

NBER WORKING PAPER SERIES

THE EFFECTS OF HIGH-SKILLED IMMIGRATION POLICY ON FIRMS:  
EVIDENCE FROM VISA LOTTERIES

Kirk Doran  
Alexander Gelber  
Adam Isen

Working Paper 20668  
<http://www.nber.org/papers/w20668>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
November 2014, Revised June 2022

This is a greatly revised version of NBER Working Paper 20668, previously titled “The Effect of High-Skilled Immigration on Patenting and Employment: Evidence from H-1B Visa Lotteries.” We thank U.S. Customs and Immigration Services for help with the H-1B lottery data. We thank Sunil Vidhani for outstanding research assistance. We thank Notre Dame and the Wharton School of the University of Pennsylvania for research support. We are grateful to George Borjas, John Bound, Sean Farhang, Richard Freeman, Hilary Hoynes, Jenny Hunt, Damon Jones, Larry Katz, Bill Kerr, Norman Matloff, Ankur Patel, Dina Pomeranz, Jesse Rothstein, and seminar participants at the Fed Board, HBS, NBER, CEMIR, U.S. Treasury, and WIGE for helpful comments. We thank Lee Fleming for sharing the patent data with us. We thank Danny Yagan for sharing his code to probabilistically identify natives and foreigners in the Treasury data. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the U.S. Treasury Department, any other person associated with the U.S. Treasury Department, or the National Bureau of Economic Research. All errors are our own.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at <http://www.nber.org/papers/w20668.ack>

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Kirk Doran, Alexander Gelber, and Adam Isen. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Effects of High-Skilled Immigration Policy on Firms: Evidence from Visa Lotteries  
Kirk Doran, Alexander Gelber, and Adam Isen  
NBER Working Paper No. 20668  
November 2014, Revised June 2022  
JEL No. J18,J21,J23,J24,J44,J48,J61,O3,O32,O34,O38

**ABSTRACT**

We compare winning and losing firms in lotteries for H-1B visas, matching administrative data on these lotteries to administrative tax data on U.S. firms and to approved U.S. patents. Winning one additional H-1B visa crowds out about 1.5 other workers at the firm. Additional H-1Bs have insignificant and at most modest effects on firm innovation. More general evidence from the universe of U.S. firms and the universe of H-1B visas using alternative estimation strategies is consistent with these results. Firms that hire H-1Bs grow faster and innovate more because they are different in other ways from firms that do not.

Kirk Doran  
University of Notre Dame  
3060 Jenkins Nanovic Halls  
Notre Dame, IN 46556  
kdoran@nd.edu

Adam Isen  
Office of Tax Analysis  
U.S. Department of the Treasury  
1500 Pennsylvania Ave., NW  
Washington, DC 20220  
adam.isen@gmail.com

Alexander Gelber  
Department of Economics University  
of California at San Diego 9500  
Gilman Dr.  
San Diego, CA 92093  
and NBER  
gelber@gmail.com

## 1. Introduction

A key issue in several fields of economics is how easily firms can substitute one type of worker for another. Substitutability has implications for firms' production functions, predicting the demand for labor, and the incidence of policies that affect the price or availability of different types of labor. However, it is rare to observe a setting in which the availability of one or more type of labor is truly random, allowing a sharp test of hypotheses.

In this paper, we explore what happens to firms that win lotteries for an important type of labor in the United States: highly skilled foreign workers in the H-1B visa program.<sup>1</sup> Firms often argue that H-1B workers have exceptional skills that they cannot otherwise obtain, and that obtaining these unique skills is necessary for them to grow and innovate. Others argue that H-1Bs have skills that firms could otherwise obtain, and thus additional H-1Bs would generally crowd out other workers and have more muted effects on firm outcomes like innovation. While firms that hire H-1B visa workers grow faster and innovate more than other firms, a key question is whether this is due to the causal effect of H-1Bs on the firms or because of other differences.

---

Acknowledgements: Doran: [kdoran@nd.edu](mailto:kdoran@nd.edu); Gelber: [amgelber@ucsd.edu](mailto:amgelber@ucsd.edu); Isen: [adam.isen@gmail.com](mailto:adam.isen@gmail.com). We thank U.S. Customs and Immigration Services for help with the H-1B lottery data. We thank Gita DeVaney and Sunil Vidhani for outstanding research assistance. We thank Notre Dame, the UC Berkeley Institute for Research on Labor and Employment, and the Wharton School of the University of Pennsylvania for research support. We are grateful to George Borjas, John Bound, Sean Farhang, Richard Freeman, Hilary Hoynes, Jenny Hunt, Ben Jones, Damon Jones, Larry Katz, Bill Kerr, Norman Matloff, Ankur Patel, Dina Pomeranz, Jesse Rothstein, Fabian Waldinger, and seminar participants at the Fed Board, HBS, LERA, NBER, CEMIR, U.S. Treasury, and WIGE for helpful comments. We thank Lee Fleming for sharing the patent data with us. We thank Danny Yagan for sharing his code to probabilistically identify natives and foreigners in the Treasury data, and the authors of Bell et al. (2018) for sharing their data. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the U.S. Treasury Department, or of any other person associated with the U.S. Treasury Department. All errors are our own. Edited by Michael Greenstone.

<sup>1</sup> In 2010, immigrants accounted for 16 percent of the U.S. adult population with at least a bachelor's degree, and high-skilled immigrants represent 24 percent of workers in occupations closely tied to innovation (Pekkala Kerr, Kerr, and Lincoln 2015). In recent years, prominent voices from government, business, labor, and academia have discussed significant changes to U.S. immigration law. Many proposals have envisioned changes to the largest U.S. high-skilled immigration program: H-1B visas for temporary immigration, which allow U.S. firms to employ foreign workers for three years. The path of high-skilled immigration into the United States is unusual by international standards: the H-1B program is built around written requests from individual firms for access to specific workers with ostensibly unique skills. How H-1B workers affect the firms that have applied for them is the subject of much public discussion, but little empirical work.

We answer this question by using administrative data on the entrants in these lotteries, matched to their tax filings and patenting. We find that winning a lottery for a skilled foreign worker crowds out otherwise available workers and does not increase firm patenting.<sup>2</sup> Winning firms neither grow faster nor innovate more than losing firms.

This result applies more generally than the setting of these lotteries. Regression discontinuity estimates based on application submission dates that compare firms which just did and did not miss out on being rationed, estimates based on the winners of large lotteries compared to a proxy control group, and shift-share estimates based on variation in visa caps all reveal that H-1B visa workers crowd out otherwise available workers at the firm level, leaving measures of firm innovation roughly constant. Likewise, adding firm fixed effects to naïve regressions of firm employment on new firm H-1B visas among the universe of U.S. firms produces similar results.

In particular, lottery point estimates show that winning one additional H-1B visa worker crowds out approximately 1.5 otherwise available workers. Regression discontinuity point estimates show crowd out of roughly 1.5 other workers, point estimates from the winners of a larger quasi-lottery show crowd out of about 0.7 other workers, shift share point estimates show crowd out of 1.25 other workers, and firm fixed effect point estimates show crowd of 0.9 other workers.<sup>3</sup>

Overall, our results are not supportive of the narrative that H-1Bs have unique skills that firms cannot otherwise obtain, in contrast to what is suggested by lawmakers.<sup>4</sup> Rather, they are

---

<sup>2</sup> We also find that they decrease the probability of hiring contractors and have no effect on the amount of the research and experimentation credit claimed.

<sup>3</sup> Note that we generally cannot reject that any particular strategy's estimates are different from another's.

<sup>4</sup> The Senate Judiciary Committee reports which accompanied legislation to expand the H-1B program in 1998 and 2000 exemplify the narrative in which H-1Bs help firms address "shortages" of special skills. These reports noted that: "Companies across America are faced with severe high-skill labor shortages that threaten their competitiveness" (Senate Judiciary Committee 1998). "America faces a serious dilemma when employers find that

more supportive of a narrative in which marginal H-1Bs crowd out other workers, are paid less than alternative workers, and increase the firm's profits—despite having little effect on measures of firm innovation. Indeed, we find some evidence from our lottery variation that additional H-1Bs increase profits, and some evidence that additional H-1Bs decrease payroll costs per employee.

Several advances over the previous literature make these analyses possible, including administrative H-1B data on *both* lottery winners and losers and linkage of this and other H-1B data to the tax filings and patent data for the universe of U.S. firms. Thus, relative to other studies on H-1Bs and other immigration programs, ours is the only one to our knowledge to leverage randomized variation or a discontinuity to estimate the effect of immigration on outcomes in the receiving economy as well as to focus on the universe of affected U.S. firms. Our paper relates to previous work on the effects of immigration on the labor market (*e.g.* Card 1990; Borjas, Freeman, and Katz 1997; Card 2001; Friedberg 2001; Borjas 2003; Edin, Fredriksson, and Åslund 2003; Lubotsky 2007; Borjas, Grogger, and Hanson 2012; see surveys in Borjas 1994; Friedberg and Hunt 1995; Freeman 2006; Dustmann, Glitz, and Frattini, 2008; Hanson 2009; and Pekkala Kerr and Kerr 2011), as well as on measures of innovation (*e.g.* Borjas and Doran 2012; Foley and Kerr 2013; Moser, Voena, and Waldinger 2014; Grogger and Hanson forthcoming; see the Kerr 2013 survey). Previous studies on the labor market or innovation impacts of the H-1B program or similar programs include Kerr and Lincoln (2010), Hunt and Gauthier-Loiselle (2010), Hunt (2011), Peri, Shih, and Sparber (2013), Pekkala Kerr,

---

they cannot grow, innovate, and compete in global markets without increased access to skilled personnel. Even apart from shortages in particular fields, in our increasingly global economy, highly skilled foreign workers are certain to be in a position to make unique contributions to the U.S. economy. A person from another country may simply be a uniquely talented individual with unique knowledge and skills. The country needs to increase its access to skilled personnel immediately in order to prevent current needs from going unfilled” (Senate Judiciary Committee 2000).

Kerr, and Lincoln (2015), and Bound *et al.* (2015). Regression analysis in the literature has found no clear evidence of crowdout of other employment, and in some cases has found crowd-in.<sup>5</sup> The literature has also found that H-1Bs lead to positive effects on patenting.

The paper is structured as follows. Section 2 describes the policy environment. Section 3 describes the data. Section 4 discusses simple statistics and overall comparisons among firms. Section 5 discusses our empirical specifications. Section 6 describes the full set of results, and Section 7 concludes.

## **2. Policy environment**

H-1Bs are sponsored by firms, which apply to the United States Citizenship and Immigration Services (USCIS) and pay a fee to obtain a visa for each H-1B worker they wish to hire. In its application for each visa, a firm must specify the identity of the worker it wishes to hire. An H-1B visa allows a skilled foreigner to enter the U.S. for three years. The H-1B is considered a “non-immigrant” visa because it allows those with H-1Bs to stay in the U.S. only temporarily. After these three years, the worker may leave the U.S. or a firm may seek to renew the worker’s H-1B visa. Firms may also sponsor the worker to be a permanent resident. The H-1B worker may also move to another firm before the initial three years are up, though several frictions pose barriers to a move: the new firm must pay USCIS application and legal fees; upon moving, an H-1B goes to the “back of the line” for gaining permanent residency; some H-1Bs may not know that they can change jobs; and in the years we study, the worker had to wait for

---

<sup>5</sup> Kerr and Lincoln (2010) find no evidence that H-1Bs crowd out other workers. Pekkala Kerr, Kerr, and Lincoln (2015) find mixed evidence on the effect of H-1Bs on total firm size. Peri, Shih, and Sparber (2013) find that H-1Bs increase native employment. However, the simulations of Bound *et al.* (2015) show that the ability to hire foreign computer scientists should reduce equilibrium employment and wages of natives, while increasing equilibrium aggregate employment and output.

several months until the new firm's H-1B application was approved, but a gap of only two weeks was allowed between jobs.<sup>6</sup>

The firm submitting the H-1B application (I-129) must attest, among other things, that: “(a) H-1B nonimmigrants will be paid at least the actual wage level paid by the employer to all other individuals with similar experience and qualifications for the specific employment in question or the prevailing wage level for the occupation in the area of employment, whichever is higher”; and “(b) The employment of H-1B non-immigrants does not adversely affect working conditions of workers similarly employed in the area of intended employment.” Firms are required to pay H-1Bs comparably with workers in one of four skill categories (defined by experience, education, and level of supervision).<sup>7</sup>

### **Caps and lotteries on H-1Bs**

The total number of H-1B visas awarded to for-profit firms in a given year is subject to a maximum number or “cap.” This cap is different for visas given to workers who have a master's degree or higher from a U.S. institution (the “Advanced Degree Exemption” (ADE) H-1B visa), and those without such a degree (the “Regular” H-1B visa). In recent years, including the years for which we have lottery data, the cap for ADE visas has been 20,000, and the cap for Regular visas has been 65,000. Much existing literature has identified the effects of H-1Bs through variation in these caps over time, especially the large increase in the caps in 1999 followed by a large decrease in 2004.

---

<sup>6</sup> Depew, Briggs, and Sorensen (2013) study a single multinational information technology firm and find that from 2003 to 2011, 22 percent of its H-1Bs quit and moved to another firm while on the H-1B.

<sup>7</sup> Employers who are “H-1B dependent”—whose workforce is comprised of a sufficiently large fraction of H-1B employees—face additional requirements to attempt to recruit, and not displace, U.S. workers. Firms may legally hire an H-1B in lieu of a worker who would have been at a higher skill level.

The number of H-1B visa applications in any given year has not always exceeded the cap, but in the two lottery years we study, the cap was reached for each of the two types of H-1B visa. Visa applications were accepted on a rolling basis once the application season began, and USCIS allocated visas by lottery only for applications submitted *on* the date when the total number of applications received exceeded the remaining available slots. In each of these lotteries, the total number of applications that won the lottery was equal to the number of remaining visas necessary to reach the cap.

Our main strategy makes use of lotteries for H-1B visas in fiscal year (FY) 2006 and 2007. (While lotteries were run in subsequent years, USCIS did not keep data on which firms won and lost the lottery.) The caps for the FY2006 Regular visa, FY2006 ADE visa, FY2007 Regular visa, and FY2007 ADE visa, were reached on August 10, 2005, January 17, 2006, May 26, 2006, and July 26, 2006, respectively. That those were the dates the cap was reached was only known *ex post*—making it effectively impossible for firms to game the system by applying on the (unknown) lottery date for more visas than they desire, on the basis of the (unknown) probability of selection. Even across the four lotteries we study, the probability that an application won varied widely. Indeed, these were the first two years USCIS used a lottery to allocate H-1Bs, and it was not announced in advance that lotteries were going to be run.<sup>8</sup> Each lottery was conducted within a month of reaching the relevant cap.

If a firm is denied a capped H-1B, it has several alternatives to hiring no one. Other than hiring U.S. citizens or foreigners who are permanent residents, firms can hire foreigners on other visas, including L-1 temporary work visas, Optional Practical Training (OPT) extensions of F-1

---

<sup>8</sup> One to two weeks prior to each lottery, USCIS publicly announced the number of applications it had received. Thus, firms may have been able to anticipate approximately when the cap might be reached, but they could not reasonably predict the exact day it would be reached or probability of selection on this day.



student visas, or H-1Bs not subject to the cap. L-1s allow multinational firms to bring a worker at a foreign branch to the U.S. temporarily. Visa lottery losers would likely not resort to bringing the same worker to the U.S. on an L-1, since a firm would have typically applied for an L-1 rather than an H-1B if the L-1 were feasible (as the L-1 is more advantageous to the firm than the H-1B). Only 11 percent of lottery participants are multinationals, further limiting the importance of the L-1 in our context. In FY2006 and FY2007, OPT extensions allowed F-1s to extend their stays for only 12 months, limiting substitutability with H-1Bs, and the majority of H-1B applicants are not eligible for a F-1 visa.

### **What job tasks do H-1Bs do?**

Before describing the data, it is worthwhile to consider what tasks H-1Bs typically do, which can help contextualize the results that follow. Data on I-129 applications show that H-1B visa workers work in a variety of occupations doing a variety of tasks, from working as researchers to working in tech support call centers. However, despite this diversity, the majority of all I-129 applications from firms list one occupation: “systems analysis and programming”.<sup>9</sup>

We break down this broad occupation category into specific occupations that reveal worker tasks by using the Department of Labor’s 2006 Occupational Employment Statistics by State, containing mean annual earnings for the 16 sub-occupations within the “Computer and mathematical occupations” category. The mean annual earnings of these different sub-occupations vary widely within each state. Thus, the 2006 salary and state of a person in the computer occupations can determine whether they are being paid like someone in a creative research role or like a worker in tech support.

---

<sup>9</sup> While this occupation only constitutes a plurality among just ADE applications, those H-1B applications total less than one-quarter of all capped H-1B slots.

We report in Appendix Table 1 differences between the mean annual salaries in each subcategory and the actual annual salaries paid to Regular H1-Bs in “systems analysis and programming”, reported in their I-129s. The occupations with the closest annual salaries to H-1Bs, adjusted by time and place, are not the creative professions of software engineers, or even the higher-level technical professions such as programmers. Rather, they are lower-level supporting jobs, such as computer support specialists, technicians, and systems administration. The wages are so much lower for these H1-Bs than for the time-and-place-adjusted wages for creative professions that it is very unlikely many of these H-1Bs are doing creative work, at least if prevailing wage restrictions are met.

### **3. Data**

This paper combines, for the first time, data on the universe of U.S. firms with data on H-1B immigrants at the firm level. It combines this data with the universe of U.S. patents to paint a thorough picture of highly-skilled immigrants and their output at U.S. firms.

**IRS data:** We use Internal Revenue Service (IRS) tax data on the universe of U.S. firms. The holder of every unique Employer Identification Number (EIN) that employs workers must generally report firm employment on a quarterly basis (Form 941). We also rely on additional tax forms for other outcomes, including business income tax returns (the 1120, 1120S, and 1065), W-2s, and 1099-MISC.

**USCIS data:** We use USCIS administrative data on the H-1B submissions for FY2006 and FY2007. The data contain the following information on each H-1B visa application in each of these years: EIN and name of the firm applying; the date the firm applied for a visa; the type of H-1B (Regular or ADE); how many of each firm’s applications won or lost the lottery; whether each application was approved by USCIS; and firm-reported worker characteristics from the I-

129 such as highest degree completed. These data are used for our main lottery analysis as well as our regression discontinuity design.

For the remaining ancillary identification strategies, we use USCIS data on the number of newly approved capped H-1Bs at each firm in the United States for each FY from 2003 through 2008 (not including renewals). In each FY, this data contains approximately 30,000 firm names (but not EINs), each with an associated number of approved H-1Bs for that FY.

**Patenting data:** We obtained the Patent Dataverse on the universe of granted U.S. patent applications from 1975 to 2013 at each firm (identified by name) based on USPTO data organized by year of application.<sup>10</sup> We also observe total patent citations. For ancillary identification strategies, we use the data in Bell et al. (2019) that links individuals in the IRS data with those listed on patent applications.

### **Match between USCIS, IRS, and patenting data**

Using EINs, we merged firms from the USCIS lottery data to their IRS records.<sup>11</sup> The IRS data gives us firm-level quarterly employment and annual net income (“profit”), wage bill, and Research and Experimentation (R&E) credit for research and development expenses. For the patent merge, as explained in Appendix 1, we performed an intentionally liberal automatic string matching procedure between the USCIS lottery and patenting datasets to obtain all plausible matches between firms and patents. We then searched through the matches by hand to detect and remove all matches that appeared spurious.

---

<sup>10</sup> See <https://dataverse.harvard.edu/dataverse/patent> (accessed 5/24/2015). Granted patents are classified by the calendar year a firm applied for the patent. For example, our measure of the number of patents at a firm in Year 0 refers to patents the firm applied for in Year 0 that were approved by 2013.

<sup>11</sup> We drop the 2.0 percent of firms in the USCIS data that did not match to the IRS EIN master list. 4.5 percent of the remaining firms did not match to the quarterly firm employment data; we treat these data as missing for employment analyses. Of the rest, 17.9 percent have missing employment data in Year -1. We try two options for these firms: first, dropping these data for the purpose of the employment specifications; and second, using a dummy variable to indicate missing pre-period employment and assigning mean pre-period employment to firms with missing data. We separately test for balance across each selection point.

For estimation strategies that rely on USCIS data on the number of newly approved H-1Bs at each firm for FYs 2003-2008, we match it to IRS data using a fuzzy match on firm names. (Since there are approximately 30,000 separate firms for each FY, we must apply a fuzzy matching technique that does not involve any by-hand work, as explained in Appendix 1.) The resulting matched data contains over 10 million for-profit firms (the universe of for-profit firms in the U.S.) of which 68,092 firms have at least one H-1B between the years 2003 and 2008.<sup>12</sup> We also construct a match between the universe of firms in the IRS data and the universe of U.S. patenting data. We start with an existing link between individuals on approved patent applications and their taxpayer IDs (Bell et al., 2019), and then use W-2 data to determine whether each individual was an employee of a firm in the year a patent application was submitted, and if so, at which firm the individual was working. This allows us to infer patent counts by firm-year. We explain in further detail in Appendix 1.

#### **4. Simple statistics and comparisons**

Table 1 shows summary statistics for the universe of U.S. firms. It is clear that firms that have ever had an H-1B have higher employment, experience larger increases in employment each year, patent more, have higher profits, and pay their workers higher salaries. It is possible that these differences in outcomes are due to positive causal effects of H-1Bs on firms. It is also possible that these differences in outcomes are artifacts of other differences between the types of firms that happen to hire H-1Bs and the types of firms that do not. Our lottery identification strategy and ancillary identification strategies can address this question.

Table 1 column (3) shows summary statistics for the FY2006 and FY2007 lottery sample. We use data on 2,750 firms (*i.e.* EINs) in the full lottery sample. In 300 cases (9.84 percent),

---

<sup>12</sup> To make the regressions tractable, we restrict to a 1% random sample of firms without an H-1B between 2003 and 2008, assigning them a weight of 100.

firms apply for at least one visa in both FY2006 and FY2007. Thus, over both lottery years, there are 3,050 firm-lottery year observations, where “year” refers to a year of the lottery, rather than a year when an outcome is observed. The mean and standard deviation of the number of employees during Q1 to Q4 in the full sample are very large. In firms with 30 or fewer, or 10 or fewer, employees in Year -1 (two representative cutoffs we use), the mean and standard deviation of Q1 to Q4 employment are lower but still large. Median employment is lower than mean. Winsorizing also reduces the mean and standard deviation.

In the FY2006 Regular lottery, the vast majority of applications lost the lottery, and in the FY2007 Regular lottery, the vast majority won. The ADE lotteries have a more even fraction of winners and losers. The fact that the vast majority either won or lost the Regular lotteries will not directly pose an issue for our estimates; such effects on precision will be reflected in the confidence intervals.

The lottery sample contains 7,243 visa applications, with an average of 2.37 H-1B applications per firm summing over both years. The average firm in our sample won 0.57 H-1B visas when aggregating across both years. The standard deviation of the number of chance lottery wins (defined below) is 0.33, and its range runs from -2.65 to 2.96. Over half of firms are in North American Industry Classification System (NAICS) code 54, representing professional, scientific, and technical services. The H-1B application data show that across all lotteries, applicant average age is around 30 (Appendix Table 2).

### **Comparison of lottery firms to other firms that applied before the last day**

Our primary identification strategy comes from comparing firms that randomly received H-1Bs to those that randomly did not. This comparison is comprised of firms that applied on the day the cap was reached, which addresses effects for marginally changing the number of H-1Bs

allowed, a question of great relevance as visa cap changes are contemplated. Nonetheless, it is worthwhile to compare this sample to the broader sample of firms applying for H-1B visas in these years. In Appendix Table 3, we regress characteristics of the firms or workers on a dummy for applying on the last day and lottery fixed effects. Applications on the last day tend to be from larger firms and those that are more likely to be in professional, scientific, and technical services industries and to have patented more in the past compared to the set of firms that applied earlier. On the last day, firms disproportionately submit applications for workers with higher educational degrees; for those with higher intended worker salaries; for “systems analysis and programming” jobs; and for younger workers. If H-1Bs hypothetically have more positive innovation effects in firms that patented more in the past and/or are in scientific industries, or among workers with advanced degrees or higher salaries, then our sample will arguably be primed to find a positive effect on innovation relative to using the full set of H-1B firms. Likewise, in Appendix Table 5, we also show that firm growth rates are not inferior for firms applying on the last day or on a day approaching it. Results from other empirical strategies will speak more directly to the generalizability of the lottery experiment.

## **5. Empirical strategies**

Our main outcomes of interest are number of employees and patenting. We also consider the effect on the R&E credit, the firm’s wage bill per employee, and profits. In the below, we first discuss our primary empirical strategy that exploits the random assignment of H-1B visas among firms in the FY2006 and FY2007 H-1B lotteries before turning to secondary strategies.

Our lottery strategy must accommodate firms that applied for multiple H-1B visas. If a firm submits  $n$  visa applications to a lottery in which  $p$  percent of applications won a visa, and  $W$  is the random number of H-1B visas given to the firm, then the average number of H-1B visas

given to the firm in expectation is  $E[W]=pn$ . If  $w$  is the random realization of  $W$ , then the number of “chance lottery wins” or “chance visas,”  $u=w-pn$ , is the random realization of the net number of wins relative to the *ex ante* statistical expectation conditional on  $p$  and  $n$ , and will be exogenous in the regression we specify below. Thus, our main independent variable is the random variable  $U$ , the net number of a firm’s chance lottery wins, which by construction has a mean of 0 and whose realization is  $u$ .

To find the causal effect of  $U$  on an outcome  $Y$ , we estimate versions of:

$$Y_{iT} = \beta_0 + \beta_1 U_{iT} + \varepsilon_{iT}. \quad (1)$$

$t$  is the number of calendar years since the lottery in question occurred; for example,  $t=0$  corresponds to Year 0.<sup>13</sup>  $T$  indexes the year of the lottery in question, *i.e.* FY2006 or FY2007.  $U_{iT}$  is the number of chance H-1B visa lottery wins for firm  $i$  in the lottery in year  $T$ .<sup>14</sup>  $\varepsilon_{iT}$  is an error term.  $\beta_1$  represents the intent-to-treat (ITT) effect of an additional chance H-1B visa win, which is relevant because firms and policymakers are interested in the raw effects on firms of allowing a marginal capped visa to the firm. To increase statistical power, we pool the Regular and ADE lotteries for a given fiscal year  $T$  and stack data from FY2006 and FY2007 in the same regression.<sup>15</sup> We cluster the standard errors at the firm level.

---

<sup>13</sup> For a given lottery year (*i.e.* FY2006 or FY2007), we refer to the calendar year the lottery occurred (*e.g.* 2005 in the case of the FY2006 lottery) as “Year 0.” The year before this calendar year is “Year -1”; the year after Year 0 is “Year 1”; *etc.* We refer to the first quarter when an H-1B employee would begin work at a firm (*e.g.* the first quarter of FY2006 in the case of the FY 2006 lottery) as “Q1”; the next quarter as “Q2”; *etc.*, which is relevant for employment, the only outcome we can observe quarterly.

<sup>14</sup> Although the randomization implies that  $U_i$  should be exogenous in (1), it is also possible to control for various pre-determined covariates. We generally control for a lagged value of an outcome variable at the firm (*e.g.* when the dependent variable is the number of employees, we can control for  $Y_{i,pre-period,T}$ , the number of employees in firm  $i$  observed in the period preceding the lottery) and for the expected number of lottery wins  $pn$ .

<sup>15</sup> Similarly, whenever we examine an outcome across multiple time periods  $t$ , we stack the data across these periods. For the 2.69 percent of firms that participate in more than one lottery in a given fiscal year  $T$  (*e.g.* a firm participates in both the 2006 Regular and ADE lotteries), we calculate  $U_{iT}$  by summing the total number of chance lottery wins across both of the lotteries that the firm enters in year  $T$ .

After a firm wins an H-1B lottery, its application may be approved, denied, or withdrawn. The total number of capped H-1B visas approved for a firm in any given year is potentially endogenous because it depends on the fraction of those that win the lottery that are approved. In practice, most are approved. Still, we can use lottery wins as an instrument for approved capped H-1B visas in a two-stage least squares (2SLS) model:

$$A_{iT} = \alpha_0 + \alpha_1 U_{iT} + v_{iT} \quad (2)$$

$$Y_{iT} = \gamma_0 + \gamma_1 A_{iT} + \eta_{iT} \quad (3)$$

$A_{iT}$  represents the number of capped H-1B visas approved for firm  $i$  in the lottery that occurred in year  $T$ . In the first stage (2), we regress  $A_{iT}$  on  $U_{iT}$  using ordinary least squares (OLS). In the second stage (3), we regress  $Y_{iT}$  on  $A_{iT}$  (instrumented using  $U_{iT}$ ) using OLS. The coefficient  $\gamma_1$  represents the local average treatment effect (LATE) of an extra approved capped H-1B visa among the compliers (*i.e.* those induced by winning the lottery to increase their number of approved capped H-1B visas).  $v_{it}$  and  $\eta_{it}$  are error terms. This specification is most relevant in the employment context where we are testing whether additional H-1Bs crowd in or out other employment.

Our employment variable measures a firm's total employment, including any H-1B workers, so we adjust our employment coefficients by 1 to subtract the H-1B worker and present effects on the employment of other workers. The key question of interest is a two-sided test of whether the coefficient on H-1B visas is significantly different from 0, as theory is ambiguous over such a relationship. If the coefficient is positive and significant, it would indicate that the extra H-1B visa increases employment of other workers at the firm—as opposed to crowding out workers that the firm would have otherwise hired, in which case the coefficient would be negative. One for one crowdout would yield a coefficient of -1. Finally, it is possible that the



new H-1B worker works more hours or works harder than others (perhaps to secure another visa or green card for continued employment in the U.S.) and therefore crowds out more than one other worker, for a coefficient of less than -1.

We investigate the effects of H-1Bs on each outcome for the three-year duration of the initial visa for consistency, but we also focus on shorter and longer durations that are most appropriate for our main employment and patenting outcomes, respectively. In particular, we focus on quarterly employment from Q1 to Q4 of the visa, when the H-1B worker is almost always working at the firm and when a coefficient's relation to 0 will therefore most reliably indicate crowd in or crowd out, and on patents from Year 0 to the latest year available in the data, Year 8, given the sometimes substantial time taken to develop patents.

To adjust for the long right tail of the employment distributions and relatively small sample sizes in the lottery sample, we take two approaches. First, in our baseline specification, we run median regressions. Second, we run mean regressions where we either include firm by lottery fixed effects and winsorize the dependent variable at the 95% level or let the dependent variable be the winsorized first difference of employment. The first difference  $\Delta Y_{itT}$  is taken from before the lottery to period  $t$  after the lottery. Such regressions would not capture large effects on employment outcomes, but we later explore this issue in depth.

Due to the long right tail of the distribution of patents, previous literature has typically examined transformations of the number of patents. We approximate the log of the number of patents using the inverse hyperbolic sine (IHS), which is defined at zero and approximates the log for larger values of its arguments (*e.g.* Burbidge, Magee, and Robb 1988, Pence 2006). The IHS of patents  $Y$  is defined as:

$$IHS(Y) = \ln(Y + \sqrt{1 + Y^2})$$

We alternatively deal with the distributional challenges posed by patents by using the negative binomial regression. This regression takes into account that patenting is a count variable.

We expect our results to be most statistically distinguishable in small and medium-sized firms, which, in the aggregate, contribute in important ways to U.S. employment and innovation (Acs and Audretsch 1990), and comprise a substantial fraction of all H-1B lottery applicants (e.g., 19% of H-1B visas come from firms with 10 or fewer employees; 34% come from firms with 30 or fewer employees). To evaluate how the effects vary across firms of different sizes, we investigate firms with 10 or fewer employees in Year -1 (roughly the 25<sup>th</sup> percentile of firm size in our sample); those with 30 or fewer employees in Year -1 (roughly the 50<sup>th</sup> percentile); many other firm size cutoffs; and firms of all sizes.

We now turn to discuss our secondary strategies, which, while relying on stronger identification assumptions, complement our main approach by providing two main benefits. First, the alternative approaches provide substantial improvements in statistical power over our lottery results, affording us the ability to more regularly recover meaningful confidence intervals from unwinsorized mean regressions and for full samples of firms.<sup>16</sup> Second, they enable us to examine contexts that sometimes provide larger-scale variation affecting more H-1B applications.

Our first alternative approach adopts a regression discontinuity (RD) design that exploits how the probability a firm had a successful application fell from near one hundred percent if submitted before the cap was reached to much lower if submitted on the day a cap was reached in FY 2006 and 2007. Importantly, it was not known in advance which days the caps were going to be reached, eliminating firms' ability to sort around the cutoff. We implement our RD

---

<sup>16</sup> As a result, we exclusively present unwinsorized employment regressions for these strategies.

approach using days relative to the cap being reached as the running variable and the optimal bandwidth of 20 days calculated by the method of Calonico, Cattaneo, and Titiunik (2014). Specifically, we tweak equations (2) and (3) by instrumenting for the number of H-1B visas approved with whether the applications were submitted before the day the cap was reached, and including a quadratic running variable to fit the relationship between number of accepted H-1B applications and number of days before the cap was reached and visa type by year fixed effects.<sup>17</sup>

We then consider the large FY2008 lottery, in which 123,480 H-1B applications were received in the first two days of the application season for only 85,000 capped slots. In this lottery, all applications were entered into the lottery, making this setting an interesting contrast to the moderately binding years of FY2006 and FY2007. The setting is difficult to analyze, however, because USCIS did not record information for the firms that lost the lottery.<sup>18</sup> Therefore, we must compare approved lottery winners with a control group constructed to proxy lottery losers. We construct this comparison group by determining the subset of Labor Condition Application (LCA) submissions that were likely to be eligible for submitting an H-1B application that would have been subject to the lottery (a prior LCA submission is required before an H-1B application can be submitted, but it does not guarantee an H-1B application submission).<sup>19</sup> Ultimately, we cannot rule out that the results are influenced by attrition from the LCA sample into the lottery, nor can we determine which losing applications were in the regular

---

<sup>17</sup> Some firms submit applications on multiple days, but keeping firms that show up in both the treatment and control group greatly attenuates any effect on visa approval. Thus, in order for there to be a first stage, we must limit our sample to firms that submitted all their applications for a particular visa category-year on one given day (76.4% of firms).

<sup>18</sup> Additionally, we can only observe approved winners.

<sup>19</sup> This consisted of all LCAs submitted by for-profit firms between March and April of 2007 in which the proposed H-1B worker would begin work between September and October of 2007. These restrictions are based on the USCIS rules that determine which I-129 applications would be subject to the FY2008 lottery, but we also tested and found robustness to variation in these restrictions, and we determine for-profit status through a merge to the IRS business entity database.

versus ADE lotteries, which had different selection probabilities.<sup>20</sup> With this proxy control group constructed and approved wins, we calculate “chance wins” and estimate equation (1).

We also utilize a shift-share strategy, which is one of the most common strategies employed when studying the impact of H-1B workers or migrants more generally (e.g., Pekkala Kerr, Kerr, and Lincoln 2015).<sup>21</sup> This strategy exploits cross-sectional reliance on H-1Bs and changes in the national H-1B cap over time, including those that occurred after 2003 when the cap decreased substantially from 195,000 in 2003 to 65,000 in 2004 while modestly increasing to 85,000 for the remaining years for which we have H-1B data (2005-2008). The shift-share strategy requires stronger identification assumptions, namely that of parallel trends. To implement the shift-share strategy, we calculate the share of new H-1Bs as a function of lagged firm employment in the first year (2003) for which we have H-1B counts. Firms that relied on many new H-1Bs in 2003 as a share of the firm’s 2002 overall employment were more likely to be affected by post-2003 visa cap changes than firms with a smaller share of H-1Bs. We run a first stage regression in which a firm’s new H-1Bs in a given year between 2003 and 2008 as a share of previous year employment are regressed on the firm’s 2003 H-1B share interacted with each year’s annual visa cap. We also control for year and firm fixed effects and an interaction of a time trend with the pre-period value of each dependent variable.

Finally, we compare H-1B firms to a large random sample of the universe of remaining U.S. firms. To do so, we augment equation (3) with firm fixed effects in which the independent variable of interest is the number of new capped H-1Bs approved for a given firm in a given year

---

<sup>20</sup> The ratio between LCA submissions and H-1B applications was approximately 1.63, raising the possibility of large amounts of bias. Related, Peri, Shih, and Sparber (2015) exploit these large lotteries at the metropolitan area level, which likely implies, due to the law of large numbers, that nearly all of the variation being exploited stems from attrition and differences in reliance on ADE versus Regular H-1Bs across areas.

<sup>21</sup> The existing literature on the impact of H-1Bs on firms has made use of smaller samples (such as portions of the COMPUSTAT database or large firms from Census data), we are the first to analyze the results from this strategy applied to the universe of U.S. firms.

between 2003 and 2008, ranging from zero (most firms) to several thousand (the most heavy H-1B users).

## 6. Results

### Validity

Appendix Table 6 verifies the validity of the randomized design by regressing variables that should not be affected by the lottery on chance lottery wins. The table confirms that none of the lagged dependent variables is significantly (or jointly) related to chance lottery wins.<sup>22</sup> Employee characteristics, as shown in Appendix Table 2, are also individually and jointly insignificantly related to lottery wins ( $p = 0.31$  in the joint test).

### Employment Results

Table 2 shows our main results from estimating equations (1)-(3). Across all specifications for small and medium sized firms the coefficients are negative and statistically significant, indicating that H-1B workers decrease the number of other workers employed at the firm. Column (1) of Table 2 displays the  $\beta_1$  coefficients from median regressions of the form outlined in equation (1) while columns (2) and (3) show the results from 2SLS mean regressions shown in equations (2) and (3).<sup>23</sup> Column (1) shows that for firms with 10 or fewer employees,

---

<sup>22</sup> We investigate the effects on Year -2 outcomes and control for the dependent variable measured in Year -1, which is the same control as in our regressions in later tables. By investigating Year -2 outcomes, we can also determine the firm size cutoffs by measuring employment in Year -1, yielding the same firms in each size category as in our later regressions. When we investigate Year -1 outcomes as the dependent variable, controlling for Year -2 observations and using firm size cutoffs from Year -2, the regressions are insignificant for all but one of the 27 dependent variables, consistent with random chance (results not shown).

<sup>23</sup> Because instrumental variables quantile regressions typically did not converge, we use the ITT for median regression. These regressions do not reflect that some H-1B lottery winners' applications are rejected, but our first stage coefficient presented in Appendix Table 9 is extremely precise and quite close to 1 (ranging from 0.88 to 0.89) such that scaling our estimates by the first stage would not alter our findings. Of course, instrumental variables quantile regressions do not rely on a Wald estimate, but in practice, in the rare median instrumental variables median regressions that converged, the coefficients on approved H-1B visas were very similar to the ITT median coefficient divided by the OLS or median first stage—i.e. only around 10 percent larger than in the ITT median regressions. Additionally, after their visas are approved by USCIS, some workers may not show up for their jobs in the U.S. However, North (2011) estimates that around the time we study, nearly all (95 percent) of those approved for H-1Bs

an additional chance H-1B visa win crowds out approximately one and a half other workers in the year after the lottery. The results are similar for firms with 30 or fewer employees or the full sample of firms. We then repeat these exercises for mean regressions with firm fixed effects (Column 2) and firm differences (Column 3), and similarly find evidence of crowdout for firms with 10 or fewer and 30 or fewer employees, albeit with somewhat more negative but less precise estimates.<sup>24</sup> (Note there is little difference between the firm fixed effects and differenced estimates). However, the mean regressions estimates for the full sample of firms are imprecise, which is unsurprising due to the high variance of large employers. Across all specifications, the full set of results remain similar whether we pool Q1 to Q4 or Q1 to Q12, with point estimates slightly increasing in absolute value in the longer time horizon.

To see more comprehensively how the estimates vary across the employer size distribution, we estimate our baseline specification from Column (1) of Table 2, varying the size threshold from 10 or fewer employees to 500 or fewer, in increments of 10.<sup>25</sup> Panel A of Figure 1 displays the results, showing that the results remain stable across the employer size distribution, with most point estimates hovering around -1.5, as seen in Table 2. In all cases, the coefficients are statistically significant, ruling out no effect on the employment of other workers at the 1 percent level. Figure 2 displays the dynamics of these effects by augmenting equation (1) and allowing the impact of chance H-1B visa wins to vary by years since the lottery. Before the lottery, the effect of a chance H-1B visa win has a point estimate that hovers around zero, in

---

end up being admitted. Thus, (further) scaling our scaled ITT or TOT estimates a bit more to reflect this would have negligible effects on our results and not affect our conclusions.

<sup>24</sup> We find no evidence that winsorizing employment in mean regressions is biasing our findings. Namely, an extra H-1B visa has an insignificant effect on the probability that the change in employment is outside the 95<sup>th</sup> (or higher) percentile or fixed size cutoffs (Appendix Table 9), and we get similar results (though not as precise) without winsorizing (Appendix Table 10).

<sup>25</sup> The necessity of keeping a sufficiently large number of firms in each category, to prevent the potential identification of any given firm, precludes us from going beyond 500 employees in increments of 10.

line with the random assignment to treatment. Following the lottery, the impact on the employment of others is negative, consistent with crowd out. Appendix 2 describes additional robustness exercises,<sup>26</sup> and Appendix 3 describes efforts to estimate the effect on employment of foreigners and nonforeigners separately. While there are limitations with this exercise, the results suggest that at least some of the crowdout may be of other non-U.S. citizens.

### **Innovation and other outcomes**

To examine the effect that H-1B visas have on other firm outcomes, we estimate regressions of the form outlined in equation (1), varying the dependent variable. We present the results in Table 3. We first examine the effect of chance H-1B visa wins on the innovative activity of the firm. Column (1) shows the results with the IHS of patents over the 8 years following the lottery as the dependent variable. We obtain a precisely estimated null effect of chance H-1B visa wins on the IHS of patents. For firms with 10 or fewer employees, the point estimate is a 0.026% increase in patents. With 95% confidence, we bound the effect between -0.42% and 0.47%, ruling out any material impact of an additional chance H-1B on patenting activity. Column (2) uses negative binomial regressions and recovers similar null effects, where with 95% confidence we can bound the impact to be between -1.08% and 0.21% for the smallest firms.

Examining robustness of the patenting results, Panel B of Figure 1 again varies the employer size threshold from 10 or fewer employees to 500 or fewer, in increments of 10. At the lower end of the firm size distribution, the point estimates hover around zero. As employer size increases to 80, we see a decline in the point estimates, yet we still cannot reject the null hypothesis of no effect at the 5% size, but a near zero effect lies towards the top of the 95%

---

<sup>26</sup> In particular, we show negative effects on employment of other workers quarter-by-quarter, where firms missing pre-lottery employment are included, across a variety of subgroups, and of contractors.

confidence interval. Figure 3 repeats a similar exercise as Figure 2, but with patents as the dependent variable. Effects are insignificant in the years before the lottery and the dynamics in response to a chance H-1B visa do not display a clear pattern, in line with our precisely estimated “pooled” zero from before. Together, the evidence indicates that an additional chance H-1B visa win results in at most a very modest increase in patenting and in the most negative case, a small decline in patenting. In Appendix 2, we also explore and confirm robustness to our patenting findings, including when examining patents weighted by citations or limited to only firms in the more highly-skilled ADE visa lotteries (Appendix Tables 16 and 18).

While patenting is one measure of the innovative output of a firm, measuring inputs into the production of innovative output is another way of capturing the firm’s innovative activity. To do this, we estimate how a chance H-1B visa win effects the intensive and extensive margin claim of the R&E credit during the duration of the visa in Columns (3) and (4), respectively. For firms with 10 or fewer employees the point estimate is negative, although the 95% confidence interval is large, only allowing us to rule out increases of greater than 4.1%. The point estimate for extensive margin use of the R&E credit is negative with the 95% confidence interval ruling out an impact greater than 0.41 percentage points. We can rule out even smaller increases for firms with 30 or fewer employees, but when all firms are used to estimate the effect, the results are less precise.

Focusing on the remaining firm outcomes, column (5) examines the impact that a chance H-1B visa win has on median payroll per employee during the duration of the visa. For smaller firms, we find declines in payroll per employee with point estimates of several thousand dollars. Column (6) shows the results for firm profits. While the results are only statistically significant for firms with 30 or fewer employees, the positive point estimates in all three cases provide



suggestive evidence for a moderate positive effect of chance H-1B wins on profits ranging from three to eleven thousand dollars.

### **Results from secondary strategies**

We now turn to estimating the impact of H-1Bs with our secondary empirical strategies. The results for all strategies, along with some reproduced experimental estimates, are presented in Table 4 across the main dependent variables of interest. Starting with the regression discontinuity design, Appendix Figure 1 visualizes the discrete and dramatic drop, relative to prior days, in the number of H-1B applications accepted that occurs on the day the cap is reached for all visa type-years but Regular 2007, which is roughly in line with their respective win rate probabilities. In column 1, we report unwinsorized mean employment results for all secondary strategies. For small and medium sized firms (and for the full set of firms in median regressions as shown in Appendix Figure 2), the RD employment estimates are precisely measured and demonstrate crowd out of other workers. These estimates are in line with the results from our experimental strategy, but, as expected, with tighter confidence intervals than the mean *winsorized* lottery estimates (and similar confidence intervals to the median lottery results).<sup>27</sup> While the effect of an H-1B on patenting is less precisely measured than for our experimental design, the effect on patenting and use of the R&E credit is negative and generally significant (columns 2 and 3). We take this as suggestive evidence that the RD approach finds little effect of H-1Bs on the innovative activity of the firm, with small negative effects possible. Results for pay per employee and profits are mixed and imprecise (columns 4 and 5).

---

<sup>27</sup> The findings are robust to variation in bandwidth, polynomial order of the running variable, firm samples, and use of median regressions, with Appendix Figure 2 showing the full set of median estimates from varying the bandwidth from as low as 10 days through as high as 100 days.

Turning to the large FY 2008 lottery winners and comparison group of non-winners presented in the third row of each panel in Table 4, the estimates we recover are precise, including for the full sample of firms, and substantially more so than the prior strategies. We find the effect of an H-1B is to similarly crowd out other workers, although the magnitude (around 0.75 worker) is smaller than what is seen in the experimental and RD results. The point estimates for patenting estimates are small and negative. The tight confidence interval allows us to rule out modest changes to the patenting activity of the firm. The results for the R&E credit rule out moderate changes in the use of the tax credit. Together these results confirm the picture painted by the experimental and RD strategies: the effect of an H-1B on the innovative activity of a firm is small. While the results for profits and payroll per employee are less robust to changes in firm size, we see some evidence of higher profits and decreases in payroll per employee, which is in line with the experimental evidence.

For the shift-share strategy, as expected, the results indicate that the variation in the visa cap over time interacted with a firm's pre-period H-1B employment share has a statistically significant impact on the firm's H-1Bs employment share, with first stage F-statistics of at least 66. As presented in the fourth row of each panel in Table 4 and starting with employment, the results are in line with our previous findings; namely, they show a decrease in the employment of other workers on the order of 1.25, and the effects are precisely estimated for all samples of firms. We find the effect of an H-1B on the patenting of the firm to be insignificant, with point estimates around zero, while effects on use of the R&E credit are negative. We also find negative but insignificant effects on payroll per employee and, somewhat surprisingly, negative effects on profit. Outside of the latter estimate, these results align with the prior results and more generally contribute to the story that H-1Bs have somewhere between a small negative and a zero impact

on the innovative activity of a firm as measured by patenting activity and use of the R&E credit.<sup>28</sup>

Finally, the OLS regressions, presented in the fifth row of each panel in Table 4, indicate that with firm fixed effects, an additional H-1B visa is associated with a decrease in the employment of other workers on the order of nearly 1 employee, regardless of firm size. Thus, the naïve difference between firms that do and do not hire H-1B workers as seen in Table 1 are an artifact of other differences. H-1Bs are associated with very small increases in patenting and no changes in the R&E credit, with confidence intervals that rule out even small effects for both measures. Finally, we find little evidence of a relationship between H-1Bs and other outcomes at the firm.

In summary, we find that alternative strategies, with a variety of different strengths, weaknesses, and degrees of generalizability, validate our experimental finding that the effect of an H-1B is to reduce the employment of other workers at the firm and that H-1Bs have no meaningful impact on the patenting activity of firms.<sup>29</sup>

## **Discussion of Results**

Our employment results consistently show decreases in the employment of other workers, thus implying that additional H-1Bs robustly crowd out other individuals who would have worked at the firm. Taken at face value, most lottery and regression discontinuity point estimates

---

<sup>28</sup> That the prior literature finds positive effects on patenting could be driven by a failure to adequately capture time-specific differences that correlate with H-1B availability, by the earlier increases in the cap not contained within our sample, or the cap failing to bind in many years. Using our specific years of data, sample of firms, and specification, we estimate positive effects on patenting when we do not control for interactions of a time trend with the pre-period values of the dependent variables. However, this sensitivity may not apply to other analyses. Related, the negative effects on profits we recover could in part be because shift share designs, which compare relative changes between the independent and dependent variables, are not well suited to dependent variables that are frequently negatively signed (in this case net profit) or could hint at violations in the parallel trend assumption.

<sup>29</sup> Appendix Figures 3 and 4 show that for these identification strategies our findings on the effect of H-1Bs on employment of other workers and patenting are robust across the firm employment distribution.

in fact imply that an H-1B worker crowds out more than one worker, although we cannot reject a null hypothesis of one-for-one crowdout; or said differently, we do not find statistically significant evidence that overall employment (including the H-1B) at the firm goes down. In considering our other strategies, the shift-share point estimates, while smaller in size than the lottery and RD estimates, similarly show greater than one-for-one crowdout (statistically significantly so), while the remaining strategies point to crowdout close to, but below, one-for-one, and their confidence intervals can generally rule out more than that one-to-one crowdout. Yet while there is some variation in the point estimates across strategies, it is also almost always the case that we cannot reject the null hypothesis that any strategy's estimates are the same as another sets. Altogether, our results are consistent with the possibility that new H-1B workers and otherwise available workers are perfect substitutes.<sup>30</sup>

Still, given the size of the coefficients for some strategies, it is worthwhile to ask: what underlying feature of H-1B employment might cause them to crowd out multiple other workers at once? We offer several hypotheses that would be consistent with this. First, it is possible that H-1Bs work harder than alternative employees do. This could arise due to employer monopsony power over opportunities for a green card or for reasons intrinsically related to the quality of the workers themselves. Second, (full-time) H-1B workers could be replacing multiple part-time workers, though our results on payroll per employee are not consistent with this hypothesis. Third, H-1Bs may indeed have special skills, but firms may use these skills to facilitate changes in production technology that lead to a different mix of outputs and inputs, and in particular, to a

---

<sup>30</sup> They do not necessarily imply perfect substitutability, as it would also depend on the degree of substitutability or complementarity of H-1B labor with capital. See Lewis (2011), which studies the interaction of immigration with capital. And if a firm faces frictions in finding a new employee that limit the degree of crowdout of other workers, the amount of crowdout we detect is all the more notable.

decrease in demand for otherwise available workers. The evidence in Section 2 about the types of tasks H-1B workers likely perform is not very consistent with this hypothesis.

Turning to patenting, we find generally small and insignificant effects.<sup>31</sup> This result is not especially surprising given the likely occupations reported in Section 2. While our lottery results are not an artifact of focusing on small non-innovative firms hiring workers with relatively low education – firms applying on the day the cap is reached are more likely than other applicants to have patented in the past, to be in scientific industries, and to apply for workers with higher educational degrees – the ability to focus on lottery effects in large or especially innovative firms where most patenting is concentrated is limited by statistical power. However, this is not the case for our other strategies.

In Table 5 Panel A, we restrict the sample to innovative firms by dropping those that have never patented in prior periods and examining effects using all of our identification strategies. The results confirm that several strategies (in particular, the 2008 quasi-lottery, shift-share, and OLS firm fixed effect regressions) are indeed sufficiently powered to produce precise estimates for the small subsample of innovative firms, showing no increase in patenting due to an additional H1B. In Panel B, we focus only on firms with over 30 employees in the preperiod, given that patenting is more prevalent among larger firms, and we find similar results. Thus, our findings apply even when just considering more innovative firms. Additionally, the R&E credit results tend to be clustered around zero or somewhat negative. While not always as precise as they are for patenting, the combined R&E and patenting results paint the picture that both firm inputs into the innovative process and innovative output do not see meaningful increases from additional H-1Bs.

---

<sup>31</sup> The extremely small positive estimates from firm fixed effects regressions are offset by the modest decreases in patenting found in some 2008 quasi lottery and shift-share estimates

Finally, we find some evidence of a decrease in payroll per employee and increase in profits in our lottery setting. The impact on payroll per employee could be coming from multiple mechanisms. First, employers may pay H-1Bs less than the average wage. This could be due to monopsony power that the employer has due to the individual's H-1B status. Second, employment of H-1Bs could reduce wages paid to H-1B workers, which combined with lower employment of other workers, suggests that the effect of H-1Bs is to lower the demand for other workers. In turn, the positive indication on profits along with payroll per employee results suggest that the firm is able pay its employees less without a similar sized drop in output, that is, the firm is able to extract (additional) rents from labor. That said, the results on profits and payroll per employee are less precise and robust across strategies than our main outcomes and thus more speculative.

## **7. Conclusion**

The effect of highly skilled immigrants on firms is one of the centrally important U.S. immigration policy questions. We examine the impact of the United States' largest high-skill immigration program, the H-1B program, and find that new H-1Bs crowd out other workers associated with similar observable levels of innovation. This result is informative for understanding the effects of these individual workers on firms, and for understanding the labor market for high-skilled technical workers in the U.S. and nature of labor market substitution. We bring several new advances to the literature, including randomized visa lotteries and IRS data on the universe of U.S. firms. We apply this IRS data not only to the new randomized visa lotteries strategy, but also to a new regression discontinuity design, as well as other strategies that the prior literature has relied upon that we now apply to the firm level, drawing on the universe of firms.

Taken together, these identification strategies run the gamut from precise causal identification on small subsets of firms to more general correlative statements about the universe of U.S. firms, and from an analysis of years in which demand for H-1Bs moderately outstripped supply to an analysis of years in which demand for H-1Bs greatly outstripped supply. The primary finding that across all identification strategies one additional H-1B leads to significant crowd out of other workers at the firm level holds whenever the sample we study is sufficiently powered to produce precise estimates, sometimes even when including very large firms. This is relevant in light of frequent claims that H-1Bs have unique skills that cannot easily be obtained elsewhere.

Likewise, across all identification strategies and all years, one additional H-1B does not cause any meaningful increase in patenting. In contrast, point estimates of the effects of H-1Bs on patenting are typically negative, and standard errors are typically small enough to rule out more than a small percentage or absolute effect. These findings are not an artifact of small numbers of marginal H-1Bs being denied due to the cap, as they also hold even for identification strategies such as the 2008 quasi-lottery and the 2004 cap reduction which involve substantial numbers of H-1Bs at stake in the economy as a whole and in each firm.

Consistent with firm profit maximization, we find some evidence that extra H-1B visas increase firm profits. We also find some evidence that extra H-1B visas lead to a decrease in earnings per employee, especially in the lottery identification strategy. If these findings reflect higher economic profits and/or lower pay for H-1Bs than for alternative workers, then this would suggest the existence of market frictions, such as firm labor market monopsony power.

Overall, our results are more supportive of the view that H-1Bs crowd out alternative workers, are paid less per unit of effort than the alternative workers whom they crowd out, and

thus increase the firm's profits despite no measurable effect on innovation. *Prima facie*, these results appear at odds with a chief goal of the program, as articulated by policy-makers in legislation, of providing firms with skilled workers who have unique, innovative skills that the firms cannot otherwise obtain. Even though firms attest that hiring the H-1B does not adversely affect similarly employed workers, our results raise this possibility.<sup>32</sup> Future research should investigate whether H-1B's pay is consistent with prevailing wage regulations and whether employers meet the test of being unable to hire a comparable worker. And while we find little effect on firms' quantity of innovation, assessing impacts on productivity should be a priority for further research.

Our study only estimates partial-equilibrium effects, not general equilibrium effects. In particular, we isolate the effect of additional H-1B visas allocated to a given firm on outcomes at that firm (holding all else equal). However, these should be a key determinant of the general equilibrium effects. If the crowded-out workers become employed in other firms (i.e., assuming labor demand is not perfectly inelastic) and innovate there, aggregate employment and patenting could still increase, as long as this boost to innovation does not crowd out innovation elsewhere. These or other mechanisms could reconcile positive aggregate effects with small firm-level effects. However, this mechanism for raising employment and innovation would be very different than the hypothesis that H-1Bs directly raise employment and innovation at the firm level as well, as both the business community and policy makers have claimed, and suggests smaller aggregate effects than would otherwise be expected.

---

<sup>32</sup> Our results do not necessarily imply that firms' behavior is inconsistent with their attestations, for example because the Congressional intent may have been to prevent harm to U.S. citizens specifically.



## References

- Acs, Zoltan, and David Audretsch. 1990. *Innovation and Small Firms*. MIT Press.
- Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. 2018. "Who becomes an inventor in America? The importance of exposure to innovation." *The Quarterly Journal of Economics*, 134(2), 647-713.
- Borjas, George. 1994. "The Economics of Immigration." *Journal of Economic Literature* 32: 1667–1717.
- Borjas, George. 2003. "The Labor Demand Curve Is Downward-Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics* 118: 1335-1374.
- Borjas, George, and Kirk Doran. 2012. "The Collapse of the Soviet Union and the Productivity of American Mathematicians." *Quarterly Journal of Economics* 127.3: 1143-1203.
- Borjas, George, Richard Freeman, and Lawrence Katz. 1997. "How Much Do Immigration and Trade Affect Labor Market Outcomes?" *Brookings Papers on Economic Activity*: 1-90.
- Borjas, George, Jeffrey Grogger, and Gordon Hanson. 2012. "Comment: Substitution between Immigrants, Natives, and Skill Groups." *Journal of the European Economic Association* 10.1: 198-210.
- Bound, John, Breno Braga, Joseph Golden, and Gaurav Khanna. 2015. "Recruitment of Foreigners in the Market for Computer Scientists in the U.S." *Journal of Labor Economics*, vol. 33, issue S1.
- Burbidge, John, Lonnie Magee, and A. Leslie Robb. 1988. "Alternative Transformations to Handle Extreme Values of the Dependent Variable." *Journal of the American Statistical Association* 83: 123-127.
- Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica* 82.6: 2295-2326.
- Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review* 43.2: 245-257.
- Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration." *Journal of Labor Economics* 19.1: 22-64.
- Case, Steve. 2016. "Fixing America's Broken Immigration System." *CNBC* 20 June 2012. Web 25 January.
- Depew, Briggs, Peter Norlander, and Todd Sorensen. 2013. "Flight of the H-1B: Inter-Firm Mobility and Return Migration Patterns for Skilled Guest Workers." IZA Discussion Paper No. 7456.
- Doran, Kirk, Alexander Gelber, and Adam Isen. 2014. "The Effect of High-Skilled Immigration on Patenting and Employment: Evidence from H-1B Visa Lotteries." Working Paper no. 20668 (November), NBER, Cambridge, MA.
- Dustmann, Christian, Albrecht Glitz, and Tommaso Frattini. 2008. "The Labour Market Impact of Immigration." *Oxford Review of Economic Policy* 24.3: 477-494.
- Edin, Per-Anders, Peter Fredriksson, and Olof Åslund. 2003. "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment." *The Quarterly Journal of Economics* 118.1: 329-357.
- Foley, Fritz, and William Kerr. 2013. "Ethnic Innovation and U.S. Multinational Firm Activity." *Management Science* 59.7: 1529-1544.
- Freeman, Richard. 2006. "People Flows in Globalization." *Journal of Economic Perspectives* 20.2: 145–170.
- Friedberg, Rachel M. 2001. "The Impact of Mass Migration on the Israeli Labor Market." *The Quarterly Journal of Economics* 116.4: 1373-1408.

- Friedberg, Rachel, and Jennifer Hunt. 1995. "The Impact of Immigrants on Host Country Wages, Employment and Growth." *Journal of Economic Perspectives* 9.2: 23-44.
- Grogger, Jeff, and Gordon Hanson. "Attracting Talent: Location Choices of Foreign-Born Ph.D.s in the U.S." Forthcoming, *Journal of Labor Economics*.
- Hanson, Gordon. 2009. "The Economic Consequences of International Migration." *Annual Review of Economics* 1: 179-208.
- Hunt, Jennifer. 2011. "Which Immigrants Are Most Innovative and Entrepreneurial? Distinctions by Entry Visa." *Journal of Labor Economics* 29.3: 417-457.
- Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. 2010. "How Much Does Immigration Boost Innovation." *American Economic Journal: Macroeconomics* 2.2: 31-56.
- Kerr, William R. 2013. "U.S. High-Skilled Immigration, Innovation, and Entrepreneurship: Empirical Approaches and Evidence." Harvard Business School Working Paper: 14-17.
- Kerr, William R., and William F. Lincoln. 2010. "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention." *Journal of Labor Economics* 28.3: 473-508.
- Levin, Sharon, and Paula Stephan. 1991. "Research Productivity over the Life Cycle: Evidence for Academic Scientists." *American Economic Review* 81.1: 114-132.
- Lewis, Ethan. 2011. "Immigration, Skill Mix, and Capital Skill Complementarity." *Quarterly Journal of Economics* 126.2: 1029-1069.
- Lowell, B. Lindsay. 2000. "H-1B Temporary Workers: Estimating the Population." UCSD Center for Comparative Immigration Studies Working Paper No. 12.
- Lubotsky, Darren. 2007. "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings." *Journal of Political Economy* 115.5: 820-867.
- Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German Jewish Émigrés and U.S. Invention." *American Economic Review* 104: 3222-3255.
- North, David. 2011. "Estimating the Size of the H-1B Population in the U.S." *Center for Immigration Studies Memorandum*.
- Pekkala Kerr, Sari, and William R. Kerr. 2011. "Economic Impacts of Immigration: A Survey." *Finnish Economic Papers* 24.1: 1-32.
- Pekkala Kerr, Sari, William R. Kerr, and William F. Lincoln. 2015. "Skilled Immigration and the Employment Structures of U.S. Firms." *Journal of Labor Economics* vol. 33, issue S1, S147 - S186
- Pence, Karen. 2006. "The Role of Wealth Transformations: An Application to Estimating the Effect of Tax Incentives on Saving." *The B.E. Journal of Economic Analysis & Policy* 5.1: 1-24.
- Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. "STEM Workers, H-1B Visas, and Productivity in US Cities." *Journal of Labor Economics* S1: S225-S255.
- Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. "The Effects of Foreign Skilled Workers on Natives: Evidence from the H-1B Visa Lottery." UC Davis Working Paper.
- Senate Judiciary Committee. 1998. *Report to Accompany American Competitiveness Act*, United States Senate, 105<sup>th</sup> Congress. S. Rep. No. 105-186.
- Senate Judiciary Committee. 2000. *Report to Accompany American Competitiveness in the Twenty-First Century Act*, United States Senate, 106<sup>th</sup> Congress. S. Rep. No. 106-260.
- U.S. Customs and Immigration Services. 2006, 2007. "Characteristics of Specialty Worker Occupations." Washington, D.C.: U.S. Government Printing Office.
- U.S. Patent and Trade Office. 2012. "Performance and Accountability Report: Fiscal Year 2012." Washington, D.C.: U.S. Government Printing Office.
- Yagan, Danny. 2014. "Moving to Opportunity? Migratory Insurance over the Great Recession." UC Berkeley Working Paper.

**Table 1. Summary Statistics for Full Merged IRS Data**

|   | All Firms<br>(1) | H-1B Firms<br>(2) | Lottery Firms<br>(3) |
|---|------------------|-------------------|----------------------|
| Variable                                    | Mean (SD)        | Mean (SD)         | Mean (SD)            |
| Number of employees (all)                   | 22.23 (827.55)   | 422.88 (7503.84)  | 1,877.84 (39,721.31) |
| Number of employees ( $\leq 30$ )           | 6.09 (11.36)     | 12.45 (94.91)     | 43.09 (1,904.34)     |
| Number of employees ( $\leq 10$ )           | 3.78 (9.32)      | 7.11 (105.75)     | 9.64 (55.63)         |
| Median employees (all)                      | 4                | 20                | 31                   |
| Median employees ( $\leq 30$ )              | 4                | 8                 | 10                   |
| Median employees ( $\leq 10$ )              | 3                | 5                 | 6                    |
| Winsorized emp. first diff. (all)           | 0.21 (3.57)      | 3.20 (14.66)      | 27.28 (92.39)        |
| Winsorized emp. first diff. ( $\leq 30$ )   | 0.13 (1.58)      | 1.01 (3.04)       | 4.35 (9.43)          |
| Winsorized emp. first diff. ( $\leq 10$ )   | 0.12 (1.13)      | 0.75 (1.96)       | 3.22 (6.84)          |
| Number of patents (all)                     | 0.070 (5.15)     | 2.17 (39.38)      | 4.52 (56.11)         |
| Number of patents ( $\leq 30$ )             | 0.011 (0.33)     | 0.17 (1.93)       | 0.23 (8.59)          |
| Number of patents ( $\leq 10$ )             | 0.007 (0.26)     | 0.10 (1.27)       | 0.023 (0.49)         |
| IHS of patents (all)                        | 0.012 (0.173)    | 0.23 (0.81)       | 0.15 (0.80)          |
| IHS of patents ( $\leq 30$ )                | 0.005 (0.099)    | 0.06 (0.37)       | 0.017 (0.22)         |
| IHS of patents ( $\leq 10$ )                | 0.003 (0.081)    | 0.04 (0.28)       | 0.010 (0.14)         |
| IHS of R&E (all)                            | 0.049 (0.788)    | 0.70 (2.98)       | 1.55 (4.74)          |
| IHS of R&E ( $\leq 30$ )                    | 0.014 (0.399)    | 0.24 (1.69)       | 0.15 (1.39)          |
| IHS of R&E ( $\leq 10$ )                    | 0.007 (0.290)    | 0.14 (1.31)       | 0.14 (1.22)          |
| Fraction with R&E (all)                     | 0.004 (0.063)    | 0.054 (0.225)     | 0.099 (0.30)         |
| Fraction with R&E ( $\leq 30$ )             | 0.001 (0.036)    | 0.020 (0.141)     | 0.013 (0.11)         |
| Fraction with R&E ( $\leq 10$ )             | 0.001 (0.026)    | 0.012 (0.110)     | 0.013 (0.11)         |
| Median payroll per employee (all)           | \$16,800         | \$37,944          | \$49,332             |
| Median payroll per employee ( $\leq 30$ )   | \$16,433         | \$33,930          | \$42,281             |
| Median payroll per employee ( $\leq 10$ )   | \$16,049         | \$31,075          | \$38,657             |
| Median firm profits (all)                   | \$11,308         | \$45,844          |                      |
| Median firm profits ( $\leq 200$ )          |                  |                   | \$80,250             |
| Median firm profits ( $\leq 30$ )           | \$9,822          | \$27,928          | \$43,301             |
| Median firm profits ( $\leq 10$ )           | \$8,196          | \$20,454          | \$30,398             |
| New H-1Bs as share of employ. (all)         | 0.00079 (0.08)   | 0.075 (0.75)      |                      |
| New H-1Bs as share of employ. ( $\leq 30$ ) | 0.00081 (0.08)   | 0.12 (0.95)       |                      |
| New H-1Bs as share of employ. ( $\leq 10$ ) | 0.00085 (0.09)   | 0.16 (1.19)       |                      |
| Fraction winning lottery                    |                  |                   |                      |
| 2006 Regular                                |                  |                   | 0.04                 |
| 2006 ADE                                    |                  |                   | 0.17                 |
| 2007 Regular                                |                  |                   | 0.98                 |
| 2007 ADE                                    |                  |                   | 0.55                 |
| Fraction in NAICS=54 (all)                  | 0.11 (0.31)      | 0.32 (0.47)       | 0.56                 |
| Fraction in NAICS=54 ( $\leq 30$ )          | 0.11 (0.32)      | 0.36 (0.48)       | 0.66                 |
| Fraction in NAICS=54 ( $\leq 10$ )          | 0.12 (0.32)      | 0.37 (0.48)       | 0.65                 |

Notes: The data are from IRS, USCIS, and USPTO administrative sources. The “All Firms” column refers to the universe of firm-year observations in the IRS data. The “H-1B Firms” column refers to firms listed as hiring at least one new H-1B during fiscal years 2003 through 2008. The “all” rows refer to firm-year observations of all employee sizes; the “ $\leq 30$ ” (“ $\leq 10$ ”) rows refer to those firms with 30 (10) or fewer employees in a given year. Number of patents refers to approved patents in each year from Year 0 to 2013. Employment data are observed in Q1-Q4, the first four quarters when the H-1B worker may work at the firm. R&E, payroll per employee, and firm profits are measured in Years 0 to 3, the duration of the H-1B visa. We pool and stack time periods. For profits, we use the size category with  $\leq 200$  employees; our regressions did not converge for larger thresholds. NAICS code 54 is professional, scientific, and technical services. For R&E, the sample size is also smaller because the data only measure the R&E credit for C-corporations. The fraction patenting or with the R&E refer to the mean of a yearly patenting dummy in Years 0 to 8, or to the mean of a yearly dummy for taking the R&E credit in Years 0 to 3. Here and throughout the paper dollar amounts (*e.g.* the R&E credit) are measured in real \$2014.

**Table 2. Effect of H-1B Lottery Wins on Employment of Other Workers**

| Employment in first year (Q1 through Q4):         |                            |  |   |
|---|----------------------------|--|---|
|   | (1) Median regressions     | (2) 2SLS mean regressions<br>with firm fixed effects | (3) 2SLS mean regressions<br>with differences |
| A) $\leq 10$ employees                            | -1.52<br>[-2.15, -0.89]*** | -2.09<br>[-3.85, -0.33]**                            | -2.10<br>[-3.88, -0.32]**                     |
| B) $\leq 30$ employees                            | -1.36<br>[-2.09, -0.63]*** | -2.13<br>[-4.21, -0.05]**                            | -2.26<br>[-4.25, -0.29]**                     |
| C) All  | -2.05<br>[-3.67, -0.43]**  | -12.39<br>[-65.60, 40.82]                            | -3.41<br>[-18.76, 11.94]                      |
| Employment in first three years (Q1 through Q12): |                            |  |   |
|   | (1) Median regressions     | (2) 2SLS mean regressions<br>with firm fixed effects | (3) 2SLS mean regressions<br>with differences |
| A) $\leq 10$ employees                            | -1.64<br>[-2.30, -0.88]*** | -2.60<br>[-5.38, 0.19]*                              | -2.75<br>[-5.42, -0.07]**                     |
| B) $\leq 30$ employees                            | -1.74<br>[-2.61, -0.87]*** | -2.56<br>[-5.46, 0.34]*                              | -2.98<br>[-5.66, -0.30]**                     |
| C) All  | -2.95<br>[-5.37, -0.52]*** | -29.74<br>[-98.57, 39.09]                            | -11.00<br>[-31.68, 9.67]                      |

Notes: The table shows coefficients of a chance H-1B visas, minus the H-1B visa worker the firm won in the lottery, with 95 percent confidence intervals in brackets. The first column shows median regressions of firm employment on chance lottery wins, defined as actual wins minus the expectation of wins conditional on number of applications and the probability each application wins. The next two columns show 2SLS (mean) regressions where the dependent variable in column 2 is firm employment in the quarter of question winsorized at the 95<sup>th</sup> percentile (with firm-lottery fixed effects and the observations from the first quarter of Year -1) and in column 3, the difference of firm employment from the first quarter of Year -1 to the quarter in question, winsorized at the 95<sup>th</sup> percentile. We pool and stack observations across quarters. The first panel examines employment from Q1 to Q4 while the second panel examines employment from Q1 to Q12. All specifications control for employment in this pre-period and expected lottery wins (equal to number of H-1B applications entering a lottery multiplied by the probability of winning the lottery). The 5<sup>th</sup> and 95<sup>th</sup> percentiles of the first difference in employment are -109 and 352, respectively, in the full sample; -9 and 30, respectively, among those with 30 or fewer employees; and -6 and 22, respectively, among those with 10 or fewer. In these regressions, the instrument is chance lottery wins and the endogenous variable is approved capped H-1B visas. The regressions include controls for employment from the first quarter of Year -1, and the number of expected lottery wins. See Table 1 for other notes and sample sizes. \*\*\* denotes  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.10$ . If the H-1B worker works at the firm, a coefficient of 0 corresponds to neither crowd-out or nor crowd-in of other employment, and a coefficient of -1 corresponds to one-for-one-crowdout of other employment.

**Table 3. Effect of H-1B Lottery Wins on Other Firm Outcomes**

|                  | IHS of #<br>of patents<br>(1) | # of patents<br>(negative<br>binomial)<br>(2) | Amount of<br>RHE credit<br>(IHS)<br>(3) | Claiming RHE<br>dummy<br>(4) | Payroll per<br>employee<br>(5) | Profits<br>(6)             |
|------------------|-------------------------------|---|---|------------------------------|--------------------------------|----------------------------|
| A) ≤10 employees | 0.00026<br>[-0.0042, 0.0047]  | -0.0044<br>[-0.0108, 0.0021]                  | -0.12<br>[-0.27, 0.041]                 | -0.011<br>[-0.025, 0.0041]   | -4,861<br>[-9,553, -168]**     | 6,519<br>[-6,943, 19,979]  |
| B) ≤30 employees | 0.0018<br>[-0.0094, 0.013]    | -0.0073<br>[-0.0239, 0.0093]                  | -0.065<br>[-0.15, 0.018]                | -0.0061<br>[-0.014, 0.0016]  | -2,725<br>[-5,977, 527]*       | 11,469<br>[201, 22,736]**  |
| C) All           | -0.0089<br>[-0.037, 0.019]    | -0.067<br>[-0.18, 0.04]                       | 0.19<br>[-0.33, 0.72]                   | 0.016<br>[-0.018, 0.049]     | 80<br>[-1,348, 1,509]          | 2,527<br>[-32,169, 37,222] |

Notes: The table shows regressions of patents in each year on chance H-1B lottery wins. Specification (1) is an OLS regression with IHS of patents as the dependent variable. Specification (2) is a negative binomial regression with # of patents as the dependent variable. Controlling for “prior patents” refers to controlling for the IHS of the total number of patents in Year -1. The coefficients in the IHS specifications should be interpreted as the approximate percent effect on the number of patents. In specification (3) the dependent variable is the IHS of the amount of the R&E credit claimed in each year over Years 0 to 3. In specification (4) the dependent variable is a dummy variable for whether the firm claimed any R&E credit in each year from Years 0 to 3, so that the coefficient reflects the effect on the fraction of years claiming any R&E. The “Prior R&E” control refers to controlling for the amount (in Columns 1 and 2) or presence (in Columns 3 and 4) of the R&E credit in Year -1. The IRS data only measure the R&E credit for C-corporations; other firms are excluded from the regressions. We find comparable results at other size thresholds; no significant interactions with covariates; and no significant differences across groups. The coefficients in the IHS specifications should be interpreted as the approximate percent effect on the amount of R&E taken. Specification (5) shows a median regression of payroll costs per employee in Years 0 to 3 on chance H-1B visas and controls, pooling and stacking years. Years 0 to 3 cover the duration of the H-1B visa. The effect on payroll per employee in Years 0 to 1 is comparable to the estimates shown. Payroll costs per employee in a given year is measured as total firm payroll costs in that year (in real \$2014) divided by the total number of employees in the firm in that year. We use W-2 data because median regressions using form 941 data generally did not converge. Specification (6) shows median regressions of profits in Years 0 to 3 on chance H-1B visas and controls, pooling and stacking years. Profits are measured in real \$2014. In Row C we investigate firms with 200 or fewer employees because regressions above this firm size cutoff did not reliably converge; they did not converge, for example, in the sample of firms of all sizes. Years 0 to 3 cover the duration of the H-1B visa. See Table 1 for additional notes. Standard errors are clustered by firm. \*\*\* refers to significance at the 1% level; \*\* at the 5% level, and \* at the 10% level.

**Table 4. Comparison of Identification Strategies: Estimates of Effect of H-1Bs on Firm Outcomes**

|   | Employment of<br>other workers<br>(1) | IHS patents<br>(2)                | IHS R&E tax credit<br>(3)       | Pay per emp.<br>(4)        | Profits<br>(5)                   |
|---|---------------------------------------|-----------------------------------|---------------------------------|----------------------------|----------------------------------|
| <b>Panel A: ≤10 employees</b>                   |                                       |                                   |                                 |                            |                                  |
| Experimental                                    | -1.52<br>[-2.15, -0.89]***            | 0.00026<br>[-0.0042, 0.0047]      | -0.12<br>[-0.27, 0.041]         | -4,861<br>[-9,553, -168]** | 6,519<br>[-6,943, 19,979]        |
| Regression Discontinuity                        | -1.46<br>[-2.38, -0.54]***            | -0.002<br>[-0.036, 0.031]         | -0.43<br>[-0.89, 0.04]*         | -281<br>[-3117, 2554]      | -3585<br>[-15451, 8280]          |
| 2008 Lottery Winners vs.<br>Proxy Control Group | -0.78<br>[-0.83, -0.73]***            | -0.0005<br>[-0.00010, -0.00001]** | -0.0047<br>[-0.0082, -0.0012]** | -42.79<br>[-158.41, 72.82] | 5506<br>[640, 10373]**           |
| Shift-Share                                     | -1.24<br>[-1.41, -1.08]***            | 0.0046<br>[-0.013, 0.022]         | -0.14<br>[-0.38, 0.09]          | -0.033<br>[-0.082, 0.016]  | -1.22<br>[-2.33, -0.18]**        |
| Fixed Effects                                   | -0.83<br>[-0.96, -0.70]***            | 0.00059<br>[0.00025, .00094]**    | 0.00062<br>[-0.0019, 0.0032]    | 76.90<br>[-0.12, 153.92]*  | 8,442<br>[-6,801, 23,685]        |
| <b>Panel B: ≤30 employees</b>                   |                                       |                                   |                                 |                            |                                  |
| Experimental                                    | -1.36<br>[-2.09, -0.63]***            | 0.0018<br>[-0.0094, 0.013]        | -0.065<br>[-0.15, 0.018]        | -2,725<br>[-5,977, 527]*   | 11,469<br>[201, 22,736]**        |
| Regression Discontinuity                        | -1.60<br>[-2.90, -0.29]**             | -0.003<br>[-0.033, 0.026]         | -0.32<br>[-0.76, 0.12]          | 417<br>[-1451, 2286]       | -1080<br>[-14029, 11868]         |
| 2008 Lottery Winners vs.<br>Proxy Control Group | 0.73<br>[-0.79, -0.69]***             | -0.0015<br>[-0.0025, -0.0005]***  | 0.0016<br>[-0.0090, 0.0122]     | -55.38<br>[-157.94, 47.19] | 49008<br>[-36216, 134233]        |
| Shift-Share                                     | -1.27<br>[-1.41, -1.13]***            | 0.0027<br>[-0.013, 0.018]         | -0.10<br>[-0.24, 0.04]          | -0.028<br>[-0.068, 0.013]  | -1.05<br>[-1.93, -0.17]**        |
| Fixed Effects                                   | -0.87<br>[-0.98, -0.75]***            | 0.00077<br>[0.00041, 0.00024]**   | 0.00045<br>[-0.0015, 0.0024]    | -5.57<br>[-58.55, 47.40]   | -11,027<br>[-56,701, 34,647]     |
| <b>Panel C: All</b>                             |                                       |                                   |                                 |                            |                                  |
| Experimental                                    | -2.05<br>[-3.67, -0.43]**             | -0.0089<br>[-0.037, 0.019]        | 0.19<br>[-0.33, 0.72]           | 80<br>[-1,348, 1,509]      | 2,527<br>[-32,169, 37,222]       |
| Regression Discontinuity                        | -29.89<br>[-82.04, 22.25]             | -0.020<br>[-0.049, 0.009]         | 0.18<br>[-0.30, 0.66]           | 529<br>[-518, 1575]        | 10689<br>[-2023, 23402]*         |
| 2008 Lottery Winners vs.<br>Proxy Control Group | -0.64<br>[-1.42, 0.14]*               | -0.000015<br>[-0.00058, 0.00055]  | 0.012<br>[0.0030, 0.021]**      | -0.20<br>[-5.81, 5.42]     | -5958330<br>[-1.19e+07, 28514]*  |
| Shift-Share                                     | -1.29<br>[-1.41, -1.17]***            | -0.022<br>[-0.040, -0.004]**      | -0.34<br>[-0.59, -0.09]***      | -0.018<br>[-0.053, 0.017]  | -0.90<br>[-1.65, -0.15]**        |
| First-Differences                               | -0.97<br>[-1.30, -0.64]**             | 0.00012<br>[-0.00001, 0.00024]*   | 0.00095<br>[-0.0018, 0.0037]    | -0.86<br>[-2.6, 0.88]      | 1,957,148<br>[-1631617, 5545912] |

Notes: The table shows estimates of the effect of new H-1Bs on firm outcomes using various identification strategies. The rows labeled “Experimental” are taken from Tables 2 and 3. For the experimental identification strategy, specification (1) refers to column (1) of Table 2 Panel A, while specifications (2)-(5) come from Table 3. The dependent variables for all other strategies are the same as for the base lottery estimates, except we use unwinsorized means given the sufficient statistical power afforded by the other approaches. The employment regression coefficients all subtract 1 to represent the effect on employment of other workers. For the rows labeled “Regression Discontinuity,” the number of new H-1Bs is instrumented by the firm submitted its visa application before the day the cap was reached and the (quadratic) running variable is the day in which the firm submitted its visa application(s). The effect of a firm applying before the day the cap is reached on the # of H-1Bs approved for that firm is 0.601 (0.050)\*\*\* for the 2006 ADE lottery, 0.976 (0.025)\*\*\* for the 2006 regular lottery, and 0.459 (0.057)\*\*\* for the 2007 ADE lottery. Similar first stages are observed when examining firms with thirty or fewer or ten or fewer employees. We restrict

the sample to the 76.4% of all firms (73.32% of applications) that submitted all their applications in any given visa category-year on one day (in the absence of this restriction, there is no discontinuity at the *firm* level). All regressions use the optimal bandwidth of 20 days. All specifications include visa category by year fixed effects, lagged dependent variable, and a constant. The number of observations varies from 1,881 (Panel A, Column 6) to 58,452 (Panel C, Column 3). The rows labeled “2008 Lottery Winners vs. Proxy Control Group” are estimated via regressions that control for expected wins based on a control group proxy of LCAs, lagged dependent variable, and a constant. The numbers of observations vary from 4,977 (Column 1, Panel A) to 79,870 (Column 2, Panel C). The rows labeled “Shift-Share” show the results from the shift-share strategy which controls for firm fixed effects, indicators for year, an interaction of a time trend with the pre-period value of each dependent variable, and a constant and uses the IHS of the dependent variable for interpretation as a percent change on a percent change. Given the shift-share setup, the dependent variables are contemporaneous to the number of approved H-1B. F-statistics in the first stages range from 66 to 143. The numbers of observations vary from 9,826 (Column 4, Panel A) to 77,930 (Column 2, Panel C). The rows labeled “Firm fixed effects” show OLS regressions of # of new H-1Bs in that firm on firm outcomes. All regressions include year and firm fixed effects and a constant. The numbers of observations vary from 247,759 (Column 5, Panel A.B) to 2,882,375 (Column 2, Panel C.E). Standard errors are clustered at the firm level in all regressions. \*\*\* denotes  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.10$ .

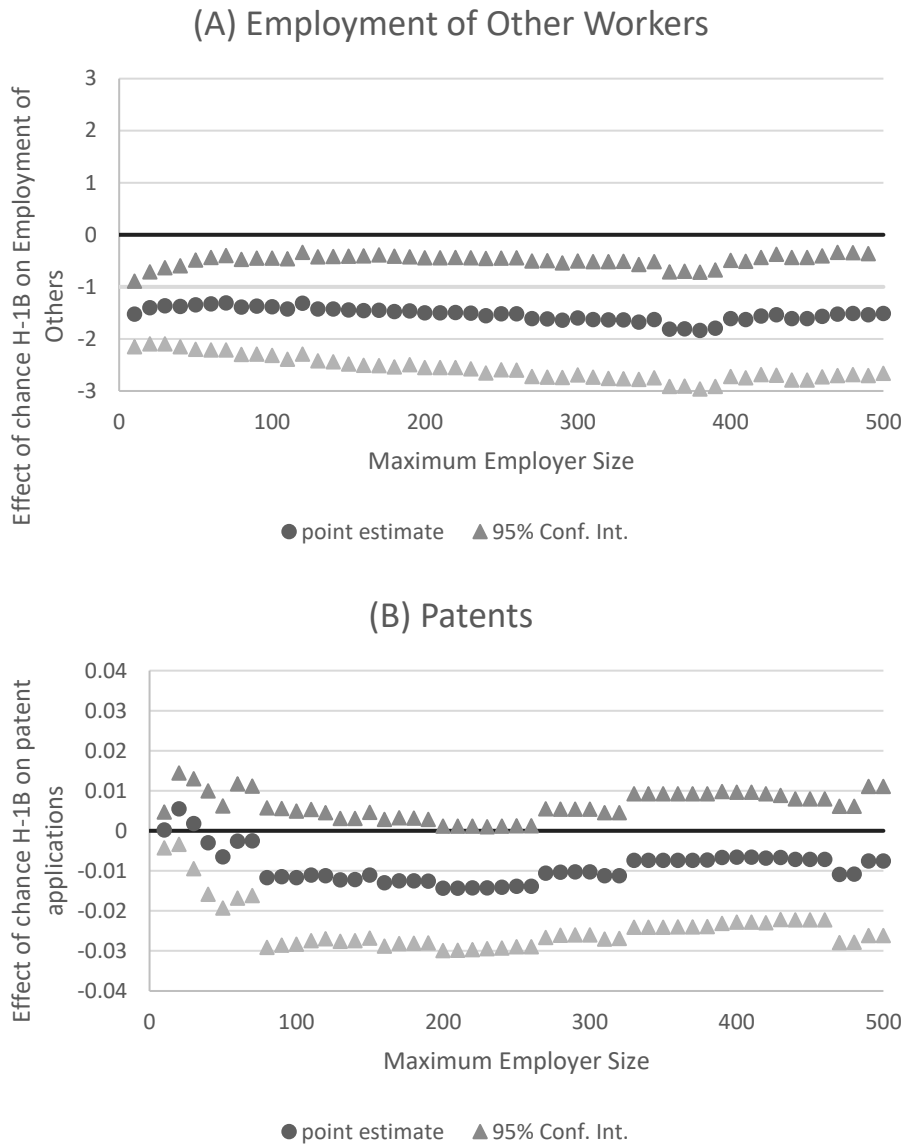
**Table 5. Effect of H-1Bs on Patents in Large or Innovative Firms**

| FY 2006/2007 lotteries<br>(1)                                    | FY 2006/2007 RD<br>(2)    | FY 2008 quasi-lottery<br>(3)     | Shift-share<br>(4)      | OLS firm fixed effects<br>(5)     |
|--|---------------------------|----------------------------------|-------------------------|-----------------------------------|
| <b>Panel A: Firms with at least one patent in the pre-period</b> |                           |                                  |                         |                                   |
| -0.057<br>[-0.30, 0.19]  | -0.17<br>[-0.46, 0.12]    | -0.00064<br>[-0.0013, 1.49e-06]* | -0.034<br>[-0.11, 0.04] | 0.000070<br>[-0.000063, 0.000203] |
| <b>Panel B: Large firms (i.e., pre-employment &gt; 30)</b>       |                           |                                  |                         |                                   |
| -0.023<br>[-0.073, 0.027]  | -0.030<br>[-0.074, 0.013] | -0.00038<br>[-0.00094, 0.00019]  | -0.15<br>[-0.41, 0.10]  | 0.000088<br>[-0.000021, 0.000196] |

Notes: The table shows the results of five identification strategies applied to firms that are either innovative (Panel A, with at least one patent in the pre-period) or large (Panel B, with pre-employment > 30). The specification in each column corresponds with the baseline specification in each case; see Table 4 for more details. The numbers of observations vary from 1,557 (Column 1, Panel A) to 636,605 (Column 5, Panel B). The standard errors are clustered at the firm level.

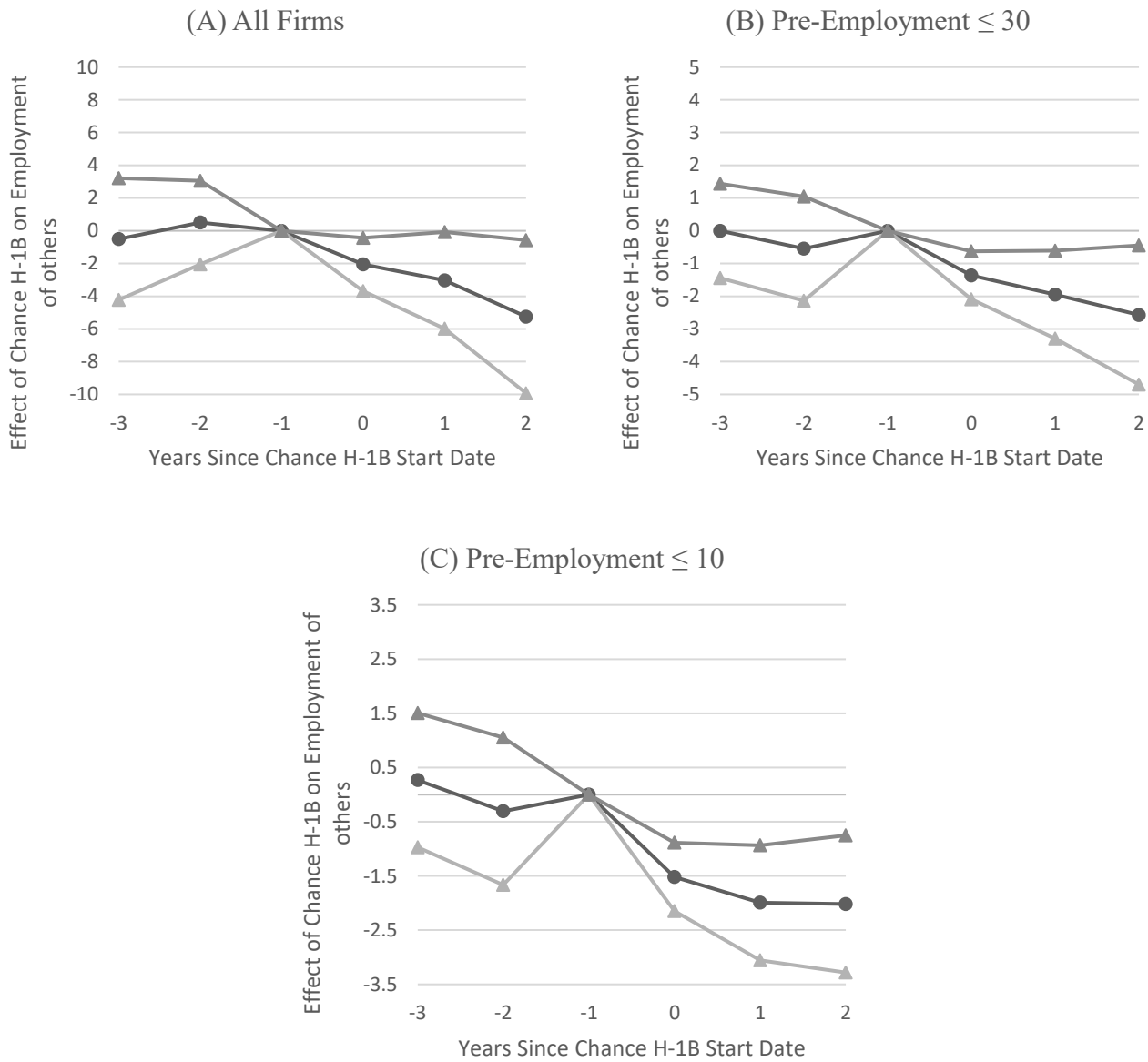


**Figure 1.** *Effect of H-1B Visas on Firm Outcomes, by Employer Size*



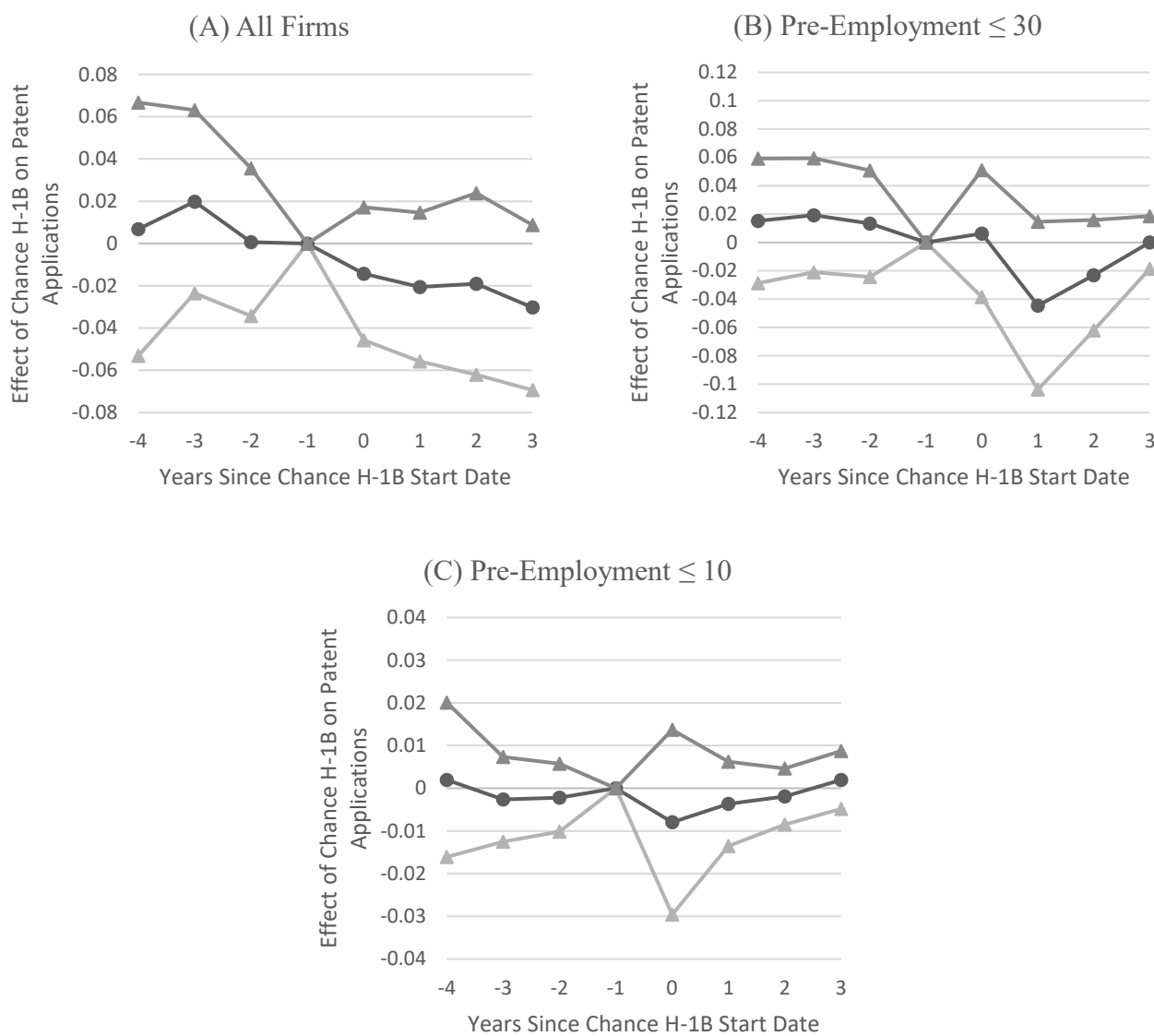
Notes: Panel A of this figure shows the coefficient and 95 percent confidence interval on a chance lottery win, minus the chance H-1B visa worker the firm won in the lottery, from median regressions, pooling together Quarters 1-4 of the first fiscal year that an employee can work at the firm in the regression, among employers of the indicated size or smaller in Year -1 (where the maximum employer size in each case is shown on the *x*-axis). The horizontal line at 0 on the *y*-axis corresponds to the case where hiring an extra H-1B visa worker leaves other employment unchanged. The horizontal line at -1 on the *y*-axis corresponds to the case where hiring an extra H-1B visa worker crowds out other workers one-for-one. Panel B of the figure shows the coefficient and 95 percent confidence interval on chance H-1B visas when the dependent variable is the IHS of patents in each year over Years 0 to 8, among employers of the indicated sizes or smaller in Year -1 (where the maximum employer size in each case is shown on the *x*-axis). After multiplying by 100, the coefficient should be interpreted as the approximate percentage increase in firm patenting due to a chance H-1B visa lottery win. We show the coefficient for employers of each size ranging from 0-10 to 0-500, with the upper bound of the size range in increments of 10. Note that the samples overlap across different regressions; for example, firms with 10 or fewer employees are included in the samples in all 50 regressions shown. We use the baseline employment specification, in which we control for the lagged dependent variable and expected lottery wins.

**Figure 2.** *Effect of Chance H-1B Visa on Employment of Other Workers, by Years since H-1B Lottery and by Employer Size*



Notes: This figure shows the effect of a chance H-1B on firm employment from median regressions, by the years since the H-1B lottery, minus in years 0 through 2 the chance H-1B worker the firm won in the lottery. Because the employment data is quarterly, year 0 includes all quarters in the first year of the visa, year 2 reflects the third year of the visa, and all coefficients are estimated relative to base year -1. Firm employment in years 0-2 represents average quarterly employment each year. Each point estimate (blue) is a coefficient on chance H-1Bs from the baseline specification in Table 2, with the dependent variable varying from year -3 to year 2. The blue lines represent the point estimates, and the gray lines represent 95% confidence intervals.

**Figure 3. Effect of Chance H-1B Visa on Patents, by Years Since H-1B Lottery and by Employer Size**



Notes: This figure shows the effect of a chance H-1B on the inverse hyperbolic sine of firm patenting, by the years since the H-1B lottery. Because the patent data is by calendar year and the visa is on a fiscal year basis, year 0 includes only one treated quarter. All coefficients are estimated relative to base year -1. Each point estimate (blue) is a coefficient on chance H-1Bs from the baseline specification in Table 3, with the dependent variable varying from year -4 to year 3. The blue lines represent the point estimates, and the gray lines represent 95% confidence intervals.

## Appendices (for online publication)

### Road Map of Appendixes:

In Appendix 1, we describe the matching between our various datasets. In Appendix 2, we report robustness checks and heterogeneity of our main lottery estimates across subsamples of firms. In Appendix 3, we discuss the limited data we have on the immigration status of workers within a firm.

### Appendix 1. Description of matching procedures

#### (a) Matching between USCIS lottery data and patents

As noted in the main text, we performed an intentionally liberal automatic matching procedure between the USCIS and patenting datasets to obtain all plausible matches between companies and patents. We then searched through the matches by hand to detect and remove all matches that appeared spurious.

The automatic matching procedure proceeded as follows. First, we assigned clearly related firm names to single categories (*i.e.*, “Sony”, “Sony Co.”, “Sony Corporation”, *etc.*). Then we searched for complete string matches between the name categories in the patenting data, using the full period from 1975 to 2013, and the name categories in the USCIS H-1B visa lottery data, and we classified these as matches between the datasets. After all such matches were made, we then searched for complete string matches between these two sets of name categories with all spaces in the names removed and also classified these as matches. Finally, we performed a “fuzzy” match between USPTO and USCIS firm names. The fuzzy matching procedure calculated a “distance” between words in each list by determining how many characters in the words need to be edited to transform a word from one list into a word in the other. This is necessary to identify all matches because, for example, firm names are occasionally misspelled. Pairs of words in firm name categories were classified as non-matching if the number of characters that differed between the words was more than one for words with six or fewer characters, or when the number of characters that differed between the words was more than two for words with seven or more characters (using the word as spelled in the USCIS data to determine the number of characters in the word). Otherwise, this pair of words was classified as a possible match. If at least 75 percent of the pairs of words in the firm name were possible matches, then the entire firm name was classified as a possible match.

We intentionally designed this “liberal” procedure so that it is liable to classify many non-matches as matches (but not the reverse); thus, if a firm did not match at all between the two datasets according to the fuzzy match, we can be rather certain that it was not granted any US patents between 1975 and 2013. The goal of this automatic matching procedure was to generate a list of all *potential* matches, which we could then winnow by hand in the next step.

Once this automatic matching procedure was complete, all of the resulting matches were checked by hand to determine whether they appeared to be a possible match. Of the 668 companies in the USCIS lottery list that obtained at least one automatic match in the patenting data, we identified 208 cases in which all of that company’s matches were clearly incorrect through by-hand inspection. We further identified 392 cases in which all of that company’s matches were clearly correct (legitimate variations on the correct company name) through by-hand inspection. Finally, we identified 63 cases in which the matches were ambiguous; in our judgment the match is possibly correct, but we cannot be fully confident that it is correct. We

assume that both unmatched companies and those that received clearly incorrect matches did not patent at all between 1975 and 2013. (A firm would not match between the datasets if it did not patent during this time period; thus, under any of our ways of determining which companies were non-matches, we code the non-matching firms as having zero patents.)

In the results that we report in the main tables, we exclude the 63 possible matches from the list of matched companies. In Appendix Table 17, we show that we find comparable results when assuming that the possible matches were in fact matches. The results are also robust to alternative assumptions and similar alternative matching procedures.

### **(b) Matching between IRS data and USCIS annual new H-1B counts by firm**

We match IRS data at the firm-year level with USCIS data on annual new H-1B counts by firm. This USCIS data does not have EINs, so we match it to IRS data using a fuzzy match on firm names. We cannot apply the same technique as the match between USCIS lottery data and U.S. patenting data described above, because that technique was predicated on one of the two datasets being quite small (the USCIS data). Since the USCIS complete count of H-1Bs by firm contains approximately 30,000 separate firm names for each fiscal year, we must apply an alternative fuzzy matching technique that does not involve any by-hand work.

We first put the firm names in the USCIS data into a filter in which each name undergoes the following changes: all letters are changed to upper case; all commas and trailing spaces are removed; common trailing letters are removed (e.g. LP); some common words are removed (e.g. corporation); and other common words are standardized (e.g. “technology” standardized to “tech” and “institute” standardized to “inst”). We then perform an exact match to a similarly filtered firm name in the IRS data. In the resulting match, just under 15% of H-1Bs do not match to any firm name in contemporaneous IRS data. We use the same procedure when matching LCA applications to the IRS data.

### **(c) Matching between IRS data and patent data**

We construct a match between the universe of firms in the IRS data and the universe of U.S. patenting data. We start with a match between individual people in the IRS data and individual patents from (Bell et al., 2019). The authors linked data from individuals in the tax data (both those who filed for taxes and those who did not) with U.S. patents by inventor’s name, city, and state. The authors report that “88% of individuals who applied for or were granted a patent were successfully linked, with higher match rates in more recent years since information returns are unavailable prior to 1999.”

We received permission from the authors to make use of this person-patent match, and we then used tax data to determine whether each inventor was an employee of a firm in the patent application year, and, if so, for which firm each inventor was working. In particular, we identified each person’s employer through their Form W-2. This allows us to determine patent counts by firm and year.

## **Appendix 2. Lottery-based Ancillary Results**

### **First stage considerations**

We find that chance H-1B wins in earlier lotteries have no significant effect on future H-1B applications. In both the cases of FY2006 and FY2007 visas, the Regular visa lottery chronologically occurred on a date before the ADE cap was reached. When we pool FY2006 and

FY2007 and regress total ADE H-1B visa approvals in a given year on chance lottery wins in the Regular lottery in that year, the coefficient on chance lottery wins is -0.20, with a standard error of 0.18 ( $p=0.26$ ). Additionally, chance lottery wins in 2006 have no effect on approved 2007 visas; for example, when regress total FY2007 Regular and ADE approvals (summed) on chance lottery wins in the FY2006 Regular and ADE lotteries combined, the coefficient on chance lottery wins is -0.05, with a standard error of 1.45 ( $p=0.97$ ).

As noted, it is possible that new H-1Bs crowd out other H-1Bs who would have worked at the firm (e.g. H-1Bs not subject to the cap), or other visa types such as L-1s or those participating in OPT. We find an insignificant impact (coefficient -0.03,  $p=0.25$ ) of chance H-1Bs on the number of approved H-1B visas for applications received after the cap was reached. Such applications include those not subject to the cap, e.g. those for citizens of the five countries not subject to the cap. As L-1s are only available to multinationals, it is relevant that our results are similar when we remove multinationals from the sample. OPT applies to workers already in the U.S.; the majority of H-1B applications were for workers previously locating outside the U.S. (USCIS 2006, 2007), though a substantial minority were for those previously in the U.S.<sup>1</sup> The ITT results again are policy-relevant effects of interest, regardless of whether these H-1Bs crowd out other visas.

### **Ancillary employment results**

We perform a number of variations on our basic specifications. Appendix Table 8 shows that in each individual quarter from Q1 to Q4, we typically rule out a coefficient of 0, particularly in smaller firms. Appendix Table 9 shows that chance lottery wins have a precise zero effect on a dummy for being above the 95th, 99th, and other percentiles of employment—demonstrating no effect on being a “star” employer. Appendix Tables 10 and 11 show the robustness of our results to varying controls and treatment of missing data. We also find similar results when we control for fixed effects for each of the four lotteries or a firm’s number of applications in a given lottery interacted with lottery fixed effects, i.e. conditioning on the risk set to which each firm is exposed, or reweighting our sample to match the full population of firms that applied for H-1Bs in FY 2006 and 2007 (results not shown). Finally, we found no evidence that winning a chance H-1B affects whether the firm stays in business (not shown).

In Appendix Table 12, we examine whether there is heterogeneity in the effect on employment across type of lottery or type of industry, using our baseline employment specification in Q1 to Q4 with median regressions. There are no significant differences across the different samples. We find no evidence of significantly different effects across the Regular *vs.* ADE lotteries; professional, scientific, and technical services firms *vs.* firms in other industries; or firms like Infosys or Wipro in industries that often offer outsourcing for temporary support services (often specifically for temporary technical support services) *vs.* other firms; firms that made H-1B applications only on the day of the lottery *vs.* firms that made H-1B applications both on the day of the lottery and on earlier days (relevant to whether the last day shows unique effects); and firms in which the average age of the H-1B for which they are applying is under 27 *vs.* at least 27 at the time the visa begins. Focusing however on point

---

<sup>1</sup> Young H-1Bs could be more substitutable with OPT workers (who are typically students and therefore young) than older H-1Bs, for example if firms denied an H-1B often hire the same worker through an OPT visa. Interestingly, Appendix Table 12 shows more evidence of employment crowdout among H-1Bs over the age of 27 when they begin the H-1B than those under 27. This is not what we might predict if crowdout consisted largely of OPTs substituting for H-1Bs. However, these results do not rule out that such substitution can occur

estimates, they are more negative for the Regular lotteries than for the ADE lotteries, and they are more negative for scientific services (*i.e.* NAICS code 54) than for other industries. The point estimates are also more negative in likely “temporary support services” employers relative to other six-digit industries.

In Appendix Table 13, we explore the effects on a proxy for foreign born vs. native employment, as well as U.S. citizen vs. non-citizen employment; we discuss these results in more detail in Appendix 3 below. In Appendix Table 14, we explore the effects on contract workers, showing that winning the lottery likely decreased the employment of contract workers.

### **Ancillary patenting results**

We perform a number of variations on our basic patenting specifications. We find that the effects are comparable when examining patents during just the duration of the visa or a later period (Appendix Table 15); on patenting within subgroups (Appendix Table 16); when we assume that possible matches between the USCIS and patenting data did match, instead of the baseline assumption that they did not (Appendix Table 17); and when controlling for risk sets for each of the four lotteries, using the log of 1+patents instead of the IHS, and when reweighting our sample to match the full population of H-1B firms (results not shown).

Appendix Table 18 weights each patent by its number of citations, *i.e.* the dependent variable is patent citations. The results rule out more than a small percentage increase in citations, suggesting that the quality of patenting does not increase. Appendix Table 19 shows that the effect on a dummy for whether the firm patented is at most small.

## **Appendix 3. Estimating effects on employment of foreigners and non-foreigners**

### **Measures of foreigners and non-foreigners**

In an exploratory analysis, we investigate how additional new H-1Bs affect employment of other foreigners, and separately affect employment of non-foreigners by using two different approaches that each face their own limitations.

First, citizenship status is available through IRS data, but these data only have information on the individual’s citizenship status *most recently reported* to SSA, as opposed to always being measured in the year in question in our regressions (*e.g.* Year 0 or Year 1). While we use this classification, that data on past citizenship status is not directly available is a relevant limitation because a substantial fraction of H-1Bs go on to become permanent residents and in many cases citizens (Lowell 2000). We use W-2 data for this analysis (rather than form 941 data) because the individual-level W-2 data can be linked to information on citizenship, whereas the form 941 data has no individual employee information available. Using this measure, we find that foreigners constitute a majority (56.30 percent) of workers in our sample of firms.

Given the limitation of the first measure, it is desirable to use a second, unrelated method to probabilistically determine whether individuals are natives or non-natives. Using an algorithm developed in conjunction with Yagan (2014), we identify individuals as natives or non-natives on the basis of individuals’ Social Security Numbers (SSNs) in the data. Prior to 2011, SSNs were assigned in a way that makes it possible to determine with some, but ultimately unknown, degree of confidence whether a given individual is an immigrant to the U.S. or a native. SSNs consisted of: 1) a three-digit “Area Number” representing the area where an individual applied for the SSN; 2) a two-digit “Group Number” that is assigned in a specified sequence *within* each area number; and 3) a four-digit “Serial Number” that is assigned sequentially within each Group

Number.<sup>2</sup> Thus, within a given geographic area associated with the Area Number, it is possible to determine on the basis of the Group Number and the Serial Number whether the individual applied for the SSN at an earlier or a later date. A majority of H-1Bs arrive when they are aged in their late 20s and early 30s. If they eventually apply for an SSN, they will do so well later in life than natives whose applications are typically submitted very early in their lives. Individuals whose SSNs indicate that they applied for the SSN late in life have a substantial probability of being foreign-born, but those whose SSNs indicate that they applied early in life have a much smaller probability of being foreign-born. We follow Yagan (2014) in probabilistically classifying individuals as immigrants when their SSNs indicate that they were in the oldest 10 percent of a given set of SSNs applicants within an Area Number.<sup>3</sup>

### **Estimated effect on employment of foreigners and non-foreigners**

We estimate the effect on employment of foreigners vs. natives in Appendix Table 13. To make the time period investigated with these yearly W-2 data as comparable as possible to the quarterly data shown elsewhere (where we investigate Q1 through Q4 of the first fiscal year, corresponding to observations from both calendar years straddled by Q1 through Q4), we pool Year 0 with Year 1.<sup>4</sup> For these exercises, we do not subtract 1 to reflect the H-1B worker.

In Rows A and B, we measure citizenship using the most recent measure of citizenship in the IRS data. When the dependent variable is the number of non-citizens employed at the firm, in all cases we are able to rule out a coefficient of one or higher—suggesting that new H-1Bs do at least partially crowd out other non-citizens. We are unable to rule out that there is no effect of chance lottery wins on the median number of citizens, but we are always able to rule out that the median number of citizens decreases by one. Thus, we find evidence for crowdout of non-citizens, do not find evidence for crowdout of U.S. citizens, and are able to rule out one-for-one crowdout of citizens (though our results are at the same time consistent with substantial crowdout of citizens). For the reasons described earlier, this measure will overstate (understate) crowdout of noncitizens (citizens). The results could mean that new H-1Bs do not crowd out citizens, but could also mean that H-1Bs sometimes go on to become citizens later. Likewise, the results could indicate that new H-1Bs crowd out other non-citizens, or they could mean that new H-1Bs sometimes become citizens later.

Using our second approach, we show in rows C and D results where we probabilistically identify natives and non-natives using their SSNs. Just as when we use the baseline employment specification, we find evidence for crowdout of non-natives (*i.e.* can rule out a coefficient of 1), do not find definitive evidence for crowdout of natives (*i.e.* the coefficient is insignificantly different from zero in this case), and are able to rule out one-for-one crowdout of natives (*i.e.* can rule out a coefficient of -1)—though the results are also consistent with substantial crowdout of natives. This concordance of results between two very different methods (in Rows A and B vs. C and D) increases our confidence that new H-1Bs at least partially crowd out other foreigners. However, note that whether an individual is a native is not the same as whether s/he is a citizen, so the results are not directly comparable across the two measures.

---

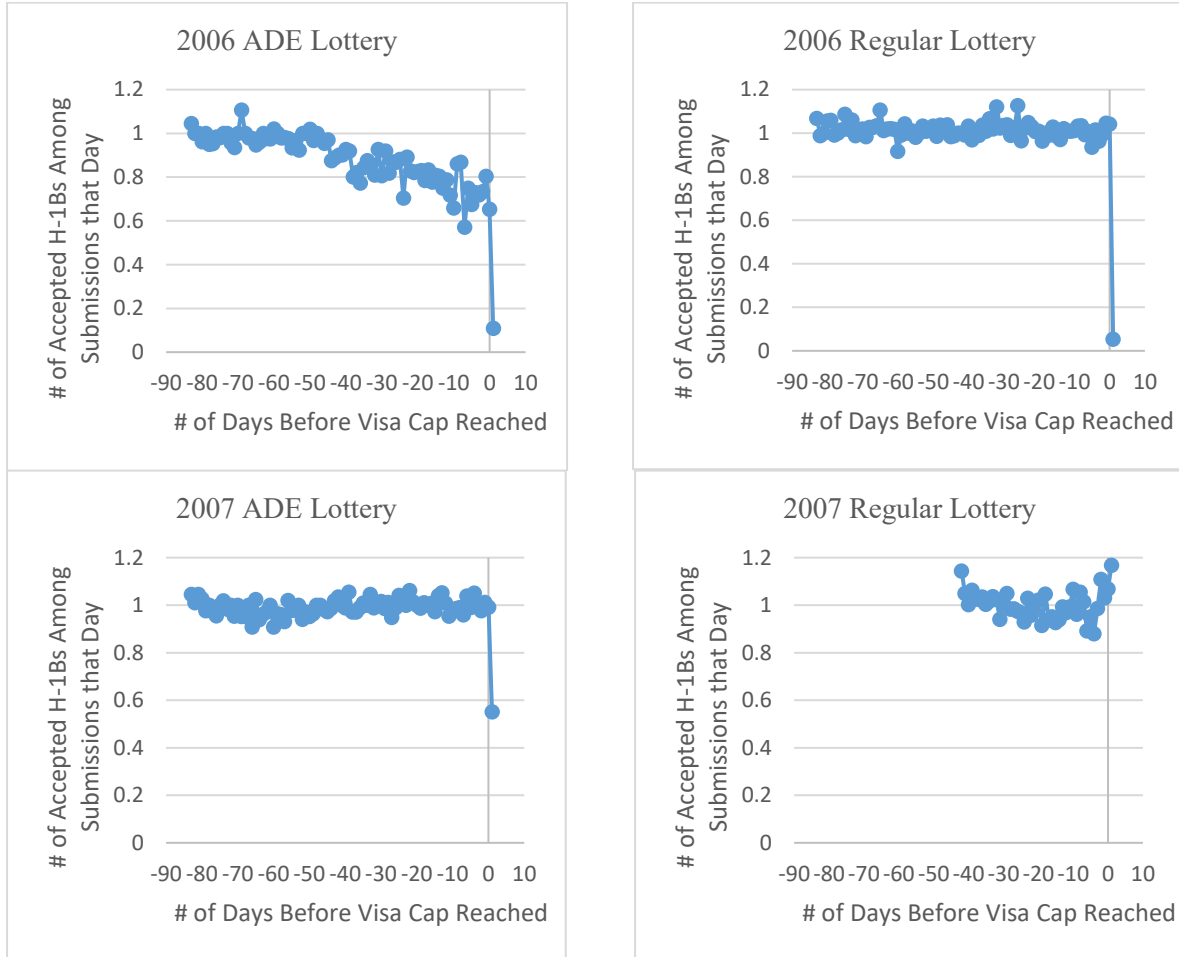
<sup>2</sup> See <http://www.ssa.gov/history/ssn/geocard.html>.

<sup>3</sup> Even if both were perfectly measured, citizenship at the time of the lotteries could be different than whether an individual is a native—in cases in which a non-native became a citizen prior to the time of the lotteries. Thus, there is no presumption that regressions.

<sup>4</sup> Of course, in interpreting the median regressions, we must recognize that the effects across separate regressions for foreigners and non-foreigners do not “add” to the median effect on overall employment.

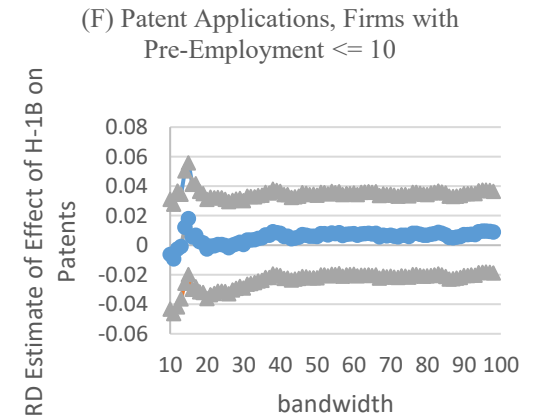
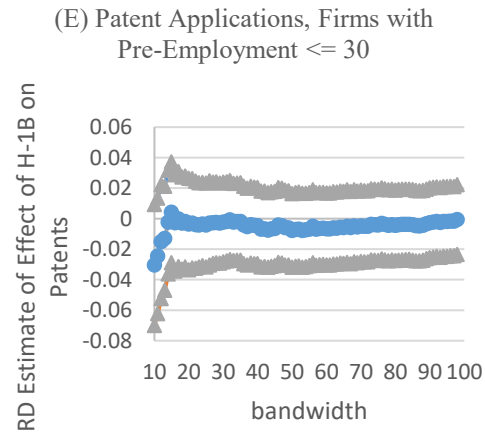
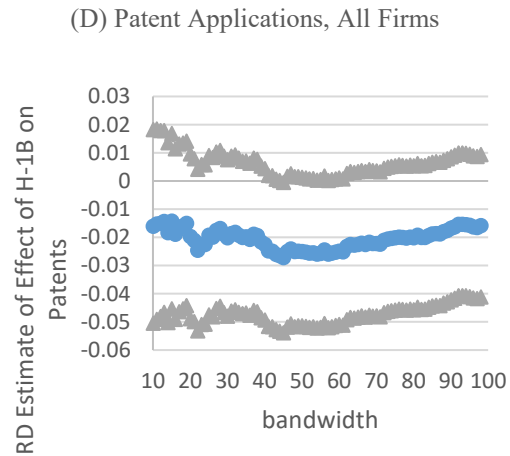
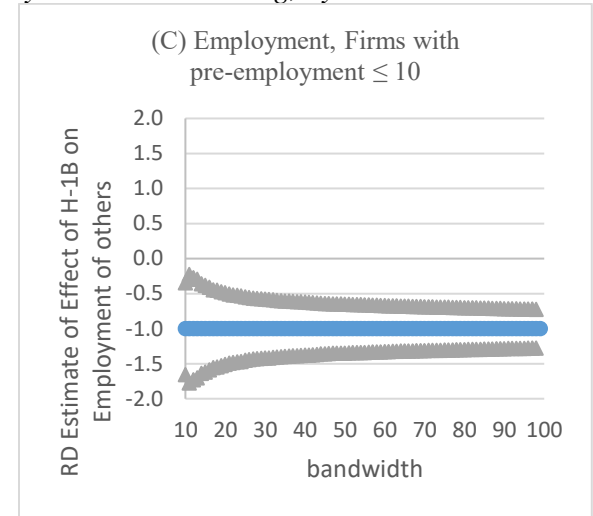
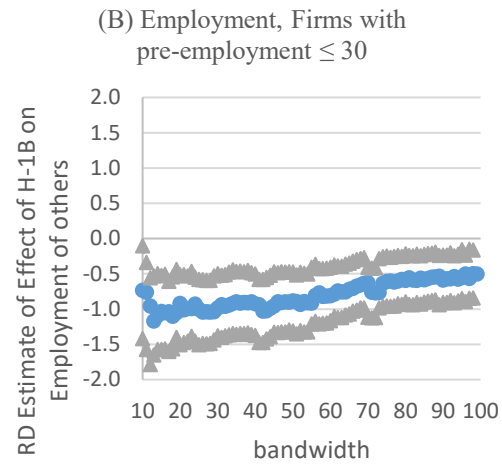
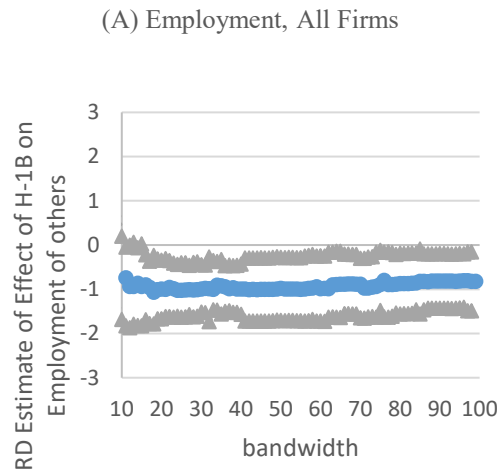


**Appendix Figure 1. Number of Accepted H-1Bs by Date of H-1B Application**



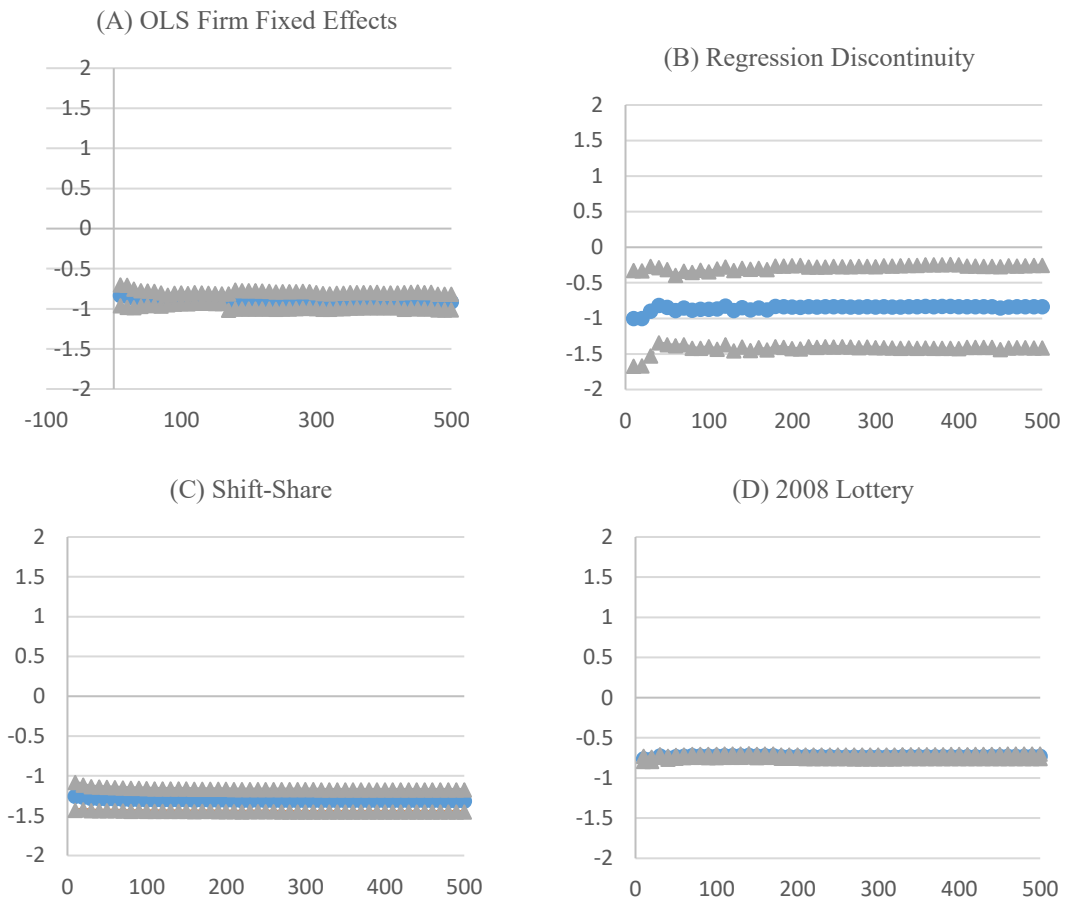
Notes: The figure shows the # of H-1Bs a firm receives by the day on which that firm submitted the application(s). The days are counted backward from the day the cap was reached for each category (so -40 means the firm(s) submitted their H-1B application(s) 40 days before the day the visa cap was reached for that category). We restrict the sample to firms that submitted all of their applications for any given visa category on one day (this represents 76.4% of all firms, 73.32% of applications).

**Appendix Figure 2. Regression Discontinuity Estimates of Effects of Chance H-1B on Employment and Patenting, by Bandwidth**



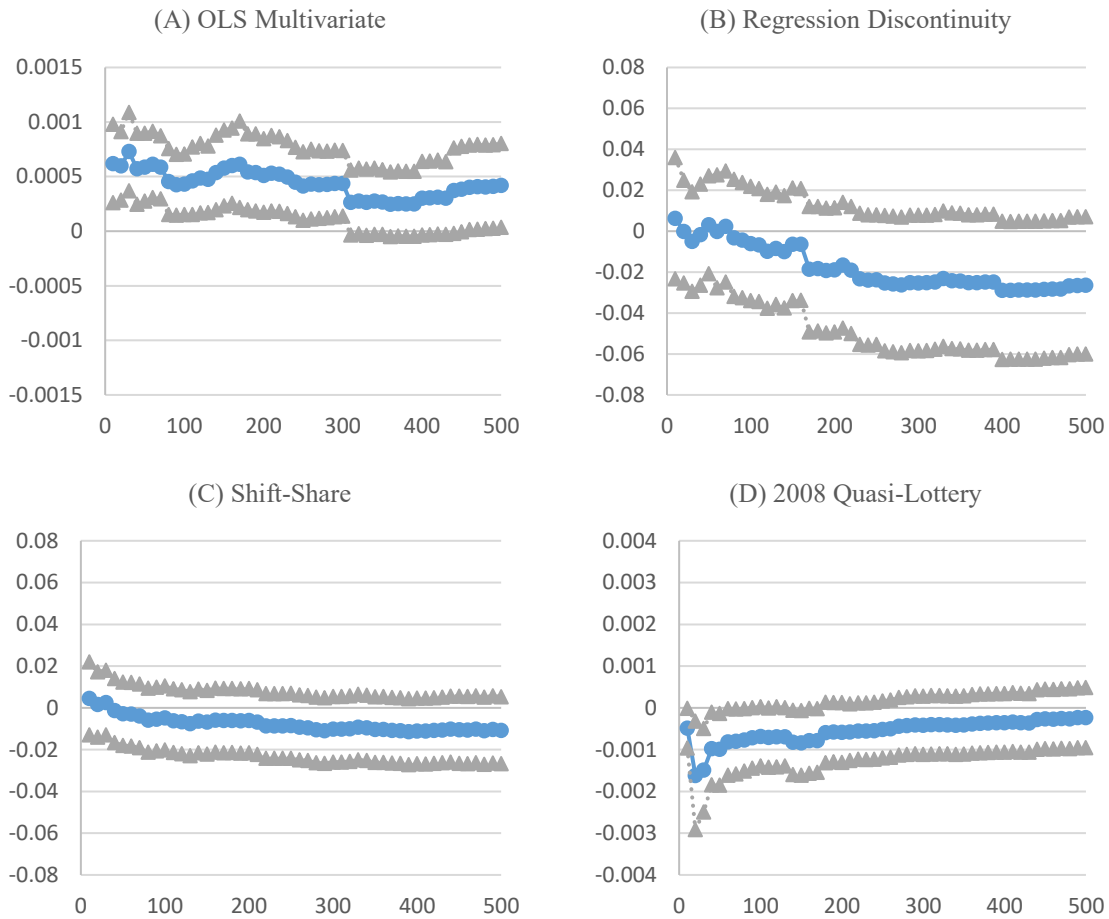
Notes: This figure shows how regression discontinuity estimates of median employment of other workers and mean IHS patenting vary with bandwidth. The blue lines represent the coefficient estimates, and the gray lines represent the 95% confidence intervals.

**Appendix Figure 3.** *Effect of one H-1B on Firm Employment of Other Workers, by Firm Size, Ancillary Strategies*



Notes: The figure shows the coefficient and 95% confidence interval from the employment regressions reported in Table 4 or Figure Appendix Figure 2. We show the coefficient for employers of each size range from 0-10 to 0-500, with the upper bound of the size ranging in increments of 10. Note that the samples overlap across different regressions; for example, firms with 10 or fewer employees are included in the samples in all 50 regressions shown. The dependent variable in each case is number of employees in the firm in a given year. In the case of the OLS firm fixed effects regressions, the regression discontinuity, 2008 quasi-lottery, the independent variable is number of new H-1Bs in the firm in that year. In the shift-share, the independent variable is the number of new H-1Bs scaled by lagged employment in the firm.

**Appendix Figure 4. Effect of one H-1B on Firm Patents, by Firm Size, Ancillary Strategies**



Notes: The figure shows the coefficient and 95% confidence interval from the patenting regressions reported in Table 4. We show the coefficient for employers of each size range from 0-10 to 0-500, with the upper bound of the size ranging in increments of 10. Note that the samples overlap across different regressions; for example, firms with 10 or fewer employees are included in the samples in all 50 regressions shown. The dependent variable in each case is IHS of patents in the firm over six years. In the case of the OLS multivariate regression, the regression discontinuity, 2008 quasi-lottery, the independent variable is number of new H-1Bs in the firm in that year. In the shift-share, the independent variable is the number of new H-1Bs scaled by lagged employment in the firm.

**Appendix Table 1.** *Mean difference between the mean annual salaries for each job title in a state and the actual annual salaries paid to H1-Bs in that state:*

---

|  |             |
|--|-------------|
| Computer support specialists                     | -\$7,579.43 |
| Mathematical technicians                         | \$11,901.26 |
| Network and computer systems administrators      | \$13,271.47 |
| Database administrators                          | \$15,237.22 |
| Computer specialists, all other                  | \$15,438.71 |
| Statisticians                                    | \$15,776.49 |
| Network systems and data communications analysts | \$16,787.21 |
| Computer and mathematical occupations            | \$18,211.92 |
| Computer programmers                             | \$18,455.78 |
| Operations research analysts                     | \$20,156.83 |
| Computer systems analysts                        | \$21,189.35 |
| Computer software engineers, applications        | \$28,614.63 |
| Computer software engineers, systems software    | \$34,617.26 |
| Actuaries  | \$35,698.44 |
| Mathematicians                                   | \$38,378.18 |
| Computer and information scientists, research    | \$43,798.35 |

---

Notes: We use data from the Department of Labor’s 2006 Occupational Employment Statistics by State, containing mean annual earnings for the above sub-occupations within the “Computer and mathematical occupations” category. For each individual H-1B applicant, their I-129 application reports the state they will be working in and their proposed salary. For each Regular H1-B visa worker in “systems analysis and programming” in each state in 2006, we calculate the mean annual salaries in each state, and subtract them from the mean annual earnings in each of the job categories above in that state. In each of these job categories, we then take the mean of these differences across states and report them above. Thus, H-1Bs are typically paid \$11,901 less than Mathematical Technicians in the same state and year. The results show that the wages are so much lower for H1-Bs than for the time-and-place-adjusted wages for creative professions that it is very unlikely that the typical H1-B visa workers are working in creative professions, at least if the prevailing wage restriction is being met to some degree.

**Appendix Table 2. Worker Characteristics: Means and Standard Deviations, and Validity of Randomization**

|  | Mean (SD)    | <i>n</i> | Coefficient (SE) on<br>Chance Wins |
|--|--------------|----------|------------------------------------|
|  | (1)          | (2)      | (3)                                |
| Age (all)  | 30.31 (6.44) | 2,966    | -0.14 (0.24)                       |
| Age ( $\leq 30$ )                                      | 30.20 (6.29) | 1,233    | 0.37 (0.48)                        |
| Age ( $\leq 10$ )                                      | 30.19 (6.12) | 727      | -0.51 (0.57)                       |
| Log intended salary (all)                              | 10.43 (2.14) | 2,966    | 0.12 (0.086)                       |
| Log intended salary ( $\leq 30$ )                      | 10.24 (2.38) | 1,233    | 0.04 (0.22)                        |
| Log intended salary ( $\leq 10$ )                      | 10.25 (2.33) | 727      | 0.27 (0.30)                        |
| Fraction in computer-related occupations (all)         | 0.53 (0.49)  | 2,966    | -0.030 (0.026)                     |
| Fraction in computer-related occupations ( $\leq 30$ ) | 0.56 (0.49)  | 1,233    | -0.03 (0.046)                      |
| Fraction in computer-related occupations ( $\leq 10$ ) | 0.58 (0.49)  | 727      | 0.01 (0.06)                        |
| Fraction with higher degree (all)                      | 0.25 (0.39)  | 2,966    | 0.0026 (0.018)                     |
| Fraction with higher degree ( $\leq 30$ )              | 0.26 (0.39)  | 1,233    | 0.04 (0.03)                        |
| Fraction with higher degree ( $\leq 10$ )              | 0.25 (0.39)  | 727      | -0.002 (0.039)                     |

Notes: Column 1 of the table shows means and standard deviations of worker characteristics from I-129s of firms applying on the day of the lottery. We report means and standard deviations of means at the firm level; *n*'s in Column 2 refer to number of firms, not number of applications. Column 3 shows that these predetermined worker characteristics are uncorrelated with chance lottery wins, by regressing the worker characteristics on chance lottery wins. See Tables 1 and 3 for further information.

**Appendix Table 3. Comparison of Applications on Day of Lottery to Other Applications**

| Dependent Variable                                   | Coefficient (SE) on<br>“Last Day” Dummy<br>(1) | <i>n</i><br>(2) |
|--|--|-----------------|
| <b>Panel A: Comparison of firm characteristics</b>   |  |                 |
| A) IHS of employment in Year -1                      | 0.10<br>(0.052)**                              | 41,849          |
| B) Fraction in NAICS=54                              | 0.17<br>(0.0097)***                            | 46,706          |
| C) IHS of patents in Year -1                         | 0.054<br>(0.018)***                            | 51,483          |
| D) Fraction patenting in Year -1                     | 0.011<br>(0.0054)**                            | 51,483          |
| <b>Panel B: Comparison of worker characteristics</b> |  |                 |
| E) Fraction with superior degree                     | 0.040<br>(0.0069)***                           | 51,483          |
| F) Log intended salary                               | 0.043<br>(0.0069)***                           | 50,272          |
| G) Fraction in “systems analysis and programming”    | 0.22<br>(0.0090)***                            | 51,483          |
| H) Age   | -0.71<br>(0.12)***                             | 51,466          |

Notes: Panel A compares characteristics of firms that applied on the day the cap was reached (so they are subject to the lottery) to all firms whose applications reached USCIS (including others that applied before the cap was reached). We report the coefficient and standard error on the dummy for applying on the last day, from an OLS regression of the dependent variable (shown in the first column) on a dummy for applying on the last day, plus dummies for each of the four lotteries (FY06 Regular, FY06 ADE, FY07 Regular, FY07 ADE). Observations on firms that applied on both the last day and prior to the last day are included in both the sample of firms applying on the last day and the sample applying prior to the last day; thus, the table effectively compares firms that applied only on the last day to firms that applied only on one or more days before the last day. Panel B compares worker characteristics from firm applications on the last day to those from firm applications on other days, using firm-reported information on worker characteristics from I-129s, and reporting the same specification as Panel A. “Superior degree” is defined as a master’s, professional, or Ph.D. degree for the Regular lottery, and is defined as a Ph.D. for the ADE lottery (and the results are similar with alternative definitions). These degrees refer to the highest degree completed in any country (not just the U.S.). Age is measured in years. NAICS code 54 is professional, scientific, and technical services. Sample sizes differ across regressions because some outcomes are missing in some cases (for example, Year -1 employment is missing in some cases because the firm did not exist in Year -1). The sample size is far below the number of total visa applications received across these lotteries primarily because a small number of firms apply for many visas, with a very skewed distribution. Standard errors are clustered by firm. *n*’s refer to the number of firm-lottery years; the number of firms is around 75 percent as large. \*\*\* refers to significance at the 1% level; \*\* at the 5% level; and \* at the 10% level.

**Appendix Table 4.** *Comparison of Worker Characteristics from Applications from Day of Lottery to Worker Characteristics on Applications from Previous Days, Among Firms Applying both on Last Day of Lottery and Previous Days*

| Dependent Variable                                | Coefficient (SE) on<br>“Last Day” Dummy<br>(1) | <i>n</i><br>(2) |
|---|--|-----------------|
| A) Fraction with superior degree                  | -0.0089<br>(0.013)                             | 3,856           |
| B) Log intended salary                            | -0.0041<br>(0.0091)                            | 3,806           |
| C) Fraction in “systems analysis and programming” | -0.0082<br>(0.0089)                            | 3,855           |
| D) Age  | 0.33<br>(0.17)**                               | 3,854           |

Notes: See notes to Appendix Table 3. Appendix Table 4 is identical to Appendix Table 3, except that in Appendix Table 4, we examine only firms applying *both* on day of lottery and on previous days, and we compare worker characteristics between these samples within firms by adding firm fixed effects to the regression. The table shows that in this sample, workers whose applications were submitted on the last day are generally similar to other workers, except that those on the last day are around 1/3 of a year older on average.

**Appendix Table 5.** *Relationship between date of H1b application submission and growth rate of firm*

|   | Employment(t-1) / Employment(t-2)<br>(1) | Employment(t) / Employment(t-1)<br>(2) |
|---|--|--|
| <b>Panel A: Firms with ≤ 10 employees in the pre-period</b> |  |  |
| Date of submission – date quota was reached                 | -0.00043<br>[-0.0089, 0.0081]            | 0.0012<br>[0.0002, 0.0021]***          |
| N   | 14,857                                   | 10,854                                 |
| <b>Panel B: Firms with ≤ 30 employees in the pre-period</b> |  |  |
| Date of submission – date quota was reached                 | 0.0010<br>[-0.0034, 0.0055]              | 0.0016<br>[0.0009, 0.0023]***          |
| N   | 28,263                                   | 23,404                                 |
| <b>Panel C: All firms</b>                                   |  |  |
| Date of submission – date quota was reached                 | 0.0018<br>[0.0004, 0.0033]***            | -0.0025<br>[-0.011, 0.006]             |
| N   | 73,442                                   | 67,961                                 |

Notes: The dependent variable in specification (1) is the employment growth rate of the firm measured as time -2 relative to time -1. The dependent variable in specification (2) is the employment growth rate of the firm measured as time 0 relative to time -1. All regressions control for visa category fixed effects and a constant. Standard errors are clustered at the firm level. \*\*\* refers to significance at the 1% level; \*\* at the 5% level; and \* at the 10% level.



**Appendix Table 6. Validity of the Randomized Design**

| <u>Dependent Variable</u>  | <u>Coefficient (SE) on Chance Lottery Wins</u> |
|--|--|
| Lottery data has firm information                                | 0.0028 (0.0032)                                |
| Whether match to tax master file                                 | 0.0080 (0.0079)                                |
| Whether match to quarterly employment data                       | -0.0031 (0.0096)                               |
| Employment in Year -2 (all, median)                              | 0.50 (1.30)                                    |
| Employment in Year -2 ( $\leq 30$ , median)                      | -0.55 (0.81)                                   |
| Employment in Year -2 ( $\leq 10$ , median)                      | -0.31 (0.69)                                   |
| Employment in Year -2 (all, winsorized first-difference)         | 0.082 (9.71)                                   |
| Employment in Year -2 ( $\leq 30$ , winsorized first-difference) | 0.56 (0.89)                                    |
| Employment in Year -2 ( $\leq 10$ , winsorized first-difference) | -0.091 (0.57)                                  |
| Number of patents in Year -2 (all)                               | 0.011 (0.093)                                  |
| Number of patents in Year -2 ( $\leq 30$ )                       | -0.004 (0.011)                                 |
| Number of patents in Year -2 ( $\leq 10$ )                       | -0.003 (0.003)                                 |
| IHS of patents in Year -2 (all)                                  | 0.0019 (0.019)                                 |
| IHS of patents in Year -2 ( $\leq 30$ )                          | -0.013 (0.019)                                 |
| IHS of patents in Year -2 ( $\leq 10$ )                          | -0.0028 (0.0044)                               |
| IHS of R&E in Year -2 (all)                                      | -0.30 (0.28)                                   |
| IHS of R&E in Year -2 ( $\leq 30$ )                              | -0.0037 (0.015)                                |
| IHS of R&E in Year -2 ( $\leq 10$ )                              | -0.0040 (0.0034)                               |
| Payroll per employee in Year -2 (all, median)                    | 91.01 (594.95)                                 |
| Payroll per employee in Year -2 ( $\leq 30$ , median)            | 1,591.82 (1,519.61)                            |
| Payroll per employee in Year -2 ( $\leq 10$ , median)            | 1,645.07 (3,141.91)                            |
| Profits in Year -2 ( $\leq 200$ , median)                        | -6,268.96 (4,528.82)                           |
| Profits in Year -2 ( $\leq 30$ , median)                         | -8,027.92 (5,498.00)                           |
| Profits in Year -2 ( $\leq 10$ , median)                         | -20,306.35 (19,756.56)                         |
| Dummy for NAICS=54 (all)   | 0.007 (0.03)                                   |
| Dummy for NAICS=54 ( $\leq 30$ )                                 | -0.033 (0.043)                                 |
| Dummy for NAICS=54 ( $\leq 10$ )                                 | 0.010 (0.058)                                  |
| F-Test on mean outcomes  | P=0.64   |
| F-Test on median outcomes  | P=0.38   |

Notes: The table regresses placebo outcomes on chance H-1B lottery wins. We run OLS regressions for outcomes when main regressions in later tables are OLS (*i.e.* for patenting, R&E, winsorized employment, the NAICS=54 dummy, and the match dummies in the first three rows), and median regressions when our main regressions are median (*i.e.* for employment, earnings per employee, and profits). In the first three rows, the dependent variables are dummies for: whether the USCIS data contain the firm's EIN; whether a firm's EIN in the USCIS data matches an EIN in the IRS universe of U.S. EINs; and whether a firm's EIN in the USCIS data matches an EIN in the IRS form 941 data. Dummies for whether R&E, profits, or payroll match are also insignificant. We investigate the effects on Year -2 outcomes because we can then control for the dependent variable measured in Year -1, which is the same control as in our regressions in later tables. Moreover, we can then determine the firm size cutoffs by measuring employment in Year -1, yielding the same firms in each size category as in later regressions. "Winsorized first-difference" means that the dependent variable is the first-difference of employment between Year -2 and Year -1, winsorized at the 5<sup>th</sup> and 95<sup>th</sup> percentiles. Standard errors clustered by firm. Table 2 shows sample sizes. \*\*\* means  $p < 1\%$ ; \*\*  $p < 5\%$ ; and \*  $p < 10\%$ .

**Appendix Table 7. First Stage Regressions**

| <u>Sample</u> | <u>Employment First Stage</u>                        |                                    | <u>Patenting First Stage</u>                         |                                    |
|---------------|--|------------------------------------|--|------------------------------------|
|               | Coefficient (SE) on<br>Chance Lottery<br>Wins<br>(1) | First-stage F-<br>statistic<br>(2) | Coefficient (SE) on<br>Chance Lottery<br>Wins<br>(3) | First-stage F-<br>statistic<br>(4) |
| A) All        | 0.88<br>(0.029)***                                   | 935.14                             | 0.87<br>(0.027)***                                   | 1053.65                            |
| B) ≤30        | 0.89<br>(0.040)***                                   | 495.51                             | 0.88<br>(0.042)***                                   | 435.14                             |
| C) ≤10        | 0.88<br>(0.052)***                                   | 281.57                             | 0.86<br>(0.059)***                                   | 214.04                             |

Notes: The table shows the first stage regression of the number of approved H-1Bs on the number of chance lottery wins. The first stage is slightly different for employment than for patenting because the sample sizes differ; for employment there are a small number of missing observations in the data (as noted in the main text). See Table 1 for other notes and sample sizes. \*\*\* denotes  $p < 0.01$ ; \*\* denotes  $p < 0.05$ ; \* denotes  $p < 0.10$ .

**Appendix Table 8. Other Worker Employment Regressions by Quarter in Q1 to Q4**

|                               | Median Regressions         |                            | Two-stage least squares<br>(mean regressions) |                            |
|-------------------------------|----------------------------|----------------------------|---|----------------------------|
|                               | (1)                        | (2)                        | (3)   | (4)                        |
| <b>Panel A: ≤10 employees</b> |                            |                            |   |                            |
| A) Q1                         | -1.00<br>[-2.28, 0.28]     | -1.03<br>[-2.64, 0.58]     | -0.93<br>[-2.24, 0.39]                        | -1.15<br>[-3.15, 0.86]     |
| B) Q2                         | -1.00<br>[-1.68, -0.32]*** | -1.41<br>[-2.17, -0.64]*** | -1.80<br>[-3.34, -0.25]**                     | -2.46<br>[-4.29, -0.64]*** |
| C) Q3                         | -1.78<br>[-1.78, -0.77]*** | -1.53<br>[-2.42, -0.64]*** | -1.66<br>[-3.40, 0.08]*                       | -2.33<br>[-4.47, -0.20]**  |
| D) Q4                         | -1.76<br>[-3.05, -0.49]*** | -1.61<br>[-2.79, -1.57]*** | -1.90<br>[-4.12, 0.31]*                       | -2.72<br>[-5.52, 0.08]*    |
| <b>Panel B: ≤30 employees</b> |                            |                            |   |                            |
| E) Q1                         | -1.35<br>[-2.41, -0.28]*** | -1.32<br>[-2.38, -0.27]**  | -2.05<br>[-4.17, 0.06]*                       | -2.31<br>[-4.47, -0.15]**  |
| F) Q2                         | -1.22<br>[-2.08, -0.35]*** | -1.17<br>[-2.11, -0.22]**  | -1.73<br>[-3.57, 0.10]*                       | -1.95<br>[-3.90, 0.00]*    |
| G) Q3                         | -1.95<br>[-3.17, -0.73]*** | -1.76<br>[-2.83, -0.69]*** | -2.00<br>[-4.23, 0.23]*                       | -2.33<br>[-4.62, -0.04]**  |
| H) Q4                         | -1.53<br>[-2.82, -0.24]*** | -1.53<br>[-2.85, -0.21]**  | -1.92<br>[-4.51, 0.67]                        | -2.25<br>[-4.99, 0.49]     |
| <b>Panel C: All</b>           |                            |                            |   |                            |
| I) Q1                         | -2.41<br>[-4.40, -0.42]*** | -2.67<br>[-3.89, -0.46]**  | -63.10<br>[-769.40, 643.19]                   | -10.40<br>[-23.73, 2.92]   |
| J) Q2                         | -2.35<br>[-4.72, 0.02]*    | -2.00<br>[-4.11, 0.12]*    | -18.32<br>[-181.09, 146.44]                   | -3.75<br>[-19.09, 13.58]   |
| K) Q3                         | -1.06<br>[-4.15, 2.03]     | -0.75<br>[-3.33, 1.83]     | 3.76<br>[-73.71, 81.24]                       | 3.43<br>[-16.97, 23.83]    |
| L) Q4                         | 0.36<br>[-5.80, 1.07]      | -1.31<br>[-4.64, 2.01]     | -14.70<br>[-192.01, 162.60]                   | -0.96<br>[-22.57, 20.64]   |
| Prior employment              | X                          | X                          | X   | X                          |
| E[wins]                       |                            | X                          |   | X                          |

Notes: See other notes to Table 2. See Table 1 for sample sizes. \*\*\* refers to significance at the 1% level; \*\* at the 5% level, and \* at the 10% level.

**Appendix Table 9: Effect of chance H1-B on increases in employment above a given threshold**

|   | Employment increase of         |                                  |                                  |                                 |                                 |                                 |                                 |
|---|--------------------------------|----------------------------------|----------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
|   | ≥ 100<br>(1)                   | ≥ 1,000<br>(2)                   | ≥ 10,000<br>(3)                  | ≥ 95 <sup>th</sup> ptile<br>(4) | ≥ 99 <sup>th</sup> ptile<br>(5) | ≥ 99.5 <sup>th</sup> pt.<br>(6) | ≥ 99.9 <sup>th</sup> pt.<br>(7) |
| <b>Panel A: Firms with ≤ 10 employees in the pre-period</b> |                                |                                  |                                  |                                 |                                 |                                 |                                 |
| Chance wins   | -0.00076<br>[-0.0020, 0.00044] | -0.00020<br>[-0.00060, 0.00021]  | n/a                              | -0.027<br>[-0.068, 0.014]       | -0.0058<br>[-0.0108, 0.0066]    | 0.0003<br>[-0.0047, 0.0054]     | -0.00077<br>[-0.00199, 0.00045] |
| <b>Panel B: Firms with ≤ 30 employees in the pre-period</b> |                                |                                  |                                  |                                 |                                 |                                 |                                 |
| Chance wins   | -0.0015<br>[-0.018, 0.015]     | -0.000032<br>[-0.00037, 0.00029] | 0.000096<br>[-0.000099, 0.00029] | -0.016<br>[-0.058, 0.024]       | -0.0042<br>[-0.031, 0.023]      | 0.0008<br>[-0.024, 0.026]       | -0.0049<br>[-0.0144, 0.0046]    |
| <b>Panel C: All firms</b>                                   |                                |                                  |                                  |                                 |                                 |                                 |                                 |
| Chance wins   | -0.011<br>[-0.053, 0.031]      | -0.0025<br>[-0.022, 0.017]       | -0.0012<br>[-0.0049, 0.0024]     | -0.0033<br>[-0.0351, 0.0285]    | 0.0042<br>[-0.0087, 0.0170]     | -0.0010<br>[-0.0107, 0.0086]    | -0.0015<br>[-0.0038, 0.00077]   |

Notes: The table shows the results of OLS regressions in which the dependent variable is a dummy for measures of whether the firm's change in employment was above a given number or percentile, and the independent variables are chance lottery wins, expected wins, and prior employment (as in the baseline 2SLS specification in Table 2, shown in the third column). The table shows coefficients on chance H-1B visas, with 95 percent confidence intervals in brackets. The table shows that there is no significant effect of chance lottery wins on the probability of being above these limits. Panel A Column 3 is blank because no firms with ten or fewer employees experienced an increase in employment of at least 10,000 See other notes to Table 2. \*\*\* refers to significance at the 1% level; \*\* at the 5% level, and \* at the 10% level.

**Appendix Table 10. Effect of H-1B Lottery Wins on Employment of Other Workers, Varying Controls**

| Employment in first year (Q1 through Q4):         |                            |                            |                                 |                                 |                            |                           |
|---|----------------------------|----------------------------|---------------------------------|---------------------------------|----------------------------|---------------------------|
|   | Median regressions         |                            | 2SLS (mean regressions)         |                                 |                            |                           |
|   | (1)                        | (2)                        | (3)                             | (4)                             | (5)                        | (6)                       |
| A) $\leq 10$ employees                            | -2.08<br>[-4.24, 0.24]*    | -1.52<br>[-2.15, -0.89]*** | -3.80<br>[-6.96, -0.65]**       | -3.58<br>[-6.48, -0.69]**       | -2.46<br>[-4.61, -0.32]**  | -2.10<br>[-3.88, -0.32]** |
| B) $\leq 30$ employees                            | -4.51<br>[-6.57, -2.44]*** | -1.36<br>[-2.09, -0.63]*** | 3.20<br>[-17.20, 23.59]         | 10.63<br>[-18.67, 39.94]        | -4.84<br>[-7.88, -1.80]*** | -2.26<br>[-4.25, -0.29]** |
| C) All  | -2<br>[-16.78, 12.77]      | -2.05<br>[-3.67, -0.43]**  | -1407.80<br>[-4102.92, 1289.33] | -1503.06<br>[-4021.05, 1012.93] | 23.69<br>[-180.31, 229.68] | -3.41<br>[-18.76, 11.94]  |
| Employment in first three years (Q1 through Q12): |                            |                            |                                 |                                 |                            |                           |
|   | Median regressions         |                            | 2SLS (mean regressions)         |                                 |                            |                           |
|   | (1)                        | (2)                        | (3)                             | (4)                             | (5)                        | (6)                       |
| A) $\leq 10$ employees                            | -3.12<br>[-5.07, -1.16]*** | -1.64<br>[-2.30, -0.88]*** | -2.69<br>[-7.38, 1.99]          | -2.42<br>[-6.90, 2.04]          | -3.10<br>[-6.02, -0.18]**  | -2.75<br>[-5.42, -0.07]** |
| B) $\leq 30$ employees                            | -4.86<br>[-6.52, -3.12]*** | -1.74<br>[-2.61, -0.87]*** | -6.55<br>[-17.77, 4.67]         | -1.44<br>[-14.39, 11.51]        | -5.73<br>[-9.07, -2.37]*** | -2.98<br>[-5.66, -0.30]** |
| C) All  | -4<br>[-17.77, 9.77]       | -3.95<br>[-5.37, -0.52]*** | -823.84<br>[-2719.29, 1071.60]  | -707.81<br>[-2198.10, 782.47]   | 11.28<br>[-223.32, 247.88] | -11.00<br>[-29.68, 9.67]  |
| Prior employment                                  |                            | X                          |                                 | X                               |                            | X                         |
| 95% Winsorization                                 |                            |                            |                                 |                                 | X                          | X                         |

Notes: The table shows coefficients on chance H-1B visas, with 95 percent confidence intervals in brackets, minus the H-1B worker the firm won. The first two columns show median regressions of firm employment. The next four columns show 2SLS (mean) regressions. We pool and stack observations across quarters. All specifications control for expected lottery wins, while only some include a control for prior employment or winsorize the change in employment. In the 2SLS regressions, the instrument is chance lottery wins and the endogenous variable is approved capped H-1B visas. See Table 2 for other notes and sample sizes. \*\*\* denotes  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.10$ . If the H-1B worker works at the firm, a coefficient of 0 corresponds to neither crowd-out nor crowd-in of other employment, and a coefficient of -1 corresponds to one-for-one-crowdout of other employment.

**Appendix Table 11: Effect of Chance H-1Bs on Employment of Other Workers, including versus excluding firms with missing pre-period employment**

|   | Employment, 1<br>Year incl missing<br>(1) | Employment, 1<br>Year excl missing<br>(2) | Employment, 3<br>Years incl missing<br>(3) | Employment, 3<br>Years excl missing<br>(4) |
|---|---|---|--|--|
| <b>Panel A: Median Regressions</b>      |   |   |  |  |
| Chance wins                             | -1.75<br>[-3.39, -0.12]**                 | -2.05<br>[-3.67, -0.43]**                 | -2.51<br>[-4.64, -0.38]**                  | -2.95<br>[-5.37, -0.52]***                 |
| N                                       | 10799                                     | 9803                                      | 31797                                      | 28644                                      |
| <b>Panel B: 2SLS (Mean) Regressions</b> |   |   |  |  |
| Chance wins                             | -1.90<br>[-60.12, 58.31]                  | -3.41<br>[-18.76, 11.94]                  | -9.97<br>[-75.06, 55.13]                   | -11.00<br>[-31.68, 9.67]                   |
| N                                       | 10799                                     | 9803                                      | 31797                                      | 28644                                      |

Notes: Dependent variables are employment within the firm over one year or three years. Controls and sample restrictions are identical to Table 2 of the paper, except in (1) and (3) where firms with missing pre-period employment are included, with a dummy variable for missing pre-period employment added to the regression. By construction, we cannot examine smaller firms while including missing employment values in the pre-period because we do not know which firms meet any particular cutoff. Standard errors are clustered at the firm level. \*\*\* refers to significance at the 1% level; \*\* at the 5% level, and \* at the 10% level.

**Appendix Table 12. Effect of Chance Lottery Wins on Other Worker Employment in Subgroups**

|   | ≤10 employees<br>(1)                | ≤30 employees<br>(2)                  | All firm sizes<br>(3)                |
|---|-------------------------------------|---------------------------------------|--------------------------------------|
| A) Regular  | -1.41<br>[-2.10, -0.73]***<br>{651} | -1.59<br>[-2.46, -0.72]***<br>{1,069} | -2.26<br>[-4.33, -0.19]**<br>{1,969} |
| B) ADE  | -1.02<br>[-2.36, 0.36]<br>{67}      | -0.48<br>[-2.51, 1.55]<br>{134}       | 0.38<br>[-6.63, 7.39]<br>{400}       |
| C) Professional, sci., and<br>tech. services                          | -1.58<br>[-2.54, -0.61]***<br>{456} | -1.72<br>[-2.92, -0.52]***<br>{759}   | -2.46<br>[-4.60, -0.33]**<br>{1,275} |
| D) Industries other than<br>professional, sci., and<br>tech. services | -0.64<br>[-1.50, 0.22]<br>{257}     | -0.35<br>[-1.36, 0.65]<br>{426}       | 0.16<br>[-3.74, 4.05]<br>{1,015}     |
| E) “Temporary support<br>services” industries                         | -2.56<br>[-6.70, 1.57]<br>{384}     | -1.68<br>[-3.09, -0.27]**<br>{628}    | -2.54<br>[-5.03, -0.05]**<br>{4,738} |
| F) Non-“temporary<br>support services”<br>industries                  | 0.65<br>[-1.42, 0.72]<br>{330}      | -1.00<br>[-1.95, -0.05]**<br>{560}    | -0.86<br>[-3.46, 1.74]<br>{1,265}    |
| G) Applied on last day<br>and before                                  | -1.80<br>[-3.05, -0.56]***<br>{377} | -1.76<br>[-3.31, -0.20]**<br>{627}    | -2.17<br>[-5.18, 0.83]<br>{1,299}    |
| H) Applied only on last<br>day  | -1.00<br>[-1.60, -0.40]***<br>{340} | -1.07<br>[-1.78, -0.36]***<br>{563}   | -1.38<br>[-2.77, 0.00]*<br>{997}     |
| I) Average age of<br>applications < 27                                | -1.13<br>[-3.40, 1.14]<br>{202}     | -1.79<br>[-4.74, 1.16]<br>{341}       | 1.79<br>[-16.82, 20.40]<br>{683}     |
| J) Average age of<br>applications ≥ 27                                | -1.52<br>[-2.26, -0.77]***<br>{500} | -1.35<br>[-2.15, -0.54]***<br>{830}   | -2.71<br>[-5.37, -0.04]**<br>{1,618} |

Notes: The table shows the effect of chance lottery wins on employment, minus the H-1B worker won in the lottery, displaying point estimates and 95% confidence intervals in [square brackets] for median regressions of employment in Q1-Q4 on chance lottery wins.  $n$ 's in {curly brackets} show the total number of firms. All specifications have the baseline controls: employment in the pre-period and expected lottery wins. “Temporary consulting industries” refers to six-digit NAICS codes 541511, 541519, 541600, 541330, 519100, 423600, and 541512; “non-temp industries” refers to all others. “Professional, scientific, and technical services” refers to NAICS code 54. The number of observations is in {curly brackets} below the confidence intervals in [square brackets]. See Tables 1 and 2 for additional notes. Some firms participate in both the Regular and ADE lotteries in a given year; in these cases, we classify the firms as participating in the Regular (not ADE) lottery, though the results are extremely similar when classifying them as participating in the ADE lottery instead. Total sample sizes differ slightly in Rows A+B, Rows C+D, Rows E+F, Rows G+H, and Rows I+J because whether firms are in the ADE vs. Regular lottery, and firms’ industries, differ slightly across years. Total sample sizes in each of these combined groups also differ slightly from those reported in Table 1 because Table 1 reports  $n$ 's at the firm-lottery year level. \*\*\* denotes  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.10$

**Appendix Table 13. Effect of Chance Lottery Wins on Employment of Foreigners and non-Foreigners**

| Outcome                                     | All<br>( <i>n</i> =2,143)<br>(1) | ≤30 employees<br>( <i>n</i> =1,198)<br>(2) | ≤10 employees<br>( <i>n</i> =723)<br>(3) |
|---|----------------------------------|--|--|
| A) U.S. citizen employment, IRS measure     | -0.012<br>[-0.41, 0.39]***       | 0.00<br>[-0.15, 0.15]***                   | 0.00<br>[-0.19, 0.19]***                 |
| B) Non-U.S. citizen employment, IRS measure | -0.55<br>[-1.89, 0.79]***        | -0.12<br>[-0.97, 0.72]***                  | -0.26<br>[-1.14, 0.62]***                |
| C) Native employment, SSN-based measure     | -0.073<br>[-0.72, 0.58]***       | 0.11<br>[-0.47, 0.69]***                   | 0.018<br>[-0.41, 0.44]                   |
| D) Non-native employment, SSN-based measure | -0.37<br>[-1.32, 0.59]***        | -0.065<br>[-0.80, 0.67]***                 | -0.16<br>[-1.34, 1.03]*                  |

Notes: The table shows the effect of chance lottery wins on employment of foreigners or non-foreigners, displaying point estimates of the coefficient on chance lottery wins and 95% confidence intervals from median regressions. “IRS measure” refers to a specification in which we measure employment using IRS data on the most recent measure of citizenship (the only measure of citizenship immediately available in the data). “SSN-based measure” refers to a measure of nativity using an algorithm developed in conjunction with Yagan (2014), identifying individuals as natives and non-natives on the basis of individuals’ Social Security Numbers (SSNs) in the data. The table shows that the results are similar under both measures. All specifications control for employment in the pre-period and expected lottery wins, as in the baseline. The measure of a firm’s employment is taken from the W-2, because the W-2 data can be matched to citizenship. The results are similar when we measure employment as the total number of employees observed at the firm over the year from the W-2 data. To make the time period investigated as comparable as possible to the quarterly data shown elsewhere (where we investigate Q1 to Q4), we pool the snapshot from Year 0 with Year 1. *n*’s refer to the number of firms. See Table 2 for additional notes. For Rows A and C (regressions for non-foreigners), \*\*\* denotes estimates that are significantly different from -1 at the 1% level; \*\* at the 5% level; \* at the 10% level. For Rows B and D (regressions for foreigners), the number of stars instead denotes the significance test for difference from 1. The reason for the difference is that in the case of foreigners, we are primarily interested in testing whether the additional H-1B crowds out employment of other foreigners—which corresponds to the test of a difference from 1 because if the H-1B works at the firm, the coefficient should be 1. In the case of non-foreigners, we are interested in testing whether the H-1B crowds out non-foreigners one-for-one—which corresponds to the test of whether the coefficient is different from -1. None of the estimates is significantly different from zero at any conventional significance level.



**Appendix Table 14. Effect of Chance Lottery Wins on Contract Workers**

|                      | Any Contractors<br>(1)        | IHS \$ on contractors<br>(2)  |
|----------------------|-------------------------------|-------------------------------|
| <b>≤10 employees</b> | -0.033<br>[-0.144, 0.078]     | -0.409<br>[-1.878, 1.061]     |
| <b>≤30 employees</b> | -0.067<br>[-0.145, 0.011]*    | -0.753<br>[-1.766, 0.261]     |
| <b>All</b>           | -0.054<br>[-0.088, -0.019]*** | -0.663<br>[-1.147, -0.178]*** |

Notes: This table investigates the effects of chance H-1Bs on whether firms hire any contract workers and the inverse hyperbolic sine of total dollars spent on contract workers. The independent variables are chance lottery wins, expected wins, and prior contracting. The table shows coefficients on chance H-1B visas, with 95 percent confidence intervals in brackets. (1) and (2) are both mean regressions. \*\*\* refers to significance at the 1% level; \*\* at the 5% level, and \* at the 10% level.

**Appendix Table 15. Effects of Chance H-1B Lottery Wins on Patenting in Years 0-3 and 4-8**

| Patenting in first 3 years: |                               |                               |                                  |                              |
|-----------------------------|-------------------------------|-------------------------------|----------------------------------|------------------------------|
|                             | IHS of # of patents           |                               | # of patents (negative binomial) |                              |
|                             | (1)                           | (2)                           | (3)                              | (4)                          |
| A) ≤10 employees            | -0.00033<br>[-0.0090, 0.0084] | -0.00015<br>[-0.0082, 0.0079] | -0.0106<br>[-0.0287, 0.0074]     | -0.0089<br>[-0.0203, 0.0026] |
| B) ≤30 employees            | -0.00053<br>[-0.018, 0.017]   | -0.00030<br>[-0.018, 0.017]   | -0.0161<br>[-0.0444, 0.0122]     | -0.0138<br>[-0.0386, 0.0110] |
| C) All                      | -0.021<br>[-0.052, 0.010]     | -0.021<br>[-0.052, 0.010]     | -0.0627<br>[-0.1847, 0.0593]     | -0.0840<br>[-0.2423, 0.0742] |
| Patenting in years 4-8:     |                               |                               |                                  |                              |
|                             | IHS of # of patents           |                               | # of patents (negative binomial) |                              |
|                             | (1)                           | (2)                           | (3)                              | (4)                          |
| A) ≤10 employees            | 0.000041<br>[-0.0022, 0.0023] | 0.000022<br>[-0.0021, 0.0021] | -0.0002<br>[-0.0021, 0.0018]     | -0.0003<br>[-0.0021, 0.0015] |
| B) ≤30 employees            | 0.0043<br>[-0.0054, 0.014]    | 0.0044<br>[-0.0053, 0.014]    | -0.0012<br>[-0.0123, 0.0099]     | -0.0012<br>[-0.0122, 0.0098] |
| C) All                      | -0.00081<br>[-0.033, 0.031]   | -0.0017<br>[-0.033, 0.029]    | -0.0488<br>[-0.1177, 0.0201]     | -0.0502<br>[-0.1279, 0.0275] |
| Prior employment            | X                             | X                             | X                                | X                            |
| E[wins]                     |                               | X                             |                                  | X                            |

Notes: The table shows the effect of an extra chance H-1B visa on patent outcomes over the indicated years. The table is identical to Table 3, except that the dependent variable is the IHS of patents in each year over Years 4 to 8. See Tables 1 and 3 for additional notes and sample sizes. Standard errors are clustered by firm. \*\*\* refers to significance at the 1% level; \*\* at the 5% level; and \* at the 10% level.

**Appendix Table 16. Effect of Chance H-1B Lottery Wins on Patenting in Subgroups**

|   | ≤10 employees<br>(1)                   | ≤30 employees<br>(2)                  | All firm sizes<br>(3)                  |
|---|--|---------------------------------------|--|
| A) Regular  | 0.0017<br>[-0.0040, 0.0074]<br>{654}   | 0.0045<br>[-0.011, 0.020]<br>{1,062}  | 0.0070<br>[-0.011, 0.025]<br>{2,327}   |
| B) ADE  | -0.0038<br>[-0.012, 0.0038]<br>{67}    | 0.00076<br>[-0.0087, 0.010]<br>{137}  | -0.031<br>[-0.11, 0.046]<br>{494}      |
| C) Professional, sci., and tech. services                       | -0.0010<br>[-0.0046, 0.0026]<br>{459}  | 0.0021<br>[-0.012, 0.017]<br>{762}    | -0.010<br>[-0.041, 0.021]<br>{1,486}   |
| D) Industries other than professional, sci., and tech. services | 0.0011<br>[-0.0057, 0.0080]<br>{261}   | 0.0018<br>[-0.0075, 0.011]<br>{432}   | -0.0087<br>[-0.066, 0.049]<br>{1,273}  |
| E) “Temporary support services” industries                      | -0.0015<br>[-0.0057, 0.0028]<br>{388}  | 0.0048<br>[-0.012, 0.021]<br>{632}    | -0.010<br>[-0.044, 0.024]<br>{1,191}   |
| F) Non-“temporary support services” industries                  | 0.0014<br>[-0.0042, 0.0070]<br>{333}   | -0.0015<br>[-0.0085, 0.055]<br>{565}  | -0.0051<br>[-0.056, 0.046]<br>{1,572}  |
| G) Applied on last day and before                               | -0.00026<br>[-0.0070, 0.0065]<br>{379} | 0.0017<br>[-0.014, 0.017]<br>{629}    | -0.0036<br>[-0.037, 0.030]<br>{1,502}  |
| H) Applied only on last day                                     | 0.0011<br>[-0.0015, 0.0038]<br>{345}   | 0.00050<br>[-0.0082, 0.0092]<br>{570} | -0.037<br>[-0.080, 0.0055]*<br>{1,271} |
| I) Average age of applications < 27                             | -0.0022<br>[-0.011, 0.0071]<br>{206}   | -0.0098<br>[-0.029, 0.0091]<br>{347}  | -0.049<br>[-0.12, 0.023]<br>{828}      |
| J) Average age of applications ≥ 27                             | 0.0018<br>[-0.0048, 0.0084]<br>{503}   | 0.0087<br>[-0.0079, 0.025]<br>{833}   | 0.010<br>[-0.019, 0.039]<br>{1,940}    |

Notes: The table shows OLS regressions of the IHS of patents in each year from Year 0 to Year 8 on chance H-1B lottery wins. All specifications control for patents in the pre-period and expected lottery wins, as in the baseline. The results are comparable when we investigate the patenting dummy or the number of patents as the dependent variable. “Temporary consulting industries” refers to six-digit NAICS codes 541511, 541519, 541600, 541330, 519100, 423600, and 541512; “non-temp industries” refers to all others. “Professional, scientific, and technical services” refers to NAICS code 54. The number of observations is in {curly brackets} below the confidence intervals in [square brackets]. See Tables 3 for additional notes. Some firms participate in both the Regular and ADE lotteries in a given year; in these cases, we classify the firms as participating in the Regular (not ADE) lottery, though the results are extremely similar when classifying them as participating in the ADE lottery instead. Total sample sizes differ slightly in Rows A+B, Rows C+D, Rows E+F, Rows G+H, and Rows I+J because whether firms are in the ADE vs. Regular lottery, and firms’ industries, differ slightly across years. Total sample sizes in each of these combined groups also differ slightly from those reported in Table 1 because Table 1 reports  $n$ ’s at the firm-lottery year level, whereas Appendix Table 16 reports them at the firm level. Standard errors are clustered by firm. \*\*\* refers to significance at the 1% level; \*\* at the 5% level; and \* at the 10% level.

**Appendix Table 17. Effect of Chance H-1B Lottery Wins on Patenting, using Alternative Matching Procedure**

|                  | IHS of number of patents    |                             | # of patents (negative binomial) |                              |
|------------------|-----------------------------|-----------------------------|----------------------------------|------------------------------|
|                  | (1)                         | (2)                         | (3)                              | (4)                          |
| A) ≤10 employees | -0.011<br>[-0.029, 0.0065]  | -0.011<br>[-0.029, 0.0068]  | -0.019<br>[-0.035, -0.003]**     | -0.018<br>[-0.030, -0.004]** |
| B) ≤30 employees | -0.0090<br>[-0.026, 0.0084] | -0.0088<br>[-0.026, 0.0085] | -0.021<br>[-0.048, 0.005]        | -0.020<br>[-0.045, 0.004]    |
| C) All           | -0.028<br>[-0.067, 0.011]   | -0.027<br>[-0.066, 0.011]   | -0.086<br>[-0.195, 0.023]        | -0.096<br>[-0.222, 0.030]    |
| Prior patents    | X                           | X                           | X                                | X                            |
| E[wins]          |                             | X                           |                                  | X                            |

Notes: See notes to Table 3. The table is similar to Table 3, except in defining the firms that match between the USCIS data and the Patent Dataverse, this table includes those firms that are “possible” matches (whereas Table 3 excludes those firms). The table examines patenting in each year from Year 0 to Year 8,. \*\*\* refers to significance at the 1% level; \*\* at the 5% level; and \* at the 10% level.

**Appendix Table 18. Effect of Chance H-1B Visas on Patent Citations**

|                  | (1)                        | (2)                        |
|------------------|----------------------------|----------------------------|
| A) ≤10 employees | -0.0059<br>[-0.023, 0.011] | -0.0057<br>[-0.022, 0.010] |
| B) ≤30 employees | -0.0053<br>[-0.032, 0.022] | -0.0049<br>[-0.032, 0.022] |
| C) All           | -0.022<br>[-0.071, 0.028]  | -0.025<br>[-0.074, 0.023]  |
| Prior citations  | X                          | X                          |
| E[wins]          |                            | X                          |

Notes: The table investigates the effect on patents when we weight each patent by its number of citations, *i.e.* the dependent variable is patent citations. “Prior citations” is measured using patents from Year -1, to parallel Table 3. Otherwise, the specification is the same as in the Table 3 IHS specifications. The mean of citations in the ≤10 employees, ≤30 employees, and “all” groups is 2.27, 8.94, and 40.77, respectively. The mean of the IHS of citations in these three groups is 0.22, 0.045, and 0.33, respectively. \*\*\* refers to significance at the 1% level; \*\* at the 5% level, and \* at the 10% level.

**Appendix Table 19. Effect of Chance H-1B Lottery Wins on Patenting Dummy**

|                   | (1)                          | (2)                          |
|-------------------|------------------------------|------------------------------|
| A) ≤10 employees  | -0.0010<br>[-0.0042, 0.0022] | -0.0010<br>[-0.0041, 0.0020] |
| B) ≤30 employees  | -0.0029<br>[-0.0095, 0.0038] | -0.0028<br>[-0.0094, 0.0038] |
| C) All firm sizes | -0.0014<br>[-0.014, 0.011]   | -0.0012<br>[-0.014, 0.011]   |
| Prior patents     | X                            | X                            |
| E[wins]           |                              | X                            |

Notes: See notes to Table 3. The table runs the same specification as Table 3, except that in Appendix Table 21 the dependent variable is a dummy for whether the firm patented in each year, so that the coefficient reflects the effect on the fraction of years that the firm has at least one patent, and we run a linear probability (OLS) model. We control for a dummy for whether the firm patented in a pre-period. \*\*\* refers to significance at the 1% level; \*\* at the 5% level; and \* at the 10% level.