

THREE ESSAYS IN IMMIGRATION ECONOMICS

Ashlee Li

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy
in
Economics and Public Policy



Carnegie Mellon University

March 1, 2022

Dissertation Committee:

Brian K. Kovak

Rebecca Lessem

Lowell Taylor

Erina Ytsma

Abstract

The first chapter of the dissertation examines the role of migration networks in the location choices of Mexican-born immigrants, who are found to be highly mobile in response to labor demand shocks. We rely on the sizable variation in labor demand declines across states during the Great Recession to identify migration responses to demand shocks and use a novel set of data, the *Matrícula Consular de Alta Seguridad* (MCAS) data, to construct migration network measures. We find that migration networks indeed play an important part in Mexican migrants' responsiveness to local demand shocks. In particular, migrants respond to local economic conditions as well as conditions in network-connected locations when making location decisions.

In recent years, the competition for the H-1B visas has become so intense that a random lottery has been put in place to meet the specific quotas. In the second chapter, I examine the role of the H-1B lottery in the self-selection of the skilled immigrant workers in the U.S. Building on the basis of the Roy-Borjas Model, I construct a self-selection framework with the uncertainty in obtaining legal admission into the host country incorporated. As the proposed framework suggests, a decrease in the probability of obtaining the H-1B visas will likely reinforce the direction of the self-selection of international workers, be it positive or negative. Empirically, I use the Form I-129 data from USCIS and find evidence that the presence of the H-1B lottery and its decreasing odds have contributed to the deterioration in the quality of skilled international workers applying for H-1B visas in the U.S.

As we know, most developed countries have engaged in a race for global talent, as skilled

immigration becomes increasingly imperative in maintaining competitive advantages and galvanizing economic growth through various channels. In the last chapter, I study a novel firm-level channel, in which skilled immigrant workers potentially drive growth by improving start-ups' access to international venture capital funding. My first empirical approach makes use of the exogenous variation in firms' access to high-skilled foreign workers as a result of the H-1B lotteries. To supplement in terms of external validity, I extend the period of analysis and conduct a panel study using ex-post changes in the start-ups' approved skilled workers. In both approaches, I find empirical evidence that having high-skilled foreign workers from a particular country increases firms' access to venture capital funding from that country. Through home-country professional networks, common language, and cultural familiarity, skilled immigrants reduce investment frictions and information gaps and better facilitate cross-border venture capital investments from their origins, especially in the case of young firms or inexperienced venture capitalists. Lastly, I construct a framework of firms' hiring decisions to incorporate both the standard labor market and foreign venture capital channels. With a numerical exercise, I find that almost one-third of the displacement effect of high-skilled immigrants can be mitigated by the venture capital channel.

Contents

1	Migration Networks and Mexican Migrants' Spatial Mobility in the U.S.	1
1.1	Introduction	2
1.2	Empirical Framework	4
1.3	Data Sources and Empirical Approach	6
1.4	Estimation Results	9
1.5	Conclusion	12
2	H-1B Visa Lottery and the Selection of Skilled Immigrant Workers in the U.S.	23
2.1	Introduction	24
2.2	Overview of H-1B	25
2.3	Proposed Framework	27
2.4	Data	31
2.5	Empirical Approach and Results	33
2.6	Conclusion	35
3	High-skilled Immigrant Workers and U.S. Firms' Access to Foreign Venture Capital	44
3.1	Introduction	45
3.2	Background on H-1B Visas	48
3.3	Data	50

3.3.1	Data on VC investments	50
3.3.2	Data on high-skilled immigrant workers	51
3.4	Empirical Approaches and Results	53
3.4.1	H-1B random lottery approach	53
3.4.2	Panel study approach	58
3.5	Proposed Framework	63
3.5.1	Setup	64
3.5.2	Steady state	65
3.5.3	Response to immigration shocks	67
3.5.4	Numerical exercise	68
3.6	Conclusion	70
3.7	Appendix	71
3.7.1	VentureXpert from Thomson ONE	71
3.7.2	Form I-129 data from USCIS	71
3.7.3	LCA data from Department of Labor	71
3.7.4	Construction	71

List of Figures

1.1	Distribution of MCAS Card Issuances for Migrants Born in Tiquicheo	16
1.2	Distribution of MCAS Card Issuances for Migrants Born in Hidalgo	17
2.1	Changes in H-1B Caps	39
2.2	Simulation Results with Positive Selection	40
2.3	Simulation Results with Negative Selection	40
3.1	Changes in H-1B Caps	78
3.2	The average number of VC deals received by firms from a given source country	79
3.3	Estimated coefficients from LPM and logit specifications, and when individual countries are dropped in the win rate approach	80
3.4	Estimated coefficients when individual countries are dropped in the FE approach	81

List of Tables

1.1	Descriptive Statistics	18
1.2	Mexican-born Population Response to Labor Demand Shocks	19
1.3	Mexican-born Population Response to Labor Demand Shocks: Bartik (1991) IV Estimates	20
1.4	White Natives' Population Response to Labor Demand Shocks	21
1.5	Non-Mexican Immigrants' Population Response to Labor Demand Shocks . .	22
2.1	Final Receipt Dates and H-1B Lottery Details	41
2.2	Summary Statistics	42
2.3	Quality of H-1B Applicants in the Random Lottery	43
3.1	Final Receipt Dates and H-1B Lottery Details	82
3.2	VC Deal Volumes by U.S. and Non-U.S. Funds	83
3.3	Top Investor Countries for VC Activities in the U.S.	84
3.4	Top Source Countries for H-1B Workers in the U.S.	85
3.5	H-1B Lottery Win Rate and Foreign Venture Capital - Multiple Imputation .	86
3.6	H-1B Lottery Win Rate and Foreign Venture Capital - Dichotomous Win Rates	87
3.7	Skilled Immigrants and Foreign Venture Capital - FE Main Effects	88
3.8	Skilled Immigrants and Foreign Venture Capital - Mechanism	89
3.9	Skilled Immigrants and Foreign Venture Capital - Heterogeneous Effects . . .	90
3.10	Skilled Immigrants and Foreign Venture Capital - Robustness	91

Chapter 1

Migration Networks and Mexican Migrants' Spatial Mobility in the U.S.

(joint with Brian C. Cadena¹, Brian K. Kovak², and Rebecca Lessem³)

¹University of Colorado - Boulder and IZA

²Carnegie Mellon University, NBER, and IZA

³Carnegie Mellon University

1.1 Introduction

Given the substantial influx of immigrants to the United States over the past decade, the nature and determinants of migrants' location choices are among the key topics in the immigration-related discussion. In particular, social networks are found to be central in migrants' location choices as these connections likely reduce moving costs and improve labor market outcomes in the destination. Another important finding in the recent related literature is that migrants in the U.S., especially Mexican-born ones, respond very strongly to changes in the labor market conditions and are much more geographically mobile than natives (Cadena and Kovak 2016). However, the mechanisms through which features of migration networks facilitate this mobility remain largely understudied due to the lack of adequate data with extensive geographic coverage and detailed network information.

With a novel set of data, the *Matrícula Consular de Alta Seguridad* (MCAS) data, we aim to examine the role of migration networks in Mexican-born migrants' mobility responses to local demand changes. The MCAS data provides extremely detailed and comprehensive source-destination information on Mexican immigrants in the United States. Caballero, Cadena and Kovak (2017) establishes the quality and representativeness of the MCAS data through the comparisons with various data sources including the Mexican census, American Community Surveys (ACS), the *Encuesta Nacional de la Dinámica Demográfica* (ENADID), and the *Encuesta sobre Migración en la Frontera Norte* (EMIF). Our paper is the first to use the MCAS data in empirical estimation and the migration network measures calculated with this data present more reliable and representative information on Mexican immigrants' network connections in the U.S.

In this paper, we focus on the mobility responses of low-skilled Mexican-born immigrants in the depth of the Great Recession, following the finding by Cadena and Kovak (2016) that this group is highly sensitive to changes in the labor market conditions. We first develop a static location choice model, from which we derive our estimation equation. In the estimation, we use the cross-sectional network information right before the Great Recession from the MCAS data to construct network measures and include various control variables to rule out the effect of preexisting migration patterns. We also use a Bartik (1991) instrument to conduct IV estimations to test if the results are robust to different approaches to identify labor demand changes. However, the results are not statistically

significant potentially due to the issue of weak instruments. We demonstrate that migrants from different Mexican counties (*municipios*) within the same Mexican state (*entidad*) vary vastly in their destination choices. This feature helps allay concerns that location choices are merely driven by regional preferences over destinations' characteristics rather than social networks present there. Several placebo tests are included to confirm that other ethnic groups are insensitive to Mexican migrants' social networks. Our analysis shows that social networks indeed influence migrants' location decisions. In particular, Mexican migrants respond to both local economic conditions as well as demand shocks in places where they have connections.

The existing literature documents a wide spectrum of migration network effects in relation to international migrants' experiences. For instance, a large literature confirms that by lowering the costs associated with migration, networks increase the rates of international migration.⁴ Bartel (1989), Dunlevy (1991) and Jaeger (2000) observe that international migrants tend to locate in places with high concentrations of immigrants of the same ethnicity and Patel and Vella (2013) finds that they further clustered in particular occupations.⁵ In addition, these social connections tend to improve migrants' labor market outcomes, as found by Munshi (2003) and Mundra and Rios-Avila (2016). The novel contribution of our paper is that we document the effect of migration networks on migrants' responsiveness to labor demand shocks. This enriches our understanding of both immigrants' location choices as well as social network effects.

This paper proceeds as follows. The next section provides an empirical framework with a static discrete choice model. Section 3 describes the data used in this paper and lays out the estimation specification and procedure. The estimation results are discussed in Section 4 and the last section concludes.

⁴For literature on network effects on the probability of international migration, see, for example, Massey (1986); Massey and Espinosa (1997); Palloni, Massey, Ceballos, Espinosa and Spittel (2001); Winters, de Janvry and Sadoulet (2001); Curran and Rivero-Fuentes (2003); Fussell and Massey (2004); Liu (2013); and Garip and Asad (2016).

⁵Other related works include Bauer, Epstein and Gang (2002); Diaz McConnell (2008); and Lafortune and Tessada.

1.2 Empirical Framework

This section proposes a static discrete choice model of an individual's location choices. The value of living in a location depends on the strength of the labor market, s , and the strength of the social network, n . We assume there are no moving costs incurred. Individuals receive random, independent and identically distributed (i.i.d.) payoff shocks, η , to living in each location, which are drawn from a type I extreme value distribution. In this model, people are from Mexican source locations k and they choose from a set of U.S. destination locations j to live. Thus, the utility of choosing destination j for a migrant from hometown k , u_{jk} , is as specified in equation (1).⁶

$$u_{jk} = \alpha s_j + \beta n_{jk} + \eta_{jk} \quad (1.1)$$

Since we assume the payoff shocks are following an extreme value distribution, the probability that such an individual chooses location j takes a logit form as in the following formula (McFadden 1973 and Rust 1987):

$$P_k(j) = \frac{\exp(\alpha s_j + \beta n_{jk})}{\exp(\alpha s_j + \beta n_{jk}) + \sum_{i \neq j} \exp(\alpha s_i + \beta n_{ik})}. \quad (1.2)$$

We then derive the the following comparative statics to assess the effects of employment shocks in location j and an alternative location, h , on the probability of choosing location j :

$$\begin{aligned} \frac{\partial P_k(j)}{\partial s_j} &= \alpha \frac{\exp(\alpha s_j + \beta n_{jk}) \left[\sum_{i \neq j} \exp(\alpha s_i + \beta n_{ik}) \right]}{\left[\exp(\alpha s_j + \beta n_{jk}) + \sum_{i \neq j} \exp(\alpha s_i + \beta n_{ik}) \right]^2} \\ &= \alpha P_k(j) (1 - P_k(j)) \end{aligned} \quad (1.3)$$

⁶As all individuals are assumed to be identical except for their source locations and current residence, individual subscripts are suppressed.

$$\begin{aligned}\frac{\partial P_k(j)}{\partial s_h} &= -\alpha \frac{\exp(\alpha s_j + \beta n_{jk}) \exp(\alpha s_h + \beta n_{hk})}{\left[\exp(\alpha s_j + \beta n_{jk}) + \sum_{i \neq j} \exp(\alpha s_i + \beta n_{ik}) \right]^2} \\ &= -\alpha P_k(j) P_k(h).\end{aligned}\tag{1.4}$$

We denote the Mexican population from source k living in destination location j as M_{jk} . The total Mexican population in destination location j , $M_j = \sum_k M_{jk}$, is the sum of M_{jk} over all k sources. Taking the total derivative of the total Mexican population in destination location j , we have that the change in the total Mexican population in location j depends on the respective changes in the Mexican population from different sources:

$$dM_j = \sum_k dM_{jk}.\tag{1.5}$$

If M_k is the total migrant population from source location k , the Mexican population from source k living in destination location j can be calculated by multiplying M_k with the probability of a person from source k choosing location j :

$$M_{jk} = P_k(j) M_k.\tag{1.6}$$

We assume that the total migrant population from source location k does not vary with shocks, so we can express equation (5) as follows, where the change in total Mexican population in destination j is a function of the total migrant population from source location k and $dP_k(j)$:

$$dM_j = \sum_k M_k dP_k(j).\tag{1.7}$$

As we assume the existing strength of migration networks to be constant, the total derivative of the choice probability depends on the employment shocks in destination j and the alternative locations, as specified below:

$$dP_k(j) = \frac{\partial P_k(j)}{\partial s_j} ds_j + \sum_{h \neq j} \frac{\partial P_k(j)}{\partial s_h} ds_h.\tag{1.8}$$

Combining equations (7) and (8), we can express the change in the total Mexican population living in location j as in equation (7) and arrive at a formula for the proportional change in total Mexican

population living in location j .

$$dM_j = \sum_k M_k \left(\frac{\partial P_k(j)}{\partial s_j} ds_j + \sum_{h \neq j} \frac{\partial P_k(j)}{\partial s_h} ds_h \right) \quad (1.9)$$

$$= \sum_k M_k \left(\alpha P_k(j) (1 - P_k(j)) ds_j - \sum_{h \neq j} \alpha P_k(j) P_k(h) ds_h \right) \quad (1.10)$$

$$= \alpha \sum_k \underbrace{M_k P_k(j)}_{=M_{jk}} \left((1 - P_k(j)) ds_j - \sum_{h \neq j} P_k(h) ds_h \right) \quad (1.11)$$

$$\frac{dM_j}{M_j} = \alpha \sum_k \frac{M_{jk}}{M_j} \left((1 - P_k(j)) ds_j - \sum_{h \neq j} P_k(h) ds_h \right) \quad (1.12)$$

Equation (12) relates the proportional change in Mexican population in destination location j to the vector of employment shocks ds across destination locations. This relationship depends first upon the source mix of Mexicans in j , given by M_{jk}/M_j . The effect of the employment shocks also depends upon the destination mix of migrants from the relevant sources given by $P_k(j)$ and the $P_k(h) \forall h \neq j$. The shock in the reference location j matters more when a smaller share of migrants from source k chooses that location. That is, if a migrant in location j has many other options in the U.S., $(1 - P_k(j))$ is large and the response to local shocks will be relatively volatile. The shock in another location $h \neq j$ matters more when more people from source k choose location h , i.e. when $P_k(h)$ is larger.

1.3 Data Sources and Empirical Approach

As our empirical framework suggests, we study the changes in a state's Mexican-born population as a function of the local demand shock and migration networks in the depth of the Great Recession. We use data from the American Community Survey (ACS) to calculate the changes in the natural log of the group's population from 2006 to 2010, which serves as our dependent variable.⁷ Our main sample includes Mexican-born individuals aged 18-64 with no college education, not currently

⁷ACS data was obtained from the Integrated Public Use Microdata Series (IPUMS) (Ruggles et al. 2010).

enrolled in school and not living in group quarters.⁸

Our first independent variable is the change in the natural log of employment in a given state from 2006 to 2010, which we derive using employment information from County Business Patterns (CBP) data. We identify changes in labor market conditions as the shocks to the employment levels in different states. This takes reference from previous literature, which finds that employers respond to negative demand shocks through layoffs rather than wage cuts during the Great Recession.⁹ As wages are downward-rigid and all regions face declining labor demand, changes in employment fully capture the shifts in labor demand. We also use an instrument in Bartik (1991) as an alternative way of identifying labor demand shocks.

The other independent variable is the migration network measure of interest. We assume two individuals are connected if both come from the same source *municipio* in Mexico. We use migration network-related information in 2006, before the start of the Great Recession.¹⁰ These network measures are obtained from the *Matrícula Consular de Alta Seguridad* (MCAS) data. Administered by the Mexican government, the MCAS program issues identity cards to Mexican citizens living in the United States, and all records of approved applications are stored in a centralized database. The recorded information from each issuance includes the recipient's *municipio* (similar to county) of birth in Mexico and the U.S. state and consular area of current residence. Caballero, Cadena and Kovak (2017) establishes the quality and representativeness of the MCAS data by showing strong accordance on place-to-place migrant distribution patterns between MCAS and a variety of standard data sources, including the Mexican Census and American Community Surveys (ACS). The MCAS data contains more comprehensive migration network information than existing data sources such as the *Encuesta Nacional de la Dinámica Demográfica* (ENADID) and the *Encuesta sobre Migración en la Frontera Norte* (EMIF) as the latter two suffer from small sample issues and relatively aggregate location information (Caballero, Cadena and Kovak 2017). Another popular data source for migration network literature, the Mexican Migration Project (MMP), only covers a

⁸We are focusing on this group of individuals because they are more responsive to changes in labor market conditions, according to Cadena and Kovak (2016).

⁹Rothstein (2012) finds that changes in employment levels were substantially larger than adjustments in average wages during the Great Recession.

¹⁰We only use pre-shock network information as we are unable to confirm that changes in the networks necessarily correlate with changes in the distribution of Mexicans in the U.S. Also, we do not want to include endogenous network shifts in the independent variables.

small number of Mexican communities, whereas the MCAS data identifies 70 U.S. destinations and more than 2,000 source *municipios* in Mexico with a very large sample size (Massey and Zenteno 2000).

We rearrange equation (12) to derive an estimating equation. We define local labor market shocks as in Cadena and Kovak (2016): ds_j is the change in log employment in destination j , which we will write as $d\ln(emp_j)$. Network measures include the share of immigrants from source k choosing destination j , M_{jk}/M_j , and the probabilities of immigrants from source k choosing a particular location, the $P_k(\cdot)$ terms, where we calculate $P_k(j)$ as M_{jk}/M_k . Finally, we arrive at equation (13) by combining $P_k(j)ds_j$ and $\sum_{h \neq j} P_k(h)ds_h$ in equation (12):

$$d\ln(M_j) = \alpha d\ln(emp_j) - \alpha \sum_k \frac{M_{jk}^{2006}}{M_j^{2006}} \left(\sum_h \frac{M_{hk}^{2006}}{M_k^{2006}} d\ln(emp_h) \right). \quad (1.13)$$

The first term on the right hand side is the employment shock in the reference location j . The second term is the migration network measure of interest, taking the form of the average destination employment shock faced by the mix of migrants from different sources who live in location j . We generalize this expression by allowing the coefficients on these two terms to vary, adding an intercept term δ , and including an error term ε_j to account for sampling variation in the Mexican population estimate in location j .

$$d\ln(M_j) = \delta + \alpha d\ln(emp_j) - \gamma \sum_k \frac{M_{jk}^{2006}}{M_j^{2006}} \left(\sum_h \frac{M_{hk}^{2006}}{M_k^{2006}} d\ln(emp_h) \right) + \varepsilon_j \quad (1.14)$$

Equation (14) is our estimating equation and we expect a positive coefficient on the direct shock term and a negative coefficient on the network measure term. A statistically significant coefficient on the migration network measure indicates the relevance of migration networks in driving location choices of Mexican migrants within the U.S.

We have also taken some additional steps to allow our analysis to be more representative. First, although the CBP data gives us the universal coverage of establishments in most of the industries, it does not cover employment in agricultural production, private household services, or the government services. We have therefore complemented information obtained from CBP with

employment calculations using the ACS data for the missing industries. Second, we construct weights for the states to account for the heteroskedasticity in comparing proportional population changes across labor markets of different sizes. Third, we introduce several controls to account for factors that may affect the Mexican mobility and correlate with local changes in demand at the same time. Specifically, we use the Mexican-born share of each state’s population in 2000 to control for the potential decline in the value of traditional enclaves (Card and Lewis 2007) and we include indicators for anti-immigrant employment legislations and new 287(g) agreements that allow local officials to receive delegated authority for immigration enforcement.¹¹

Table 1 provides the summary statistics for our key variables.

1.4 Estimation Results

We first reproduce the estimation in Cadena and Kovak (2016) at the U.S. state level for Mexican-born immigrants, and the results are shown in the first three columns of Table 2. Column (1) reports the results from the basic specification with local demand changes only, while the next two columns progressively add the enclave control and policy controls. These estimated coefficients confirm that Mexican-born immigrants were highly sensitive to labor demand changes during the Great Recession. Our results on local demand shocks differ slightly from those in Cadena and Kovak (2016), mainly due to the fact Cadena and Kovak (2016) conducts analysis at the level of Metropolitan Statistical Areas (MSAs) while we aggregate observations to the state-level. In spite of this change in geographic unit of analysis, the results are remarkably similar.

The results with the migration network measure included are presented in column (4) to column (6) in Table 2. Column (4) reports the coefficients for the basic regression specification in equation (15). In line with our expectation, the estimated coefficient on local demand shock is positive and statistically significant. This represents the elasticity of the Mexican-born population with respect to labor demand, holding other things constant, and we find that Mexican-born workers are highly responsive to local demand shocks. To illustrate, we compare Mexican mobility in California, where the demand shock was at the 25 percentile of the shock distribution, and that in

¹¹Information on immigration policies is obtained from the immigration policy database in Santillano and Bohn (2012).

Wyoming, where the demand change was at the 75 percentile. Wyoming, facing a 0.119 percentage point higher employment growth than California, is expected to experience 0.103 percentage point larger increase in Mexican-born population, holding other things constant. On the other hand, the coefficient on the migration network measure is negative and statistically significant. This suggests that given demand shocks, Mexican-born immigrants with access to better outside options are more likely to move away from current place of residence. Massachusetts and Oklahoma stand at the 25 percentile and 75 percentile of the network measure distribution respectively. Our estimation implies that Oklahoma, having a 0.0133 percentage point higher network measure, will experience 0.0217 percentage point larger decline in its Mexican-born population.

We then include several controls to account for the possibility that the observed mobility is driven by other determinants of location choice that are associated with local employment changes. First, we add the enclave control that accounts for the decline in value of traditional migrant enclaves, and the results are shown in column (5). We also control for anti-immigrant employment legislation in relevant states and new 287(g) agreements which might encourage migrants to relocate elsewhere. Column (6) presents the estimates when both the enclave control as well as the policy controls added. In both specifications, the coefficients on local demand changes remain statistically significant and comparable to those in column (4). Notably, the coefficient on the network measure becomes larger in magnitude when the controls are included, implying a more significant impact of migration networks on the Mexican-born mobility. One likely explanation is that the additional controls successfully absorb the variation in locations where Mexican-born population prefer to live regardless of the social networks present there.

Another potential source of bias with our specification is that there still might be unmeasured factors driving population changes, which in turn influence employment, possibly by affecting consumer demand or the number of available workers. We attempt to address this issue by using the standard "Bartik instrument" (Bartik 1991). This measure predicts local demand shocks by proportionally allocating changes in industrial employment at the national level across states, based on each state's industry composition of employment as of 2006.¹² This Bartik instrument relates

¹²Other examples of the Bartik instrument include Bound and Holzer (2000); Blanchard and Katz (1992); Autor and Duggan (2003); Wozniak (2010); Notowidigdo (2013); Charles, Hurst, and Notowidigdo (2013); and Cadena and Kovak (2016).

closely to employment changes and plausibly exogenous to other potential sources of fluctuations in population. We calculate the instrument as $\psi_j = \sum_i \varphi_{ij}^{t_0} d\ln L_i$, where $\varphi_{ij}^{t_0}$ is the fraction of state j employment in industry i in 2006, and $d\ln L_i$ is the proportional change in national employment in industry i . As employment changes across states are also used in the calculation of the network measure, we construct another instrument for the network measure, which is derived by replacing the change in log of employment with the Bartik instrument in the calculation of migration networks:

$$Network_k^j^{IV} = \sum_k \frac{M_{jk}^{2006}}{M_j^{2006}} \left(\sum_h \frac{M_{hk}^{2006}}{M_k^{2006}} \psi_h \right). \quad (1.15)$$

The results using the Bartik instrument and the network instrument are shown in Table 3. Column (1) present the IV estimates and the first-stage statistics from the basic specification, while column (2) and (3) add in the effects of enclave and policy controls progressively. Although the IV estimates are statistically insignificant, the coefficients on employment changes and the network measure are similar to those OLS estimates. We also notice that the Kpleibergen-Paap Wald F statistics associated with these IV estimations are very small, suggesting that we might face an issue of weak instruments that ultimately leads to the insignificant results.

There also remains another concern with our estimation, as Mexican migrants originating from the same region might be choosing destinations simply due to some common preferences over certain characteristics of these locations such as climate or proximity to home. These preferences likely correlate with both the network measure and the population changes, thus introducing bias to the results. To rule out the possibility that the destination distribution only reflects a source region's preferences over the persistent features of destination locations, we compare migration networks of *municipios* in the same source state. Caballero, Cadena and Kovak (2017) shows that it is very common for two *municipios* within the same state to have very different destination distributions. Figures 1 and 2 illustrate an example of such, where two *municipios*, *Ciudad Hidalgo* and *Tiquicheo*, located in the same Mexican state (*Michoacán*) and only 3 hours' drive apart, differ sharply in their immigrant destination distributions. In particular, migrants from *Hidalgo* are mostly clustered in Illinois, while those from *Tiquicheo* settle primarily in Texas. Immigrants from these two *municipios* should have very similar preferences due to geographical proximity and the differences in their

location choices can be largely attributed to existing social networks in the destination.

Finally, we conduct placebo tests with the group of male white natives as well as non-Mexican male immigrants with similar education background and present the results in Tables 4 and 5. The specifications from column (1) to (6) in both tables resemble those in Table 2, which relate group-specific population changes to group-specific employment changes, Mexican migrants' network measure (the last three columns) and a combination of controls. The results in the first three columns in both tables are largely comparable to those in Cadena and Kovak (2016), confirming that Mexican-born immigrants are much more mobile than these two groups. In line with our expectation, the coefficients on the network measures are small and none are statistically significant, suggesting that neither group reacts to Mexican migrants' network measure.

1.5 Conclusion

In this paper, we demonstrate that migration networks play an important role in Mexican-born migrants' location choices in the U.S. during the Great Recession, where they were likely hit by local demand shocks. We discover that both the local economic conditions and the conditions in destinations where Mexican-born immigrants have connections matter in their location choices .

We develop a static location choice model to arrive at our estimation equation, which relates population changes to employment changes and a network measure. Using a novel set of data, the MCAS data, we calculate the Mexican migration network measure of interest, which encompasses both the number of connections as well as the demand shocks in the connected destinations. To reduce potential bias, we introduce several controls to account for factors such as the decline in the value of traditional enclaves and immigrant-related policies in different states. We further test the robustness of our results by using the "Bartik instrument" to identify local labor demand shocks and used that to construct an instrument for the network measure. We also conduct placebo tests with other groups and find that neither white native nor non-Mexican immigrants react to Mexican immigrants' network measure.

Our analysis documents the effect of migration networks on migrants' responsiveness to local demand changes. The findings in this paper add a novel contribution to the current literature on

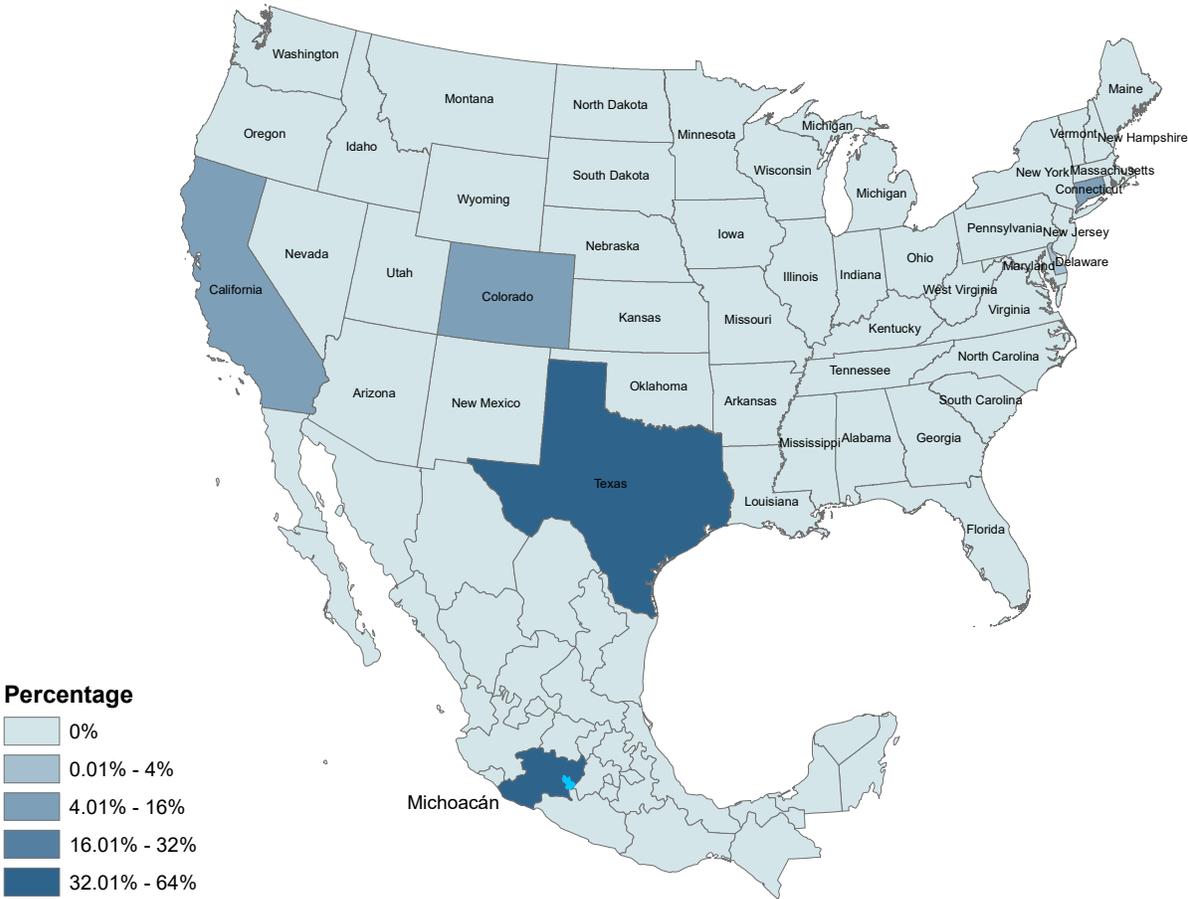
both migrant location choices as well as social network effects.

Bibliography

- Bartel, A. P. (1989), ‘Where do the new us immigrants live?’, *Journal of Labor Economics* **7**(4), 371–391.
- Bauer, T. K., Epstein, G. S. & Gang, I. N. (2002), ‘Herd effects or migration networks? the location choice of mexican immigrants in the us’.
- Bohn, S. & Santillano, R. (2012), ‘Do local immigration laws impact employment and wages? evidence from the 287 (g) program’, *Public Policy Institute of California* .
- Caballero, M. E., Cadena, B. C. & Kovak, B. K. (2017), ‘Measuring sub-national migration networks using matriculas consulares’.
- Cadena, B. C. & Kovak, B. K. (2016), ‘Immigrants equilibrate local labor markets: Evidence from the great recession’, *American Economic Journal: Applied Economics* **8**(1), 257–290.
- Card, D. & Lewis, E. G. (2007), The diffusion of mexican immigrants during the 1990s: Explanations and impacts, in ‘Mexican immigration to the United States’, University of Chicago Press, pp. 193–228.
- Curran, S. R. & Rivero-Fuentes, E. (2003), ‘Engendering migrant networks: The case of mexican migration’, *Demography* **40**(2), 289–307.
- Fussell, E. & Massey, D. S. (2004), ‘The limits to cumulative causation: International migration from mexican urban areas’, *Demography* **41**(1), 151–171.
- Garip, F. & Asad, A. L. (2016), ‘Network effects in mexico–us migration: Disentangling the underlying social mechanisms’, *American Behavioral Scientist* **60**(10), 1168–1193.
- Jaeger, D. A. (2000), ‘Local labor markets, admission categories, and immigrant location choice’, *Manuscript, College of William and Mary* .
- Lafortune, J., Tessada, J. et al. (2010), Smooth (er) landing? the dynamic role of networks in the location and occupational choice of immigrants, Technical report, Mimeo, October.
- Liu, M.-M. (2013), ‘Migrant networks and international migration: Testing weak ties’, *Demography* **50**(4), 1243–1277.
- Massey, D. S. (1986), ‘The social organization of mexican migration to the united states’, *The Annals of the American Academy of Political and Social Science* **487**(1), 102–113.
- Massey, D. S. & Espinosa, K. E. (1997), ‘What’s driving mexico-us migration? a theoretical, empirical, and policy analysis’, *American journal of sociology* **102**(4), 939–999.
- McConnell, E. D. (2008), ‘The us destinations of contemporary mexican immigrants’, *International Migration Review* **42**(4), 767–802.
- McFadden, D. et al. (1973), ‘Conditional logit analysis of qualitative choice behavior’.
- Mundra, K. & Rios Avila, F. (2016), ‘Immigrant birthcountry networks and unemployment duration: Evidence around the great recession’.
- Munshi, K. (2003), ‘Networks in the modern economy: Mexican migrants in the us labor market’, *The Quarterly Journal of Economics* **118**(2), 549–599.
- Palloni, A., Massey, D. S., Ceballos, M., Espinosa, K. & Spittel, M. (2001), ‘Social capital and international migration: A test using information on family networks’, *American Journal of*

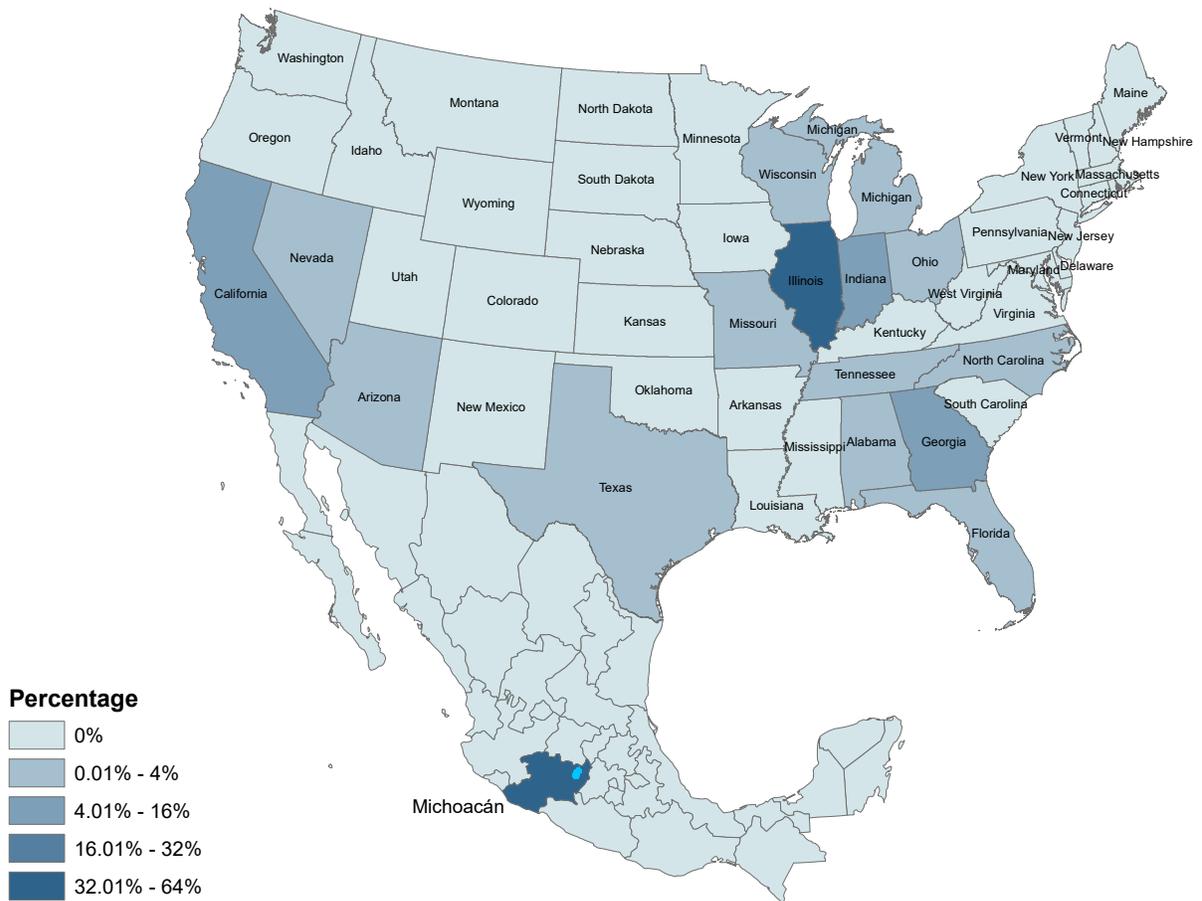
- Sociology* **106**(5), 1262–1298.
- Patel, K. & Vella, F. (2013), ‘Immigrant networks and their implications for occupational choice and wages’, *Review of Economics and Statistics* **95**(4), 1249–1277.
- Rothstein, J. (2012), ‘The labor market four years into the crisis: Assessing structural explanations’, *ILR Review* **65**(3), 467–500.
- Rust, J. (1987), ‘Optimal replacement of gmc bus engines: An empirical model of harold zurcher’, *Econometrica: Journal of the Econometric Society* pp. 999–1033.
- Winters, P., De Janvry, A. & Sadoulet, E. (2001), ‘Family and community networks in mexico-us migration’, *Journal of human resources* pp. 159–184.

Figure 1.1: Distribution of MCAS Card Issuances for Migrants Born in Tiquicheo



This map shows the destination distribution of the share of MCAS identity cards issued to migrants born in the *municipio* of Tiquicheo (shown in bright blue) who had current addresses in each US state. The data source is the universe of MCAS identity cards issued during the 2006-2010 time period. Source: Caballero, Cadena and Kovak (2017).

Figure 1.2: Distribution of MCAS Card Issuances for Migrants Born in Hidalgo



This map shows the destination distribution of the share of MCAS identity cards issued to migrants born in the *municipio* of Hidalgo (shown in bright blue) who had current addresses in each US state. The data source is the universe of MCAS identity cards issued during the 2006-2010 time period. Source: Caballero, Cadena and Kovak (2017).

Table 1.1: Descriptive Statistics

	Mean	Std. Dev	Min	Max	Observations
Change in log of Mexican-born population	0.085	0.314	-0.725	1.003	49
Change in log of Mexican-born employment	0.474	0.119	0.200	0.769	49
Network measure	0.443	0.015	0.408	0.496	49
Enclave control	0.026	0.037	0	0.166	49
New 287(g) control	0.184	0.391	0	1	49
Employment policy control	0.224	0.422	0	1	49

The sample includes male Mexican-born working age immigrants with high school or less education. Enclave control is the Mexican-born share of each state's population in 2000. New 287(g) indicator equals 1 if the state has effected the new 287(g) agreement, 0 otherwise. Employment policy indicator equals 1 if the state has enacted anti-immigration employment legislation, 0 otherwise

Table 1.2: Mexican-born Population Response to Labor Demand Shocks

	Dependent variable: Change in log of population					
	(1)	(2)	(3)	(4)	(5)	(6)
Change in log of Mexican-born employment	0.516** (0.195)	0.504** (0.200)	0.396* (0.206)	0.870*** (0.213)	1.081*** (0.281)	0.968*** (0.323)
Network measure				-1.636** (0.706)	-2.691*** (0.818)	-2.664*** (0.984)
Enclave control	No	Yes	Yes	No	Yes	Yes
Policy controls	No	No	Yes	No	No	Yes
Observations	49	49	49	49	49	49
R-squared	0.235	0.318	0.409	0.276	0.419	0.507

The sample includes male Mexican-born working-age immigrants with high school or less education. It relates the change in $\log(\text{population})$ for the group (2006-2010, using the American Community Surveys) to change in $\log(\text{group-specific employment})$ from County Business Patterns over the same time period, using the demographic group's industry mix. The specifications include a constant term. There are 49 state observations in total due to missing change in $\log(\text{group-specific employment})$ for Vermont and missing change in $\log(\text{population})$ for the group for Montana. Observations are weighted by the inverse of the estimated sampling variance of the dependent variable. Heteroskedasticity robust standard errors are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 1.3: Mexican-born Population Response to Labor Demand Shocks: Bartik (1991) IV Estimates

	Dependent variable: Change in log of Mexican-born population		
	(1)	(2)	(3)
<i>IV estimate</i>			
Change in log of Mexican-born employment	0.753 (0.779)	0.915 (0.833)	0.661 (0.675)
Network measure	-1.053 (2.746)	-2.636 (3.042)	-2.143 (2.619)
<i>First stage</i>			
Predicted change in log of employment	-3.654 (5.500)	-7.307 (5.771)	-7.882 (5.446)
Predicted network measure	40.28*** (14.78)	45.91*** (16.92)	46.14** (17.24)
Kleibergen-Paap Wald F statistic	4.206	2.492	2.642
Enclave control	No	Yes	Yes
Policy controls	No	No	Yes
Observations	49	49	49

The sample includes male Mexican-born working age immigrants with high school or less education. It relates the change in log(population) for the group (2006-2010, using the American Community Surveys) to change in log(group-specific employment) from County Business Patterns over the same time period, using the demographic group's industry mix. The specification includes a constant term. There are 49 state observations in total due to missing change in log(group-specific employment) for Vermont and missing change in log(population) for the group for Montana. Observations are weighted by the inverse of the estimated sampling variance of the dependent variable. The excluded instruments are predicted change in log(employment) and predicted network measure, based on Bartik (1991) and described in the text. The first-stage coefficients on the instrument and the Kleibergen-Paap rk Wald F statistic are reported below the corresponding IV estimate. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.4: White Natives' Population Response to Labor Demand Shocks

	Dependent variable: Change in log of population					
	(1)	(2)	(3)	(4)	(5)	(6)
Change in log of white natives' employment	0.107 (0.0823)	0.108 (0.0829)	0.113 (0.0902)	0.136 (0.127)	0.138 (0.128)	0.134 (0.138)
Network measure				-0.163 (0.346)	-0.164 (0.349)	-0.113 (0.371)
Enclave control	No	Yes	Yes	No	Yes	Yes
Policy controls	No	No	Yes	No	No	Yes
Observations	51	51	51	50	50	50
R-squared	0.0477	0.0513	0.0986	0.0528	0.0568	0.1009

The sample includes male working-age natives with high school or less education. It relates the change in log(population) for the group (2006-2010, using the American Community Surveys) to change in log(group-specific employment) from County Business Patterns over the same time period, using the demographic group's industry mix. The specification includes migration network measure, enclave and policy controls. There are only 50 state observations from column (4) to (6) due to missing value in the migration network measure for Vermont. Observations are weighted by the inverse of the estimated sampling variance of the dependent variable. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.5: Non-Mexican Immigrants' Population Response to Labor Demand Shocks

	Dependent variable: Change in log of population					
	(1)	(2)	(3)	(4)	(5)	(6)
Change in log of non-Mexican immigrants' employment	-0.0101 (0.296)	-0.0396 (0.327)	-0.0375 (0.330)	-0.194 (0.282)	-0.150 (0.325)	-0.152 (0.346)
Network measure				0.920 (1.095)	0.573 (1.420)	0.604 (1.508)
Enclave control	No	Yes	Yes	No	Yes	Yes
Policy controls	No	No	Yes	No	No	Yes
Observations	51	51	51	50	50	50
R-squared	0.0000	0.0201	0.0239	0.0116	0.0237	0.0279

The sample includes male working-age non-Mexican immigrants with high school or less education. It relates the change in log(population) for the group (2006-2010, using the American Community Surveys) to change in log(group-specific employment) from County Business Patterns over the same time period, using the demographic group's industry mix. The specification includes migration network measure, enclave and policy controls. There are only 50 state observations from column (4) to (6) due to missing value in the migration network measure for Vermont. Observations are weighted by the inverse of the estimated sampling variance of the dependent variable. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 2

H-1B Visa Lottery and the Selection of Skilled Immigrant Workers in the U.S.

2.1 Introduction

The H-1B visa program has been the main channel for foreign-born skilled professionals to enter the U.S. workforce for almost three decades. Being one of the most important immigration policies in the U.S., this program has been the focus of considerable scrutiny and debate, especially with the Trump administration. The current policymakers have shown their determination in pushing for an H-1B reform, most likely to make it more restrictive than ever.¹ These impending political changes necessitate better understanding of the implications of a more restrictive visa program, particularly on the immigrant profile that it potentially attracts. Insights can be perhaps drawn from past changes to the design of the H-1B visa program. For instance, the competition for H-1B visas has become so intense in recent years that a computerized random lottery has been put in place to allocate these visas. As the H-1B lottery entails uncertainty in gaining work permissions, application behavior may deviate and the profile of H-1B applicants may be altered accordingly.

By exploiting the relatively new feature of the H-1B visa program, the H-1B lottery, this paper examines the impact of varying the probability of obtaining an H-1B visa on the quality of the prospective applicants, especially through the mechanism of self-selection. In this paper, I propose a theoretical framework to model self-selection patterns with uncertainty in obtaining admission into the host country and find empirical evidence that given lower chances of securing a work visa, the quality of H-1B workers deteriorated. These results indicate that H-1B applicants are negatively selected and a toughened H-1B policy will reinforce this unfavorable self-selectivity.

Following the seminal Borjas (1987) paper, researchers have been exploring self-selection patterns of immigrants into the U.S. For instance, Chiquiar and Hanson (2005), Fernandez-Huertas (2011) and Kaestner and Malamud (2014) all find Mexican immigrants in the U.S. are negatively selected. Borjas (2008) argues that the migration flow from Puerto Rico to the U.S. was unfavorably selected. However, the self-selection discussion has been primarily on source countries typically sending low-skilled workers. The self-selection patterns of more educated immigrants in the U.S. have not yet received adequate attention.

On the other hand, previous works on the H-1B visa program focus on a spectrum of con-

¹The Trump administration issued 45 percent more challenges to H-1B visa applications in 2017, as compared to 2016 (Glennon 2018).

sequences of policy changes, including those on sectoral employment,² career choices (Amuedo-Dorantes and Furtado 2016), foreign affiliate activity (Glennon 2018), firms' innovation activities,³ and natives' labor market outcomes.⁴ Only a handful of papers touch on the issue of immigrant quality, among which Kato and Sparber (2013) and Shih (2016) respectively relate H-1B quota reduction to lower quality of incoming international students and lower enrollment of students with higher expected returns to working in the U.S. Very limited work has been done on policy's effect on the quality of the H-1B workers themselves.

This paper contributes to the existing literature in several ways. Firstly, the proposed framework presents a novel extension of the Roy-Borjas model of self-selection and provides a theoretical basis for analysis involving immigration choices with uncertainty in admission. Secondly, it enriches the discussion on H-1B policy impact by examining implications of an H-1B reform from a fresh angle - the quality of the H-1B workers. Thirdly, it enhances our knowledge on the selection patterns of more skilled immigrants, which lacks adequate discussion in existing literature.

The rest of the paper is organized as follows. The next section provides an overview of the H-1B visa program, followed by Section 3, which details a theoretical framework of self-selection outcomes with the H-1B lottery element. Section 4 describes the data used for empirical estimation. The empirical approach and results are discussed in Section 5 and the last section concludes.

2.2 Overview of H-1B

Incepted in 1990, the H-1B visa program allows foreign-born college-educated professionals to work temporarily in specialty occupations in the U.S.⁵ In particular, this program has been extensively used by those in the fields of science, technology, engineering, and mathematics (STEM), to enter

²See Kerr and Lincoln (2010), and Mayda et al (2018).

³See Kerr and Lincoln (2010), and Doran et al (2014).

⁴See Zavodny (2013), and Peri et al (2015)

⁵To qualify for an H-1B visa, "the position must meet one of the following requirements: (1) a bachelor's or higher degree or its equivalent is normally the minimum entry requirement for the position; (2) the degree requirement is common to the industry in parallel positions among similar organizations or, in the alternative, the position is so complex or unique that it can be performed only by an individual with a degree; (3) the employer normally requires a degree or its equivalent for the position; or (4) the nature of the specific duties is so specialized and complex that the knowledge required to perform the duties is usually associated with attainment of a bachelor's or higher degree."

the U.S. workforce.⁶

A foreign-born individual intending to work in the U.S. with an H-1B visa must find a sponsoring firm that offers him a job in qualifying occupations. This firm must first submit a Labor Condition Application (LCA) with the Department of Labor, demonstrating accordance with the U.S. law. Upon LCA approval, the firm proceeds to complete an I-129 application with U.S. Citizenship and Immigration Services (USCIS) during the filing period.⁷ An H-1B visa is valid for a period of three years, extendable once to a maximum of six years in total. Nonetheless, H-1B visa holders can possibly remain in the U.S. after the six-year period as these visas allow for dual intent, where workers can simultaneously apply for permanent residency while working on a nonimmigrant basis.

The total number of new H-1B visa issuances is subject to a cap and the cap has changed several times as depicted in Figure 1. Specifically, the original limit of 65,000 H-1B visas was increased to 115,000 for FYs 1999 and 2000 and lifted again for FYs 2001 to 2003 to 195,000. For FY 2004, the cap was reverted to 65,000 regular visas (hereinafter "regular cap") with added 20,000 visas for workers with U.S. postgraduate degrees (hereinafter "postgraduate cap") and these limits have prevailed since then. Several groups are exempted from those limits. From FY 2001 onwards, individuals on cap-subject H-1Bs who are switching jobs or renewing their visas are excluded from the caps. An uncapped category for workers in colleges, universities, and non-profit research institutions was also created in FY 2001. In addition, a suite of Free Trade Agreements (FTAs) have made special provisions for citizens of Australia, Canada, Chile, Mexico and Singapore with regard to temporary work visas in the U.S.⁸

USCIS normally grants H-1B visas in a first-come, first-serve fashion until the caps are exhausted, beyond which no petitions are processed. The end of the application period is demarcated by the "final receipt date", which is the date on which USCIS receive enough petitions to fill the

⁶According to Peri, Shih, and Sparber (2015), H-1B workers have contributed to approximately half of the growth registered in the college-educated STEM workforce since 1990.

⁷The filing period starts on the first working day of April for petitions to start work in the next fiscal year (FY), which is from October 1 in the same year to September 30 in the following year.

⁸Chilean and Singaporean nationals are permitted to work temporarily in the U.S. on a nonimmigrant basis in specialty occupations under the H-1B1 program, active from FY 2004 onwards. Although the H-1B1 program specifies caps of 1,400 and 5,400 H-1B1 visas for Chileans and Singaporeans respectively, the caps have never been binding up-to-date. Australians are permitted to work temporarily in the U.S. on a nonimmigrant basis in specialty occupations under the E-3 program, applicable from May 2005. Citizens of Canada and Mexico are allowed to work in the U.S. in prearranged business activities for U.S. or foreign firms under the TN program, effective from 1994.

remaining quotas. Prior to FY 2004, the H-1B cap was never filled. Beginning FY 2004, the time it took to reach the limit steadily shortened. Notably, for FYs 2008 and 2009, USCIS received overwhelming volumes of applications within the first week of the filing period, which compelled USCIS to resort to a computerized random lottery to allocate all cap-subject H-1B visas. The demand for H-1B visas subsided for a few years until it surged remarkably again for FY 2014 and it has remained strong since then. Table 1 summarizes the final receipt dates and specifics of the H-1B lotteries over the years. If a lottery is put in place, USCIS follows a two-tiered procedure, where a first lottery is implemented to allocate the 20,000 postgraduate cap-subject H-1B visas, and unsuccessful applications are pooled with the regular cap-subject ones to enter a second lottery for the 65,000 regular visas.

2.3 Proposed Framework

This section proposes a two-country framework built from the Roy-Borjas model of self-selection, incorporating uncertainty in obtaining the permission to work in the destination country. Workers of a source country (country 0) make a decision of whether to apply for a visa to work in the host country (country 1). If a worker decides to apply, there is a probability of p that he will be granted a work visa and he will start employment in the host country. Otherwise, he will remain in the source country to work. Workers of the source country thus face the following earnings opportunities:

$$\log w_0 = \mu_0 + v_0, \tag{2.1}$$

$$\log w_1 = p(\mu_1 + v_1) + (1 - p)(\mu_0 + v_0), \tag{2.2}$$

where μ_0 and μ_1 are the mean (log) earnings in the source and the host country respectively, and the random variables v_0 and v_1 are i.i.d, measuring deviations from mean earnings and assumed to follow normal distribution with means zero and variances σ_0^2 and σ_1^2 . The parameter ρ_{01} gives the correlation between the source and the host country earnings. For convenience, all individual indexed subscripts are suppressed.

I assume that the visa application decision in this framework is entirely based on a comparison

of the earnings opportunities⁹, net of a fixed amount of opportunity cost, π . As such, an individual in the source country will choose to apply if the expected earnings in the host country is higher than the wage that he would be getting at home, that is when

$$p(\mu_1 + v_1) + (1 - p)(\mu_0 + v_0) - \pi > \mu_0 + v_0. \quad (2.3)$$

The probability of a randomly chosen worker in the source country to apply for a work visa is given by:

$$Pr(Apply) = Pr\left(v > \mu_0 - \mu_1 + \frac{\pi}{p}\right) = \Phi(-z). \quad (2.4)$$

where $v = v_1 - v_0$; $z = (\mu_0 - \mu_1 + \frac{\pi}{p}) / \sigma_v$; and Φ is the standard normal distribution function.

Only a subset of the worker population in the source country will find it worthwhile to apply for a visa and work in the host country. The average earnings in both the source and the host country for the self-selected sample of applicants are given by the following:

$$Q_0 = E(\log w_0 | Apply) = \mu_0 + \frac{\lambda(-z)}{\sigma_v} (\sigma_{01} - \sigma_0^2), \quad (2.5)$$

$$Q_1 = E(\log w_1 | Apply) = p\mu_1 + (1 - p)\mu_0 + \frac{\lambda(-z)}{\sigma_v} (p\sigma_v^2 + \sigma_{01} - \sigma_0^2), \quad (2.6)$$

where $\lambda(-z) = \phi(-z) / (1 - \Phi(-z))$, and ϕ is the density function of the standard normal distribution.

The self-selection of visa applicants is considered positive if they have above-average earnings in both source and host countries. That is, $Q_0 > \mu_0$ and $Q_1 > p\mu_1 + (1 - p)\mu_0$, which simplify to the the following conditions:

$$\sigma_{01} > \sigma_0^2, \text{ and } p > \frac{\sigma_0^2 - \sigma_{01}}{\sigma_v^2}, \quad (2.7)$$

where $\sigma_{01} = Cov(v_0, v_1)$ and the latter condition is automatically satisfied by the first one. The variances σ_0^2 and σ_1^2 can be thought as the rewards to skills. As a result, applicants are favorably selected when the source country offers a lower return to skills, as compared to the host country. In such cases, high-skill workers are heavily taxed to provide for the social security benefits to the

⁹This takes reference from Hicks(1932), where Sir John Hicks argues that migration decisions are primarily driven by wage differentials.

low-skill workers in the source country, thus more motivated to migrate.

Negative selection, on the other hand, describes a situation when the applicants are lower-skilled than the average workers in both source and host countries ($Q_0 < \mu_0$ and $Q_1 < p\mu_1 + (1-p)\mu_0$), that is when:

$$\sigma_{01} < \sigma_0^2, \text{ and } p < \frac{\sigma_0^2 - \sigma_{01}}{\sigma_v^2}. \quad (2.8)$$

To see how the self-selection patterns of the immigrant workers react to immigration policy changes, one should examine how the average earnings of the self-selected applicants change in response to a change in the probability of being awarded a work visa, mathematically given by $\frac{\partial Q_0}{\partial p}$ and $\frac{\partial Q_1}{\partial p}$:

$$\frac{\partial Q_0}{\partial p} = \frac{(\sigma_{01} - \sigma_0^2)}{\sigma_v} \lambda'(-z) \quad (2.9)$$

$$\frac{\partial Q_1}{\partial p} = \mu_1 - \mu_0 + \frac{\lambda(-z)}{\sigma_v} \sigma_v^2 + \frac{\lambda'(-z)}{\sigma_v} (p\sigma_v^2 + \sigma_{01} - \sigma_0^2) \quad (2.10)$$

where

$$\lambda'(-z) = \frac{-\lambda(-z)(\lambda(-z) - z)\pi}{p^2\sigma_v}. \quad (2.11)$$

As the inverse mills ratio is the expectation of the truncated normal, we have $\lambda(-z) = E(Z|Z > z) > z$ and thus $\lambda'(-z) < 0$. Hence, Equation (9) is negative in the case of positive selection, where $\sigma_{01} > \sigma_0^2$, and positive in the case of negative selection, where $\sigma_{01} < \sigma_0^2$. When the probability of getting a work visa is reduced, the average source country earnings of the favorably selected applicants will be higher and that of the unfavorably selected applicants will be lower.

Next, I turn to how expected host country earnings of the positively and negatively selected applicants react to a change in the chance of getting a work visa respectively. However, it is slightly more complex to ascertain how Q_1 reacts to a change in p , which is the direction of Equation (10). First, following Autor (2003)'s notes on the Roy Model, we can assume $\mu_0 \approx \mu_1$ to focus on self-selection rather than mean effects. Then the derivative in Equation (10) can be expressed as follows,

$$\frac{\partial Q_1}{\partial p} = \frac{\lambda(-z)}{\sigma_v} \sigma_v^2 + \frac{\lambda'(-z)}{\sigma_v} (p\sigma_v^2 + \sigma_{01} - \sigma_0^2). \quad (2.12)$$

In the case of negative selection, there is no ambiguity about the sign of $\frac{\partial Q_1}{\partial p}$. As $\sigma_{01} < \sigma_0^2$ when applicants are unfavorably selected, both terms on the right-hand side are positive. Thus, in the case of negative selection,

$$\frac{\partial Q_0}{\partial p}, \frac{\partial Q_1}{\partial p} > 0, \quad (2.13)$$

where both source and host expected earnings will drop when p decreases. On the other hand, when foreign workers are positively selected, the direction of $\frac{\partial Q_1}{\partial p}$ is obscured by the fact that the first term on the left-hand side of Equation (12) is positive while the second term is negative. If $\frac{\partial Q_1}{\partial p}$ is to move in the same direction of $\frac{\partial Q_0}{\partial p}$, that is

$$\frac{\partial Q_0}{\partial p}, \frac{\partial Q_1}{\partial p} < 0, \quad (2.14)$$

then the following condition must be met:

$$\frac{p^2 \sigma_v^3}{p \sigma_v^2 + \sigma_{01} - \sigma_0^2} < (\lambda(-z) - z) \pi. \quad (2.15)$$

In this case, the expected earnings for the positively selected applicants will be even higher, given a stricter work visa policy, where p decreases.

Given the ambiguity in the case of positive selection, I have conducted a simulation exercise to assess the likely effect of varying p on the self-selection patterns of the applicants. Simulation results in the cases of positive and negative selection are included in Figure 2 and Figure 3 respectively.¹⁰ With varying parameters including ρ_{01} , σ_0 and σ_1 , the simulations consistently suggest two main effects of changing p . Firstly, a higher p will result in less correlation between $\log w_0$ and $\log w_1$, as the source country earnings play a less important role in the expected earnings in the host country. Secondly, better chances of getting a work visa will encourage marginal workers near the cutoff line to apply. As a result, in the scenario of a positive selection, workers that are lower-skilled than the existing applicants are induced to apply with a higher p , thus suggesting a decline in the average quality of the applicants. On the other hand, in the case of negative selection, a higher better chances of getting a visa will motivate workers higher in the wage distribution than the existing

¹⁰In total, 20 simulations each were completed in the cases of positive and negative selection respectively.

pool of applicants and thus pushing up the average quality of applicants.

To summarize, the self-selection of the applicants will be reinforced with a more restrictive immigration visa policy. If the applicants are positively selected to begin with, lower chances of getting a visa will make the applicant pool more favorably selected. If the applicants are negatively selected before the policy change, policy restrictions will lead to an even more negative selection with a decrease in the average applicant quality.

2.4 Data

I use individual-level information from I-129 forms of processed H-1B petitions, obtained from USCIS through a Freedom of Information Act (FOIA) request made in 2017 (hereinafter, "USCIS dataset"). This dataset comprises 2.9 million H-1B applications submitted between FYs 2004 and 2014, detailing the beneficiary's country of birth, occupation, proposed compensation, current visa status, final decision, the employer's specifics and more administrative information. This is an ideal period to analyze the effects of the H-1B lottery on the quality of the H-1B applicants. Firstly, there were no changes in the specific H-1B quotas following the 66% (from 195,000 to 65,000) reduction in the regular quota in FY 2004. Secondly, there was exceptionally heightened demand resulting in random lotteries for H-1B visas in FY 2008, FY 2009 and FY 2014 during this period, providing the necessary variation in the probability of obtaining an H-1B visa.

I first limit the sample to include H-1B petitions for new employment only. Petitions that seek to extend an existing H-1B visa or to change jobs within the validity of an existing H-1B visa are not subject to any cap. The quality of these applicants is thus likely insensitive to the winning probability for the current FY but possibly dependent on the winning probability a few years ago, potentially obscuring the analysis. To avoid outliers and reduce measurement error, petitions with offered annual compensation of more than \$500,000 are not considered in the analysis, which account for approximately 0.5% of the sample. Additionally, to limit the sample to full-time employees, petitions with less than \$15,080 in annual compensation are removed.¹¹

Next I proceed to generate the independent variable of interest, the probability of obtaining

¹¹This takes reference from the definition of full-time employment using the federal minimum hourly wage of \$7.25 and 52 work weeks of 40 hours per week.

an H-1B visa, for the observations in the sample. There are two types of petitions in the sample, those who are subject to the random H-1B lottery if there is one and those who are not. The latter consists of (1) those for non-profit firms, universities, and research institutions, and (2) citizens of Australia, Canada, Chile, Mexico and Singapore. The winning probability is always one for these individuals as they successfully obtain H-1B visas regardless if there exists an H-1B lottery. Applications associated with employment in colleges, universities and research institutions are dropped as compensation packages for this group can be designed very differently from the rest. Table 2 present the descriptive statistics of the sample.

As for the rest of the sample, which are subject to the random lottery, the winning probability is calculated using the information published by USCIS. For the period we are examining, other than the three FYs where a computer-generated random lottery was adopted, petitions were approved on a first-come, first-serve basis and thus a winning probability of one is assigned to those FYs. For FY 2008, FY 2009 and FY 2014, USCIS stated in its press releases that it received approximately 150,000, 163,000, and 124,000 cap-subject H-1B petitions respectively. Taking into account of the 20,000 permits granted to new H-1B applicants with advanced degrees from U.S. institutions, the probability of obtaining a regular H-1B visa for cap-subject applicants was 0.5 in FY 2008, 0.45 in FY 2009 and 0.63 in FY 2014.¹² It would be ideal to distinguish petitions involving an advanced U.S. degree as they were likely associated with a different winning probability, a result of the two-tiered random lottery process. However, the USCIS dataset does not allow the identification of those petitions.

Although compensation information is included in the USCIS dataset and can serve as our dependent variable, variation in compensation can be potentially driven by occupation-specific factors. To isolate the effects of the random lottery from these noises, a second dataset, the Occupational Employment Statistics (OES) survey by the Bureau of Labor Statistics (hereinafter, "BLS dataset") is used to generate alternative dependent variables. This dataset produces annual cross-sectional estimates of total employment and wage information for all defined occupations, based on survey

¹²If the number of new H-1B workers with an advanced U.S. degree exceeds 20,000, a random lottery will be conducted to allocate the permits and those unselected will be put into the main H-1B pool that is subject to the cap of 65,000. Hence, there were 130,000, 143,000, and 104,000 applicants competing for 65,000 H-1B visas in FY 2008, FY 2009, and FY 2014 respectively.

results obtained every May. The estimates are available at the national, state and metropolitan levels. I match the occupational codes between the BLS dataset and the USCIS dataset to generate annual national average wages for the list of occupations included in the USCIS dataset.¹³

2.5 Empirical Approach and Results

I use the difference in differences (DID) approach for empirical estimation. The control group in the analysis consists of citizens of Australia, Canada, Chile, Mexico, and Singapore in the sample, who were exempted from the H-1B caps. The rest of the applicants in the sample forms the treatment group, who were subject to the H-1B caps. FY 2008, FY 2009 and FY 2010 are the treatment periods, where a computerized random lottery was implemented to ration H-1B visas.

The main regression specification follows a generalized DID design with treatment intensity incorporated:

$$\ln(wage_{ijt}) = \beta_0 + \beta_1 treat_{ijt} + \beta_2 DID_{ijt}(1 - p_t) + \lambda_t \cdot \delta_j + \epsilon_{ijt}, \quad (2.16)$$

in which the subscripts i , j , and t denote individual, occupation and year respectively.

The dependent variable $\ln(wage_{ijt})$ is the natural log of the proposed annual compensation associated with an H-1B application and is used as a measure for the quality of the applicant. The indicator variable $treat_{ijt}$ takes the value of one if the individual was subject to the H-1B permit cap, and takes the value of zero otherwise. The key DID variable, DID_{ijt} , equals one if the application was cap-subject and submitted during the treatment periods, and zero otherwise. The fall in the probability of being awarded an H-1B visa is taken as the treatment intensity, and p_t takes the value of 0.5 for applications submitted for FY 2008, 0.45 for FY 2009, and 0.63 for FY 2014, and equals one otherwise. Lastly, $\lambda_t \cdot \delta_j$ represents year-occupation fixed effects and ϵ_{ijt} is the residual. The coefficient of interest, β_2 represents the proportional change in annual wages of the H-1B applicants, given a change in the winning probability in the H-1B random lottery.

¹³The BLS dataset uses Standard Occupational Classification (SOC) codes while the USCIS dataset adopts three-digit Dictionary of Occupational Titles (DOT) codes. The mapping between 2000 SOC and DOT can be obtained from USCIS. As OES implemented 2010 SOC in its May 2012 estimates, manual mapping is required for 2012 and thereafter.

The estimation results for Equation (12) without and with treatment intensity are detailed in the first and second columns of Table 3. Additional control variables such as state fixed effects and whether the applicant was physically present in the U.S. at the point of application are included to improve the precision of the estimates. The estimates for the DID coefficient are negative and statistically significant in both specifications, suggesting deteriorating applicant quality under a more restrictive immigration policy. Specifically, the host country earnings for H-1B applicants are on average decreased by 0.8% when there was a rationing lottery for H-1B visas.

As the majority of the H-1B visas were granted to STEM workers and the U.S. has been long in need of a stronger STEM workforce, I further zoom in to this specific group for the change in self-selection patterns and the associated results are included in Column (3). The finding aligns with that of the general H-1B applicant pool that on average, the U.S. wages for H-1B STEM workers are 2.0% lower when a visa lottery is in place.

In addition, wages are likely influenced by occupation-specific factors such as unbalanced growths across industries, biased technological changes, or simply shifts in the demand and supply of in certain occupations over time. As such, I use the occupation-specific national means obtained from the BLS dataset to calculate demeaned dependent variable as alternatives in the estimation:

$$wage_{ijt} - mean_{jt} = \beta_0 + \beta_1 treat_{ijt} + \beta_2 DID_{ijt}(1 - p_t) + \lambda_t \cdot \delta_j + \epsilon_{ijt}. \quad (2.17)$$

$$\ln(wage_{ijt}) - \ln(mean_{jt}) = \beta_0 + \beta_1 treat_{ijt} + \beta_2 DID_{ijt}(1 - p_t) + \lambda_t \cdot \delta_j + \epsilon_{ijt}. \quad (2.18)$$

The estimation results from focusing on the deviation of the H-1B wages from the national occupation-specific mean wages are reported in Column (4) & (5). The estimated DID coefficients remain negative and remain similar in magnitude to those from the previous specifications, suggesting a consistent story that given visa uncertainty, the H-1B applicants become lower-skilled.

As treatment periods FY 2008 and FY 2009 happened to be in the depth of the Great Recession, there might be concerns of possible omitted variable bias, where the recession influenced both the offered H-1B wages and demand for H-1B visas, resulting in an H-1B lottery. In the original specification, the inclusion of year-occupation effects should have alleviated the issue of omitted variable bias, accounting for possible occupation-specific effects of the Great Recession. To ensure

the robustness of the results, I have included two more specifications to examine applicant quality within and outside of the Great Recession period, using the simple DID design:

$$\ln(wage_{ijt}) = \alpha_0 + \alpha_1 treat_{ijt} + \alpha_2 DID_{ijt} + \alpha_3 post_{ijt} + v_{ijt}, \quad (2.19)$$

where DID_{ijt} is the interaction term of $treat_{ijt}$ and the treatment period indicator, $post_{ijt}$. In particular, results in Column (6) are obtained from comparing FY 2006 and FY 2009, both in the depth of the Great Recession. Column (7) reports the estimates from a DID comparison of FY 2011 and FY 2014. The estimated coefficients from both specifications are consistently negative and statistically significant.

In summary, the effect of a more restrictive H-1B visa program on applicant quality is consistently estimated to be negative. The presence of the H-1B lottery reduced the quality of the applicants, measured by the offered annual compensation, by between 0.8% and 3.5% on average, holding others constant. According to the proposed framework, the results suggest that the H-1B applicants are negatively selected and H-1B visa restrictions have reinforced the unfavorable selection of the applicants.

2.6 Conclusion

In this paper, I examine the impact of a change in the probability of obtaining an H-1B visa on the quality of H-1B applicants. I first propose a theoretical framework incorporating uncertainty in getting legal permission to work in the host country. The framework indicates that a more restrictive immigration visa policy will reinforce the selection of the prospective applicants. By exploiting the heterogeneity in the winning probability in the H-1B visa for different FYs and different applicant groups during the period of FY 2004 to FY 2014, I estimate the effect of the H-1B lottery on the quality of the applicants.

I discover that the quality of H-1B workers deteriorated when the chances of getting H-1B visa were lower. Specifically on average, the quality of the applicants is reduced by between 0.8% and 3.5% when a random lottery was implemented in FY 2008, FY 2009 and FY 2014. Taking reference

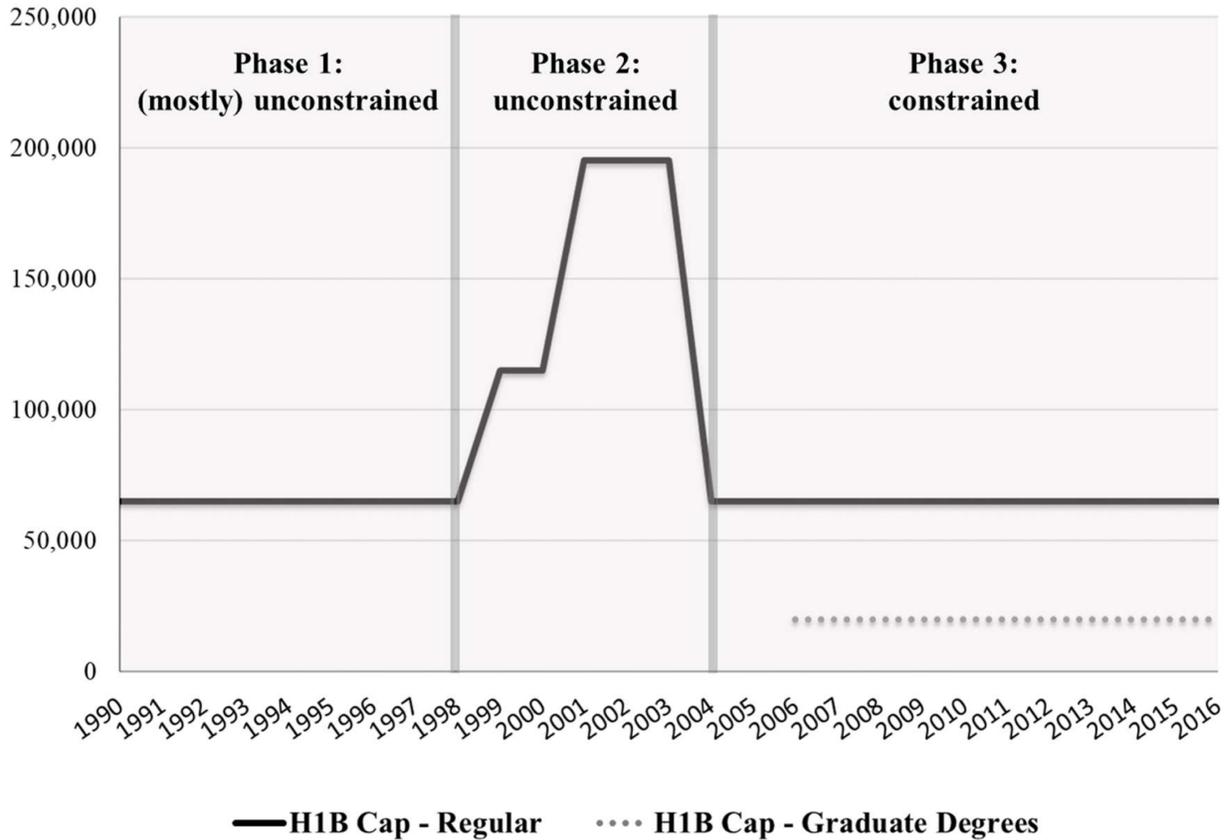
from the proposed framework, the results indicate that the pool of H-1B applicant are negatively selected and this unfavorable selection is strengthened with a more restrictive H-1B policy.

Bibliography

- Amuedo-Dorantes, C. & Furtado, D. (2017), ‘Settling for academia? h-1b visas and the career choices of international students in the united states’, *Journal of Human Resources* pp. 0816–8167r1.
- Aydemir, A. (2011), ‘Immigrant selection and short-term labor market outcomes by visa category’, *Journal of Population Economics* **24**(2), 451–475.
- Borjas, G. J. (1987), ‘Self-selection and the earnings of immigrants’.
- Borjas, G. J. (2008), ‘Labor outflows and labor inflows in puerto rico’, *Journal of Human Capital* **2**(1), 32–68.
- Borjas, G. J. (2014), *Immigration economics*, Harvard University Press.
- Bound, J., Khanna, G. & Morales, N. (2017), Understanding the economic impact of the h-1b program on the us, in ‘High-Skilled Migration to the United States and its Economic Consequences’, University of Chicago Press.
- Chiquiar, D. & Hanson, G. H. (2005), ‘International migration, self-selection, and the distribution of wages: Evidence from mexico and the united states’, *Journal of political Economy* **113**(2), 239–281.
- Chiswick, B. R. (2000), ‘Are immigrants favorably self-selected’, *Migration theory: Talking across disciplines* pp. 61–76.
- Doran, K., Gelber, A. & Isen, A. (2014), The effects of high-skilled immigration policy on firms: Evidence from h-1b visa lotteries, Technical report, National Bureau of Economic Research.
- Glennon, B. (2018), How Do Restrictions on High-Skilled Immigration Affect Offshoring? Evidence from the H-1B Program, PhD thesis, Carnegie Mellon University.
- Grogger, J. & Hanson, G. H. (2011), ‘Income maximization and the selection and sorting of international migrants’, *Journal of Development Economics* **95**(1), 42–57.
- Ibarraran, P. & Lubotsky, D. (2007), Mexican immigration and self-selection: New evidence from the 2000 mexican census, in ‘Mexican immigration to the United States’, University of Chicago Press, pp. 159–192.
- Kaestner, R. & Malamud, O. (2014), ‘Self-selection and international migration: New evidence from mexico’, *Review of Economics and Statistics* **96**(1), 78–91.
- Kato, T. & Sparber, C. (2013), ‘Quotas and quality: The effect of h-1b visa restrictions on the pool of prospective undergraduate students from abroad’, *Review of Economics and Statistics* **95**(1), 109–126.
- 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**(3), 473–508.
- Massey, C. G. (2016), ‘Immigration quotas and immigrant selection’, *Explorations in Economic History* **60**, 21–40.
- Mayda, A. M., Ortega, F., Peri, G., Shih, K. & Sparber, C. (2018), ‘The effect of the h-1b quota on the employment and selection of foreign-born labor’, *European Economic Review* **108**, 105–128.
- McKenzie, D. & Rapoport, H. (2010), ‘Self-selection patterns in mexico-us migration: the role of

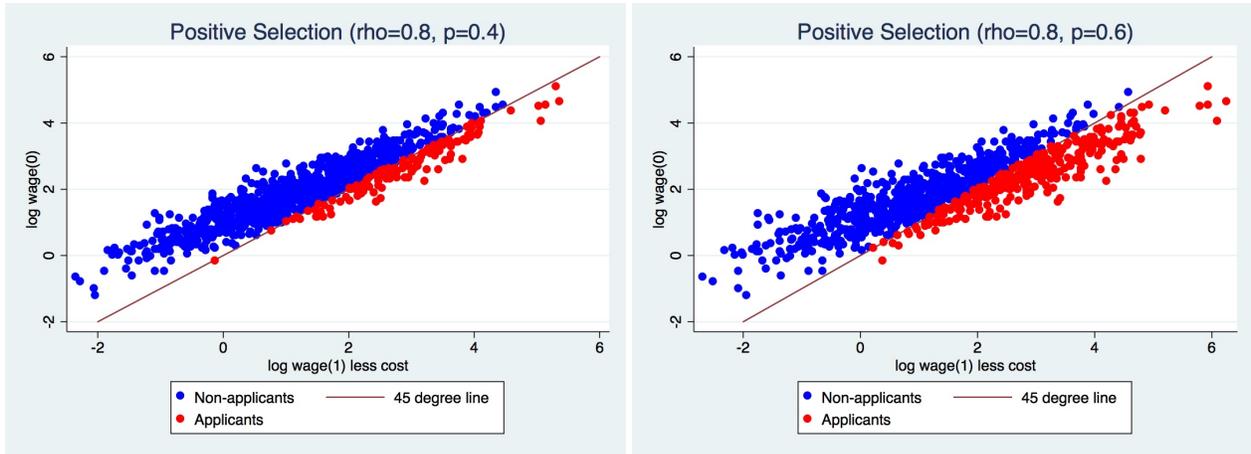
- migration networks', *The Review of Economics and Statistics* **92**(4), 811–821.
- Miller, P. W. (1999), 'Immigration policy and immigrant quality: The Australian points system', *American Economic Review* **89**(2), 192–197.
- Orrenius, P. M. & Zavodny, M. (2015), 'Does immigration affect whether US natives major in science and engineering?', *Journal of Labor Economics* **33**(S1), S79–S108.
- Peri, G., Shih, K. & Sparber, C. (2015a), 'Foreign and native skilled workers: What can we learn from the H-1B visa lotteries?'.
- Peri, G., Shih, K. & Sparber, C. (2015b), 'STEM workers, H-1B visas, and productivity in US cities', *Journal of Labor Economics* **33**(S1), S225–S255.
- Peri, G., Shih, K. Y. & Sparber, C. (2014), Foreign STEM workers and native wages and employment in US cities, Technical report, National Bureau of Economic Research.
- Shih, K. (2016), 'Labor market openness, H-1B visa policy, and the scale of international student enrollment in the United States', *Economic Inquiry* **54**(1), 121–138.
- Sparber, C. (2018), 'Choosing skilled foreign-born workers: Evaluating alternative methods for allocating H-1B work permits', *Industrial Relations: A Journal of Economy and Society* **57**(1), 3–34.
- Sperry, W. (2017), Are H1B Visa Workers Paid Less than Similarly Employed Natives?, PhD thesis.
- Watts, J. R. (2001), 'The H-1B visa: Free market solutions for business and labor', *Population Research and Policy Review* **20**(1-2), 143–156.
- Zavodny, M. & VThe, H. (2003), 'The H-1B program and its effects on information technology workers', *Economic Review-Federal Reserve Bank of Atlanta* **88**(3), 33–44.

Figure 2.1: Changes in H-1B Caps



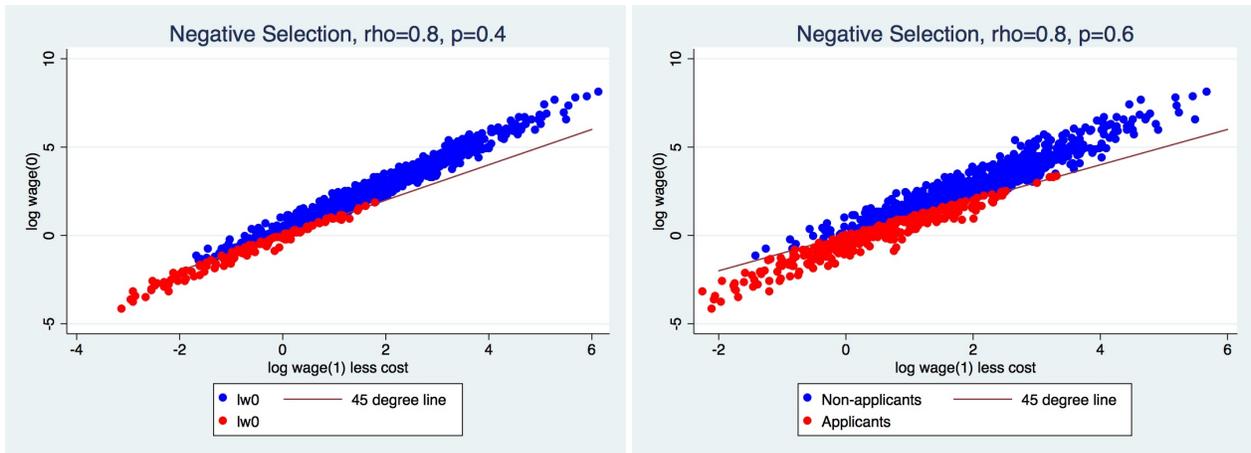
Notes: This figure plots the cap on the number of H-1B visas by fiscal year. since the Immigration Act of 1990, there has been an annual cap on the number of new H-1B visas that can be issued to private sector businesses. This cap is set by Congress and the President. Throughout most of the 1990s, the cap was set at 65,000 visas and applications rarely outstripped supply. It was increased to 195,000 visas by the American Competitiveness and Workforce Improvement Act of 1998 and the American Competitiveness in the Twenty-First Century Act of 2000 (AC21). During this period, the cap limits were never reached. The AC21 stipulated that this reversion would happen in the absence of any additional legislation, but, despite a trend towards less restrictive labor laws, no legislation was enacted, and the cap level reverted back to 65,000. It was raised by 20,000 in 2006, but those additional 20,000 could only be used for applicants with a graduate degree. Since then, the cap has not changed, and it has been binding in every year. The identification in this paper exploits the sharp reduction in the annual H-1B cap in fiscal year 2004; in the early 2000s, the cap was not binding. After the cap change, it has been binding in every year. Source: Glennon (2018).

Figure 2.2: Simulation Results with Positive Selection



Notes: This figure plots simulated log source country earnings and log host country earnings less a fixed opportunity cost in the case of positive selection. Workers in the source country make the decision of whether to apply for the work visa based on the comparison of the log source country earnings and log host country earnings less a fixed opportunity cost. Workers who decide to apply are marked in red and those who decide to remain in the source country are marked in blue. The correlation between the source and host country earnings is assumed to be 0.8.

Figure 2.3: Simulation Results with Negative Selection



Notes: This figure plots simulated log source country earnings and log host country earnings less a fixed opportunity cost in the case of negative selection. Workers in the source country make the decision of whether to apply for the work visa based on the comparison of the log source country earnings and log host country earnings less a fixed opportunity cost. Workers who decide to apply are marked in red and those who decide to remain in the source country are marked in blue. The correlation between the source and host country earnings is assumed to be 0.8.

Table 2.1: Final Receipt Dates and H-1B Lottery Details

Fiscal Year	Final Receipt Date	Days in Filing Period	Lottery	Number of Lottery-subject Petitions Received
2004	February 17, 2004	323	No	N.A.
2005	October 1, 2004	184	No	N.A.
2006	August 10, 2005	132	No	N.A.
2007	May 26, 2006	56	No	N.A.
2008	April 3, 2007	3	Yes	150,000
2009	April 7, 2008	7	Yes	163,000
2010	December 21, 2009	265	No	N.A.
2011	January 26, 2011	301	No	N.A.
2012	November 22, 2011	236	No	N.A.
2013	June 11, 2012	72	No	N.A.
2014	April 7, 2013	7	Yes	124,000
2015	April 7, 2014	7	Yes	172,500
2016	April 7, 2015	7	Yes	233,000
2017	April 7, 2016	7	Yes	236,000
2018	April 7, 2017	5	Yes	199,000
2019	April 6, 2018	5	Yes	190,098

Notes: USCIS begins accepting H-1B petitions on the first business day of April every year for permits counting towards the next fiscal year. Permits are usually granted on a first-come-first-serve basis until the caps are reached, beyond which no petitions are accepted. The end of the filing period is demarcated by the "final receipt date", which is the date on which USCIS receive enough petitions to fill remaining quotas. In FY 2008, FY 2009, and from FY 2014 onwards, USCIS received so many petitions within the first week that a computerized random lottery was used to allocate the permits. All information in this table is gathered from the press releases of USCIS.

Table 2.2: Summary Statistics

	N	Mean Wage	s.d	Top 3 Occupations	Top 3 Countries
Cap-subject applicants	1,140,847	63,324	36,958	Computer-related (57.9%) Architecture, Engineering and Surveying (10.4%) Administrative (9.0%)	India (60.9%) China (8.0%) Philippines (3.1%)
Cap-exempt applicants	56,414	83,666	52,677	Computer-related (20.8%) Architecture, Engineering and Surveying (18.7%) Administrative (14.2%)	Canada (68.0%) Mexico (25.7%) Singapore (4.1%)

Notes: The sample includes H-1B petitions for new employment from FY 2004 to FY 2014. All petitions with annual compensation less than \$15,080 or more than \$500,000 are excluded in the sample to restrict the analysis to full-time employees and to remove outliers. Petitions associated with colleges, universities, and research institutions are removed.

Table 2.3: Quality of H-1B Applicants in the Random Lottery

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DID	-0.0079* (0.0042)	-0.0139* (0.0084)	-0.0204** (0.0095)	-2706*** (874.1)	-0.0167* (0.0094)	-0.0133* (0.0081)	-0.0351*** (0.0096)
Treatment group	-0.2490*** (0.0022)	-0.2493*** (0.0022)	-0.2306*** (0.0025)	17152*** (220.2)	-0.2339*** (0.0024)	-0.2513*** (0.0058)	-0.2256*** (0.0065)
In U.S. dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.2384	0.2384	0.2898	0.1707	0.1896	0.1825	0.1780

Notes: The sample includes H-1B petitions for new employment from FY 2004 to FY 2014. All petitions with annual compensation less than \$15,080 or more than \$500,000 are excluded in the sample to restrict the analysis to full-time employees and to remove outliers. Petitions associated with colleges, universities, and research institutions are removed. The dependent variable used is the natural log of offered H-1B compensation stated in the application, except for Column (4) & Column (5), where the offered H-1B compensation less the national occupation-specific average wage and the natural log of this demeaned compensation are used respectively. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 3

High-skilled Immigrant Workers and U.S. Firms' Access to Foreign Venture Capital

3.1 Introduction

Skilled immigration has been an increasingly important element reshaping competitive edges by filling critical labor gaps and contributing specialized skills, especially in the high-tech industries. Countries around the world have embraced targeted immigration programs to attract skilled immigrants, for their perceived potential to propel growth in the host countries through various channels beyond the labor markets, such as innovation, entrepreneurship, and trade. In this paper, I examine a novel firm-level channel through which skilled immigration might increase growth - attracting inflows of international venture capital (VC) from their home countries and in turn enabling host country start-ups to scale up.

In this paper, I propose that skilled immigrants play a key role in overcoming cross-border investment frictions through home-country professional networks, common language and culture, and U.S. start-ups with these workers have better access to VC funding from their origin countries. Empirically, I make use of the exogenous variation in firms' country-specific H-1B visa lottery outcomes to examine the causal relationship between the access to skilled foreign workers and firms' ability to obtain VC financing from the respective source countries. I also complement the analysis by extending the period of analysis and exploiting firms' ex-post employment of skilled immigrant workers with a panel study. I find empirical evidence via both approaches that having skilled workers from a particular origin country improves start-ups' same-origin VC outcomes. Lastly, I construct a framework of firms' hiring decisions incorporating financial market frictions to demonstrate how this VC mechanism can be explored together with the standard labor market effects of immigration. A numerical exercise ensues, showing that almost one-third of the displacement effects of skilled immigration can be mitigated through the VC channel.

In the first empirical strategy, I exploit the exogenous variation in firms' access to skilled labor from a given country in Fiscal Year (FY) 2009 and FY 2014, where random lotteries were implemented to allocate H-1B visas. I use supervised machine learning¹ to create a unique firm-year-country dataset by merging VC investment data from VentureXpert, H-1B approval data from U.S. Citizenship and Immigration Services (USCIS), and H-1B application data from Labor Condi-

¹See Appendix ?? for details.

tion Applications (LCA). There is, however, an inherent data limitation: the nationality information for unselected workers is unavailable (see Clemens, 2013), and as a result, country-specific lottery win rates cannot be directly observed.

Two techniques are deployed to deal with the data issue. Firstly, as the H-1B lotteries entirely determine the availability of the nationality information, the data is missing completely at random (MCAR). Multiple imputation can then be used to generate statistically consistent estimates (Rubin, 2004). Missing data is stochastically imputed multiple times, and the combined estimates suggest 0.07 more VC deals and a 6 percentage points increase in the chance of getting VC financing from a source country when a firm has access to skilled labor from that country. This translates to an average of \$864,000² additional VC funding, which can be sizable for a single start-up. The second technique focuses on firms with overall win rates of either one or zero. All relevant country-specific win rates are one for firms with overall win rates of one, and zero for those with overall win rates of zero. Although nationality information remains absent for the latter group, the dichotomous win rates and the data characteristics allow me to produce a lower-bound estimate, which suggests 0.03 more same-origin VC deals (equivalent to \$405,000) when firms succeed in H-1B lotteries.

Although the randomness guarantees the internal validity of the results, the external validity in the first strategy might be constrained by the limited time frame (only FY 2009 and FY 2014). An alternative strategy using an extended period of analysis is thus used to supplement the first. Focusing on the top VC investor countries, I run a panel study using firm-year-country observations on firms' approved skilled immigrant workers and foreign VC deal outcomes from FY 2005 to FY 2014. I find that having skilled workers from a country brings in 0.08 more same-origin VC deals and increases the likelihood of receiving same-origin VC deals by 3.2 percentage points. This positive externality is evident for both new and follow-on VC deals, suggesting that skilled immigrants are helpful in both attracting new investments and facilitating on-going projects. On average, an additional skilled immigrant worker brings in 0.03 more same-origin VC deals for the start-up, but this effect diminishes with higher number of same-origin workers. The results survive alternative fixed effects specifications, and withstand changes in the period of analysis or inclusion of investor

²The average VC deal size is \$12.7 million, according to National Venture Capital Association (NVCA) 2019 report.

countries.

I also examine the underlying mechanism of how skilled immigrants influence same-origin VC outcomes and explore heterogeneous effects across firms and VC funds. Firstly, I find that skilled workers from a given country increases this country's share of the firm's foreign VC deals and its share of the firm's overall VC deals by 14.3 percentage points and 2.5 percentage points respectively. This shows that apart from the increased overall human capital contributed by these immigrants, there must exist certain country-specific factors driving the mechanism. Home-country professional networks likely play a key role here, as skilled immigrant workers relocating directly from their home countries are found to be much more influential, especially with regard to new VC investments. On the other hand, recent foreign-born graduates from U.S. institutions have very limited influence, as they unlikely maintain strong networks in their home countries. In addition, the effect is significantly stronger for non-English speaking groups and countries with more distinct cultures from the U.S., signaling the importance of common language and culture. Also, the effect diminishes with both the age of VC funds and firms, suggesting that experienced VC funds and mature firms are less reliant on personal connections in sourcing VC deals.

This paper sheds more light on the impact of migration at the organizational level beyond the macro lenses such as labor markets and global trade. Despite the apparent relevance, the impact of migration is rather under-explored at the organizational level, which is one of the primary venues that immigrants exert influence and get influenced. Nevertheless, there is an emerging body of research examining the role of migration in explaining firm-level phenomena (e.g., Foley & Kerr 2013; Hernandez 2014; Wang 2015; Kulchina, 2016; Choudhury & Kim 2019; Balachandran & Hernandez 2019; Glennon 2020). The organizational-level discussion of immigration greatly enriches and deepens our understanding of the fundamental mechanisms through which immigrants interact with the host country economy and society. This paper contributes to this young but growing firm-level literature on migration by providing evidence and examining the underlying mechanism of skilled immigrant workers improving their employers' access to VC from their home countries.

This paper also takes a novel angle to study skilled immigration and inbound VC investments from abroad, at firm-level granularity. Several papers take a closer look at the relationship between VC and skilled immigration. For instance, Dimmock et al (2019) use overall (both U.S. and foreign)

follow-on VC funding as an indicator of firm performance, and find that firms with better access to H-1B workers perform better. Other papers focus on cross-border VC activities specifically. Madhavan & Iriyama (2009) and Pandya & Leblang (2011) use aggregate-level analysis and argue that skilled migrants can provide U.S. VC investors with additional information on source-country investment opportunities. A recent paper by Balachandran & Hernandez (2019) proposes that U.S. VCs invest at home in immigrant entrepreneurs and gain access to their knowledge and networks, which facilitate subsequent investments in the entrepreneurs' home countries. These studies look at outbound investment activities to the sending countries. Little scrutiny has been given to VC inflows to the host countries. This paper complements the discussion on skilled immigration and cross-border VC activities by examining this alternative direction of investment flows.

Lastly, this paper proposes a novel framework of firms' hiring decisions to incorporate both the traditional labor market and foreign VC channels by weaving in credit frictions. In this new framework, when a firm hires skilled foreign-born workers, its access to VC funding from the workers' origin countries improves, easing financing constraints and enabling the firm to hire more. Through a numerical exercise, I find that almost one-third of the displacement effect of skilled immigrants on natives can be mitigated by the foreign VC mechanism. This provides a fresh angle to look more holistically at how immigrants influence natives' labor market outcomes, beyond the immediate substitution and scale effects.

The rest of the paper is organized as follows. The next section provides an overview of the H-1B visa program, followed by Section 3, which describes the data used for empirical estimation. The empirical approaches and results are discussed in Section 4. Section 5 includes a proposed framework to combine the labor market and VC market channels in relation to high-skilled immigration and the last section concludes.

3.2 Background on H-1B Visas

Introduced in 1990, the H-1B visa program allows foreign-born college-educated professionals to work temporarily in specialty occupations in the U.S.³ In particular, this program has been exten-

³To qualify for an H-1B visa, "the position must meet one of the following requirements: (1) a bachelor's or higher degree or its equivalent is normally the minimum entry requirement for the position; (2) the degree

sively used by those in the fields of science, technology, engineering, and mathematics (STEM), to enter the U.S. workforce.⁴

A foreign-born individual intending to work in the U.S. with an H-1B visa must find a sponsoring firm that offers him a job in qualifying occupations. This firm must first submit a Labor Condition Application (LCA) with the Department of Labor, demonstrating accordance with the U.S. law (see footnote 10). Upon LCA approval, the firm proceeds to complete an I-129 application with U.S. Citizenship and Immigration Services (USCIS) during the filing period.⁵ An H-1B visa is valid for a period of three years, extendable once to a maximum of six years in total. Nonetheless, H-1B visa holders can possibly remain in the U.S. after the six-year period as these visas allow for dual intent, where workers can simultaneously apply for permanent residency while working on a nonimmigrant basis.

The total number of new H-1B visa issuances is subject to a cap, and the cap has changed several times as depicted in Figure 3.1. Specifically, the original limit of 65,000 H-1B visas was increased to 115,000 for FYs 1999 and 2000 and lifted again for FYs 2001 to 2003 to 195,000. For FY 2004, the cap was reverted to 65,000 regular visas (hereinafter "regular cap") with added 20,000 visas for workers with U.S. postgraduate degrees (hereinafter "postgraduate cap"), and these limits have prevailed since then. Several groups are exempted. From FY 2001 onwards, individuals on cap-subject H-1Bs who are switching jobs or renewing their visas are excluded from the caps. An uncapped category for workers in colleges, universities, and non-profit research institutions was also created in FY 2001. In addition, a suite of Free Trade Agreements (FTAs) have made special provisions for citizens of Australia, Canada, Chile, Mexico and Singapore with regard to temporary work visas in the U.S.⁶

requirement is common to the industry in parallel positions among similar organizations or, in the alternative, the position is so complex or unique that it can be performed only by an individual with a degree; (3) the employer normally requires a degree or its equivalent for the position; or (4) the nature of the specific duties is so specialized and complex that the knowledge required to perform the duties is usually associated with attainment of a bachelor's or higher degree."

⁴According to Peri, Shih, and Sparber (2015), H-1B workers have contributed to approximately half of the growth registered in the college-educated STEM workforce since 1990.

⁵The filing period starts on the first working day of April for petitions to start work in the next fiscal year (FY), which is from October 1 in the same year to September 30 in the following year.

⁶Chilean and Singaporean nationals are permitted to work temporarily in the U.S. on a nonimmigrant basis in specialty occupations under the H-1B1 program, active from FY 2004 onwards. Although the H-1B1 program specifies caps of 1,400 and 5,400 H-1B1 visas for Chileans and Singaporeans respectively, the caps have never been binding up-to-date. Australians are permitted to work temporarily in the U.S. on a

USCIS normally grants H-1B visas in a first-come, first-serve fashion until the caps are exhausted. The end of the application period is demarcated by the "final receipt date", which is the date on which USCIS receive enough petitions to fill the remaining quotas. Prior to FY 2004, the cap was never filled. Beginning FY 2004, the time it took to reach the limit steadily shortened. Notably, for FYs 2008 and 2009, USCIS received overwhelming volumes of applications within the first week of the filing period, which compelled USCIS to resort to a computerized random lottery to allocate all cap-subject H-1B visas. The demand for H-1B visas subsided for a few years until it surged remarkably again for FY 2014, and it has remained strong since then. Random lotteries have been implemented from FY 2014 onwards and remained as a key feature of the H-1B program. Table 3.1 summarizes the final receipt dates and specifics of the H-1B lotteries over the years.

3.3 Data

This paper identifies the effect of having high-skilled immigrant workers from a source country on the VC investments received from the same country, examined at the firm-level. I combine data from multiple sources to permit the analysis of such link. Specifically, I obtain data on VC activities from Thomson Financial's VentureXpert, and data on H-1B applications and approvals from the Department of Labor and USCIS respectively. Figure 3.2 shows that conditional on receiving VC funding from a given source country, firms with skilled workers of the same origin tend to have more VC deals from that country.

3.3.1 Data on VC investments

I utilize data from Thomson Financial's VentureXpert database, the most comprehensive source for in-depth information covering worldwide venture transactions. It provides rich details at both the firm-level and deal-level, including names and locations of funds and firms, industry, initial public offering (IPO) information, time of investment, and round series, etc. Being officially endorsed by the National Venture Capital Association (NVCA), VentureXpert is one of the most widely

nonimmigrant basis in specialty occupations under the E-3 program, applicable from May 2005. Citizens of Canada and Mexico are allowed to work in the U.S. in prearranged business activities for U.S. or foreign firms under the TN program, effective from 1994.

used data source for the VC-related research. A number of prior studies, including Kaplan et al (2002), Maats et al (2011) and Kaplan & Lerner (2017), advocate the coverage and accuracy of VentureXpert through comparisons with other VC databases.

I extract from VentureXpert the universe of venture investments on firms located in the U.S. for the period of 2004 to 2017 (hereinafter, "VX dataset"). The extracted dataset contains 167,031 venture investments received by 24,982 U.S. firms, out of which 16,453 deals on 5,685 U.S. firms were funded by foreign venture capitalists. Table 3.2 presents the number of VC investments made by U.S. and non-U.S. funds on U.S. firms respectively over the years. As Table 3.2 suggests, although domestic players still dominate the VC space in the U.S., foreign investors have been increasing their presence, accounting for almost 14% of all VC deals on U.S. firms in recent years, up from around 8% during the Great Recession years. Deals funded by foreign investors experience a compound annual growth rate (CAGR) of 8% during the post Great Recession years, dwarfing the 0.5% CAGR for deals funded by U.S. venture capitalists. Table 3.3 lists the overall top investor countries and their shares of VC deals in the U.S. for the period of 2004 to 2017.

I then construct the dependent variables of interest, the various VC deal outcomes, including number of deals received, the share of deals received out of all deals received, and the probability of getting VC investments, all aggregated at the firm-year-country level.⁷ Despite all its strengths, VentureXpert lacks comprehensive coverage on deal sizes, otherwise it would be ideal to have firm-year-country-level deal amount as another outcome variable. I also obtain several control variables from the VX dataset, such as firm age, IPO outcome, and the number of U.S deals received at firm-year level.

3.3.2 Data on high-skilled immigrant workers

The first set of data on skilled immigration is the individual-level information from I-129 forms of processed petitions from USCIS (hereinafter, "USCIS dataset").⁸ This dataset comprises about 3.3 million approved applications submitted between FYs 2004 and 2014, detailing the beneficiary's

⁷To align with the H-1B datasets, I adopt the FY system and all subsequent 'year' refers to FY. Also, as H-1B visas are valid for 3 years, all deal outcomes are measured for a three-year period following the award of the H-1B visas.

⁸This dataset was obtained through a Freedom of Information Act (FOIA) request by Britta Glennon. See Glennon (2020).

country of birth, occupation, proposed compensation, current visa status, the employer's specifics and more administrative information. The USCIS dataset covers a wide spectrum of visa types for high-skilled immigrants, but for the purpose of this paper, the most relevant is the H-1B visa, totaled at approximately 2.9 million petitions. Table 3.4 lists the top source countries for H-1B workers. India accounts for more than half of all H-1B approved applications, followed by China, which accounts for 8.5%.

Since USCIS does not retain records of the unselected H-1B petitions (Clemens, 2013), I gather information on worker-level petition records and create measures of firm-level demand for foreign-born high-skilled workers from the LCA data (hereinafter, "LCA dataset"), made available by the U.S. Department of Labor.⁹ Prior to filing a petition with USCIS, the firm must first file an LCA with the Department of Labor to attest to the employment details and affirm that the employment will be in accordance with U.S. Law.¹⁰ The LCA data contains detailed information for each prospective foreign worker, including job-specific details such as salary and the intended start dates, employer information such as firm's name, address and industry, and the application status, i.e. whether it is certified, withdrawn, or denied. I restrict the sample to include firms with non-zero LCAs filed in FY 2009 and/or FY 2014, as they were the ones participating in the H-1B lotteries.

I merge the VX dataset with the USCIS and LCA datasets using firm information. Supervised machine learning algorithms are used to improve the matching of misspelled firm names. Out of the 24,982 U.S. start-ups appeared in the VX dataset, 27.4% once hired H-1B workers during the period of FY 2005 to FY 2014. Firms applied for 3.5 H-1B visas on average in a year, with 56.0% of them applying for a single cap-subject visa and only 9.0% for more than five. The average win rate calculated for the sample is approximately 59%, which is comparable to the overall win rate of 58.4% from published statistics by USCIS.¹¹

⁹It is available at <https://www.dol.gov/agencies/eta/foreign-labor/performance>.

¹⁰There are four main labor conditions that they are required to meet: (1) recipients of the visa must receive the same or better wages and benefits as other similar company employees and as similar employees in the geographic area, (2) working conditions must be similar for all employees, (3) there must not be a "strike, lockout, or work stoppage" at the employment location when the LCA is signed and submitted, (4) any employee bargaining representatives must be notified of every application submitted.

¹¹The overall lottery win rate is calculated using reported numbers of cap-subject H-1B petitions from past USCIS press releases. See <http://www.uscis.gov/archive>.

The first empirical approach exploits the random lottery feature of the H-1B program in FYs 2009 and 2014, and the resultant exogenous variation in firm-year-country-specific win rates. However, there exist two data limitations that complicate the calculation of such win rates. Specifically, the LCA data neither contains information on whether a particular H-1B petition would be subject to the random lottery nor records the origins of the applicants.

To deal with the first data limitation, I measure a firm's demand for cap-subject H-1B workers in a FY with the proxy method devised in Dimmock et al (2019), which is the number of certified LCAs for H-1B visas filed by a firm in February and March with a start date that is five to seven months in the future. As Dimmock et al (2019) have verified, these petitions are likely to be for new employment and thus subject to the H-1B caps.

The second data limitation is the missing origin information of unselected workers from the H-1B lotteries. I use two methods to handle this missing data issue - multiple imputation and using dichotomous win rates, which I cover in greater details in the following section.

3.4 Empirical Approaches and Results

I examine the impact of skilled immigrant workers on U.S. start-ups' access to same-origin VC funding through two empirical approaches. The first approach makes use of the computer-generated H-1B lotteries in FYs 2009 and 2014, which created random shocks to firms' access to foreign-born skilled labor. The exogenous variation in firms' country-specific H-1B lottery win rates guarantees the internal validity of the results, but the period of analysis is limited to only two FYs. Hence, in the second empirical approach, using the panel data for the period from FY 2005 to 2014, I exploit the ex-post changes in firms' approved skilled workers and compare their same-origin VC financing outcomes with extensive fixed effects.

3.4.1 H-1B random lottery approach

The identification in this empirical strategy relies on the random variation in the allocation of H-1B workers of various origins across U.S. start-ups, a result of the H-1B computerized lotteries in FYs 2009 and 2014. In both years, the number of cap-subject H-1B visa applications submitted

surged and quickly exhausted the quotas. A computerized random lottery was thus put in place to allocate these visas, in which only the selected petitions were processed. This generated random shocks to firms' access to H-1B workers, as some firms were successful in getting H-1B workers from certain countries, while other were not. In this approach, I utilize the random variation in firms' country-specific H-1B lottery win rates to establish causality. The main regression equation in this approach is as follows:

$$deal_{ict} = \alpha + \beta winrate_{ict} + \gamma_c + \delta_t + \varepsilon_{ict} \quad (3.1)$$

The key independent variable is $winrate_{ict}$, which measures the country-specific win rate of the firm in getting H-1B workers from the H-1B lottery for FY t . All outcome variables are measured for a period of three years, during which the approved H-1B workers are allowed to work for the same firms. That is, $deal_{ict}$ represents the various VC outcomes for firm i from country c , aggregated for the period from FY t to FY $t + 2$. The parameter β captures the relationship between a firm's immigrant worker from a source country and its VC investments received from the same source country. A positive value of β indicates that having workers from a particular country on average increased the firm's access to VC funding from this country. Country-level fixed effects, γ_c , and year fixed effects, δ_t , are included to absorb any unobserved country-specific and year-specific characteristics. Additional control variables such as firm age, average salary of workers, and the number of LCAs filed are also included.

By exogenously varying the country-specific supply of H-1B visas across firms, these lotteries enable us to isolate the effect of skilled foreign labor on firms' VC deal outcomes from the respective origins. However, the data limitation, where the worker nationality information is missing in the LCA data, presents challenges in measuring the country-specific win rates. The combined dataset from USCIS and LCA only reveals the origins of approved H-1B workers, not the unsuccessful ones, thus making it impossible to directly observe the country-specific win rates. I deal with the missing data issue in two ways, multiple imputation and a lower-bound exercise using dichotomous lottery win rates.

Multiple imputation

Unselected petitions from the H-1B lotteries were returned to applicants unopened, and the origins of these workers are thus unknown to USCIS. Whether a petition is successful or not is completely random from a computer-randomized lottery, thus making the nationality information missing completely at random (MCAR). That is, the unobserved value of the nationality variable does not predict whether a nationality value is missing and the subset of cases with missing data is randomly determined by the H-1B lottery.¹² The fulfillment of the MCAR condition allows for the use of multiple imputation to produce statistically consistent results (see Rubin, 2004).

I first subtract the number of cap-subject approved H-1B petitions from the total number of cap-subject LCAs filed for each firm-year to obtain the unselected petitions. The missing nationality information for this group is stochastically imputed fifty times, resulting in fifty imputed datasets.¹³ Country-specific win rates for each firm-year are then calculated accordingly for each of the imputed dataset. Each of the fifty complete datasets is analyzed using the regression specification in Equation (3.1). Finally, the parameter estimates obtained from individual imputed datasets are then combined into an overall estimate and covariance matrix using Rubin's rule for inference.

Table 3.5 shows the regression results. Column (1) and (2) list the estimated coefficients when the number of VC deals from a certain source country is regressed on the country-specific win rate, without and with additional controls. Having a win rate of one in getting workers from a certain country on average leads to 0.07 more VC deals from that country, compared to a firm that fails to get any worker from that country, as shown by Column (2). Considering that the size of VC deals averages at \$12.7 million¹⁴, the estimate actually suggests an additional inflow of \$864,000 from this source country to a single start-up, which can be seen as substantial. Column (3) includes the results when the probability of getting any VC investments from a certain country is used as the outcome variable. The estimated coefficient is 0.06, representing a 6 percentage point increase in the probability of being funded by venture capitalists from a given country when the firm is successful

¹²Citizens of Canada, Mexico, Singapore, Chile and Australia on TN, H-1B1 and E3 visas are not subject to the lotteries, and these cap-exempt petitions are removed from this exercise.

¹³There is no consensus on the optimal number of imputations. The Stata documentation suggests at least 20, while White et al (2011) argue that the number of imputations should be at least equal to the percentage of cases with missing values. In this case, 30.2% of the observations have missing values.

¹⁴National Venture Capital Association 2019 Report

in getting workers from that country, as compared to a firm that fails to get any.

Firms with one or zero win rates

Instead of going full-scaled with multiple imputation, there are also two special groups of observations that carry additional information that permits a lower-bound estimation.

Overall firm-level win rates can be directly obtained by dividing the total number of approved H-1B workers by the total number of LCAs filed. Firms with overall win rates of one in a given FY succeeded in getting all demanded workers with complete nationality information from the USCIS dataset. All firm-year-country H-1B lottery win rates involved are therefore one for these firms. Firms with overall win rates of zero, on the other hand, failed in getting any foreign workers in the lotteries. In other words, if firm i wants to hire workers from c_1 and c_2 but is unlucky in the H-1B lottery with $winrate_{it}$ of zero, $winrate_{ict}$ for both countries would for sure be zero. However, available data does not reveal c_1 or c_2 , which are the countries that the firm wishes to get workers from.

I limit the analysis to look at firm-year observations with overall win rates of either one or zero in FYs 2009 and 2014. These firm-year observations account for approximately 72% of all observations in the two years of H-1B lotteries, making up a decent sample size.¹⁵ Win rates here are dichotomous and the estimated coefficient on win rate given a simple linear model will take the form as follows:

$$\hat{\beta} = \overline{deal}_1 - \overline{deal}_0 \tag{3.2}$$

where $\overline{deal}_1 = (\overline{deal} \mid winrate_i = 1)$, and $\overline{deal}_0 = (\overline{deal} \mid winrate_i = 0)$.

For firms with overall zero win rates, I need to assume that they applied for H-1B workers from certain countries but failed in getting any. However, this is subject to error as firms might be recorded as having failed in getting workers from some origins when they did not apply at all. Suppose there are n_1 firm-year observations with country-specific win rates truly equal to one, n_0 firm-year observations with country-specific win rates truly equal to zero, and another n_k firm-year observations with country-specific win rates erroneously recorded as zero, the estimated coefficient

¹⁵43.2% of all firm-year observations had overall win rates of one, and 28.7% had win rates of zero.

is biased due to observational error:

$$\hat{\beta}^* = \overline{deal}_1 - \overline{deal}_0^*, \text{ where } \overline{deal}_0^* = \frac{\overline{deal}_0 n_0 + \overline{deal}_k n_k}{n_0 + n_k}. \quad (3.3)$$

The estimated coefficient on *winrate* is thus biased downwards if $\overline{deal}_k > \overline{deal}_0$, and upwards otherwise. In the case of inevitable bias, a negative bias is preferred as this will produce a lower-bound estimate. The list of imputed H-1B origin countries for the zero-win-rate firms will include both correctly imputed and erroneously imputed ones. To ensure a negative bias, the erroneously imputed observations should have better VC outcomes than the correctly imputed ones. This is highly plausible given the data characteristics, where the top VC investors are quite distinct from the top H-1B sending countries.

I thus assume that firms with win rates of zero applied for workers from two groups of countries, the top H-1B source countries and the top VC investor countries. The first group, India, China¹⁶, Philippines, South Korea and Taiwan, are likely the corrected imputed cases as they make up almost 90% of all H-1B applications.¹⁷ The second group, Canada, UK, Switzerland, Israel, Germany, Japan, Australia, France and Singapore, are probably the erroneously imputed ones, accounting for a negligible portion of H-1B applications.¹⁸ Furthermore, citizens from Canada, Singapore and Australia can apply for cap-exempt work visas such as TN, H-1B1, and E3. As a result, it is highly probable that the coefficient of interest is biased downward, resulting in a lower bound for the effect.

In Table 3.6 Column (1), I regress the number of same-origin VC investments on the country-specific H-1B win rates, accounting for country fixed effects. Having succeeded in getting immigrant workers from a certain country on average brings 0.03 more deals from that country. Considering the average deal size, it means an additional inflow of \$405,000 VC funding from the source country to a U.S. firm. Several controls are then added, such as the total number of U.S. VC deals received, H-1B workers' salaries, and number of LCAs filed, and the results are included in Table 3.6 Column (2). The estimated coefficient on *winrate* remains statistically significant and comparable in terms

¹⁶China only became one of the top VC investor countries from FY 2015 onwards, outside of the analysis period in this approach.

¹⁷Source: USCIS reports.

¹⁸For instance, the top investors who are the most active in sending H-1B workers, Canada, United Kingdom and France, only accounts for 1.1%, 0.5% and 0.4% of all petitions respectively.

of magnitude to Column (1).

Table 3.6 Column (3) to (5) present the results with alternative outcome variables. Column (3) suggests a 1.7 percentage point increase in the probability of receiving VC from a given country when the firm is successful in getting workers from there. Table 3.6 Column (4) and (5) show that being successful in hiring workers from a source country on average increases this country's share of the firm's foreign VC funding and its share of the firm's total VC deals received by 11.0 and 1.3 percentage points respectively. As expected, the results from this approach are generally smaller in magnitude, as compared to those from multiple imputation.

With the key independent variable being dichotomous, alternative models such as linear probability and logit models are also used for robustness checks. In addition, individual countries are dropped one at a time to rule out the possibility of that the results are driven by any particular country. Figure 3.3 shows the coefficients from alternative models and modified samples by removing individual countries. The results from these robustness checks remain statistically significant and withstands changes in the source country composition.

3.4.2 Panel study approach

The exogenous shocks created by the H-1B lotteries permit the examination of the causal relationship between skilled immigrants and access to same-origin VC financing. However, the period of analysis is rather limited (only FYs 2009 and 2014), which may weaken the external validity of the results. Hence, I extend the time frame and conduct a panel study using firms' ex-post employment of approved skilled-immigrants from FY 2005 to 2014.

I frame the scope of the analysis to look at the top ten venture investor countries in this approach as they count for more than 70% of all foreign venture investments made on U.S. firms during the period of analysis (see Table 3.3). The sample includes all U.S. firms that received any venture investment during the period of 2004 to 2017. When these firms had successful IPOs, subsequent firm-year observations are excluded from the analysis, as they would not be eligible for VC funding anymore. All observations are aggregated at firm-country-year level. The baseline

estimation equation is as follows:

$$deal_{ict} = \alpha + \beta immigrant_{ict} + \theta_{it} + \gamma_c + \varepsilon_{ict} \quad (3.4)$$

where the key independent variable, $immigrant_{ict}$, reflects firm i 's employment of skilled immigrant workers from country c approved in FY t .¹⁹ The dependent variables in this approach are the same as the H-1B lottery approach: the various VC deal outcomes of firm i from country c in FY t . Again, all outcome variables are aggregated for a period of three years from FY t to FY $t + 2$, to reflect the 3-year visa validity.²⁰ Firm-year fixed effects, θ_{it} control for any unobserved firm-year characteristics that may introduce omitted variable bias into the analysis, such as the case where the firm was hiring more workers and receiving more funding from everywhere simply due to good performance. The inclusion of country fixed effects, γ_c , alleviates the concern that results are driven by any country-specific factors such as geographic proximity.

Main effects

For the first two specifications in Table 3.7, I regress the number of VC investments received from a particular country on the number of approved same-origin skilled immigrant workers, and add a quadratic term subsequently to explore any non-linear relationship. Column (1) tells us that an additional skilled immigrant worker on average brings in 0.03 more same-origin VC deals, but this effect is found to diminish with higher number of same-origin skilled immigrant workers, as the quadratic term in Column (2) carries a negative coefficient.

From Column (3) onwards, the regressor of interest is an indicator variable that equals one when firm i has any skilled immigrant workers from country c approved in FY t and zero otherwise. The estimated coefficient from Table 3.7 Column (3) shows that on average, having skilled immigrants from a given country results in 0.08 more VC deals from this country. This translates to an additional inflow of over \$1 million from this source country to a U.S. start-up. The effect remains statistically significant in the cases of both new and follow-on VC deals, as shown by Column (4) and (5), proving

¹⁹Skilled immigrant workers are defined as those with H-1B, L, O, TN, E3, and H-1B1 visa status. Like the H-1B program, the L, O, TN, E3, and H1B1 programs require applicants to possess specialized skills.

²⁰Similar to H-1B visas, L, O, and TN visas are valid for an initial stay of 3 years upon issuance. E3 and H-1B1 visas are valid for shorter period of time but can be renewed indefinitely.

that skilled immigrant workers play a role in both attracting new investments and facilitating ongoing VC projects.

I also calculate the probability of receiving VC funding from a particular country as an outcome variable, and the results are shown in Column (6). The estimated coefficient is 0.03 and is statistically significant, meaning that having skilled workers from an origin on average increases the likelihood of receiving same-origin VC funding by 3 percentage points. This effect is not trivial when we consider the fact that the average probability of receiving deals from a particular foreign country is merely at 1.5%, conditional on the firm having received at least one VC investment from some country (including the U.S.) in a given year.

Mechanism and heterogeneous effects

I then explore the underlying mechanism of skilled immigrant works improving same-origin VC investments, as well as heterogeneous effects across firms and VC funds. Researchers have agreed that skilled immigrants contribute to general human capital, level of innovation, and thus overall performance, making their employers more attractive to potential investors. On top of this general human capital story, I would like to examine if there are any country-specific factors driving the mechanism.

I calculate a given country's share of all foreign VC investments and its share of all VC investments received by a firm in a year as the dependent variables. Table 3.8 Column (1) & (2) show that skilled immigrants from a given country are on average associated with 14.3 percentage points increase in this country's share of foreign VC funding, and 2.5 percentage points increase in this country's share of total VC deals, for a given firm-year. The second estimated effect is substantially lower possibly due to the fact that U.S. funds remain the dominant players in the U.S. VC scene and all foreign VC deals together only account for less than 20% of all deals received by U.S. firms. The use of shares rules out the possibility that the positive effects of skilled immigration are merely driven by firms attracting more VC deals from everywhere with good performance. The coefficient estimates confirms the presence of country-specific factors in how skilled immigration impacts same-origin VC funding.

In Column (3), I look at the group of skilled immigrant workers converting from F1 student

status. These workers recently graduated from U.S. institutions and likely have limited professional network. As compared to the baseline effect in Column (2), the effect is much smaller in the cases of these fresh-out-of-school immigrant workers. On the other hand, skilled immigrant workers relocating directly from their home countries with no prior U.S. status exert a much greater influence on same-origin VC investments, as shown in Column (4). This group likely still have strong ties with their home countries and professional connections back home. The results from these two columns together suggest that home-country professional networks play a key role in driving the country-specific effect of skilled immigration on VC financing outcomes. In particular, as shown in Column (5), newcomers to the U.S. are particularly effective in facilitating new VC investments from their home countries, as compared to the baseline results in Table 3.7 Column (4).

Home-country professional network is not the only factor driving the mechanism, as skilled immigrants who have stayed in the U.S. for some time still improve their employers' VC outcomes from their home countries. In Column (6), by including an interaction term with an indicator variable of whether the country is English-speaking²¹, I find that having skilled immigrant workers from non-English speaking countries brings in 0.03 more same-origin VC deals than in the case of English-speaking countries. In Column (7), I interact the key independent variable with the indicator variable of whether the country has a distinct culture from the U.S.²² The estimated coefficient in Column (7) shows that skilled immigrant workers from countries with distinct cultures from the U.S. exert greater influence on same-origin VC financing (0.03 more deals) as compared to those from countries that are culturally similar to the U.S. These two columns suggest that skilled immigrants improve firms' access to same-origin VC through common language and culture.

It would be also interesting to assess the heterogeneity in the effect magnitudes across firms and VC funds. In Table 3.9 Column (1), I examine how the effect varies with the experience of the VC funds, which is proxied by the age of the VC funds in months at the time of financing. Since observations are aggregated at firm-origin-year level and multiple funds may invest in the same year, average values are used. It turns out that the effect diminishes for more experienced VC funds, as

²¹Out of the top investor countries included in the analysis, Canada, Australia, Singapore, and the U.K. are considered English-speaking, and Germany, Israel, Japan, China, France, and Switzerland are considered non-English speaking.

²²Japan, China, Singapore, and Israel are considered as having distinct cultures from the U.S. Canada, Australia, Germany, France, Switzerland, and the U.K are considered culturally similar to the U.S.

shown by the negative and statistically significant coefficient on the interaction term. This is in line with the expectation that experienced VC funds are unlikely to rely heavily on personal connections to source and make investment decisions. In Column (2), I include an interaction term with the VC fund size and find no differential effect across VC fund sizes. This is no surprise as there is no apparent relationship between fund size and fund experience. It might also be partly due to the fact that the VentureXpert database only reports fund size for a small fraction of funds.

In Table 3.9 Column (3) to Column (5), I analyze the heterogeneous effects across firms of different age. I first include an interaction term with firm age in months at the time of financing and as Column (3) suggests, skilled immigrants have a smaller impact on same-origin VC outcomes when the firms that they work for are older. This aligns with the intuition that with more years in business and a record of past performance, start-ups would rely less on its employee's personal connections to source for investments. From a comparison of the coefficients in Column (4) and (5), it can be inferred that the effect is particularly strong for young firms in their first 3 years of business, and effect magnitude is more than halved for mature firms with more than 10 years in business.

Robustness checks

Several additional specifications are used to ensure the robustness of results. By regressing the probability of getting U.S. VC deals at the firm-year level on the indicator variable of having immigrant workers, I show that having immigrant workers do not jeopardize a firm's chance of attracting U.S. investments, as suggested by Table 3.10 Column (1).

An alternative set of fixed effects, firm-country and year fixed effects are used to ensure the bias will not arise from any firm-country-specific or macroeconomic characteristics. For instance, some firms might have existing connections with a particular country through subsidiaries or affiliates, and as a result they are more inclined to hire workers and more likely to receive investments from this country. In Table 3.10 Column (2), the effect of skilled immigration on same-origin VC funding survives different specification of fixed effects.

Some may argue that these skilled immigrants bring in innovation or increase the firms' general human capital and in turn these firms are more favored by VC funds in general. Although the

shares analysis already address this concern to some extent, I also include the number of U.S. deals received to proxy for the firms' general capability or level of innovation in Column (3), and the results hold.

As count data is used in this specification, Poisson model is also used to improve fit and the results are included in Column (4). Although the interpretation is difficult with a poisson model, the coefficient remains positive and statistically significant.

Table 3.10 Column (5) presents the results when the analysis is limited to a shorter and more recent period of FY 2010 to FY 2014, to avoid the influence of the Great Recession. Column (6) on the other hand, includes the results when all countries are included in the analysis, instead of including the top investor countries only. The estimated coefficients are positive and statistically significant in both settings.

Some may also worry that this effect of skilled immigration is simply shifting VC investments from one firm to another. Column (7) looks at the aggregate levels of skilled immigration from a particular country and same-origin VC financing and proves otherwise. The whole pie is bigger with more skilled immigrants from a particular country.

The last robustness check is to rule out the possibility that the effect observed is driven by some individual country. Figure 3.4 plots the estimated coefficients and standard errors from a series of regressions where I drop individual source countries, one at a time. All estimates are statistically significant and largely comparable to that of the main specification. An interesting finding is that the estimated coefficient becomes smaller when Israel is removed from the analysis, suggesting that Israeli VC funds might be particularly more inclined to invest in U.S. firms with Israeli workers. On the other hand, the coefficient is larger when India is dropped, which is in line with the fact that India is a major H1B sending country but accounts for very minimal VC activities in the U.S..

3.5 Proposed Framework

From the empirical analysis, I find evidence that high-skilled immigrant workers exert a positive impact on same-origin VC investments received by their employers. This highlights another mechanism in which immigration influences native's labor market outcomes. According to theory, when

immigration flows in, natives are potentially displaced through an interaction of substitution and scale effects. Now with the new VC mechanism, immigration has an indirect scale effect, in which the firm expands production with higher VC financing from immigrant workers' home countries and in turn demands more labor. In this section, I would like to propose a framework to capture the aforementioned mechanisms. The effect of receiving additional VC investments can be thought as easing the firm's financing constraints, which affects a firm's hiring decisions.

3.5.1 Setup

In this framework, a firm lives infinitely and produces output using a two-level nested aggregated production of native labor N , immigrant labor M , and capital K . Similar to the setup in Ottaviano & Peri (2012) and Ozden & Wagner (2014), the first level is Cobb-Douglas and the second level features constant elasticity of substitution (CES):

$$Y_t = F(K_t, L_t) = K_t^\alpha L_t^{1-\alpha} \quad (3.5)$$

$$L_t = (\rho_N N_t^\gamma + \rho_M M_t^\gamma)^{\frac{1}{\gamma}} \quad (3.6)$$

where $\rho_N + \rho_M = 1$ and capital accumulates according to the following law of motion:

$$K_{t+1} = I_t + (1 - \delta)K_t. \quad (3.7)$$

There is only one source country for immigration. Similar to the setup in Borjas (2003), immigrants and natives are imperfect substitutes but individuals within groups are assumed to be perfect substitutes, with σ_{nm} being the elasticity of substitution between native and immigrant labor, where $\sigma_{nm} = \frac{1}{1-\gamma}$.

A typical start-up usually faces challenges in securing bank financing due to the lack of traditional collateral and track record (Metrick & Yasuda, 2010). I thus assume the firm issues no inter-temporal debt and capital investment is financed fully out of dividends:

$$\Pi_t = Y_t - w_t L_t - I_t. \quad (3.8)$$

Next, I introduce a working capital constraint, where the firm must pay labor prior to production. Since start-ups lack access to traditional financing means, the firm finances this labor payment with VC funding. For simplicity, I assume venture capitalists earn no interest for supporting start-ups with intra-temporal financing, similar to the case of interest-free intraperiod loan. This is in line with the common observation that venture capitalists mainly profit from successful exits, not from firm operations prior to going public. Venture capitalists value the firm at its current capital stock and invest Q_t , which is expressed as a fraction of the capital stock. This can be viewed as a costly enforcement setup, where the firm can only borrow up to a fraction of its capital stock (Kiyotaki & Moore, 1997):

$$w_t L_t \leq \xi_t K_t = Q_t, \quad (3.9)$$

where the financing limit $\xi_t < 1$. From empirical evidence, skilled immigration leads to higher VC funding from the source country. That is, Q_t and ξ_t are strictly increasing in M_t :

$$\xi_t = G(M_t), \text{ where } G_M > 0. \quad (3.10)$$

To add the labor supply side of the story, I borrow from the standard RBC literature and define the household problem as

$$\max \sum_{t=0}^{\infty} \beta^t \left\{ \ln C_t - \theta \frac{L_t^{1+\chi}}{1+\chi} \right\}, \quad (3.11)$$

subject to $C_t \leq w_t L_t + \Pi_t$.

3.5.2 Steady state

The household's Lagrangian equation can be expressed as below:

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \left\{ \ln C_t - \theta \frac{L_t^{1+\chi}}{1+\chi} + \lambda_t (w_t L_t + \Pi_t - C_t) \right\}, \quad (3.12)$$

from which I derive the labor supply condition as follows:

$$\theta L^\chi = \frac{1}{C} w. \quad (3.13)$$

On the other hand, the firm's Lagrangian is:

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \{ K_t^\alpha L_t^{1-\alpha} - w_t L_t - K_{t+1} + (1-\delta)K_t + \mu_t(\xi_t K_t - w_t L_t) \}, \quad (3.14)$$

and by solving the first-order conditions, I get two steady state conditions:

$$w = \frac{1}{1+\mu}(1-\alpha) \left(\frac{K}{L} \right)^\alpha, \quad (3.15)$$

$$\frac{K}{L} = \left(\frac{\alpha}{\frac{1}{\beta} - (1-\delta) - \mu\xi} \right)^{\frac{1}{1-\alpha}}. \quad (3.16)$$

The multiplier, μ , is either 0 (non-binding) or positive (binding). Note that from Equation (3.16), if the constraint binds, the capital-labor ratio will be larger than in the standard case without any financial friction. So a more binding constraint will tend to reduce L , which drives up the capital-labor ratio. I am only interested in the case where the constraint binds. Solving for the multiplier, μ , I get:

$$\mu = \frac{1-\alpha}{\xi} \left(\frac{1}{\beta} - (1-\delta) \right) - \alpha. \quad (3.17)$$

As suggested by the equation, the multiplier is decreasing in ξ . The higher the financing limit, the less binding the constraint. Then the threshold value of ξ for which the constraint is binding ($\mu > 0$) is thus given by:

$$\xi < \frac{1-\alpha}{\alpha} \left(\frac{1}{\beta} - (1-\delta) \right). \quad (3.18)$$

Finally, I combine the labor supply (Equation 3.13) and labor demand (Equation 3.15) conditions evaluated in the steady state to solve for L :

$$L = \left[\frac{1}{\theta} (1+\mu)^{-1} \frac{(1-\alpha) \left(\frac{K}{L} \right)^\alpha}{\left(\frac{K}{L} \right)^\alpha - \delta \left(\frac{K}{L} \right)} \right]^{\frac{1}{1+\alpha}}. \quad (3.19)$$

According to Equation (3.19), ξ affect L through a direct effect and an indirect effect. A higher ξ makes μ smaller and in turn makes $(1+\mu)^{-1}$ bigger, which increases L . A higher ξ also makes $\frac{K}{L}$ bigger, which works in the same direction. Therefore, L is increasing in ξ .

3.5.3 Response to immigration shocks

Next I introduce an exogenous inflow of immigrants from the source country, which results in two labor market effects and one financial market effect. On the labor market side, there is first a substitution effect, where for a given level of output, firms will substitute immigrant workers for native workers. The second labor market effect is a scale effect, where the decline in the cost of production results in output expansion and hence, for a given relative wage, firms will employ more native workers. Finally, the firm receives more VC funding from the source country due to the hiring of immigrant workers and the financing limit ξ is higher.

The timeline of the events is specified as follows. First, the firm faces the immigration shock and makes the initial decision on the number of natives to hire. Then the firm observes the shock on the financing constraint, which is induced by the immigration inflow. With a more relaxed financing constraint, the firm will expand production and hire more. To avoid the issue of endogeneity, I assume that the only intraperiod source for additional labor is the native workforce. This is a reasonable assumption as immigrant workers often require work visas, which are usually associated with long processing times.

Using the first-order conditions with respect to N and M from the firm's Lagrangian and given $\sigma_{nm} = \frac{1}{1-\gamma}$, I get:

$$\ln w_{m,t} + \ln(1 + \mu_t) = \ln w_t + \ln \rho_m + \frac{1}{\sigma_{nm}} \ln L_t - \frac{1}{\sigma_{nm}} \ln M_t, \quad (3.20)$$

$$\ln w_{n,t} + \ln(1 + \mu_t) = \ln w_t + \ln \rho_n + \frac{1}{\sigma_{nm}} \ln L_t - \frac{1}{\sigma_{nm}} \ln N_t, \quad (3.21)$$

where w_n and w_m are the wages of natives and immigrants respectively. With constant returns to scale, I express their wage bill shares, s_n and s_m respectively, as below:

$$s_{m,t} = \frac{w_{m,t}M_t}{w_tL_t} = \frac{F_M M_t}{L_t}, \quad (3.22)$$

$$s_{n,t} = \frac{w_{n,t}N_t}{w_tL_t} = \frac{F_N N_t}{L_t}. \quad (3.23)$$

I define $\eta = -\frac{d \ln L}{d \ln w}$ as the elasticity of labor demand and $\phi_n = \frac{d \ln N}{d \ln w_n}$ as the elasticity of native

labor supply. After algebra, I have the effect of an immigration shock on native wage:

$$\frac{d \ln w_{n,t}}{d \ln M_t} = \frac{s_{m,t}(\eta - \sigma_{mn})}{\sigma_{mn}\eta + \phi_n(s_{m,t}\eta + s_{n,t}\sigma_{mn})}, \quad (3.24)$$

and the effect of an immigration shock on native employment is given by:

$$\frac{d \ln N_t}{d \ln M_t} = \frac{d \ln N_t}{d \ln w_{n,t}} \frac{d \ln w_{n,t}}{d \ln M_t} = \phi_n \frac{s_{m,t}(\eta - \sigma_{mn})}{\sigma_{mn}\eta + \phi_n(s_{m,t}\eta + s_{n,t}\sigma_{mn})}. \quad (3.25)$$

I then calculate the the number of natives potentially displaced by the immigration shock, absent the financial market effect of immigration:

$$\left. \frac{dN_t}{dM_t} \right|_{d\xi=0} = \frac{d \ln N_t}{d \ln M_t} \frac{N_t}{M_t} = \frac{\phi_n s_{m,t}(\eta - \sigma_{mn})}{\sigma_{mn}\eta + \phi_n(s_{m,t}\eta + s_{n,t}\sigma_{mn})} \frac{N_t}{M_t}. \quad (3.26)$$

Next I consider the effect when the financing limit has been adjusted. The firm observes the new financing limit and optimizes to reach a new steady state. As natives are the only source for additional labor at this stage, dL evaluated in steady state is essentially dN :

$$\frac{dN}{dM} = \frac{dL}{dM} = \frac{dL}{d\xi} \frac{d\xi}{dM} \quad (3.27)$$

where $\frac{dL}{d\xi}$ can be obtained from Equation (3.19) and $\frac{d\xi}{dM}$ can be obtained empirically.

By calibrating and comparing the magnitudes of the effects spelled out in Equation (3.26) and (3.27), we can use this framework to examine the resultant effects of immigration on the firm's hiring of natives, both through the labor market as well as the VC mechanism.

3.5.4 Numerical exercise

First, I take parameter values from literature to gauge the labor market effects of skilled immigration. The wage bill shares of native and immigrant workers are approximately at 0.85 and 0.15, using data from American Community Surveys.²³ According to a research review on labor supply elasticities by McClelland & Mok (2012), the combined elasticities of labor supply in the U.S. range from 0.1

²³Source: Economic Policy Institute analysis of American Community Survey (2009-2011)

to 0.3. For this exercise, I take the midpoint of 0.2 as the parameter value for ϕ_n , the elasticity of native labor supply. For the elasticity of labor demand, Hamermesh (1996) concludes that the value is bracketed by $[-0.75;-0.15]$, with his best guess being -0.30. A more recent meta-regression analysis of labor demand (Lichter et al, 2015) produces a rather close point estimate of the labor demand elasticity at -0.246, which I will take as the parameter value for η in this exercise.

There is expansive literature exploring the effect of immigration on U.S. labor markets and therefore the elasticity of substitution between native and immigrant workers (see, e.g., Card, 2001; Ottaviano and Peri, 2012; Borjas, Grogger, and Hanson, 2012; Clemens, Lewis, and Postel, 2018; Piyapromdee, 2021). The estimated elasticities ranges from 1.2 to 500 and my preferred estimate is from Piyapromdee (2021), which is at 6.93 for high-skilled immigrants.²⁴ Evaluating Equation (3.26) using the various parameter values, the estimated displacement effect of the representative firm hiring high-skilled immigrant workers is 0.297.

Taking reference from the RBC literature, I first assume the following parameter values for the evaluation of the steady state: $\beta = 0.99$, $\alpha = 1/3$, $\delta = 0.02$, $\theta = 7.71$, and $\chi = 1$. With these parameter values, the resultant cutoff value of ξ where the constraint binds is 0.0602. From the empirical analysis, the average probability of a U.S. firm getting any VC funding from a country is 0.0083. This can be viewed as the borrowing limit. That is, $\xi = 0.0083$, which means that in the steady state the constraint binds. From the first empirical approach, high-skilled labor from a given country leads to 0.0595 increase in the firm's chance of receiving same-origin VC investments, making ξ now at 0.0678 and the constraint non-binding. In this case, $\mu = 0$, and the increase in L through the VC mechanism of hiring immigrant workers from a source country is 0.0917. From the second approach, when a firm hires high-skilled immigrants from a particular country, the probability of receiving same-origin VC funding increases by 0.0319, which means ξ is increased to 0.0402 and the borrowing constraint still binds. The change in L after hiring of immigrant workers from the source country is therefore 0.0804 in this case.

This suggests that almost one-third of the displacement effect of high-skilled immigrants can be mitigated by the financial market mechanism. In addition, this is likely an underestimation

²⁴Allen et al (2018) also takes the estimated elasticity of substitution between native and immigrant workers from Piyapromdee (2021) as the preferred parameter value.

as in reality, venture capitalists not only support the portfolio firms financially, but also provide them with mentoring and network resources, which the proposed framework fails to capture. The parameter values used in evaluating the financial impact are also likely an understatement of the true effects, as discussed in Section 4.

3.6 Conclusion

In this paper, I examine the impact of skilled immigration on U.S. firm's VC funding outcomes from the immigrants' source countries. I first exploit the exogenous variation in firms' access to skilled labor from different source countries. This approach makes use of the random lottery feature of the H-1B program, and two methods are deployed to handle missing nationality information for unselected H-1B applicants. I then supplement in terms of external validity using a panel study with extensive fixed effects. I also propose a framework of firms' hiring decisions to incorporate both the labor market and VC channels. High-skilled immigrants might displace their native counterparts through the interaction of substitution and scale effects, but they also potentially enhance native labor market outcomes through an indirect scale effect, where the firms expand production and hire more natives with the additional VC funding made possible by immigrants.

I find that having high-skilled immigrant workers from a given source country is associated with better access to VC funding from the same source country along various dimensions. High-skilled immigrants from a given country are associated with increase of 0.03 to 0.08 on average in the number of deals from this country. Factoring in the average size of VC investments, these estimates suggest an effect of \$405,000 to \$1,039,000 more funding from this source country for a single start-up, which can be seen as substantial. Through home-country professional networks, common language, and cultural familiarity, skilled immigrants reduce investment frictions and information gaps and better facilitate cross-border venture capital investments from their origins, especially in the case of young firms or inexperienced venture capitalists. With a numerical exercise, I find that almost one-third of the displacement effect of high-skilled immigrants can be mitigated by the financial market mechanism, and this is likely an underestimation.

3.7 Appendix

3.7.1 VentureXpert from Thomson ONE

I extract all venture capital investments on U.S. firms for the period from 2004 to 2017 from VentureXpert data, available via the Thomson ONE database. I only keep transactions with reported origins of investments. This cause me to drop 14,313 observations, which make up approximately 8% of the extracted dataset.

3.7.2 Form I-129 data from USCIS

I obtain all approved H-1B petitions filed by U.S. firms for the period from Fiscal Year (FY) 2004 to FY 2014 from USCIS. I need at least both the firm name and the state to identify a unique firm. I thus drop 697 observations with missing firm names and 220 observations with missing state information.

3.7.3 LCA data from Department of Labor

I obtain all LCAs filed with the Department of Labor for the period of FY 2009-2018. An LCA must be filed and certified prior to the submission of an H-1B petition. I remove all non-certified LCAs, as they are not linked to any actual H-1B petitions. This cause me to drop 630,147 observations that are either withdrawn or denied LCAs.

3.7.4 Construction

Using unique firm names and state information, I aggregate observations in the VentureXpert dataset to generate the number of venture capital deals at the firm-year-country level. These unique firms in the VentureXpert dataset are the start-ups for analysis in this paper. However, both the USCIS data and LCA data suffer from spelling variations in firm names and addresses, which present challenges in directly generating firm-year-country H-1B applied and approved applications.

I use a supervised machine learning algorithm²⁵ to cluster observations that refer to the same

²⁵Gregg, Forest and Derek Eder. 2015. Dedupe. <https://github.com/dedupeio/dedupe>.

firm in the USCIS and LCA datasets respectively. This algorithm combines the string metric of Affine Gap Distance with active learning, which counts the number of substitutions that must be made to turn one string into another and relearns the weights and blocking rules through user labeling. I then use a similar algorithm to match pairwise records across datasets to produce a panel dataset covering firm-year-country investments, H-1B applications, and H-1B approvals for the period from FY2005 to FY 2014.

To validate the effectiveness of this fuzzy-matching algorithm, I obtain another dataset from USCIS, the H-1B Employer Data for FY 2009 to FY 2014.²⁶ This dataset contains all H-1B petitions processed by USCIS with firm-level information such as firm names, state, zip codes, and the last four digits of firms' tax identification numbers. I match the records to the fuzzy-matching output via machine learning based on firm names and state information, and verify if the firm names clustered together by the algorithm have the same last four digits of tax identification numbers. The algorithm is proven to be reliable for almost 85% of the cases. The remaining 15% of the cases are often large corporations that have established several legal entities, thus having different tax identification numbers. For such cases, I take observations to be referring to the same firm if they share the exact same address. It turns out that the algorithm has a success rate of 95%. The actual success rate should be even higher, considering spelling variations in addresses.

²⁶The H-1B Employer Data is available for FY 2009 onwards at <https://www.uscis.gov/tools/reports-and-studies/h-1b-employer-data-hub/h-1b-employer-data-hub-files>.

Bibliography

- Acs, Z. J. & Szerb, L. (2007), ‘Entrepreneurship, economic growth and public policy’, *Small business economics* **28**(2), 109–122.
- Aizenman, J. & Kendall, J. (2012), ‘The internationalization of venture capital’, *Journal of Economic Studies* .
- Akcigit, U., Grigsby, J. & Nicholas, T. (2017), ‘Immigration and the rise of american ingenuity’, *American Economic Review* **107**(5), 327–31.
- Allen, T., de Castro Dobbin, C. & Morten, M. (2018), Border walls, Technical report, National Bureau of Economic Research.
- Alvarez-Garrido, E. & Guler, I. (2018), ‘Status in a strange land? context-dependent value of status in cross-border venture capital’, *Strategic Management Journal* **39**(7), 1887–1911.
- Amit, R., Brander, J. & Zott, C. (1998), ‘Why do venture capital firms exist? theory and canadian evidence’, *Journal of business Venturing* **13**(6), 441–466.
- Amornsiripanitch, N., Gompers, P. A., Hu, G. & Vasudevan, K. (2021), Getting schooled: The role of universities in attracting immigrant entrepreneurs, Technical report, National Bureau of Economic Research.
- Balachandran, S. & Hernandez, E. (2019), ‘Mi casa es tu casa: immigrant entrepreneurs as pathways to foreign venture capital investments’, *Strategic Management Journal* .
- Baron, J. N. & Hannan, M. T. (2002), ‘Organizational blueprints for success in high-tech startups: Lessons from the stanford project on emerging companies’, *California Management Review* **44**(3), 8–36.
- Bergemann, D. & Hege, U. (1998), ‘Venture capital financing, moral hazard, and learning’, *Journal of Banking & Finance* **22**(6-8), 703–735.
- Bernstein, S., Diamond, R., McQuade, T., Pousada, B. et al. (2018), ‘The contribution of high-skilled immigrants to innovation in the united states’, *Stanford Graduate School of Business Working Paper* **3748**, 202019–20.
- Bernstein, S., Giroud, X. & Townsend, R. R. (2016), ‘The impact of venture capital monitoring’, *The Journal of Finance* **71**(4), 1591–1622.
- Borjas, G. J. (2003), ‘The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market’, *The quarterly journal of economics* **118**(4), 1335–1374.
- Borjas, G. J., Grogger, J. & Hanson, G. H. (2012), ‘Comment: On estimating elasticities of substitution’, *Journal of the European Economic Association* **10**(1), 198–210.
- Bradley, W. A., Duruflé, G., Hellmann, T. F. & Wilson, K. E. (2019), ‘Cross-border venture capital investments: what is the role of public policy?’, *Journal of Risk and Financial Management* **12**(3), 112.
- Burchardi, K. B., Chaney, T. & Hassan, T. A. (2019), ‘Migrants, ancestors, and foreign investments’, *The Review of Economic Studies* **86**(4), 1448–1486.
- Card, D. (2001), ‘Immigrant inflows, native outflows, and the local labor market impacts of higher immigration’, *Journal of Labor Economics* **19**(1), 22–64.

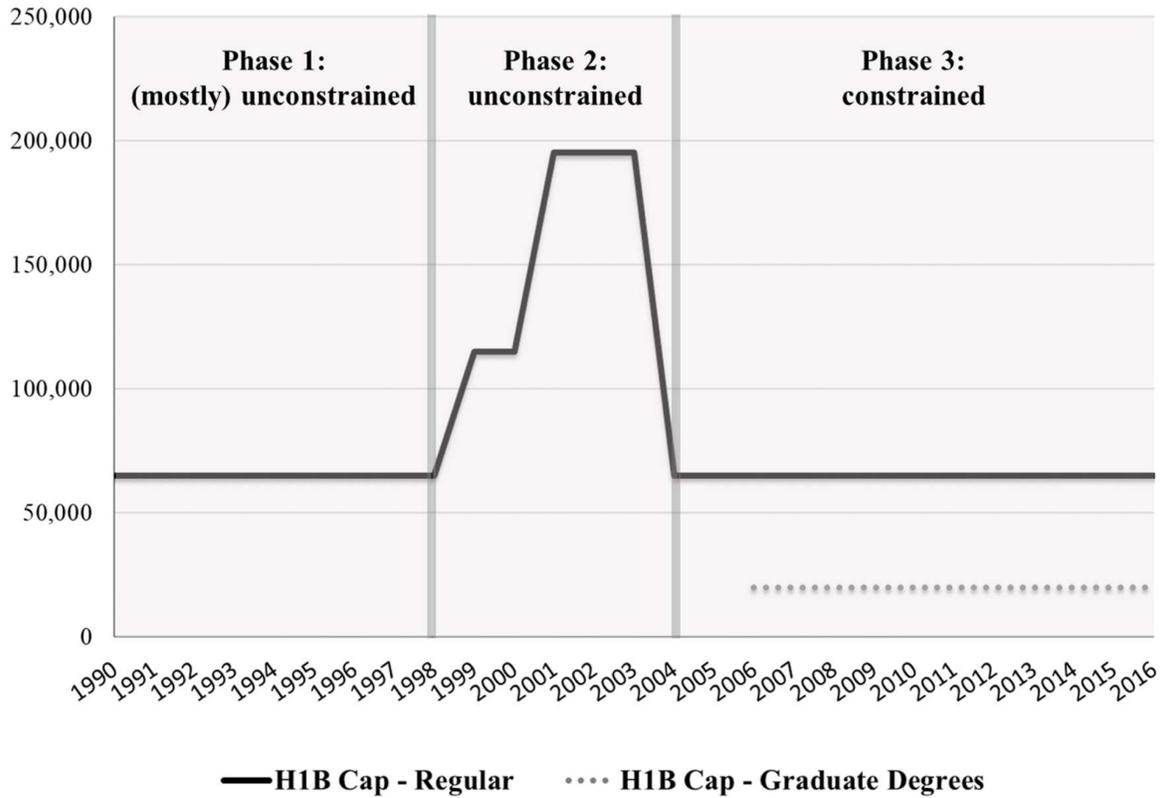
- Chemmanur, T. & Chen, Z. (2003), ‘Angels, venture capitalists, and entrepreneurs: a dynamic model of private equity financing’, *Unpublished working paper, Boston College, Boston, MA* .
- Chen, H., Gompers, P., Kovner, A. & Lerner, J. (2010), ‘Buy local? the geography of venture capital’, *Journal of Urban Economics* **67**(1), 90–102.
- Choudhury, P. & Kim, D. Y. (2019), ‘The ethnic migrant inventor effect: Codification and recombination of knowledge across borders’, *Strategic Management Journal* **40**(2), 203–229.
- Claes, K. & Vissa, B. (2020), ‘Does social similarity pay off? homophily and venture capitalists’ deal valuation, downside risk protection, and financial returns in india’, *Organization Science* **31**(3), 576–603.
- Clemens, M. A. (2013), ‘Why do programmers earn more in houston than hyderabad? evidence from randomized processing of us visas’, *American Economic Review* **103**(3), 198–202.
- Clemens, M. A., Lewis, E. G. & Postel, H. M. (2018), ‘Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion’, *American Economic Review* **108**(6), 1468–87.
- Coleman, J. S. (1988), ‘Social capital in the creation of human capital’, *American journal of sociology* **94**, S95–S120.
- Colombo, M. G. & Grilli, L. (2010), ‘On growth drivers of high-tech start-ups: Exploring the role of founders’ human capital and venture capital’, *Journal of business venturing* **25**(6), 610–626.
- De Bettignies, J.-E. & Brander, J. A. (2007), ‘Financing entrepreneurship: Bank finance versus venture capital’, *Journal of Business Venturing* **22**(6), 808–832.
- Dimmock, S. G., Huang, J. & Weisbenner, S. J. (2019), Give me your tired, your poor, your high-skilled labor: H-1b lottery outcomes and entrepreneurial success, Technical report, National Bureau of Economic Research.
- Dimov, D. & Gedajlovic, E. (2010), ‘A property rights perspective on venture capital investment decisions’, *Journal of Management Studies* **47**(7), 1248–1271.
- Docquier, F. & Lodigiani, E. (2010), ‘Skilled migration and business networks’, *Open Economies Review* **21**(4), 565–588.
- Docquier, F. & Rapoport, H. (2012), ‘Globalization, brain drain, and development’, *Journal of economic literature* **50**(3), 681–730.
- Doran, K., Gelber, A. & Isen, A. (2014), The effects of high-skilled immigration policy on firms: Evidence from h-1b visa lotteries, Technical report, National Bureau of Economic Research.
- Evans, D. S. & Jovanovic, B. (1989), ‘An estimated model of entrepreneurial choice under liquidity constraints’, *Journal of political economy* **97**(4), 808–827.
- Ewens, M., Gorbenko, A. & Korteweg, A. (2021), ‘Venture capital contracts’, *Journal of Financial Economics* .
- Florin, J. (2005), ‘Is venture capital worth it? effects on firm performance and founder returns’, *Journal of Business Venturing* **20**(1), 113–135.
- Foley, C. F. & Kerr, W. R. (2013), ‘Ethnic innovation and us multinational firm activity’, *Management Science* **59**(7), 1529–1544.
- Glennon, B. (2020), How do restrictions on high-skilled immigration affect offshoring? evidence from the h-1b program, Technical report, National Bureau of Economic Research.
- Gompers, P. A., Gornall, W., Kaplan, S. N. & Strebulaev, I. A. (2020), ‘How do venture capitalists make decisions?’, *Journal of Financial Economics* **135**(1), 169–190.
- Gornall, W. & Strebulaev, I. A. (2020), ‘Squaring venture capital valuations with reality’, *Journal of Financial Economics* **135**(1), 120–143.
- Gould, D. M. (1994), ‘Immigrant links to the home country: empirical implications for us bilateral trade flows’, *The Review of Economics and Statistics* pp. 302–316.

- Guler, I. & Guillen, M. F. (2010), ‘Home country networks and foreign expansion: Evidence from the venture capital industry’, *Academy of Management Journal* **53**(2), 390–410.
- Hall, J. & Hofer, C. W. (1993), ‘Venture capitalists’ decision criteria in new venture evaluation’, *Journal of business venturing* **8**(1), 25–42.
- Hallen, B. L. (2008), ‘The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments?’, *Administrative Science Quarterly* **53**(4), 685–718.
- Hamermesh, D. S. (1996), *Labor demand*, Princeton University press.
- Hellmann, T. & Puri, M. (2000), ‘The interaction between product market and financing strategy: The role of venture capital’, *The review of financial studies* **13**(4), 959–984.
- Hellmann, T. & Puri, M. (2002), ‘Venture capital and the professionalization of start-up firms: Empirical evidence’, *The journal of finance* **57**(1), 169–197.
- Hernandez, E. (2014), ‘Finding a home away from home: Effects of immigrants on firms’ foreign location choice and performance’, *Administrative Science Quarterly* **59**(1), 73–108.
- Hochberg, Y. V., Ljungqvist, A. & Lu, Y. (2007), ‘Whom you know matters: Venture capital networks and investment performance’, *The Journal of Finance* **62**(1), 251–301.
- Hochberg, Y. V., Ljungqvist, A. & Lu, Y. (2010), ‘Networking as a barrier to entry and the competitive supply of venture capital’, *The Journal of Finance* **65**(3), 829–859.
- Honig, B., Drori, I. & Carmichael, B. (2010), *Transnational and immigrant entrepreneurship in a globalized world*, University of Toronto Press Toronto.
- Hunt, J. & Gauthier-Loiselle, M. (2010), ‘How much does immigration boost innovation?’, *American Economic Journal: Macroeconomics* **2**(2), 31–56.
- Iriyama, A., Li, Y. & Madhavan, R. (2010), ‘Spiky globalization of venture capital investments: The influence of prior human networks’, *Strategic Entrepreneurship Journal* **4**(2), 128–145.
- Kaplan, S. N. & Lerner, J. (2017), *Venture Capital Data*, Vol. 75, University of Chicago Press.
- Kaplan, S. N., Strömberg, P. & Sensoy, B. A. (2002), ‘How well do venture capital databases reflect actual investments?’, *Available at SSRN 939073*.
- Kerr, S. P. & Kerr, W. R. (2011), ‘Economic impacts of immigration: A survey’.
- Kerr, W. R. (2008), ‘Ethnic scientific communities and international technology diffusion’, *The Review of Economics and Statistics* **90**(3), 518–537.
- 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**(3), 473–508.
- Kiyotaki, N. & Moore, J. (1997), ‘Credit cycles’, *Journal of political economy* **105**(2), 211–248.
- Kugler, M. & Rapoport, H. (2005), ‘Skilled emigration, business networks and foreign direct investment’, *Business Networks and Foreign Direct Investment (April 2005)*.
- Kulchina, E. (2016), ‘A path to value creation for foreign entrepreneurs’, *Strategic Management Journal* **37**(7), 1240–1262.
- Leblang, D. (2010), ‘Familiarity breeds investment: Diaspora networks and international investment’, *American political science review* **104**(3), 584–600.
- Li, Y., Hernandez, E. & Gwon, S. (2019), ‘When do ethnic communities affect foreign location choice? dual entry strategies of Korean banks in China’, *Academy of Management Journal* **62**(1), 172–195.
- Lichter, A., Peichl, A. & Sieglöcher, S. (2015), ‘The own-wage elasticity of labor demand: A meta-regression analysis’, *European Economic Review* **80**, 94–119.
- Little, R. J. & Rubin, D. B. (2019), *Statistical analysis with missing data*, Vol. 793, John Wiley & Sons.
- Maats, F., Metrick, A., Yasuda, A., Hinkes, B. & Vershovski, S. (2011), ‘On the consistency and

- reliability of venture capital databases’, *Unpublished working paper* .
- Madhavan, R. & Iriyama, A. (2009), ‘Understanding global flows of venture capital: Human networks as the “carrier wave” of globalization’, *Journal of International Business Studies* **40**(8), 1241–1259.
- Matusik, S. F. & Fitza, M. A. (2012), ‘Diversification in the venture capital industry: leveraging knowledge under uncertainty’, *Strategic Management Journal* **33**(4), 407–426.
- McClelland, R. & Mok, S. (2012), ‘A review of recent research on labor supply elasticities’.
- Metrick, A. & Yasuda, A. (2010), ‘The economics of private equity funds’, *The Review of Financial Studies* **23**(6), 2303–2341.
- Nanda, R. & Khanna, T. (2010), ‘Diasporas and domestic entrepreneurs: Evidence from the indian software industry’, *Journal of Economics & Management Strategy* **19**(4), 991–1012.
- Nathan, M. (2014), ‘The wider economic impacts of high-skilled migrants: a survey of the literature for receiving countries’, *IZA Journal of Migration* **3**(1), 1–20.
- Ottaviano, G. I. & Peri, G. (2012), ‘Rethinking the effect of immigration on wages’, *Journal of the European economic association* **10**(1), 152–197.
- Ozden, C. & Wagner, M. C. (2014), ‘Immigrant versus natives? displacement and job creation’, *Displacement and Job Creation (June 1, 2014)*. *World Bank Policy Research Working Paper* (6900).
- Pandya, S. & Leblang, D. (2011), Deal or no deal: The growth of international venture capital investment, Technical report, Citeseer.
- Park, H. D. & Steensma, H. K. (2012), ‘When does corporate venture capital add value for new ventures?’, *Strategic Management Journal* **33**(1), 1–22.
- Peri, G., Shih, K. & Sparber, C. (2015), ‘Stem workers, h-1b visas, and productivity in us cities’, *Journal of Labor Economics* **33**(S1), S225–S255.
- Piyapromdee, S. (2021), ‘The impact of immigration on wages, internal migration, and welfare’, *The Review of Economic Studies* **88**(1), 406–453.
- Rangan, S. (2000), ‘The problem of search and deliberation in economic action: When social networks really matter’, *Academy of Management Review* **25**(4), 813–828.
- Reagans, R. & McEvily, B. (2003), ‘Network structure and knowledge transfer: The effects of cohesion and range’, *Administrative science quarterly* **48**(2), 240–267.
- Rubin, D. B. (2004), *Multiple imputation for nonresponse in surveys*, Vol. 81, John Wiley & Sons.
- Sahlman, W. A. (1990), ‘The structure and governance of venture-capital organizations’, *Journal of financial economics* **27**(2), 473–521.
- Sapienza, H. J. (1992), ‘When do venture capitalists add value?’, *Journal of business venturing* **7**(1), 9–27.
- Sørensen, M. (2007), ‘How smart is smart money? a two-sided matching model of venture capital’, *The Journal of Finance* **62**(6), 2725–2762.
- Sorenson, O. & Stuart, T. E. (2001), ‘Syndication networks and the spatial distribution of venture capital investments’, *American journal of sociology* **106**(6), 1546–1588.
- Sorenson, O. & Stuart, T. E. (2008), ‘Bringing the context back in: Settings and the search for syndicate partners in venture capital investment networks’, *Administrative Science Quarterly* **53**(2), 266–294.
- Tyebjee, T. T. & Bruno, A. V. (1984), ‘A model of venture capitalist investment activity’, *Management science* **30**(9), 1051–1066.
- Vanacker, T., Collewaert, V. & Paeleman, I. (2013), ‘The relationship between slack resources and the performance of entrepreneurial firms: The role of venture capital and angel investors’, *Journal of Management Studies* **50**(6), 1070–1096.

- Wang, D. (2015), 'Activating cross-border brokerage: Interorganizational knowledge transfer through skilled return migration', *Administrative Science Quarterly* **60**(1), 133–176.
- Wang, L. & Wang, S. (2012), 'Economic freedom and cross-border venture capital performance', *Journal of Empirical Finance* **19**(1), 26–50.
- White, I. R., Royston, P. & Wood, A. M. (2011), 'Multiple imputation using chained equations: issues and guidance for practice', *Statistics in medicine* **30**(4), 377–399.
- Zaheer, S. (1995), 'Overcoming the liability of foreignness', *Academy of Management journal* **38**(2), 341–363.
- Zhang, J., Soh, P.-h. & Wong, P.-k. (2010), 'Entrepreneurial resource acquisition through indirect ties: Compensatory effects of prior knowledge', *Journal of Management* **36**(2), 511–536.

Figure 3.1: Changes in H-1B Caps



Notes: This figure plots the cap on the number of H-1B visas by fiscal year. since the Immigration Act of 1990, there has been an annual cap on the number of new H-1B visas that can be issued to private sector businesses. This cap is set by Congress and the President. Throughout most of the 1990s, the cap was set at 65,000 visas and applications rarely outstripped supply. It was increased to 195,000 visas by the American Competitiveness and Workforce Improvement Act of 1998 and the American Competitiveness in the Twenty-First Century Act of 2000 (AC21). During this period, the cap limits were never reached. The AC21 stipulated that this reversion would happen in the absence of any additional legislation, but, despite a trend towards less restrictive labor laws, no legislation was enacted, and the cap level reverted back to 65,000. It was raised by 20,000 in 2006, but those additional 20,000 could only be used for applicants with a graduate degree. Since then, the cap has not changed, and it has been binding in every year. The identification in this paper exploits the sharp reduction in the annual H-1B cap in fiscal year 2004; in the early 2000s, the cap was not binding. After the cap change, it has been binding in every year. Source: Glennon (2018).

Figure 3.2: The average number of VC deals received by firms from a given source country

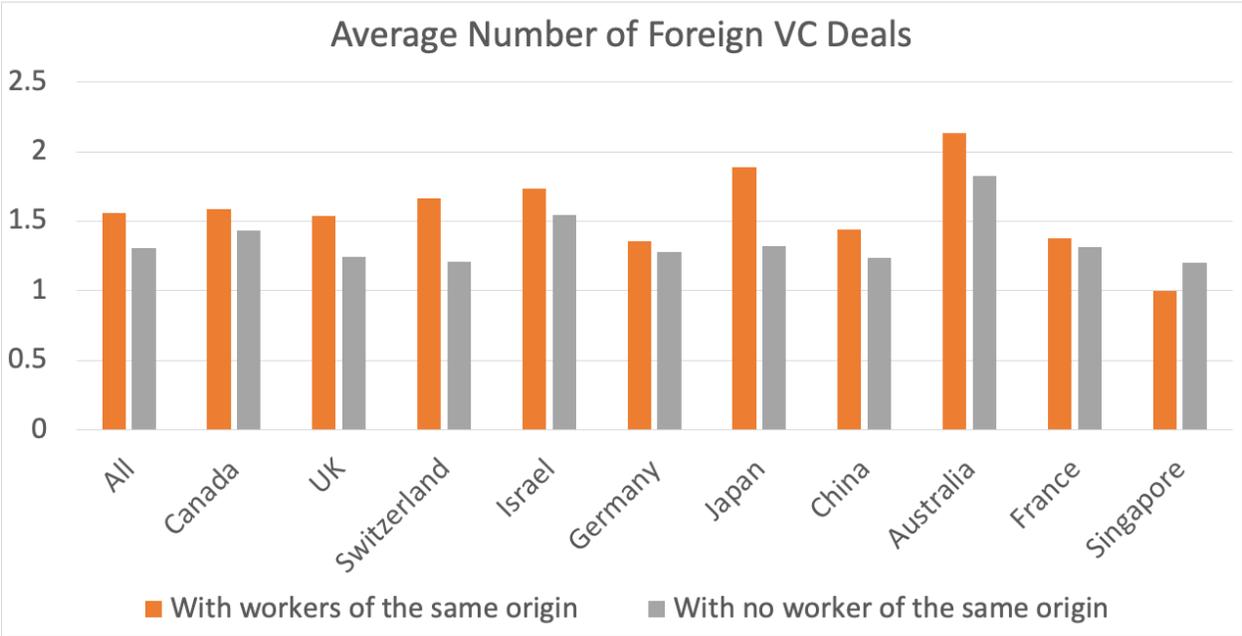


Figure 3.3: Estimated coefficients from LPM and logit specifications, and when individual countries are dropped in the win rate approach

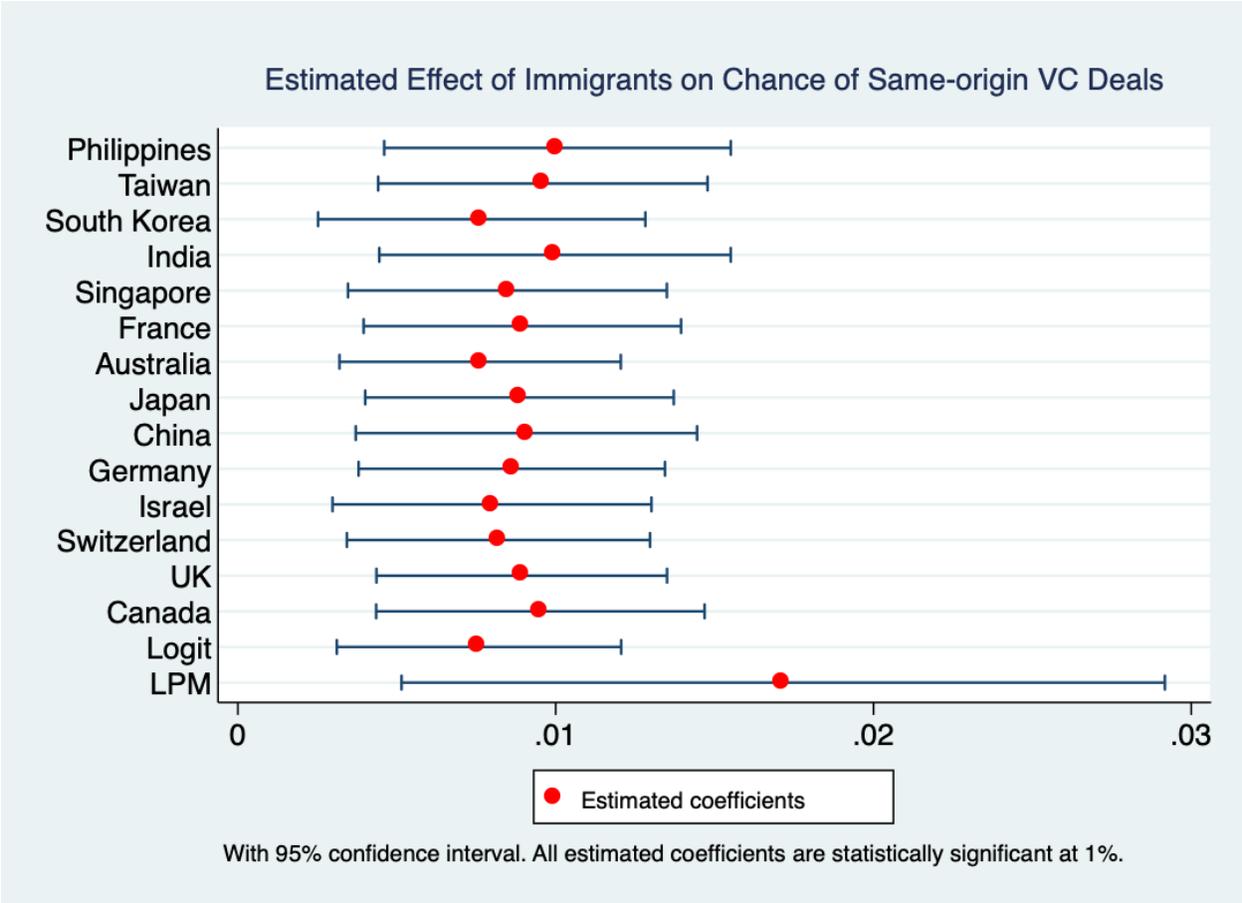


Figure 3.4: Estimated coefficients when individual countries are dropped in the FE approach

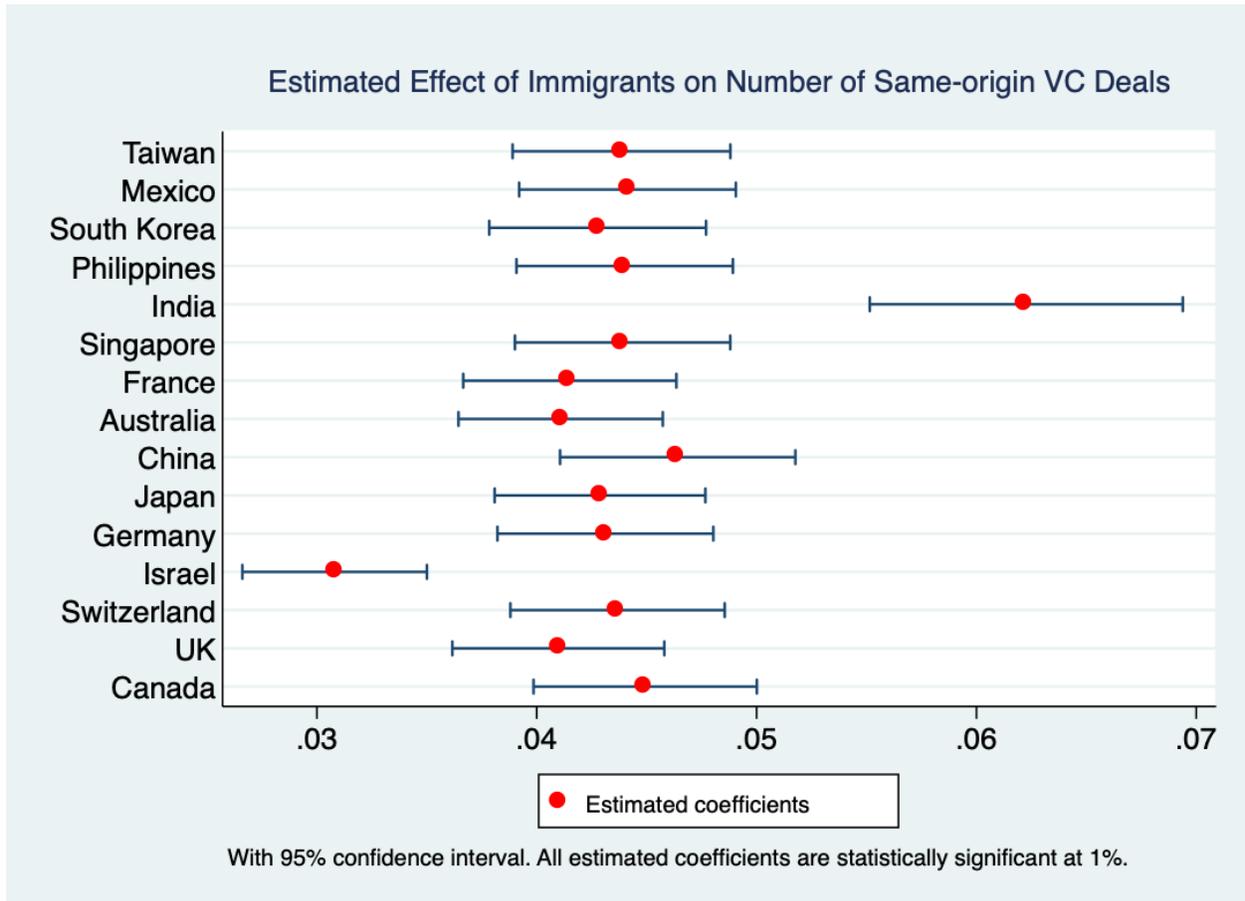


Table 3.1: Final Receipt Dates and H-1B Lottery Details

Fiscal Year	Final Receipt Date	Days in Filing Period	Lottery	Number of Lottery-subject Petitions Received
2004	February 17, 2004	323	No	N.A.
2005	October 1, 2004	184	No	N.A.
2006	August 10, 2005	132	No	N.A.
2007	May 26, 2006	56	No	N.A.
2008	April 3, 2007	3	Yes	150,000
2009	April 7, 2008	7	Yes	163,000
2010	December 21, 2009	265	No	N.A.
2011	January 26, 2011	301	No	N.A.
2012	November 22, 2011	236	No	N.A.
2013	June 11, 2012	72	No	N.A.
2014	April 7, 2013	7	Yes	124,000
2015	April 7, 2014	7	Yes	172,500
2016	April 7, 2015	7	Yes	233,000
2017	April 7, 2016	7	Yes	236,000
2018	April 7, 2017	5	Yes	199,000
2019	April 6, 2018	5	Yes	190,098

Notes: USCIS begins accepting H-1B petitions on the first business day of April every year for permits counting towards the next fiscal year. Permits are usually granted on a first-come-first-serve basis until the caps are reached, beyond which no petitions are accepted. The end of the filing period is demarcated by the "final receipt date", which is the date on which USCIS receive enough petitions to fill remaining quotas. In FY 2008, FY 2009, and from FY 2014 onwards, USCIS received so many petitions within the first week that a computerized random lottery was used to allocate the permits. All information in this table is gathered from the press releases of USCIS.

Table 3.2: VC Deal Volumes by U.S. and Non-U.S. Funds

Year	Deals by U.S. Funds	Growth Rate	Deals by Non-U.S. Funds	Growth Rate
2004	9,541	NA	1,244	NA
2005	10,011	4.9%	1,102	-11.4%
2006	11,451	14.4%	1,178	6.9%
2007	12,957	13.2%	1,317	11.8%
2008	12,814	-1.1%	1,331	1.1%
2009	8,264	-35.5%	811	-39.1%
2010	9,594	16.1%	912	12.5%
2011	10,991	14.6%	1,003	10.0%
2012	10,889	-0.9%	1,003	0.0%
2013	11,189	2.8%	1,019	1.6%
2014	11,762	5.1%	1,195	17.3%
2015	12,041	2.4%	1,388	16.2%
2016	9,135	-24.1%	1,385	-0.2%
2017	9,939	8.8%	1,565	13.0%

Source: VentureXpert. Growth rates are calculated on a year-on-year basis. The compound annual growth rate (CAGR) for the period from 2004 to 2017 for deals funded by U.S. venture capitalists is 0.3% while that for deals funded by non-U.S. venture capitalists is 1.8%. Taking the post Great Recession period from 2010 onwards, the CAGR for deals funded by U.S. venture capitalists is 0.5% while that for deals funded by non-U.S. venture capitalists is 8.0%.

Table 3.3: Top Investor Countries for VC Activities in the U.S.

Rank	Country	Number of Deals	Share of All Foreign Deals
1	Canada	2,179	13.2%
2	United Kingdom	1,949	11.9%
3	Switzerland	1,308	8.0%
4	Israel	1,265	7.7%
5	Germany	1,186	7.2%
6	Japan	979	6.0%
7	China	905	5.5%
8	Australia	869	5.3%
9	France	712	4.3%
10	Singapore	455	2.7%

Source: VentureXpert.

Table 3.4: Top Source Countries for H-1B Workers in the U.S.

Rank	Country	Number of H-1B Workers	Share of All H-1B workers
1	India	1,551,627	54.4%
2	China	241,262	8.5%
3	Canada	106,747	3.7%
4	Philippines	94,861	3.3%
5	South Korea	75,330	2.6%
6	United Kingdom	53,708	1.9%
7	Mexico	37,190	1.3%
8	Taiwan	36,809	1.3%
9	Pakistan	33,399	1.2%
10	France	27,983	1.0%

Source: I-129 forms from USCIS for FY 2004 to FY 2014.

Table 3.5: H-1B Lottery Win Rate and Foreign Venture Capital - Multiple Imputation

	Number of deals from the origin		Pr. of funding
	(1)	(2)	(3)
Win rate	0.0296*** (0.0068)	0.0680*** (0.0104)	0.0595*** (0.0090)
Total U.S. deals		0.0029 (0.0032)	-0.0005 (0.0014)
log(salary)		-0.0619*** (0.0143)	-0.0721*** (0.0116)
Number of LCAs	No	Yes	Yes
Firm age	No	Yes	Yes
Country FEs	Yes	Yes	Yes
Year FEs	No	Yes	Yes
N	3,180	3,180	3,180

Notes: The sample includes all U.S. firms that received at least one VC investment from 2004 to 2017 and submitted LCAs in FY09 or FY14. Observations are at firm-source country-fiscal year level. Outcome variables are measured over a period of three fiscal years following the approval of the H-1B application. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.6: H-1B Lottery Win Rate and Foreign Venture Capital - Dichotomous Win Rates

	Number of deals from the origin		Pr. of funding	Share of foreign deals	Share of total deals
	(1)	(2)	(3)	(4)	(5)
Win rate	0.0319*** (0.0130)	0.0315*** (0.0130)	0.0172*** (0.0061)	0.1098*** (0.0343)	0.0128*** (0.0055)
Total U.S. deals		0.0101*** (0.0028)	0.0045*** (0.0010)	0.0004 (0.0013)	-0.0005* (0.0003)
log(salary)		0.0223*** (0.0064)	0.0105*** (0.0033)	0.0121 (0.0106)	0.0059*** (0.0022)
Number of LCAs	No	Yes	Yes	Yes	Yes
Firm age	No	Yes	Yes	Yes	Yes
Country FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	No	Yes	Yes	Yes	Yes
N	9,001	9,001	9,001	1,168	4,350

Notes: The sample includes all U.S. firms that received at least one VC investment from 2004 to 2017 and submitted LCAs in FY09 or FY14. Observations are at firm-source country-fiscal year level. Outcome variables are measured over a period of three fiscal years following the approval of the H-1B application. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.7: Skilled Immigrants and Foreign Venture Capital - FE Main Effects

	Number of same-origin deals					Pr. of funding (6)
	All types			New (4)	Follow-on (5)	
	(1)	(2)	(3)			
Migrant	0.0330*** (0.0042)	0.0538*** (0.0048)				
Migrant2		-0.0017*** (0.0003)				
D(Migrant)			0.0818*** (0.0063)	0.0227*** (0.0027)	0.0591*** (0.0049)	0.0319*** (0.0022)
Firm-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Country FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	1.5 mil	1.5 mil	1.5 mil	1.5 mil	1.5 mil	1.5 mil

Notes: The sample includes all U.S. firms that received at least one VC investment from 2004 to 2017. Migrant is the number of same-origin skilled workers, Migrant2 is the quadratic term, and D(Migrant) is a dummy variable indicating the presence of any same-origin skilled workers. Observations are at firm-source country-fiscal year level. The period of analysis is from Fiscal Year (FY) 2005 to FY 2014. Outcome variables are measured over a period of three fiscal years following the approval of the visa application. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.8: Skilled Immigrants and Foreign Venture Capital - Mechanism

	Share of same-origin deals in				Number of same-origin deals		
	Foreign (1)	(2)	All deals (3)	(4)	New (5)	All types (6)	(7)
D(Migrant)	0.1426*** (0.0097)	0.0251*** (0.0019)				0.0613*** (0.0091)	0.0666*** (0.0079)
D(F1)			0.0125*** (0.0035)				
D(NewtoUS)				0.0553*** (0.0053)	0.0438*** (0.0061)		
D(NonEng)						-0.0008 (0.0007)	
D(Migrant) *D(NonEng)						0.0347*** (0.0119)	
D(Culture)							-0.0066*** (0.0006)
D(Migrant) *D(Culture)							0.0328*** (0.0121)
Firm-year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	140,000	690,990	690,990	690,990	1.5 mil	1.5 mil	1.5 mil

Notes: The sample includes all U.S. firms that received at least one VC investment from 2004 to 2017. D(Migrant) is a dummy variable indicating the presence of any same-origin skilled workers, D(F1) is a dummy variable indicating the presence of any same-origin skilled workers who were previously on F-1 student visas, and D(NewtoUS) is a dummy variable indicating the presence of any same-origin skilled workers who had no prior status in the U.S. Observations are at firm-source country-fiscal year level. The period of analysis is from Fiscal Year (FY) 2005 to FY 2014. Outcome variables are measured over a period of three fiscal years following the approval of the visa application. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.9: Skilled Immigrants and Foreign Venture Capital - Heterogeneous Effects

	Number of same-origin deals				
	(1)	(2)	(3)	Young (4)	Mature (5)
D(Migrant)	0.7636*** (0.2583)	1.2301*** (0.4449)	0.1063*** (0.0082)	0.1275*** (0.0177)	0.0348*** (0.0073)
Fund Age	-0.0003 (0.0002)				
D(Migrant) *Fund Age	-0.0016*** (0.0008)				
Fund size		0.0009* (0.0006)			
D(Migrant) *Fund Size		-0.0035 (0.0022)			
D(Migrant) *Firm age			-0.0022*** (0.0003)		
Firm-year FEs	Yes	Yes	Yes	Yes	Yes
Country FEs	Yes	Yes	Yes	Yes	Yes
N	5,733	2,653	1.4 mil	425,430	585,220

Notes: The sample includes all U.S. firms that received at least one VC investment from 2004 to 2017. D(Migrant) is a dummy variable indicating the presence of any same-origin skilled workers. Fund age and fund size are the average age in months and size of the VC funds at the time of financing respectively. Observations are at firm-source country-fiscal year level. The period of analysis is from Fiscal Year (FY) 2005 to FY 2014. Outcome variables are measured over a period of three fiscal years following the approval of the visa application. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.10: Skilled Immigrants and Foreign Venture Capital - Robustness

	Pr. of US deals (1)	Number of same-origin deals					
		(2)	(3)	Poisson (4)	FY10-14 (5)	All origins (6) (7)	
D(Migrant)	0.0678*** (0.0049)	0.0234*** (0.0050)	0.0213*** (0.0050)	1.4062*** (0.0726)	0.1083*** (0.0096)	0.0437*** (0.0026)	
Total U.S. deals			0.0047*** (0.0002)				
Total migrants							0.0133*** (0.0061)
Firm-country FEs	Yes	Yes	Yes	No	No	No	No
Year FEs	Yes	Yes	Yes	No	No	No	Yes
Firm-year FEs	No	No	No	Yes	Yes	Yes	No
Country FEs	No	No	No	Yes	Yes	Yes	Yes
N	1.5 mil	1.5 mil	1.5 mil	105,790	747,060	12.7 mil	820

Notes: The sample includes all U.S. firms that received at least one VC investment from 2004 to 2017. D(Migrant) is a dummy variable indicating the presence of any same-origin skilled workers. In (3), the total number of U.S. VC deals received at the firm-fiscal year level is used as a covariate. Observations are at firm-source country-fiscal year level except for (7), where the aggregate number of VC deals from a country in a given year is regressed on the aggregate number of skilled immigrants from the same country in the same year. Outcome variables are measured over a period of three fiscal years following the approval of the H-1B application. Heteroskedasticity robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.