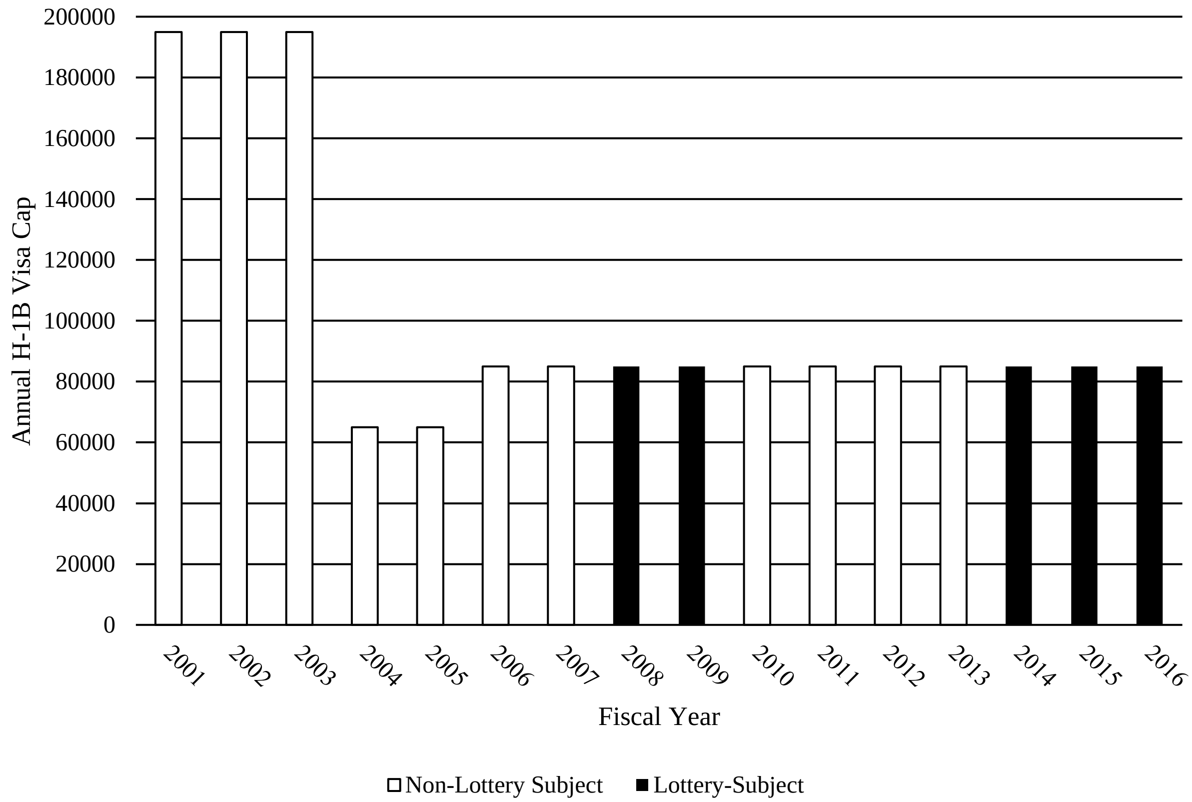

Fig. A1. Timeline of annual H-1B visa cap. This figure plots the annual H-1B visa cap from 2001 to 2016.

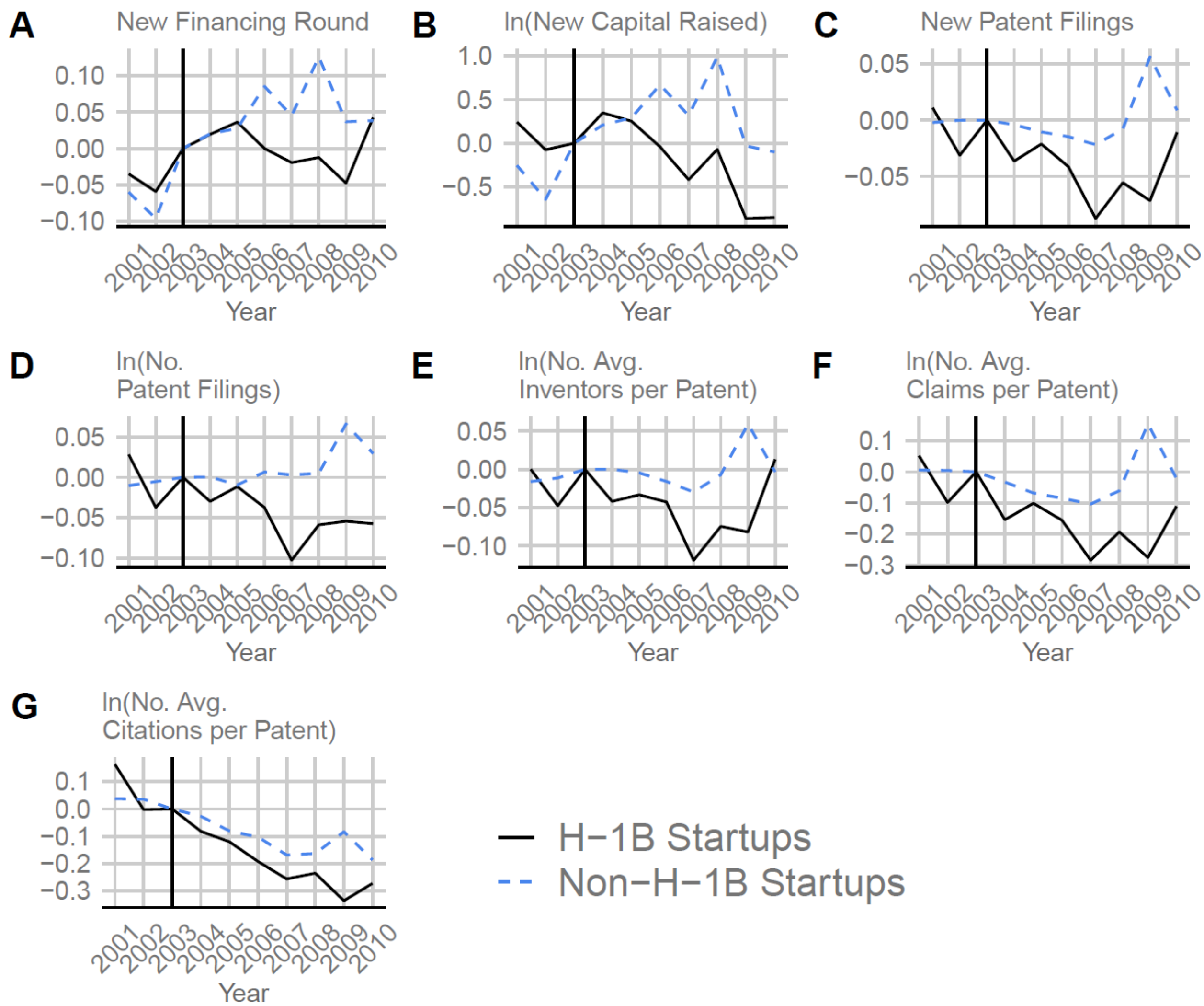

Fig. A2. Trends of the performance of H-1B and non-H-1B startups around 2004. This figure plots the mean normalized performance of H-1B startups and their characteristics-matched non-H-1B startups in each of the ten years from 2001-2010. Each startup's performance in each year is normalized by deducting its performance in 2003. The performance measures are the same as the dependent variables of the regressions in Table 5 Panel C. See the Appendix in the manuscript for variable definition.

Table A1
H-1B visa cap and demand in lottery years.

This table presents the annual H-1B visa cap and all U.S. firms' aggregate demand for capped H-1B visas, as well as the number of capped H-1B visas demanded by and granted to VC-backed startups. We extract the approximated aggregate demand data from archived USCIS news releases at https://www.uscis.gov/archive/ (last accessed on September 20, 2020), which are available only in the years when all capped H-1B visas are allocated using lotteries.

	All firms		VC-backed startups	
Fiscal year	Cap	USCIS reported demand	No. visas granted	Estimated demand
2008	85,000	123,480	681	1,331
2009	85,000	163,000	675	1,326
2014	85,000	124,000	1,745	2,677
2015	85,000	172,500	1,288	2,915
2016	85,000	233,000	1,127	3,403

Table A2
Summary statistics of VC-backed startups.

Panel A presents summary statistics of the sample of VC-backed startups. Total Financing Rounds is the number of VC financing rounds the startup has over its lifetime. Total Capital Raised is the aggregate amount of venture funds the startup raises over its lifetime. Total VC Funds is the number of VC funds that have financed the startup over its lifetime. Granted a Patent indicates whether the startup is granted a patent over its lifetime. Total Patents is the number of patents granted to the startup over its lifetime. Total Patent Claims is the total number of patent claims in the patents granted to the startup over its lifetime. Total Patent Citations is the total number of three-year forward citations to all patents granted to the startup over its lifetime. Panel B reports additional characteristics at the time of first LCA filing for VC-backed companies that have filed an LCA over its lifetime.

Panel A: Summary statistics

	N	Mean	Std. dev.	5-%ile	25-%ile	50-%ile	75-%ile	95-%ile
Total Financing Rounds	17,458	3.678	3.057	1.000	1.000	3.000	5.000	10.000
Total Capital Raised	17,458	37.793	147.066	0.000	2.750	11.579	36.754	136.824
Total VC Funds	17,458	5.559	4.636	1.000	2.000	4.000	8.000	15.000
Granted a Patent	17,458	0.210	0.408	0.000	0.000	0.000	0.000	1.000
Total Patents	17,458	2.647	23.164	0.000	0.000	0.000	0.000	12.000
Total Patent Inventors	17,458	8.337	69.689	0.000	0.000	0.000	0.000	37.000
Total Patent Claims	17,458	58.560	482.768	0.000	0.000	0.000	0.000	272.000
Total Patent Citations	17,458	23.750	190.674	0.000	0.000	0.000	0.000	75.000
IPO	17,458	0.043	0.202	0.000	0.000	0.000	0.000	0.000
Acquired	17,458	0.042	0.202	0.000	0.000	0.000	0.000	0.000
Founding year	17,458	2006.6	5.2	1999	2002	2007	2011	2015

Panel B: Characteristics of LCA-filing startups at the time of their first LCA filing

	N	Mean	Std. dev.	5-%ile	25-%ile	50-%ile	75-%ile	95-%ile
Company Age	6,156	3.356	2.726	0.000	1.000	3.000	4.000	9.000
Cum. Financing Rounds	6,156	2.217	2.145	0.000	1.000	2.000	3.000	6.000
Cum. Capital Raised	6,156	22.302	83.136	0.000	0.800	8.647	23.868	79.614
Cum. Patent Count	6,156	1.720	6.480	0.000	0.000	0.000	1.000	9.000
Cum. Patent Inventors Count	6,156	5.240	21.053	0.000	0.000	0.000	2.000	26.000
Cum. Patent Claims Count	6,156	41.725	155.411	0.000	0.000	0.000	16.000	225.000
Cum. Patent Citations Count	6,156	23.548	119.358	0.000	0.000	0.000	0.000	105.000

Table A3
Fraction of H-1B visa demand met in lotteries and startup performance.

This table presents estimation results for Eq. (1) with alternative model specifications. These results are discussed in Section 3.3 in the manuscript. In Panel A, we alter the model in four ways. In Panel B, we replace the outcome window over years [t, t+2] with three other windows. In Panel C, we estimate the model over two sub-periods. Panel D presents the weighted least-squares regression results of Eq. (1) using three different weights described below. To conserve space, we present only the coefficient on the fraction of the startup's demand for capped H-1B visas that is met. See the Appendix in the manuscript for variable definition. ***, **, and * correspond to statistical significance at the 1%, 5% and 10% levels, respectively. Standard errors in parentheses are clustered by company.

Panel A has four additional tests. The first test relates to over-petitioning. Firms might apply for more H-1B visas than they need in order to secure enough visas from the lotteries. To test whether this over-petitioning strategy affects our estimation results, we add visa demand to the regression as an additional control and observe that the main results barely change. To understand the effect of winning an additional H-1B visa on startup performance, the second test replaces the fraction of demand met with the number of visas the startup demands and the number of visas it wins in lotteries. The estimated effects of an additional H-1B visa are in line with the estimates from the baseline model. In the third test, we exclude startup characteristics from the regressions to check whether our results are driven by startup characteristics that correlate with the fraction of H-1B demand met. The results remain qualitatively unchanged, which is consistent with Table 3 Panel C and further reinforces that the fraction of H-1B demand met is the result of random lottery outcomes and is not driven by observed startup characteristics. In the last test, we add startup fixed effects to the regressions to test whether the effects of H-1B workers can be identified "within" repeat-appliers over time. This specification effectively excludes 66.3% of the 1,518 startups that appear in only one of the lottery years from the analysis and thus reduces the test power. As a result, including startup fixed effects in Eq. (1) will absorb much of the cross-sectional random variation in the lottery win rate, which could mask the systematic relation between the lottery win rate and startup performance.

Panel B tests whether our results are robust to alternative post-lottery windows. Our results remain robust to different post-lottery outcome windows. Panel C tests whether our results vary over sub-periods. The sub-period results are statistically weaker likely due to reduced test power stemming from a smaller sample size.

Panel D assesses the generalizability of our results based on the H-1B lottery sample to the whole population of VC-backed startups using weighted regressions, in which the weights make the lottery sample representative of the population on observed characteristics. We estimate the probability that a startup participates in the H-1B lottery in a logit model for each of the five years using the startup characteristics in the previous year as explanatory variables (the same as in Table 2). The logit model excludes fixed effects to avoid the incidental parameter problem (Chamberlain, 1980). The weighted regression results are similar if we include industry and location fixed effects and adjust for the incidental parameter bias using the methods of Fernandez-Val (2009) and Carro and Traferri (2011). We compute three types of weights for each startup in the H-1B lottery sample: the inverse probability of participation (IPP) weight, the overlap weight (one minus the probability of participation), and the stratification weight. The stratification method sorts the population into quintiles (strata) by the estimated probability of participation and assigns weights based on the proportion of participating startups within each quintile. Suppose a quintile has N startups, of which N_S are in the lottery sample. Each of these N_S startups receives a stratification weight of N/N_S. The weights indeed make the lottery sample more representative of the population. For example, the standardized difference in mean (SDM)—the difference in mean divided by the standard deviation (Austin and Stuart, 2015)—between the lottery sample and the population is on average 0.32 for the seven outcome variables, and sharply drops to less than 0.001 for each of the seven outcome variables after re-weighting the lottery sample with the IPP. The average SDM for the explanatory startup characteristics also significantly drops from 0.82 to 0.19 after re-weighting. To avoid extreme weights, we follow Crump et al. (2009) and Shadish and Steiner (2010) and cap the IPP weight in the regression at 62 (= 1/1.62%), which corresponds to the 10th percentile of the estimated probability of participation in the population (1.62%). The results are very similar if we cap the IPP weight at different cutoffs (e.g., 100 or 30). Panel D presents the weighted regression results for Eq. (1) using the IPP, overlap, and stratification weights. The results suggest that the H-1B worker effects are about 10% to 15% smaller in the population than in the lottery sample.

Panel A: Alternative model specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(No. Financing Rounds)	ln(New Capital Raised)	New Patent Filing	ln(No. Patent Filings)	ln(Avg. No. Inventors per Patent)	ln(Avg. No. Claims per Patent)	ln(Avg. No. Citations per Patent)
	Additional control of visa demand						
% H-1B Demand Met	0.055**	0.444**	0.054***	0.081***	0.058***	0.137***	0.006
	(0.022)	(0.201)	(0.019)	(0.031)	(0.022)	(0.048)	(0.026)
ln(No. Cap H-1B Demanded)	0.002	0.030	-0.002	0.108**	0.004	-0.009	-0.020
	(0.024)	(0.263)	(0.015)	(0.024)	(0.015)	(0.034)	(0.014)
	Replacing *% H-1B Demand Met* with demand and supply						
ln(No. Cap H-1B Granted)	0.053***	0.335*	0.037**	0.133***	0.050**	0.111**	0.013
	(0.020)	(0.194)	(0.019)	(0.039)	(0.023)	(0.050)	(0.024)
ln(No. Cap H-1B Demanded)	-0.037	-0.222	-0.031	0.012	-0.033	-0.092*	-0.029
	(0.023)	(0.218)	(0.020)	(0.041)	(0.027)	(0.055)	(0.022)
	No control variables except fixed effects						
% H-1B Demand Met	0.041**	0.399***	0.043**	0.071*	0.046**	0.118**	0.012
	(0.018)	(0.146)	(0.019)	(0.037)	(0.023)	(0.052)	(0.021)
	Additional control of company fixed effects						
% H-1B Demand Met	-0.009	0.145	0.037**	0.038	0.040***	0.091***	0.026*
	(0.024)	(0.263)	(0.015)	(0.024)	(0.015)	(0.034)	(0.014)

Panel B: Performance over alternative windows after H-1B visa lottery

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(No. Financing Rounds)	ln(New Capital Raised)	New Patent Filing	ln(No. Patent Filings)	ln(Avg. No. Inventors per Patent)	ln(Avg. No. Claims per Patent)	ln(Avg. No. Citations per Patent)
	Performance over years [t, t+1]						
% H-1B Demand Met	0.041*	0.371	0.050***	0.067**	0.058**	0.136***	0.004
	(0.022)	(0.232)	(0.019)	(0.027)	(0.023)	(0.051)	(0.028)
	Performance over years [t, t+2]						
% H-1B Demand Met	0.055**	0.441**	0.054***	0.071**	0.058***	0.138***	0.008
	(0.022)	(0.200)	(0.019)	(0.030)	(0.022)	(0.048)	(0.026)
	Performance over years [t, t+3]						
% H-1B Demand Met	0.067***	0.400**	0.054***	0.069**	0.050**	0.125***	0.004
	(0.022)	(0.186)	(0.019)	(0.031)	(0.021)	(0.047)	(0.026)
	Performance over years [t, t+4]						
% H-1B Demand Met	0.068***	0.345*	0.058***	0.071**	0.051**	0.132***	0.005
	(0.022)	(0.182)	(0.019)	(0.033)	(0.021)	(0.047)	(0.025)

Panel C: Regression results in sub-periods: 2008-2009 and 2014-2016

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(No. Financing Rounds)	ln(New Capital Raised)	New Patent Filing	ln(No. Patent Filings)	ln(Avg. No. Inventors per Patent)	ln(Avg. No. Claims per Patent)	ln(Avg. No. Citations per Patent)
	H-1B visa lotteries in 2008-2009						
% H-1B Demand Met	0.205***	0.851**	0.067*	0.132*	0.089*	0.206*	0.062
	(0.048)	(0.423)	(0.037)	(0.076)	(0.048)	(0.106)	(0.087)
	H-1B visa lotteries in 2014-2016						
% H-1B Demand Met	0.006	0.281	0.048**	0.053*	0.044*	0.110**	-0.005
	(0.025)	(0.229)	(0.021)	(0.030)	(0.024)	(0.054)	(0.017)

Panel D: Weighted regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(No. Financing Rounds)	ln(New Capital Raised)	New Patent Filing	ln(No. Patent Filings)	ln(Avg. No. Inventors per Patent)	ln(Avg. No. Claims per Patent)	ln(Avg. No. Citations per Patent)
	Inverse probability of participation weight						
% H-1B Demand Met	0.052**	0.492*	0.045**	0.044*	0.052***	0.110**	0.004
	(0.025)	(0.297)	(0.019)	(0.022)	(0.020)	(0.044)	(0.027)
	Overlap weight						
% H-1B Demand Met	0.050***	0.408**	0.045***	0.059**	0.054***	0.125***	0.012
	(0.018)	(0.190)	(0.017)	(0.025)	(0.019)	(0.043)	(0.024)
	Stratification weight						
% H-1B Demand Met	0.054*	0.409	0.043**	0.053**	0.050***	0.107**	0.001
	(0.031)	(0.486)	(0.018)	(0.022)	(0.019)	(0.042)	(0.026)

Table A4
Effects of the 2004 reduction in H-1B visa cap on the demand and supply of H-1B visas.

This table examines whether the 2004 H-1B visa cap reduction induced a shortage of foreign skilled labor for VC-backed startups that were dependent on H-1B visas. Panel A presents the OLS regression results of the diff-in-diff model in Eq. (3) over the 2001-2010 period for the sample of VC-backed startups that filed at least one LCA. The dependent variables are the number of H-1B visas the startup demands in year *t* (demand for H-1B visas), the number of H-1B visas it receives in year *t* (H-1B visas issued), and excess demand for H-1B visas in year *t* (demand for H-1B visas minus H-1B visas issued). *H1B* is a dummy variable equal to one if the VC-backed company received any approved I-129 petitions prior to 2004, and zero otherwise. Panel B replaces the lone interaction variable in Panel A with a set of interaction variables between the *H1B* dummy and year dummies; the interaction variable for the year of 2003 is omitted to avoid multi-collinearity and make 2003 the benchmark year. See the Appendix in the manuscript for variable definition. ***, **, and * correspond to statistical significance at the 1%, 5% and 10% levels, respectively. Standard errors in parentheses are clustered by company.

The reduced H-1B visa cap after 2004 will not harm startup performance if it does not make it more difficult for the H-1B startups to obtain H-1B visas. In this table, we test whether it becomes more difficult to obtain H-1B visas after 2004. The results suggest that it indeed becomes more difficult to obtain H-1B visas after 2004.

Panel A: Average effect

	(1)	(2)	(3)
	Demand for H-1B visas	H-1B visas issued	Demand minus issued
H1B × Year≥ 2004	-0.405	-1.152**	0.497***
	(1.709)	(0.515)	(0.089)
ln(Company Age)	2.859	-0.038	0.005
	(2.776)	(0.910)	(0.115)
Last 5Y Avg. Finan. Rds	-0.723	-0.325	0.001
	(1.238)	(0.415)	(0.068)
Last 5Y Avg. Cap. Raised	0.387	0.301	-0.004
	(0.553)	(0.283)	(0.040)
Last 5Y Avg. # Patents	0.940**	0.459***	0.027
	(0.375)	(0.151)	(0.029)
Observations	2,165	2,165	2,165
No. companies	617	617	617
Adj. R^2	0.340	0.370	0.067
Industry × Year × MSA FE	Yes	Yes	Yes
Company FE	Yes	Yes	Yes

Panel B: Dynamics of the effect

	(1)	(2)	(3)
	Demand for H-1B visas	H-1B visas issued	Demand minus issued
H1B × Year=2001	-2.298	-1.335	0.463*
	(5.181)	(1.454)	(0.246)
H1B × Year=2002	3.898	0.129	0.251*
	(3.247)	(0.670)	(0.138)
H1B × Year=2004	1.245	-1.156*	0.600***
	(2.855)	(0.663)	(0.131)
H1B × Year=2005	0.806	-0.934	0.576***
	(3.086)	(0.637)	(0.131)
H1B × Year=2006	0.816	-1.133*	0.675***
	(2.989)	(0.657)	(0.127)
H1B × Year=2007	1.155	-1.139	0.652***
	(2.898)	(0.721)	(0.135)
H1B × Year=2008	1.894	-1.293*	0.767***
	(2.844)	(0.719)	(0.140)
H1B × Year=2009	0.492	-1.971**	0.508***
	(3.017)	(0.881)	(0.156)
H1B × Year=2010	1.206	-1.884	0.535***
	(3.150)	(1.311)	(0.169)
Observations	2,165	2,165	2,165
No. companies	617	617	617
Adj. R^2	0.336	0.367	0.067
Controls	Yes	Yes	Yes
Industry × Year × MSA FE	Yes	Yes	Yes
Company FE	Yes	Yes	Yes

Table A5
Robustness checks of the diff-in-diff estimates based on the 2004 reduction in H-1B visa cap.

This table presents estimation results for Eq. (3) with alternative model specifications. To conserve space, we present only the coefficient on the key interaction variable. The dependent variables are measures of startup's venture financing and patenting activities. *H1B* is a dummy variable equal to one if the VC-backed company received any approved I-129 petitions prior to 2004, and zero otherwise. See the Appendix in the manuscript for variable definition. ***, **, and * correspond to statistical significance at the 1%, 5% and 10% levels, respectively. Standard errors in parentheses are clustered by company.

This table has five tests. (1) We test whether our results are robust to classifying a startup as an H-1B startup if it filed any LCAs before 2004. In the baseline model, we classify a startup as an H-1B startup if it was granted an H-1B visa before 2004. (2) We test whether the effects are different in the short run by estimating Eq. (3) over a shorter sample period of 2001–2007. (3) We test whether the coefficient estimates remain stable after excluding the startup characteristics and industry-year-MSA fixed effects from the regressions. (4) To ensure comparability between H-1B and non-H-1B startups, we use non-H-1B startups matched on the propensity to obtain H-1B visas as the non-H-1B group. (5) We test whether our results are robust to using all startups that were not granted an H-1B visa before 2004 as the non-H-1B group rather than using matched non-H-1B startups. We observe that our main results remain robust in these five additional tests. We also discuss these results in Section 4.2.1 of the manuscript.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	New Financing Round	ln(New Capital Raised)	New Patent Filing	ln(No. Patent Filings)	ln(Avg. No. Inventors per Patent)	ln(Avg. No. Claims per Patent)	ln(Avg. No. Citations per Patent)
	Classifying H-1B startups based on LCAs						
H1B × Year≥ 2004	-0.048*	-0.796***	-0.059***	-0.070***	-0.080***	-0.188***	-0.157***
	(0.025)	(0.230)	(0.017)	(0.025)	(0.024)	(0.053)	(0.042)
	Shorter sample period: 2001-2007						
H1B × Year≥ 2004	-0.034	-0.628***	-0.033*	-0.042*	-0.044*	-0.115**	-0.110***
	(0.026)	(0.233)	(0.018)	(0.025)	(0.025)	(0.055)	(0.042)
	Excluding startup characteristics and industry-year-MSA fixed effects						
H1B × Year≥ 2004	-0.015	-0.508***	-0.033**	-0.038*	-0.043**	-0.114**	-0.097**
	(0.021)	(0.193)	(0.015)	(0.023)	(0.022)	(0.048)	(0.039)
	Using propensity score matched non-H-1B startups						
H1B × Year≥ 2004	-0.086***	-1.025***	-0.055***	-0.056**	-0.081***	-0.206***	-0.170***
	(0.027)	(0.241)	(0.018)	(0.028)	(0.026)	(0.058)	(0.044)
	Using all non-H-1B startups instead of matched non-H-1B startups						
H1B × Year≥ 2004	-0.080***	-0.983***	-0.056***	-0.057**	-0.070***	-0.191***	-0.163***
	(0.023)	(0.207)	(0.015)	(0.023)	(0.022)	(0.047)	(0.037)

Table A6
Testing the imperfect labor substitution mechanism: alternative methods to compute the *HighSkill* and *Thin* dummies.

This table presents the OLS regression results of an augmented version of Eq. (3). Only coefficients on two interaction variables are shown to conserve space. *H1B* is a dummy variable equal to one if the VC-backed company received any approved I-129 petitions prior to 2004, and zero otherwise. *HighSkill* is a dummy equal to one if the startup's employees have high skill levels, and zero otherwise. *HighSkill* is computed using two methods from three proxies: an IT/Biotech dummy, the fraction of the startup's STEM jobs, and the startup's average H-1B worker salary. In the first method, we compute the percentile rankings of the two continuous proxies among the startups used in the regression and sum the percentile rankings together along with the IT/Biotech dummy. The *HighSkill* dummy equals one if the sum for the startup is above the median sum for all startups used in the regression, and zero otherwise. In the second method, we extract the first principal component of the three proxies and set the value of *HighSkill* to one if the first principal component for the startup is above the median value for the startups used in the regression, and zero otherwise. *Thin* is a dummy equal to one if the startup is located in an MSA with a thin local market for high skilled labor, and zero otherwise. *Thin* is computed using the same two methods based on three proxies: the number of universities, the number of STEM workers, and the number of public firms in the MSA where the startup was located in 2003. In addition to the control variables specified in Eq. (3), we also control for *HighSkill*, *Year≥2004 × HighSkill*, and *H1B × HighSkill* in Panel A, and control for *Thin*, *Year≥2004 × Thin*, and *H1B × Thin* in Panel B. The dependent variables are measures of the startup's venture financing and patenting activities. See the Appendix in the manuscript for variable definition. ***, **, and * correspond to statistical significance at the 1%, 5% and 10% levels, respectively. Standard errors in parentheses are clustered by company.

We test the imperfect local-foreign labor substitution mechanism in Table 8 of the manuscript using the indicator for startups with high employee skills (*HighSkill*) and the indicator for startups located in an MSA with thin local market for skilled labor (*Thin*). In Table 8, we construct *HighSkill* (*Thin*) using the index method, which sums up three dummies based on the three proxies for employee skill (labor market thinness). In this table, we test whether the results in Table 8 are robust to using *HighSkill* (*Thin*) indicators computed from two alternative methods: the percentile ranking method and the PCA method. We observe that the results in Table 8 are robust to these two new methods but are a bit weaker in the analysis of labor market thinness using the PCA method. We also discuss these results in Section 5.1 of the manuscript.

Panel A: H-1B worker effects in startups with high vs. low employee skills

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	New Financing Round	ln(New Capital Raised)	New Patent Filing	ln(No. Patent Filings)	ln(Avg. No. Inventors per Patent)	ln(Avg. No. Claims per Patent)	ln(Avg. No. Citations per Patent)
	High employee skills based on the percentile ranking method						
H1B × Year≥ 2004	0.130***	0.671*	0.019	0.015	0.027	0.049	0.078
	(0.043)	(0.378)	(0.032)	(0.048)	(0.045)	(0.104)	(0.082)
H1B × Year≥ 2004 × HighSkill	-0.251***	-2.007***	-0.108**	-0.129**	-0.143**	-0.344**	-0.364***
	(0.057)	(0.505)	(0.044)	(0.065)	(0.061)	(0.140)	(0.107)
Observations	22,465	22,465	22,465	22,465	22,465	22,465	22,465
No. companies	3,297	3,297	3,297	3,297	3,297	3,297	3,297
Observations with high skill	11,495	11,495	11,495	11,495	11,495	11,495	11,495
No. companies with high skill	1,561	1,561	1,561	1,561	1,561	1,561	1,561
	High employee skills based on principal component analysis						
H1B x Year≥ 2004	0.030	0.458	0.029	0.032	0.043	0.107	0.059
	(0.073)	(0.686)	(0.030)	(0.037)	(0.045)	(0.097)	(0.090)
H1B x Year≥ 2004 x HighSkill	-0.086	-1.437**	-0.085**	-0.102**	-0.120**	-0.302***	-0.221**
	(0.077)	(0.728)	(0.036)	(0.048)	(0.053)	(0.115)	(0.101)
Observations	22,465	22,465	22,465	22,465	22,465	22,465	22,465
No. companies	3,297	3,297	3,297	3,297	3,297	3,297	3,297
Observations with high skill	11,054	11,054	11,054	11,054	11,054	11,054	11,054
No. companies with high skill	1,575	1,575	1,575	1,575	1,575	1,575	1,575

Panel B: H-1B worker effects in startups located in areas with thin vs. thick local labor markets

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	New Financing Round	ln(New Capital Raised)	New Patent Filing	ln(No. Patent Filings)	ln(Avg. No. Inventors per Patent)	ln(Avg. No. Claims per Patent)	ln(Avg. No. Citations per Patent)
	Thin local labor market based on the percentile ranking method						
H1B × Year≥ 2004	0.007	-0.242	-0.018	-0.011	-0.039	-0.050	-0.043
	(0.036)	(0.333)	(0.025)	(0.036)	(0.037)	(0.080)	(0.063)
H1B × Year≥ 2004 × Thin	-0.102*	-0.968*	-0.074*	-0.127**	-0.062	-0.270**	-0.230**
	(0.056)	(0.515)	(0.044)	(0.062)	(0.061)	(0.135)	(0.105)
Observations	22,465	22,465	22,465	22,465	22,465	22,465	22,465
No. companies	3,297	3,297	3,297	3,297	3,297	3,297	3,297
Observations in thin market	11,367	11,367	11,367	11,367	11,367	11,367	11,367
No. companies in thin market	1,618	1,618	1,618	1,618	1,618	1,618	1,618
	Thin local labor market based on principal component analysis						
H1B × Year≥ 2004	-0.046*	-0.796***	-0.045**	-0.057**	-0.062**	-0.157***	-0.134***
	(0.026)	(0.237)	(0.018)	(0.026)	(0.025)	(0.055)	(0.043)
H1B × Year≥ 2004 × Thin	-0.234***	-2.254*	-0.203	-0.189*	0.045	-0.235	-0.129
	(0.076)	(1.344)	(0.169)	(0.111)	(0.220)	(0.607)	(0.611)
Observations	22,465	22,465	22,465	22,465	22,465	22,465	22,465
No. companies	3,297	3,297	3,297	3,297	3,297	3,297	3,297
Observations in thin market	11,339	11,339	11,339	11,339	11,339	11,339	11,339
No. companies in thin market	1,704	1,704	1,704	1,704	1,704	1,704	1,704

Internet Appendix B:
Measuring Demand and Supply of High-Skilled Foreign Labor Subject to H-1B Visa Cap

B.1. Demand for cap-subject high-skilled foreign labor

Companies must file LCAs and have their LCAs certified before petitioning the USCIS for H-1B visas. We therefore proxy their demand with their LCA filings, following Kerr and Lincoln (2010), Ghosh, Mayda, and Ortega (2016) and Xu (2018). Each LCA filing contains information on the intended hires, including employment start date and end date, job title, prevailing wage and worksite address. However, LCA petitioners do not indicate whether their intended hires are subject to annual H-1B visa cap or not. We determine whether or not the intended hires are cap-subject using the certification or approval date of the LCA by the U.S. Department of Labor.

LCAs certified after April are unlikely intended for cap-subject H-1B visas. The USCIS starts accepting I-129 petitions on the first business day in April; employers race to file I-129 petitions as early as possible because the annual cap is usually reached within a few days in the recent years. To win the race, employers have to obtain approved LCAs prior to April. Therefore, LCAs certified after April are unlikely intended for petitions subject to the annual H-1B visa cap.

Although employers must have their LCAs certified before April to be eligible for H-1B visa lotteries, they also have incentives to have their LCAs certified as late as possible. To illustrate this point, take the fiscal year 2009 as an example. Suppose a startup won an H-1B visa in the lotteries conducted on April 14, 2008. The earliest day the beneficiary H-1B foreign worker can start working for the startup is October 1, 2008, the start date of the government fiscal year of 2009.[1] The H-1B visa allows the worker to work for the startup for a maximum of three years; the startup has to petition the USCIS for an extension to the H-1B visa three years later if it wants to continue employing the worker after the H-1B visa expires. At the time of submitting the I-129 petition for the H-1B visa, the startup must also submit the certified LCA. The LCA itself has an employment start date, which is at most 180 days after the LCA certification date. For example, if the LCA was certified by the Department of Labor on November 1, 2007, the employment start date associated with the LCA can be any day in the 180-day window from November 1, 2007 to

[1] While H-1B workers do not actually have to start employment exactly on October 1, their visa status will change on October 1; any previously held visa status (e.g. student visas) will no longer be valid. This means whether employment begins on October 1 or not, their H-1B visa will expire in roughly three years from that date unless extended earlier.

April 29, 2008. The LCA is also valid for three years. Suppose the startup chose the LCA employment start date to be April 29, 2008 (the latest possible employment start date), the LCA would expire on April 28, 2011. But the earliest employment start date on the H-1B visa is October 1, 2008, which corresponds to an expiration date of September 30, 2011. The beneficiary H-1B worker has to stop working for the startup when the H-1B visa expires or when the LCA expires, whichever is earlier. Therefore, the H-1B worker can only work for the startup from October 1, 2008 (the H-1B employment start date) to April 28, 2011 (the LCA expiration date), which is 2 years and 7 months rather than 3 years. If the startup wants to continue to hire the H-1B worker after the LCA expires on April 28, 2011, it has to petition the USCIS for an extension to the H-1B visa, which is costly and time consuming. On the other hand, if the LCA was certified on March 31, 2008, the employment start dates of the LCA and the H-1B visa will coincide with each other (October 1, 2008). And they will expire on the same day (September 30, 2011). In this case, the H-1B worker can work for the startup for three years. To maximize the H-1B visa validity period, the startup should have the LCA certified as late as possible but before April 1, 2008 to avoid being excluded from the H-1B visa lotteries.

The example illustrates that companies have incentives to have their cap-subject LCAs certified on a day as close as possible to April 1 but before that so that they can submit their I-129 petitions on or right after April 1. Indeed, LCA filings drastically surge in March every year. Companies do try to maximize the H-1B visa validity period. Yet this does not exclude the possibility that some companies choose to file LCAs earlier than March, probably because of worries that the LCA will not be certified in time for I-129 petition starting on April 1 (it usually takes one week for the Department of Labor to certify an LCA, but the time it takes is uncertain ex ante). In addition, some companies could have filed cap-subject LCAs after April 1 in the few years after 2004. The H-1B visa cap of 65,000 was reached within 192 days after April 1 in 2004, within 132 days in 2005 and within 56 days in 2006. In these years, I-129 petitions submitted in late April (or even in May) were still eligible for cap-subject H-1B visas. But in the more recent years, only petitions submitted within a few days after April 1 are eligible for H-1B lotteries.

Taking these factors into account, we measure a company's demand for cap-subject high-skilled foreign workers with the number of intended hires in its LCAs certified between January and April. LCAs certified prior to January or after April are most likely for H-1B visa extensions or are not capped by the annual quota. In addition, the recruiting cycle of high-skilled foreign labor might start at the beginning of the year in January. We further require that the intended LCA hires must have an employment start date falling in the window of five to six months after the LCA certification date, because companies want to choose the LCA employment start date as close as possible to October 1 (the first day a cap-subject H-1B employee can legally work under H-1B) in order to maximize the validity period of the cap-subject H-1B visa.

B.2. The number of cap-subject H-1B visas a company wins in lotteries

The I-129 petitions database from the USCIS includes three types of processed H-1B petitions: 1) cap-subject petitions (notably, these include the lucky petitions that have been selected through lotteries); 2) cap-exempt petitions; and 3) petitions for extension of existing H-1B status, which are also cap-exempt. The second category includes I-129 petitions for foreign workers who hold advanced degrees and are to be employed in higher education or non-profit institutions in the U.S.; they are exempt from the annual cap.

The key to measure the supply of cap-subject H-1B visas for each startup is to determine to which of the above three categories a processed I-129 petition belongs. The employer must indicate whether an I-129 petition is for a visa extension. Therefore, it is easy to single out petitions in the third category. Yet it is more difficult to separate petitions in the first and second categories. The I-129 form does not contain any questions about whether the sought H-1B visa is cap-subject or cap-exempt until the version dated November 23, 2010. This and all subsequent versions of the I-129 form require the employer to specify whether the petition is cap-subject or cap-exempt. Although the I-129 forms after 2010 contain the information on identifying cap-subject vs. cap-exempt petitions, we only have such information for petitions after 2015 in the I-129 database. This is because USCIS continues to accept older versions of the I-129 form.

For petitions filed before 2015, we determine whether they are cap-subject using the following four criteria. First, we filter out petitions filed by non-profit organizations and higher education and government research institutions. These institutions are cap-exempt by definition. Second, we filter out I-129 petitions filed in months other than April to June. Cap-subject employees are expected to file I-129 petitions within a short period after April 1, to be eligible for working in the U.S. in the coming fiscal year that starts on October 1. Third, we require cap-subject petitions to have an employment start date after October 1. Whereas a cap-exempt applicant may very well have the leisure of belatedly submitting an I-129 petition in May to start a job in August, cap-subject applicants have to apply for H-1B visas for the next fiscal year that starts on October 1. That is, the employment start date of cap-subject petitions cannot be before October 1. Lastly, cap-subject petitions must check “new employment” for part 2 of question 2 on the I-129 form. We filter out petitions that do not check “new employment”.

The four filters accurately separate cap-exempt petitions from cap-subject ones. To verify their accuracy, we apply the four filters to the sample of processed I-129 petitions in 2015 and 2016, for which we have information on their cap status directly from the I-129 database. Among the 194,303 cap-subject petitions, with our filters we identify 193,606 of them, with a rate of 99.6%. Among the 534,162 cap-exempt petitions, we identify 528,154 of them, with a rate of 98.9%.

Our demand and supply measures turn out to be accurate. The demand measure is positively associated with the supply measure. In addition, the fraction of demand that is met, the ratio of supply to demand, is not correlated with company characteristics or past company performance. Furthermore, the fraction of capped H-1B visa demand met based on our measures is very close to the likelihood of winning H-1B lotteries disclosed by the USCIS. Had our demand or supply measure been too noisy, we would not observe these results.

Internet Appendix C:
Controlling for H-1B Worker Education

We estimate the following model (control variables omitted for brevity) in the H-1B lottery setting.

$$y = \alpha + \beta \frac{S_a + S_r}{D_a + D_r} + \epsilon, \tag{C1}$$

where D_a and D_r denote the number of advanced degree and regular H-1B workers the startup demands, respectively, and S_a and S_r are the corresponding number of H-1B visas the startup wins. Simple algebra shows that the explanatory variable is a weighted average of the lottery win rates of the two types of H-1B workers.

$$y = \alpha + \beta [\frac{S_a}{D_a} \times \frac{D_a}{D_a + D_r} + \frac{S_r}{D_r} \times \frac{D_r}{D_a + D_r}] + \epsilon. \tag{C2}$$

$\frac{S_a}{D_a}$ and $\frac{S_r}{D_r}$ are both exogenous but their weighted average is not if the weights are endogenous. Startups that demand proportionally more advanced degree H-1B workers will have more of their demand met since advanced degree H-1B petitions have a higher lottery win rate. That is, the fraction of demand met will not be random if H-1B worker education is not random across startups. Thus, H-1B worker education should be controlled in the regression. Unfortunately, it is omitted from the regression because it is unavailable in both the LCA and I-129 data. This omitted variable could bias the coefficient estimate. If advanced degree H-1B workers have a greater effect on startup performance, H-1B worker education will be positively correlated with both the fraction of demand met and startup performance. Hence, the coefficient estimate will be upward biased. See Angrist and Pischke (2008) and Roberts and Whited (2013) for derivations of the sign of the bias in coefficient estimate when there is an omitted variable.

The potential upward bias can be corrected by controlling for the fraction of the startup's demanded H-1B workers that have advanced degrees, $\frac{D_a}{D_a + D_r}$ in Eq. (C2). While H-1B worker education degree is unavailable in both the LCA and I-129 data, $\frac{D_a}{D_a + D_r}$ can be inferred for a small fraction (10.5%) of the startup-years in the H-1B lottery sample. The Department of Labor requires that the wage offered to an H-1B worker must be the prevailing wage paid to similarly employed workers in the same occupation in the area of intended employment. The employer can satisfy this requirement by using prevailing wage rates from multiple sources including the National

Prevailing Wage and Center (NPWC), the Occupational Employment Statistics program, and surveys conducted by the employer or its consultants. The NPWC prevailing wages (available since 2010) are based on job codes. For about half of these job codes, the NPWC provides information on the typical worker's education degree. This information allows us to infer whether an LCA applicant has an advanced degree if the employer chooses to use the NPWC data to determine the prevailing wage and the job code specified on the LCA has an associated education degree. Note that the inferred education degree may not be accurate since the NPWC data only provide information on the typical worker's education degree for each job code. It is possible for workers with the same job code to have different education degrees.

Based on the NPWC data, we can infer whether the applicant has an advanced degree for about 20% of the applicants in the cap-subject LCAs filed by our sample of VC-backed startups from 2014-2016. This fraction drops to 15% if we include the cap-subject LCAs filed in 2008-2009 in the calculation because the NPWC data are unavailable before 2010. In many cases, we can only infer the education degree of some of a startup's LCA applicants because many job codes do not have an associated education degree. As a result, we can compute the fraction of the startup's demanded H-1B workers that have advanced degrees for only 10.5% (= 232/2216) of the startup-years in our H-1B lottery sample. This small sample size likely does not allow for powerful statistical tests.

Nevertheless, we examine whether omitting H-1B worker education from the regression biases the coefficient on the fraction of H-1B demand met using the small sample with available data on H-1B worker education. Table C1 Panel A (on the next page) presents the regression results of Eq. (1) using this small sample. The coefficient on the fraction of H-1B demand met remains positive in six of the seven columns and is statistically significant in three columns within this small sample. In Table C1 Panel B, we additionally control for the fraction of the startup's demanded H-1B workers that have advanced degrees in the regressions. Adding H-1B worker education barely changes the coefficients on the fraction of H-1B demand met. In addition, the coefficient on the fraction of advanced degree H-1B workers is insignificant throughout the seven columns. These results, together with the evidence discussed in Section 3.3.1, suggest that the omission of H-1B worker education is unlikely to bias the coefficient on the fraction of H-1B demand met.

Table C1
Fraction of H-1B demand met in lotteries and startup performance: controlling for H-1B worker education.

Panel A presents OLS estimation results of Eq. (1) using the sample of startup-years in which we can infer H-1B workers' education degree. In Panel B, we additionally control for the fraction of the startup's demanded H-1B workers that have advanced degrees. The key independent variable is the fraction of the startup's demand for cap-subject H-1B visas that is met (or the probability of winning H-1B visas). The dependent variables are measures of the startup's venture financing activities and patenting activities over years *t* to *t+2* after the lottery. See the Appendix in the manuscript for variable definition. ***, **, and * correspond to statistical significance at the 1%, 5% and 10% levels, respectively. Standard errors in parentheses are clustered by company.

Panel A: Without controlling for H-1B worker education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(No. Financing Rounds)	ln(New Capital Raised)	New Patent Filing	ln(No. Patent Filings)	ln(Avg. No. Inventors per Patent)	ln(Avg. No. Claims per Patent)	ln(Avg. No. Citations per Patent)
% H-1B Demand Met	-0.035	0.624	0.088*	0.102	0.100*	0.264**	0.013
	(0.066)	(0.577)	(0.045)	(0.068)	(0.052)	(0.120)	(0.042)
ln(Company Age)	0.086	-0.928	-0.010	-0.086	-0.034	-0.085	-0.004
	(0.063)	(0.759)	(0.051)	(0.074)	(0.061)	(0.134)	(0.045)
Last 5Y Avg. Finan. Rds	0.326***	0.258	0.012	-0.084	-0.022	-0.035	-0.080
	(0.088)	(0.828)	(0.076)	(0.105)	(0.088)	(0.200)	(0.064)
Last 5Y Avg. Cap. Raised	-0.097	-0.121	-0.041	-0.074	-0.031	-0.060	-0.010
	(0.063)	(0.623)	(0.042)	(0.069)	(0.045)	(0.122)	(0.044)
Last 5Y Avg. # Patents	-0.008	0.457**	0.198***	0.497***	0.253***	0.551***	0.065***
	(0.024)	(0.177)	(0.016)	(0.039)	(0.022)	(0.044)	(0.024)
Observations	232	232	232	232	232	232	232
No. companies	205	205	205	205	205	205	205
Adj. R^2	0.109	0.129	0.526	0.787	0.551	0.573	0.307
Industry × Year × MSA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Controlling for H-1B worker education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(No. Financing Rounds)	ln(New Capital Raised)	New Patent Filing	ln(No. Patent Filings)	ln(Avg. No. Inventors per Patent)	ln(Avg. No. Claims per Patent)	ln(Avg. No. Citations per Patent)
% H-1B Demand Met	-0.035	0.629	0.089*	0.102	0.100*	0.265**	0.013
	(0.066)	(0.580)	(0.045)	(0.068)	(0.052)	(0.120)	(0.042)
% Adv. deg. H-1B workers	0.025	0.916	0.027	0.065	0.123	0.145	-0.002
	(0.075)	(0.753)	(0.059)	(0.085)	(0.088)	(0.159)	(0.047)
ln(Company Age)	0.088	-0.865	-0.008	-0.082	-0.026	-0.075	-0.005
	(0.064)	(0.743)	(0.051)	(0.074)	(0.060)	(0.131)	(0.045)
Last 5Y Avg. Finan. Rds	0.325***	0.206	0.010	-0.088	-0.029	-0.043	-0.080
	(0.088)	(0.804)	(0.076)	(0.104)	(0.087)	(0.200)	(0.064)
Last 5Y Avg. Cap. Raised	-0.098	-0.149	-0.042	-0.076	-0.035	-0.064	-0.010
	(0.063)	(0.625)	(0.042)	(0.068)	(0.043)	(0.118)	(0.044)
Last 5Y Avg. # Patents	-0.009	0.446**	0.197***	0.496***	0.251***	0.549***	0.065***
	(0.024)	(0.178)	(0.016)	(0.039)	(0.022)	(0.044)	(0.024)
Observations	232	232	232	232	232	232	232
No. companies	205	205	205	205	205	205	205
Adj. R^2	0.104	0.133	0.524	0.787	0.557	0.573	0.303
Industry × Year × MSA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Internet Appendix D:

Effects of Advanced Degree vs. Regular H-1B Workers on Startup Performance: Evidence from Bootstrap Simulations

The USCIS conducts two sequential random lotteries to allocate the 20,000 advanced degree H-1B visas and the 65,000 regular H-1B visas each year. The first lottery randomly grants the 20,000 advanced degree H-1B visas to the N_a petitioners with master's or more advanced degrees. The remaining (N_a – 20,000) advanced degree petitioners, who do not win a visa in the first lottery, are put into the pool of regular H-1B visa petitioners and thus have a second chance of winning one of the 65,000 regular H-1B visas. Suppose there are N_r regular petitioners. The two-step lottery procedure results in a higher probability of winning H-1B visa for advanced degree petitioners than regular petitioners. Specifically, an advanced degree petitioner's probability of winning a visa equals $20{,}000/N_a + (1 - 20{,}000/N_a)*65{,}000/(N_a - 20{,}000 + N_r)$, which is higher than a regular petitioner's counterpart probability of $65{,}000/(N_a - 20{,}000 + N_r)$ as long as the aggregate demand for H-1B visas ($N_a + N_r$) surpasses the total quota of 85,000 visas.

Ideally, we would estimate the probability of winning H-1B visa separately for advanced degrees petitioners and regular petitioners. The reason is that advanced degree H-1B workers most likely enhance startup performance by more than regular H-1B workers do (Abowd et al., 2005). However, neither the LCA data nor the I-129 data distinguish between advanced degree vs. regular H-1B workers. The data limitation affects all studies that use the data, including ours. As a result, existing studies estimate the average effect of advanced degree and regular H-1B workers on company performance using their average probability of winning H-1B visa (this is the *% H-1B Demand Met* variable in our study) over the years when all visas are allocated via lotteries.

In this appendix, we investigate what the estimated average effect could tell us about the individual effects of the two types of H-1B workers using data generated from Monte Carlo simulations. In each simulation, we generate 1,000 startups and randomly generate each startup's demand for H-1B workers. We ensure that the distribution of the randomly generated demand matches the distribution of startups' actual annual demand for H-1B workers over the five years in which all H-1B visas were allocated via lotteries. Almost all (more than 98%) startups' annual demand is fewer than 20 H-1B workers. We therefore cap the demand for H-1B visas at 20 in the simulation to prevent outliers. Among the startup-years with demand fewer than 20 H-1B visas, 52.9% of them demand one H-1B visa, 20.1% 2 visas; the percentage steadily decreases as the number of visas demanded increases and drops to 3.6% for five visas, 0.6% for ten visas, and finally merely 0.3% for 20 visas. For each simulated startup, we generate its H-1B demand by randomly drawing a number between one and 20 based on the observed probabilities of the 20 possible numbers of visas demanded. For example, the simulated startup demands one visa with the probability of 52.9%, two visas with the probability of 20.1%, and so on, and 20 visas with the probability of 0.3%.

We then randomly classify each demanded H-1B worker as an advanced degree H-1B worker with a probability of 20% and as a regular H-1B worker with a probability of 80%. The USCIS does not disclose the fraction of petitions that have advanced degrees. We use the fraction of 20% based on the estimate by a website that has been closely following H-1B visa lotteries.[2] We vary this fraction in additional simulations below. We then randomly grant an H-1B visa to each advanced degree petitioner with a probability of 70%, and to each regular petitioner with a probability of 42%. The two probabilities are also based on estimates of the aforementioned website; changing the probabilities does not qualitatively alter the simulation results. We then compute each simulated startup's probability of winning H-1B visa, P_{all}, as the ratio of the number of visas it wins to the total number of visas it demands. Note that this is the variable *% H-1B Demand Met* in our lottery-based experiment. We also compute each startup's probability of winning advanced degree H-1B visa (P_{adv}) as the number of advanced degree visas it wins divided by the number of advanced degree visas it demands and its probability of winning regular degree visa (P_{reg}) in the same way.

Next, we randomly generate startup i's performance using the following formula: $Y_i = C_{adv}P_{adv} + C_{reg}P_{reg} + e_i$. We set P_{adv} (P_{reg}) to zero if the startup does not demand any advanced degree (regular) H-1B visa. We normalize the true effect of advanced degrees H-1B workers on startup performance, C_{adv}, to one and vary the true effect of regular H-1B works, C_{reg}, from zero to one. e_i is a random number following the normal distribution with mean zero and standard deviation 0.55. We choose 0.55 because the actual probability of winning H-1B visa is 0.55 in our data. We then run the following regression using the simulated 1,000 startups, $Y_i = a + C_{all}P_{all} +$

[2] See the estimated fractions of petitioners with advanced degrees for fiscal years 2014-2016 on this website: https://redbus2us.com/h1b-historical-data-lottery-vs-85k-quota-masters-vs-regular-split/ (last accessed on September 20, 2020).

u_i. We repeat the simulation 1,000 times and compare the average estimated coefficient $\hat{C}_{all}$ to the true effects, C_{adv} and C_{reg}.

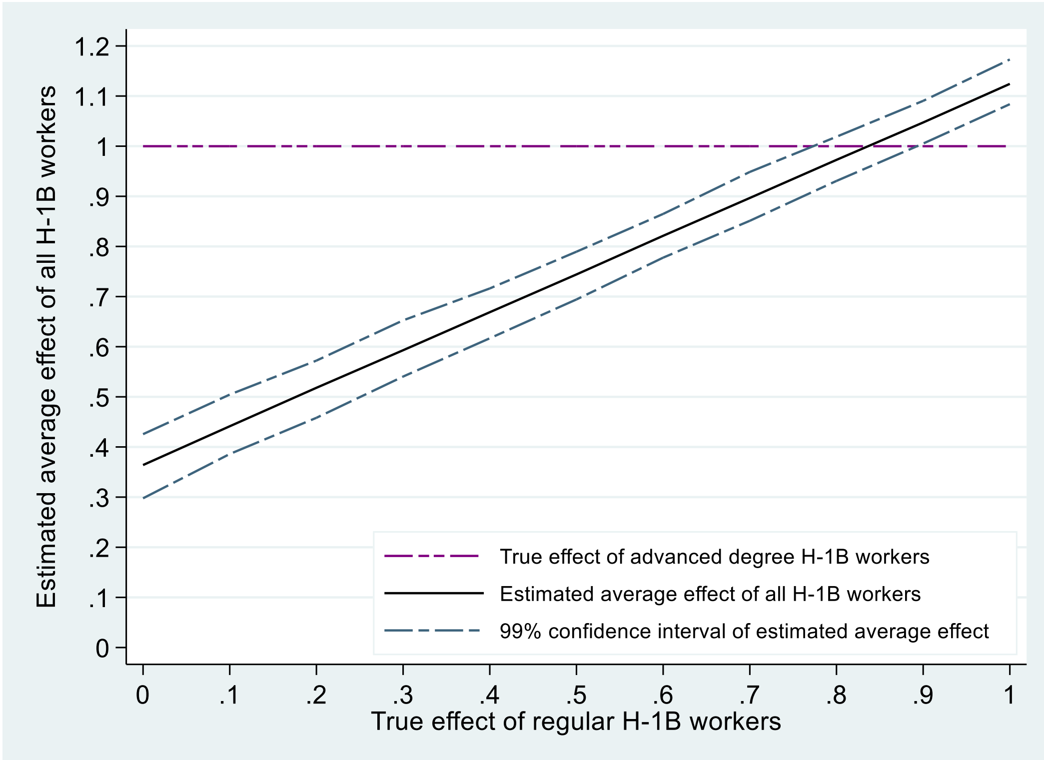

Fig. D1. This figure plots the average estimated coefficient on the startup's fraction of H-1B demand met across 1,000 bootstrap simulations, and its 1% and 99% percentiles. The true effect of advanced degree H-1B workers is normalized to one and that of regular H-1B workers varies from zero to one in different sets of simulations. The fraction of petitioners that have advanced degrees is set to 20%.

Fig. D1 depicts the average estimated coefficient $\hat{C}_{all}$ and the 99% confidence intervals across the 1,000 simulations. In the extreme case that regular H-1B workers have no effects on startup performance, the estimated coefficient overstates the effect of regular H-1B workers but underestimates the true effect of advanced degree H-1B workers by as much as 65%. The estimated coefficient monotonically increases with the true effect of regular H-1B workers. It lies in between the true effects of advanced degree H-1B workers and regular H-1B workers when the true effect of regular workers is below 0.8 and becomes one when the true effect of regular workers reaches 0.8. That is, the estimated coefficient underestimates the true effect of advanced degree workers but overstates the true effect of regular workers when the true effect of regular workers is relatively small (i.e., below 0.8); it overestimates the true effects of both advanced degree workers and regular workers if the true effect of regular workers is above 0.8. It is reasonable that the effect of regular H-1B workers on startup performance is less than 80% of the effect of advanced degree H-1B workers (Abowd et al., 2005). Therefore, the estimated coefficient on *% H-1B Demand Met* in our lottery-based experiment likely underestimates the true effect of advanced degrees H-1B workers but overestimates the effect of regular H-1B workers. The differences between the coefficient estimate and the effects of the two types of H-1B workers shrink as the true effect of regular workers converges to that of advanced degree workers.

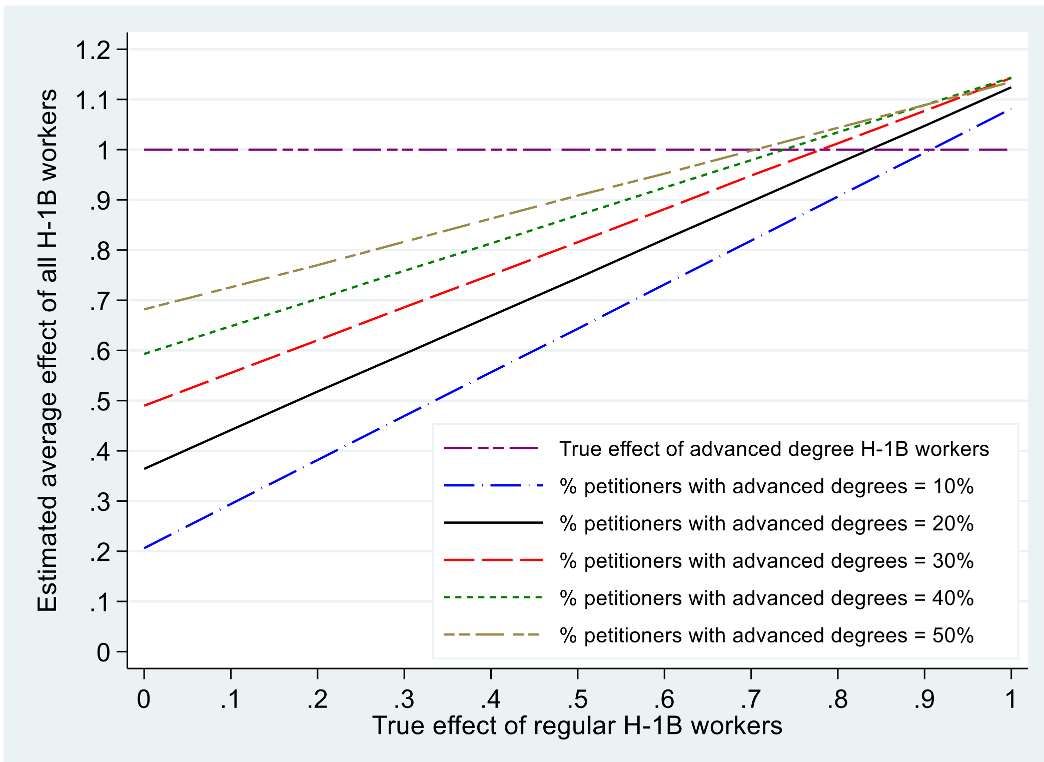

Fig. D2. This figure plots the average estimated coefficient on the startup's fraction of H-1B demand met across 1,000 bootstrap simulations. The true effect of advanced degree H-1B workers is normalized to one, and that of regular H-1B workers varies from zero to one in different sets of simulations. The fraction of petitioners that have advanced degrees is set to 10%, 20%, 30%, 40%, or 50%.

In the above simulations, we assume that 20% of H-1B visa applicants have advanced degrees, a fraction computed from the aggregate number of applicants with advanced degrees. To check the impact of the fraction of H-1B applicants with advanced degrees, we vary the fraction from 10% to 50% in additional sets of simulations. The results, depicted in Fig. D2, are consistent with the above insights. That is, the estimated coefficient underestimates the true effect of advanced degrees

H-1B workers but overestimates the effect of regular H-1B workers under reasonable parameter values.

Furthermore, we implement bootstrap simulations assuming both advanced degree and regular H-1B workers do not affect startup performance. In unreported results, we find that our lottery-based model does not produce false significant coefficient on *% H-1B Demand Met* when H-1B workers do not affect startup performance at all.

In summary, the two-step lotteries lead to higher probabilities of winning H-1B visas for advanced degree petitioners than for regular petitioners. Since advanced degree H-1B workers likely have greater effects on startup performance than regular H-1B workers, we need to estimate the probability of winning H-1B visas separately for advanced degree petitioners and for regular petitioners and use the two probabilities to estimate their different effects on startup performance. Yet, the data do not allow us to distinguish between advanced degree H-1B petitioners and regular H-1B petitioners. We therefore estimate the average effect of advanced degree H-1B workers and regular H-1B workers. Our bootstrap simulations show, unsurprisingly, that the estimated average effect could underestimate the effect of advanced degree H-1B workers but overstate that of regular H-1B workers.

References:

Abowd, J. M., Haltiwanger, J., Jarmin, R., Lane, J., Lengermann, P., McCue, K., McKinney, K., Sandusky, K., 2005. The relation among human capital, productivity, and market value: building up from micro evidence. In: Corrado, C., Haltiwanger, J., Sichel, D. (Eds.), Measuring Capital in the New Economy. University of Chicago Press, Chicago, IL, pp. 153-203.

Angrist, J.D., Pischke, J.-S., 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, NJ.

Austin, P.C., Stuart, E.A., 2015. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Statistics in Medicine 34, 3661-3679.

Carro, J.M., Traferri, A., 2014. State dependence and heterogeneity in health using a bias-corrected fixed-effects estimator. Journal of Applied Econometrics 29, 181-207.

Chamberlain, G., 1980. Analysis of covariance with qualitative data. The Review of Economic Studies 47, 225-238.

Crump, R.K., Hotz, V.J., Imbens, G.W., Mitnik, O.A., 2009. Dealing with limited overlap in estimation of average treatment effects. Biometrika 96, 187-199.

Fernández-Val, I., 2009. Fixed effects estimation of structural parameters and marginal effects in panel probit models. Journal of Econometrics 150, 71-85.

Ghosh, A., Mayda, A.M., Ortega, F., 2016. The impact of skilled foreign workers on firms: an investigation of publicly traded US firms. Unpublished working paper. Georgetown University and City University of New York.

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

Roberts, M.R., Whited, T.M., 2013. Endogeneity in empirical corporate finance. In: Constantinides, G.M., Harris, M., Stulz, R.M. (Eds.), Handbook of the Economics of Finance. Elsevier, Amsterdam, Netherlands, pp. 493-572.

Shadish, W.R., Steiner, P.M., 2010. A primer on propensity score analysis. Newborn and Infant Nursing Reviews 10, 19-26.

Xu, S.-J., 2018. Skilled labor supply and corporate investment: evidence from the H-1B visa program. Unpublished working paper. University of Alberta.