The Effects of High-Skilled Immigration Policy on Firms: Evidence from H-1B Visa Lotteries[1]

Kirk Doran
University of Notre Dame

Alexander Gelber
Goldman School of Public Policy, UC Berkeley, and NBER

Adam Isen
Office of Tax Analysis, U.S. Department of the Treasury

June 2015

Abstract

We study the effects of a firm winning an additional H-1B visa on the firm's outcomes, by comparing winning and losing firms in the Fiscal Year 2006 and 2007 H-1B visa lotteries. We match administrative data on the participants in these lotteries to the universe of approved U.S. patents, and to IRS data on the universe of U.S. firms. Winning additional H-1B visas has insignificant effects on firms' patenting and use of the research and experimentation tax credit, with confidence intervals that generally rule out more than modest effects. Additional H-1Bs cause at most a moderate increase in firms' overall employment, and these H-1Bs substantially crowd out firms' employment of other workers. There is some evidence that additional H-1Bs lead to lower average employee earnings and higher firm profits.

[1] 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, or of any other person associated with the U.S. Treasury Department. All errors are our own.

1. Introduction

What are the effects of high-skilled immigration policy on the economy receiving the immigrants? In the U.S., high-skilled immigrants represent 24 percent of workers in occupations closely tied to innovation (Pekkala Kerr, Kerr, and Lincoln forthcoming). In recent years, prominent voices from government, business, labor, and academia have discussed major changes to U.S. immigration law, often debating the impacts of changes in high-skilled immigration policy on economic outcomes. 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. How these workers affect firms is the subject of much public discussion.

One common narrative argues that H-1Bs given to a firm could lead the firm to increase innovation because H-1B workers have exceptional skills that firms cannot otherwise easily obtain. If H-1Bs have special skills, they generally would not be employed instead of others who otherwise would have worked at the firm—consistent with firms' legal obligation that the employment of H-1Bs "will not adversely affect the working conditions of workers similarly employed."[2] In fact, many have argued that extra H-1Bs lead a firm to increase employment of other workers. This is exemplified by former Microsoft Chairman Bill Gates' congressional testimony, arguing that H-1Bs have special, innovative skills and that technology firms on average hire five additional employees to support each new H-1B worker (Gates 2008, based partly on National Foundation for American Policy 2008).

In a competing, frequently encountered narrative, H-1Bs have more muted effects on firm outcomes like innovation and employment.[3] If an H-1B is employed rather than another worker, and the alternative worker otherwise would have innovated at the firm as much as the H-1B, then we would not expect innovation or employment to increase at the firm that received the H-1B. Moreover, many H-1Bs are not in scientific industries, and many H-1B workers perform jobs (*e.g.* technical support) that might be expected not to lead to innovations in the great majority of cases. Economic theory predicts that firms

[2] Immigration and Nationality Act (INA) §212(n)(1)(A)(ii).
[3] These two competing narratives do not cover all possible combinations of effects of H-1Bs on innovation, employment, profits, wages, and other outcomes, but they tend to dominate the policy debate.

will apply to hire an H-1B worker as long as this increases the firm's profit in expectation. H-1Bs could increase the firm's profit even if they have no effect on the firm's innovation and/or crowd out other workers to some extent, as in the case studies in Matloff (2003) or Hira (2010)—for example, if the H-1B is substitutable with other workers and the firm pays the H-1B less than the worker whose employment is crowded out.[4] Firms submit legal attestations that they will pay the H-1B a "prevailing wage" comparable to other similar workers, but it is possible that these regulations are ineffective in some cases. Indeed, profit-maximizing firms apply for H-1Bs even though they must pay a fee to the U.S. government to apply, suggesting that H-1Bs are paid less than alternative workers with the same marginal product of labor.

Our paper investigates these narratives by estimating the causal impact of extra H-1B visas on the receiving firm, examining outcomes relevant to assessing the narratives. We use randomized variation from the Fiscal Year (FY) 2006 and FY2007 H-1B visa lotteries. In each of these years, on the date when the cumulative number of H-1B visa applications exceeded the maximum allowed for a given visa type, the applications submitted on this day were subject to a lottery. U.S. Citizenship and Immigration Services (USCIS) randomly chose some of these visa applications to win the lottery, and the remaining applications lost the lottery. Across both years and across visa lotteries for those with and without advanced degrees, 3,050 firms applied for 7,243 visas, of which 4,180 visa applications won the lottery. We use administrative data from USCIS on the entrants in these lotteries, matched to U.S. Patent and Trademark Office (USPTO) data on the universe of patents at U.S. firms, and matched to Internal Revenue Service (IRS) microdata on the universe of U.S. firms.

We study patenting as a measure of innovation because it is an innovation outcome we can readily observe, following much previous literature (see the surveys by Nagaoka, Motohashi, and Goto 2010, and Hall and Harhoff 2012). Our patenting specifications examine the impact of additional H-1B visa wins on the firm's approved patents up to nine years after the start of the visa. The point estimates are near zero, and are insignificantly different from zero. We focus on the confidence intervals, which show

[4] Profit could also increase if H-1Bs increase a firm's productivity but not its employment of other workers.

that any increase in patenting is at most small—particularly in small and medium-sized firms, where an additional H-1B represents a meaningful change in overall employment. For example, in firms with 10 or fewer employees, we bound any increase in patenting at or below 0.47 percent, on a base mean of only 0.0098 patents per year. Such results also hold when we exclude firms that likely provide temporary technical support services. The confidence intervals similarly rule out more than a modest positive impact on firms' use of the research and experimentation (R&E) tax credit, another measure of innovative activity.

Parallel to the innovation results, we find that new H-1Bs cause no significant increase in firm employment. Our primary finding is that we can robustly rule out more than a moderate increase in overall firm employment (including employment of H-1Bs). New H-1Bs substantially crowd out employment of other workers at the firm. This evidence is again particularly strong in small and medium-sized firms.

We find some evidence that additional H-1Bs increase median profits, and some evidence that additional H-1Bs decrease median payroll costs per employee. Overall our results are more supportive of the second narrative, in which marginal H-1Bs crowd out other workers to some extent, are paid less than alternative workers, and increase the firm's profits—despite little effect on measures of the quantity of firm innovation (though our estimated effects on other measures likely related to productivity are imprecise).

Relative to other studies on H-1Bs and other immigration programs, ours is the only to our knowledge to leverage randomized variation to estimate the effect of immigration on outcomes in the receiving economy.[5] Our paper relates to previous work on the effects of immigration on measures of innovation (*e.g.* Stuen, Mobarak, and Maskus 2012; Borjas and Doran 2012; Foley and Kerr 2013; Moser, Voena, and Waldinger 2014; Grogger and Hanson forthcoming; see the Kerr 2013 survey), as well as 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 *et al.* 2008; Hanson 2009; and Pekkala Kerr and Kerr

[5] Edin, Fredriksson, and Åslund (2003) and Åslund *et al.* (2011) use variation that appears quasi-random.

2011). Previous studies in the economics literature of the innovation or labor market impacts of the H-1B program specifically or similar programs include Kerr and Lincoln (2010), Hunt and Gauthier-Loiselle (2010), Hunt (2011), Peri, Shih, and Sparber (2013), Pekkala Kerr, Kerr and Lincoln (forthcoming), and Bound *et al*. (forthcoming).[6] The literature has found that H-1Bs lead to large positive impacts on innovation (specifically patenting). Regression analysis has found no clear evidence of crowdout of other employment, and in some cases has found crowd-in.[7]

Our paper isolates the effect of additional H-1B visas given to a particular firm on outcomes at that firm (holding constant H-1Bs given to other firms).[8] As such, our findings are compatible with the possibility that an aggregate increase in H-1Bs raises firm or aggregate innovation and/or employment, as found in previous studies cited above. For example, at the firm level, our results show that new H-1B workers crowd out other workers; the crowded-out workers may find employment elsewhere (unless demand is perfectly inelastic), and they could increase innovation in these other firms relative to the counterfactual (which could lead to further positive spillovers, as in *e.g.* Bloom, Schankerman, and van Reenen 2013). If extra H-1Bs do have large positive effects on aggregate innovation or employment, then our results suggest this is not occurring because an extra H-1B visa at a given firm leads to increases in measures of these outcomes at the firm level in our context—in contrast to the first narrative above.

[6] Peri, Shih, and Sparber (2015) study H-1B visa lotteries but effectively do not rely on randomized variation; they mainly use a differences-in-differences design. Their paper uses data on the winners of the FY2008 and FY2009 H-1B visa lotteries, as well as on firms that submitted an initial H-1B application (called a Labor Condition Application (LCA)) that was approved. When their paper exploits the H-1B lottery, the paper's identification strategy assumes that conditional on having an approved LCA, selection for an H-1B is random. However, it is not random: in FY 2008 and 2009, at least 20 percent of LCAs were withdrawn *prior* to running the H-1B lottery (*e.g.* USCIS 2014, Department of Labor 2014). Thus, those results could be confounded because firms in cities experiencing negative shocks could be more likely to withdraw their applications before the lottery is run (conditional on having an approved LCA).

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

[8] Kerr and Lincoln (2010) and Pekkala Kerr, Kerr, and Lincoln (forthcoming) examine the effect of giving an additional H-1B to a firm by interacting firm characteristics with the H-1B visa cap, and as such are among the first to examine the role of firms. Changes in the aggregate H-1B cap could affect outcomes at a given firm through general equilibrium effects, including effects of the cap increase on other firms. Thus, this previous work addresses a different question of interest than ours does.

We study H-1B applications on the days the caps were reached, representing 4.3 percent of total capped H-1Bs in these years. Although these marginal H-1Bs could have different effects than other H-1Bs, our estimates address the effects on firms of marginally changing the number of capped H-1Bs they are allowed—a question of great relevance to firms and policy-makers as they actively propose and consider the consequences of modest changes in the number of capped H-1Bs. We show that firms applying on the date the cap is reached are more likely than firms applying on other dates to have patented prior to the year of the lottery, and are more likely to request workers who have higher degrees and intended salaries than those in the full sample—arguably making it more striking that we find little effect on measures of innovation even in this sample. Although a modest fraction of all H-1B applications is subject to the lottery, our results will be precise enough to rule out meaningful and relevant alternative hypotheses, including more than a modest increase in measures of innovation and employment.

The paper is structured as follows. Section 2 describes the policy environment. Section 3 discusses our empirical specification. Section 4 describes the data. Section 5 demonstrates the validity of the randomization. Section 6 presents effects on innovation. Section 7 shows effects on employment. Section 8 shows effects on payroll per employee and profits. Section 9 concludes. The Appendix contains further results and discussion.

2. Policy environment

H-1Bs are sponsored by firms, which apply to the U.S. government 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 firm submitting the H-1B application 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."[9] Firms are required to pay H-1Bs comparably with workers in one of four skill categories (defined by experience, education, and level of supervision).[10]

We study the lotteries for H-1B visas in FY2006 and FY2007. In other years, USCIS did not keep data on which firms won and lost the lottery (personal communication with USCIS, 2011). Visas for FY2006 allowed an H-1B to work from October 2005 to September 2008, and visas for FY2007 allowed an H-1B to work from October 2006 to September 2009. A fiscal year begins in October of the previous calendar year (CY), *e.g.* Q1 of FY2006 corresponds to October to December of CY2005.

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 masters 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 each of the years we study, the cap for ADE visas was 20,000, and the cap for Regular visas was 65,000. In each year and for each of the two types of H-1B visa, USCIS allocated visas by lottery for visa applications submitted on the date when the total number of applications reached the cap. 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. In a given lottery, firms sometimes applied for multiple visas; in this case, the probability that the firm won each visa was independent and equal to the number of lottery winners divided by the number of lottery entrants.

The cap does not apply to a number of H-1B visa categories, which are therefore excluded from the lotteries: visas for work at non-profit firms, including U.S. educational institutions; those applying for an extension of an existing H-1B visa; those who have an existing H-1B visa and are changing jobs during the period the existing visa covers; and citizens of five countries (Australia, Canada, Chile, Mexico, and Singapore), who are in effect not bound by H-1B limits. Our results therefore do not speak to the effects of such un-capped visas, implying that our results are not directly comparable to studies that have

[9] 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.
[10] Firms may legally hire an H-1B in lieu of a worker who would have been at a higher skill level.

examined student/trainee or temporary work visas in general (*e.g.* Hunt 2011).

Firms did not know in advance the date when the cap would be reached, and they did not know the probability that firms applying on this date would be selected for an H-1B. 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 (personal correspondence with USCIS, 2011). These dates were determined by the number of applications received on different dates in these years, which was unknown to firms at the time—making it effectively impossible for firms to game the system by applying on the lottery date for more visas than they desire, on the basis of the anticipated probability of selection. Even across the four lotteries we study, the probability that an application won varied widely, and would not have been possible to anticipate. 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. Approximately 90 percent of applications filed on dates before the lottery date were approved. Each lottery was conducted within a month of reaching the relevant cap.

Firms pay fees to USCIS for filing a visa application for initial H-1B status. The total fees range from $1,575 to $3,550 depending on firm size and whether the firm asks for expedited processing. These fees appear in firms' costs in the year of submitting the application. When applications lost the lottery, fees were refunded to firms. Firms also typically incur legal fees of several thousand dollars for submitting the applications.

The H-1B worker may stay at the initial sponsoring firm or move to another firm, 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 several months until the new firm's H-1B application was approved, but a gap of only two weeks was allowed between jobs.

If a firm is denied an H-1B, it has a number of alternatives to hiring no one. Other than hiring U.S. citizens or permanent residents, firms can hire foreigners on other visas, including L-1 temporary work visas, Optional Practical Training (OPT) extensions of F-1 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 typically considered more advantageous to the firm than the H-1B). Only 15 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 in the U.S. for only 12 months, also limiting the degree of substitutability with an H-1B.

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*.

3. Empirical strategy

Our empirical strategy exploits the random assignment of H-1B visas in the lotteries. Thus, we consider only firms that entered the FY2006 or FY2007 H-1B lotteries. Our main outcomes of interest are patenting and number of employees. We also consider the effect on the R&E tax credit, the firm's wage bill per employee, and profits.

Our 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 expected number of H-1B visas given to the firm is $E[W]=pn$. If w represents the random realization of W, then the number of "unexpected" wins $u=w-pn$ reflects the random realization of the net number of wins relative to the expected value, and will be exogenous in the regression we specify below.[11] Thus, our main independent variable is the random variable U, the net number of unexpected wins (or losses) for a given firm, whose realization is u.

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

$$Y_{itT} = \beta_0+\beta_1 U_{iT}+\varepsilon_{itT} \qquad (1)$$

t is the number of calendar years since the lottery in question occurred; for example, $t=0$ corresponds to Year 0. T indexes the year of the lottery in question, *i.e.* FY2006 or FY2007. U_{iT} is the number of unexpected H-1B visa lottery wins for firm i in the lottery

[11] By "unexpected," we are not necessarily referring to firms' actual expectations, which are unobserved.

in year T. ε_{itT} is an error term. β_1 represents the intent-to-treat (ITT) effect of an additional unexpected H-1B visa win.[12] In (1) and all other specifications, whenever we examine an outcome across multiple time periods t, we pool and stack the data across these periods in the same regression. We cluster the standard errors at the firm level.

After a firm wins an H-1B lottery, its application may be approved, denied, or withdrawn. For example, the application may not meet the eligibility criteria, leading to a denial, or the applicant firm may go out of business, leading to a withdrawal. It can be relevant to estimate the effect of an approved capped H-1B visa on firm outcomes, in addition to examining the ITT effect. 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 also approved. We can use lottery wins as an instrument for approved capped H-1B visas in a two-stage least squares (2SLS) model:

$$A_{iT} = \alpha_0+\alpha_1 U_{iT}+\nu_{iT} \qquad (2)$$

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

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}. In the second stage (3), we regress Y_{itT} on A_{iT} (instrumented using U_{iT}). The coefficient γ_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 change their number of approved capped H-1B visas). ν_{it} and η_{it} are error terms.

The ITT and LATE estimates represent different empirical objects, which are both of interest. The ITT estimates show the effects of granting another visa to a given firm. This is relevant because firms and policy-makers are interested in the raw effects on firms of allowing a marginal capped visa to the firm. Thus, for all of our main outcome variables we show our main ITT specification (1). The LATE estimates are particularly relevant when we are testing the hypothesis that additional H-1Bs crowd out other employment. This is because in the employment context we are interested in comparing the coefficient on approved capped H-1Bs to a specific non-zero level, namely to the coefficient in the scenario in which H-1Bs do not crowd out or in employment of other

[12] This specification makes a linearity assumption: moving from no visa to one has the same effect as moving from one to two, etc. We estimate insignificant coefficients on higher-order terms in visa wins.

workers—*i.e.* a coefficient of 1, because our employment data measure a firm's total employment, including H-1Bs. Thus, for employment we additionally show LATE estimates. (Doran, Gelber, and Isen 2014 show LATE estimates of effects on patenting.) The first-stage regressions (Appendix Table 1) have coefficients α_1 near 1 (ranging from 0.88 to 0.89 for employment, and from 0.86 to 0.88 for patenting), and have F-statistics in the hundreds. Thus, there is generally little difference between the ITT coefficient and standard error on unexpected lottery wins, and the LATE coefficient and standard error on approved capped H-1B visas.

In those rare cases (comprising 2.69 percent of firms) in which a firm participates 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 unexpected wins across both of the lotteries that the firm enters in year T (except for specifications in which we run separate regressions for the Regular and ADE lotteries).[13] We seek as much statistical power as possible, so we pool the FY2006 and FY2007 Regular and ADE lotteries in our baseline. In these pooled regressions, for a given firm, we stack data from the FY2006 lottery and data from the FY2007 lottery, so that we can capture the effects of winning the lottery in Year 0 on employment in each subsequent year (measured consistently as the number of years since the relevant lottery occurred).

Although the randomization implies that U_i should be exogenous in (1), it is also possible to control for various pre-determined covariates. For example, we can control for a lagged value of an outcome variable at the firm (*e.g.* when the dependent variable is the number of patents, we can control for $Y_{i,pre,T}$, the number of patents in firm i observed in a "pre-period," meaning a period before Year 0); for the expected number of lottery wins pn; or other covariates.

[13] We find that unexpected 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 unexpected lottery wins in the Regular lottery in that year, the coefficient on unexpected lottery wins is -0.20, with a standard error or 0.18 (p=0.26). Additionally, unexpected lottery wins in 2006 have no effect on approved 2007 visas; for example, when regress total FY2007 Regular and ADE approvals (summed) on unexpected lottery wins in the FY2006 Regular and ADE lotteries combined, the coefficient on unexpected lottery wins is -0.05, with a standard error of 1.45 (p=0.97). Finally, we verified that winning one lottery also does not affect the probability of winning a subsequent lottery conditional on entering the subsequent lottery.

We expect our results to be most compelling in small and medium-sized firms, where the variances of the outcomes are modest and the impact of an additional employee should be most clearly statistically distinguishable from the error term. Small and medium-sized firms in the aggregate contribute in important ways to innovation (Acs and Audretsch 1990), and comprise a substantial fraction of all H-1B lottery applicants. To evaluate how the effects vary across firms of different sizes, we investigate the sample of firms with 10 or fewer employees in Year -1 (roughly the 25th percentile of firm size in our sample); those with 30 or fewer employees in Year -1 (roughly the 50th percentile); many other firm size cutoffs; and the sample of firms of all sizes.

As noted, our measure of total employment reflects total employment at the firm and therefore *includes* the H-1B worker if the H-1B worker works at the firm; in this case, the effect of an additional H-1B visa on total firm employment will equal *one plus* the effect on employment of workers *other* than the new H-1B. One question of interest is a two-sided test of whether the coefficient β_1 on unexpected H-1B visas is significantly different from 0. If β_1 is positive and significant, it would indicate that the extra H-1B visa lottery win increases total employment at the firm—as opposed to crowding out a worker that the firm would have otherwise hired, in which case the coefficient would be 0. In principle, an extra H-1B visa could even decrease employment at the firm, for example if the new H-1B worker works more hours or works harder than others (for example, to secure another visa or green card for continued employment in the U.S.) and therefore crowds out more than one other worker.[14] Another question of interest is a two-sided test of whether β_1 is significantly different from 1. If β_1 is greater than 1, this would indicate that an additional H-1B visa leads to employing a greater number of other workers. If β_1 is less than one, this can indicate that an extra H-1B worker at least partially crowds out other worker(s) who would otherwise have worked at the firm.

Due to the long right tail of the distribution of patents, previous literature has typically examined transformations of the number of patents. Given the approximate lognormality of patents, one may wish to run a specification in which the dependent variable is log patents (*e.g.* Kerr and Lincoln 2010). In our context, this specification

[14] Hours worked is unobserved in our data, as in many administrative datasets on employment.

would lead to a problem: we would like to include firms in the regressions that have zero patents, as the majority of firms have zero patents in our context, but the log of zero is undefined.[15] Thus, we approximate the log of the number of patents using the inverse hyperbolic sine (IHS) of the number of patents, which is defined at zero and negative values (*e.g.* Burbidge, Magee, and Robb 1988, Pence 2006, or Gelber 2011). The IHS of patents Y is defined as:

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

When the IHS of patents is the dependent variable in the ITT regressions, the coefficient β_1 reflects the approximate percent increase in patents caused by an extra unexpected H-1B visa. We also show that our results are similar with a log transformation.

We additionally examine the effect of unexpected H-1Bs on a dummy for whether the firm patented. When this dummy is the outcome, we control for a dummy for whether the firm patented in a pre-period. When we investigate binary outcomes in our panel data, we run a linear probability model to avoid econometric complications relating to panel data specifications with lagged dependent variables in nonlinear contexts. (The same point applies to logits or probits in the case of binary outcomes, or to negative binomial or Poisson regressions in the case of count outcomes.)

To tailor our specifications to the relevant features of each context, our baseline specifications differ in the patenting and employment contexts. We will show that when we run exactly parallel specifications in the employment and patenting contexts, we obtain comparable results to the baseline.

To address the long right tail of the employment distribution, we use median regressions in our baseline specification. (The median value of patents is zero, so it does not make sense to run median regressions in this context.) Because instrumental variables quantile regressions typically did not converge, instead we run ITT median regressions corresponding to model (1) above.

As in the patenting context, previous literature on H-1Bs has not examined effects on the level of employment, but has instead examined transformations of employment,

[15] This is not a problem in the context of Kerr and Lincoln (2010). They examine patents at the city level, where patents are greater than zero.

such as the log, that reduce volatility (*e.g.* Pekkala Kerr, Kerr, and Lincoln forthcoming). Again, zeroes in employment imply that it is not straightforward to use the log in our context. Thus, a second way of addressing the long right tail of the employment distribution is to estimate the effect on the (first-differenced) IHS of employment. In this specification, before testing whether the coefficient on unexpected H-1B visas is equal to 1 (reflecting a scenario with no crowdout), we must transform the coefficient from the regression (which reflects the approximate percentage increase in employment) by multiplying it by the mean level of employment in a control group. We can then test whether this transformed coefficient, which should reflect the increase in the absolute level of employment for the mean firm, equals 1. However, the coefficient could instead be multiplied by any employment level other than the mean, thus generating different estimates of the implied effect on the level of employment. In light of this issue, we present the IHS employment results only in the Appendix. (In the patenting context, our interest is less in testing whether the patenting effect is different than a specific non-zero number—but in the employment context, we test for a coefficient difference from 1.)

To find another method of running mean (not median) regressions while addressing the long right tail of the employment distribution, we let the dependent variable be the winsorized first difference of employment, and we run the 2SLS regressions (2)-(3) (recall that 2SLS is most relevant in the employment but not the patenting context). The first difference ΔY_{itT} is taken from before the lottery (*i.e.* the first quarter of CY2005 for FY2006 visa applicants, and the first quarter of CY2006 for FY2007 visa applicants), to period t after the lottery. Winsorization is common in administrative data (*e.g.* Chetty *et al.* 2011) and in survey data (*e.g.* the Current Population Survey).[16] Winsorized regressions would not capture large effects on employment outcomes. However, when we run our 2SLS regressions without winsorizing, the point estimate of the effect is negative and insignificant, lessening the concern that winsorization dulls an actual positive effect. We also find that an extra H-1B

[16] We winsorize the first difference of employment and control for lagged employment, rather than winsorizing the level of employment in period t after the lottery and controlling for lagged employment, because in the context of examining firms of all sizes, winsorizing the first difference is more effective in removing large outliers than is winsorizing the level of employment. When we limit the sample to smaller firms, the two specifications show very similar point estimates and confidence intervals.

visa has an insignificant effect on the probability that the change in employment is outside the 95th percentile. Nonetheless, because of these issues, the median regressions are our primary specification in the employment context.

Parallel to examining whether the firm patented, we also examine the effect of unexpected H-1Bs on a dummy for whether the firm has a positive number of employees, a measure of whether the firm is in business.

For each outcome, the baseline time period we investigate is also chosen to be the most appropriate for that outcome. For employment, we are most interested in comparing the coefficient on unexpected H-1Bs to 1, to test the "no-crowdout" hypothesis. Thus, in our baseline we focus on the effect on employment from Q1 to Q4, when the H-1B worker is almost always working at the firm and when a coefficient below 1 will therefore most reliably indicate crowdout. (In later quarters, there is more attrition as some H-1Bs leave the initial firm.) For other outcomes, we are less interested in comparing the coefficient to any specific non-zero level; instead we are more interested in investigating periods when the H-1B likely could have had a measurable effect on the outcome. For payroll costs per employee, if H-1Bs are paid less than alternative workers, then we would expect to measure effects on payroll per employee primarily while the H-1B is usually at the firm. Thus, as a baseline for this outcome it makes sense to examine the duration of the visa, Years 0 to 3, when the H-1B is typically working at the firm. With this motivation, as a baseline we also examine the R&E credit and profits over Years 0 to 3.[17] Given the sometimes substantial time taken to develop and approve patents, it makes sense to investigate as long a time period as possible for patents. Thus, our baseline patenting specification examines patents from Year 0 to the latest year available in the data, Year 8.

Beyond the baseline period, for each outcome we also show the results in all other relevant periods. For example, we additionally show the employment, R&E, payroll per employee, and profits results until Year 8, and we show patenting for Years 0 to 3 alone.

4. Data

Match between USCIS data and patenting data

---

[17] H-1Bs typically worked at the firm for only one-quarter (*i.e.* October to December) of the calendar year in Year 0, and for three-quarters of calendar Year 3 (*i.e.* January to September).

We merge several administrative datasets. First, we use USCIS administrative data on the H-1B lotteries for FY2006 and FY2007. The data contain information on each H-1B visa application that entered the lottery in each of these years: Employer Identification Number (EIN); the date the firm applied for a visa; the type of H-1B (Regular or ADE); the name of the firm applying; 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 LCA such as highest degree completed.

We obtained the Patent Dataverse on the universe of granted U.S. patent applications from 1975 to 2013 at each firm, based on USPTO data.[18] 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. We also observe total patent citations until 2010.

The time to develop a patent can range from months to years, with substantial variance. The mean approval time reported by USPTO for patents filed in FY2008 is 32.2 months, again with substantial variance (USPTO 2012). Our data will allow us to estimate the effect on an important set of patents, namely those within up to nine years of the initial H-1B visa period, but the effect on subsequent patents is unobserved.[19]

Since the Patent Dataverse does not contain EIN, but does contain firm name, we matched firms between the Patent Dataverse from 1975 to 2013 and the USCIS lottery data using firm names. As described further in Appendix 1, to match firms between these two datasets, we performed an intentionally liberal automatic match between the datasets to obtain all plausible matches. We then searched through these matches by hand to detect and remove all matches that appeared spurious. We classified firms into three categories: (1) 392 firms that definitely matched between the datasets; (2) 63 firms for which it was ambiguous whether they matched; and (3) the remaining 2,595 firms that definitely did not match. In our main results, we classify the 63 ambiguous matches as

---

[18] See https://dataverse.harvard.edu/dataverse/patent (accessed 5/24/2015).

[19] The majority of H-1B petitions are for workers aged 25 to 34, whereas patents of academic life scientists peak around mid-career (Azoulay, Ding, and Stuart 2007), and noted innovations peak around age 40 (Jones 2010). This raises the possibility that some H-1B workers who stay in the U.S. will innovate more beyond our sample period (though Levin and Stephan 1991 find that scientists' productivity is greatest at the beginning of their careers). However, in all of these studies innovation in the 25-to-34 age range is a substantial fraction of its peak. We leave examination of effects at longer time horizons to future research.

non-matches. In the Appendix, we show that the results are comparable when assuming that the possible matches are true matches. In general, our results are robust to similar matching procedures. A firm would not match between the datasets if it did not patent during this time period, so these firms appear in our data as having zero patents.

Match between USCIS data and IRS data

Using EINs, we merged firms from the USCIS lottery data to IRS data on the universe of U.S. firms. Data from IRS form 941 contain information for each EIN on overall quarterly employment in the U.S. (where overall employment includes workers in the U.S. of both foreign and U.S. nationality), which we call "employment." Our measure of employment in Q1 (which reflects the first quarter of the fiscal year, *i.e.* the last quarter of the preceding calendar year) reflects employment as measured in mid-December of that quarter. Thus, between the time when a firm learned that it won or lost the lottery in June to August of Year -1, and the end of Q1, when workers generally begin working at the firm and when employment is measured, firms had a number of months to react. For example, firms were notified of the FY2007 Regular visa lottery results in June of CY2006, which gave firms over six months until December of CY2006. However, in the sole case of the FY2006 ADE lottery, the lottery was held on January 17, 2006, *after* Q1 of FY2006 ended. Thus, in the employment regressions, we drop data from Q1 of the FY2006 ADE lottery, since firms' decisions in Q1 could not have been influenced by the results of this lottery.

We use data from 2004 to 2013. The first available form 941 data are from the first quarter of CY2004. These data are missing in the second through fourth quarters of CY2004, so we measure employment in CY2004 using data on its first quarter.

Another measure of innovative activity is the R&E tax credit, as reported to IRS (see Hall and Van Reenen 2000 or Hall, Mairesse and Mohnen 2010 for surveys). The R&E credit goes to firms that have research and development costs in the U.S. To our knowledge, our paper is the first to investigate the effect of immigration on the R&E. In our IRS data, we observe the amount of the R&E credit claimed (not R&E expenses), and we only observe this for C-corporations. We match firms' patents to the USCIS data using a fuzzy match of firm name, and patents can take time to develop—but neither of these issues affects the R&E outcome, because we match R&E data to USCIS data using

EIN, and we can measure firms' *contemporaneous* R&E credits. We also estimate the effect on firms' yearly net income ("profit") and wage bill per employee, both as reported to IRS. In general, profits measured in the IRS data are not the same as economic profits.

We drop the 2.0 percent of firms in the USCIS data that did not match to the EIN master list in the IRS data. Pooling over all quarters, 4.5 percent of the remaining firms in the USCIS data did not match to the IRS data on quarterly firm employment; we treat these data as missing. Of the remaining firms, 17.9 percent have missing employment data in Year -1, which makes it impossible to run our specifications in which we control for Year -1 employment, and we drop these data for the purpose of the employment specifications. Of the remaining observations, pooling over Q1 to Q4, 2.2 percent are missing in a given quarter.

The USCIS data do not contain identifying information on individual H-1B applications like Tax Identification Numbers that can be linked to the IRS data.[20] Thus, we cannot distinguish the employment of a particular H-1B worker whose application entered the lottery from employment of others. Like previous literature on the effects of H-1Bs (*e.g.* Kerr and Lincoln 2010), the data also do not distinguish H-1Bs in general (whether lottery winners or other H-1Bs) either from non-H-1Bs, or from workers on other visas like the L-1. As a result, we cannot directly assess how new H-1Bs affect employment of foreign workers on other visas. In the IRS data, we do observe the most recent report to the U.S. government of a worker's citizenship status, which is an imperfect measure of whether a worker was a U.S. citizen at the time of the lotteries.

Summary statistics

Table 1 shows summary statistics. We use data on 2,750 firms (*i.e.* EINs) in the full sample. In 300 cases (9.84 percent), 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 in this context to a year of the lottery, rather than a year when an outcome is observed.[21] In the full sample, the mean (4.52) and especially standard

[20] We were given the lottery data to link firms, not workers. The LCAs cannot usefully be used to link USCIS applications to the IRS data, as this would introduce substantial measurement error.

[21] Since larger firms tend to apply in both years, the means and standard deviations tend to be moderately lower at the firm (rather than firm-lottery year) level. The results of later regressions also tend to be more precise when weighting each firm equally, strengthening our conclusions (available upon request).

deviation (56.11) of patents measured at the yearly level are large, due to a small number of firms that patent in large numbers. The mean (0.15) and standard deviation (0.80) of the IHS of patents are much lower. The means and standard deviations are smaller among the 1,276 firm-lottery years (or 1,192 firms) with 30 or fewer employees, and smaller still among the 749 firm-lottery years (or 719 firms) with 10 or fewer employees. As a result, we generally focus on such small or medium-sized firms. A modest fraction of the sample patents in a given year—*e.g.* 4.8 percent of the full sample of firms, corresponding to 9.3 percent of these firms that patented at any point during Years 0 to 8. Similarly, a modest fraction claims the R&E credit, but the mean (IHS of the) amount claimed is substantial.

Table 1 shows that the mean (1,877.84) and standard deviation (39,721.31) 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, the mean and standard deviation of Q1 to Q4 employment are much lower but still quite large. Median employment is much lower than the mean. Winsorizing also reduces the mean and standard deviation greatly.

In the FY2006 Regular lottery the vast majority of applications lost the lottery, and in the FY2007 Regular lottery the vast majority of applications won the lottery. 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: the confidence intervals will show the degree of precision of the results.

The 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 unexpected lottery wins (as defined above) 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.

LCAs show that across all lotteries, around half of applications were for computer-related jobs. Around one-tenth were for engineering-related jobs.

Comparison of lottery firms to other firms

As our regressions use firms that applied on the day the cap was reached, it is relevant to compare this sample to the broader sample of firms applying for H-1B visas in these years. In Table 2, we regress characteristics of the firms or workers on a dummy for applying on the last day and lottery fixed effects. Firms applying on the last day are more likely to have patented in the past, and patented more in the past. Similarly, firms applying on the last day are 17 percentage points more likely to be in NAICS code 54. Applications on the last day tend to be from larger firms. 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 more advanced degrees or higher salaries, then our sample will arguably be primed to find a particularly positive effect on measures of innovation.

5. Validity of the randomization

Table 3 verifies the validity of the randomized design by regressing variables that should not be affected by the lottery on unexpected lottery wins. The table confirms that none of the lagged dependent variables is significantly related to unexpected lottery wins: various measures of patenting, the R&E, employment, firm wage bill per employee, and profits.[22] Dummies for whether firms from the USCIS lottery data match to other datasets (*i.e.* the sample restrictions discussed earlier), and a dummy for NAICS 54, are also insignificantly related to unexpected wins.

6. Effects on measures of innovation

A priori, it is not clear whether H-1Bs should raise or lower patenting, or use of the R&E credit, at the firm level. H-1Bs could have special skills that raise firms' innovation, but H-1Bs could alternatively lead to lower firm innovation, for example if firms use H-1Bs in place of higher-skilled alternative workers (as in Matloff 2003). We begin by measuring effects on patenting and then turn to the R&E credit.

Effect on patenting

---

[22] 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. 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 calculated from Year -2, the regressions are insignificant for all but one of the 27 dependent variables, consistent with random chance (results available upon request).

Table 4 estimates the effect on patenting of unexpected H-1B visas, during the patenting baseline period of Years 0 to 8, as well as over the duration of the initial H-1B visa in Years 0 to 3. By "Years 0 to 8," we mean that we pool the FY2006 lottery, for which we observe Years 0 to 8, with the FY2007 lottery, for which we observe Years 0 to 7. We also examine the effect on a dummy for whether the firm patented in each year, so that β_1 represents the effect on the fraction of years that firms patented. For each of the outcomes, we show the results with two alternative sets of controls: *(a)* controlling for the number of patents in Year -1 (or for a dummy for patenting in Year -1 when the patenting dummy is the outcome); or *(b)* additionally controlling for the expected number of lottery wins (conditional on the number of H-1B applications and the probability of winning the lottery in question). The results are similar either way; we take *(b)* as a baseline. The results are also similar when we add additional controls, such as controlling additionally for the NAICS code of the firm, for the number of H-1B lottery applications *n*, or for dummies for each of the four lotteries. Finally, the results are also similar when pre-period patenting is measured over another time period rather than Year -1.

In Table 4, we estimate a precise zero effect of unexpected visas on patenting. The point estimates are generally very close to zero. As the estimates are insignificant, we focus on the confidence intervals to determine what we can rule out with statistical confidence. When the dependent variable is the IHS of the number of patents in firms with 10 or fewer employees, the upper end of the 95 percent confidence interval in the baseline rules out an increase greater than just 0.47 percent, relative to a "base" mean number of patents of only 0.0098 per year.[23] Our main conclusion—that any absolute increase in the quantity of patenting is at most modest—makes sense given our finding that only a modest fraction of the sample patents. For firms with 30 or fewer employees, in the baseline we bound the increase in patents below 1.3 percent, and in the full sample, below 1.9 percent. When the dependent variable is the patenting dummy, the confidence interval also indicates at most a small impact (*e.g.* at most an increase of 0.0020 in firms with 10 or fewer employees). We also find no evidence that H-1Bs increase high quantiles of patenting, and bound any increase below a similarly small level.

[23] We calculate the "base" mean by taking the mean number of yearly patents in Years 0 to 8 in a "control group," specifically firms whose number of unexpected wins was less than or equal to zero.

Our choices of the number of employees in our size thresholds (*e.g.* 10 or fewer) could be varied. Figure 1 plots the coefficient and confidence interval on unexpected H-1B visas when the dependent variable is the IHS of number of patents over Years 0 to 8, as a function of the employer's size, from under 10 employees to under 500, in increments of 10.[24] The upper end of the 95 percent confidence interval ranges from near 0 to just above 0.01; across all 50 choices of the employer size threshold shown, in the *most positive* case we are able to rule out an increase in patents greater than around 1.5 percent (and usually the upper bound is substantially smaller). The point estimate is positive in only three of 50 cases—notably, for size thresholds of 10, 20, and 30—though it is insignificant and very small in all of these cases. We also find no significant effects in the largest firms (over 500 employees). Overall, we find no evidence of a notable increase in patenting and robustly rule out more than a small increase.

Other specifications

An important issue is whether our results generalize to H-1B applications submitted on other days. We cannot directly address this question, but we can re-weight observations so that the weighted distribution of key firm and worker characteristics from the day of the lottery matches that among applicants for capped H-1B visas over all days that applications were submitted. Appendix Table 2 shows that these results are very similar to the baseline. Throughout the paper, the results are also similar when weighting by firms' number of H-1B applications, or by the expected number of lottery wins.

Appendix Table 3 shows that the effect on the absolute number of patents is at most small relative to the baseline variation. For example, in employers with 30 or fewer employees, the upper end of the confidence interval rules out an increase in the number of patents greater than 0.38 percent of a standard deviation. Appendix Table 4 shows that the effects in only the later period (Years 4 to 8) are comparable to those shown earlier.

Appendix Table 5 shows the effects when we weight each patent by its number of citations until 2010, *i.e.* the dependent variable is patent citations through 2010. The results still rule out more than a small percentage increase in citations.

[24] The necessity of keeping a sufficiently large number of firms in each category, to prevent the potential identification of any given firm, prevents us from going beyond 500 employees in increments of 10.

Appendix Table 6 shows that the results are comparable when we assume that possible matches between the USCIS and patenting data did match, instead of the baseline assumption that they did not. If anything, this shows more negative effects than the baseline—occasionally negative and significant at 10 percent, though not robustly so.

In Appendix Table 7, the dependent variable is the log of one plus the number of patents in each year. This is an alternative way of addressing the skewness of the outcome distribution, while recognizing that the number of patents is often zero, but without resorting to the less-known IHS transformation. However, this specification has the limitation that we add an arbitrary constant (*i.e.* 1) to patents. The results are very similar to Table 4, with slightly smaller upper bounds on the confidence intervals.

In Appendix Table 8, we examine whether there is heterogeneity in the effect on patents across type of lottery or type of industry, using our baseline specification. We find no evidence of significant, or 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. Appendix Table 9 shows that there is also no significant interaction of winning the lottery with prior firm patenting, or how with early in the application season the cap was reached. Among firms that patented at any point prior to Year 0, we also find no significant effect. Finally, there is no significant difference between the effects in the 2006 and 2007 sets of lotteries. Appendix 2 discusses these results further.

Our main focus in this section is on the effect on patents; increasing the quantity of innovation is often seen as desirable. Appendix Table 10 shows that unexpected H-1B visas also have an insignificant effect on the firm's number of patents per employee. At the same time, the results are compatible with some increase in patents per employee.

Effect on R&E Credit

Table 5 shows the effect on the R&E credit in Years 0 to 3. In firms with 10 (30) or fewer employees, the baseline rules out that an extra unexpected H-1B increases the amount of the R&E claimed by more than 4.1 (1.8) percent, and rules out that the fraction of years when taking any R&E credit increases by more than only 0.0041 (0.0016). The point estimates are negative. In the largest firms, the results are imprecise. Years 4 to 8 also show no evidence of a significant positive impact, and in a minority of cases actually show barely significant negative impacts (see Appendix Table 11). We again find comparable results at other size thresholds; no significant interactions with covariates; and no significant differences across groups.

We only observe the amount of R&E credit claimed, which could be affected by factors like firm profit: firms with higher profits will on average have higher tax rates and thus claim more credit per dollar of R&E expenditures, and will also on average claim more credit because the R&E is non-refundable. However, we later find some evidence that unexpected H-1Bs raise firm profits, and this impact on profits should push toward showing that H-1Bs *raise* R&E claims. This makes our finding of no significant increase in the R&E all the more striking. We focus less on the R&E than on patenting also because R&E is an input into innovation, not an output (Lerner and Seru 2015).

7. Effect on employment

Table 6 shows our baseline estimates of the effect of extra H-1B visas on total firm employment in Q1 to Q4. Our main finding is that we bound any increase in employment below a moderate level. In the baseline median regressions, the top end of the 95 percent confidence interval in firms with 10 or fewer employees is 0.11, indicating that an extra unexpected H-1B visa leads to an increase in total employment of at most 0.11 workers. Although the point estimate is below zero, it is insignificant. We do not conclude from the point estimate that unexpected H-1B visas decrease employment, because our confidence interval is compatible with an increase in employment; of course, this is why confidence intervals are useful in determining what we can rule out with a standard degree of statistical certainty. Similarly, in this specification in firms with 30 or fewer employees, the top end of the confidence interval is 0.37. In the full sample of firms, we can rule out an increase greater than 0.57. All of these estimates are significantly different from 1 at the 1 percent level, suggesting crowdout of other employment. In the 2SLS specification among firms with 10 or fewer employees, the top end of the confidence interval with more controls is 0.68, again significantly different from 1 at the 1 percent level, but compatible with a moderate positive effect. With 30 or fewer employees, we can rule out a coefficient of 0.71 or greater (p<0.01). In the full sample of firms, the 2SLS results are extremely imprecise.

Figure 2 plots the coefficient and confidence interval on unexpected lottery wins from the baseline median employment specification, as a function of the employer size threshold. We focus on the upper end of the 95 percent confidence interval; across all 50 choices of the employer size threshold, in the most positive case we are able to rule out an increase in employment of more than 0.6. In all cases, the estimate is significantly less than 1 at the 1 percent level. The point estimates are always negative and insignificantly different from zero. We also find no significant effect in firms with over 500 employees.

Appendix Table 12 shows that re-weighting the sample to the characteristics of the full population of firms and workers again shows comparable results to the baseline. Appendix Table 13 shows that in each quarter from Q1 to Q4 individually, we are typically able to rule out a coefficient of 1, particularly in smaller firms. Appendix Table 14 shows that several other specifications yield comparable results: winsorizing instead at the 99th percentile; letting the dependent variable be the IHS of the first difference in employment (as in the IHS patenting specifications); winsorizing the IHS of the first difference in employment at the 99th percentile (to address the long right tail further); winsorizing the IHS of the level of employment at the 99th percentile; and running median regressions when the dependent variable is the first difference of employment (rather than controlling for the lag of employment). Appendix Tables 15 and 16 verify that unexpected wins are also unrelated to whether the firm is in business.[25] Appendix 3 describes how the results are generally similar when other patenting or employment specifications are run to make the full set of specifications exactly parallel in the patenting and employment contexts. Quantiles other than the median also show no evidence of increases in employment.

Other time periods and samples

Table 7 shows employment effects in other time periods. Rows A and B show Q5 to Q8, and Q9 to Q12, respectively, *i.e.* each of the remaining two of the three years covered by the H-1B visa in question, after Q1 to Q4. We generally rule out a coefficient of 1 at the 5 percent significance level in both of these periods. Row D shows results for Q13 through Q32 (the latest quarter in the sample), when we estimate less significant

[25] Since Appendix Tables 15 and 16 measure the effect on whether the firm has employment in the U.S., these results also encompass effects on whether a firm chooses to locate in the U.S.

results. As a greater proportion of H-1Bs leave their initial firm, the interpretation of a coefficient below 1 as indicating crowdout becomes weaker; for example, many H-1Bs have left their initial firm by three years after the start of their visa. Thus, the results are less informative about crowdout in later years. As Appendix 2 discusses in detail, Appendix Tables 9 and 17 find no significant differences in the effects across the same subsamples and interactions we investigated in the patenting context.

Interpreting the estimates

Our ITT employment estimates are relevant for firms and policy-makers interested in understanding the average employment effects of granting additional capped H-1B visas to firms. We find no indication that overall firm employment will rise on average, and we find that overall firm employment will increase on average by at most a moderate amount for every additional new capped H-1B visa.

Moving beyond the policy-relevant ITT estimates, institutional features of this context are relevant to determining whether new H-1Bs crowd out employment of other workers. In the employment context, the ITT does not reflect that some H-1B lottery winners' applications are rejected, but our first stage coefficient is extremely precise and quite close to 1 (ranging from 0.88 to 0.89).[26] Meanwhile, after their visas are approved by USCIS, some workers may not show up for their jobs in the U.S., for example because they die in the meantime. However, North (2011) estimates that around the time we study, nearly all (95 percent) of those approved for H-1Bs end up being admitted. Even after accounting for both of these factors together—*i.e.* by inflating the confidence intervals by a factor around 1.05 (=1/0.95) for the 2SLS regressions, or by around 1.19 (=1/(0.95*0.88)) for the median regressions—we would still conclude that new H-1Bs partially crowd out other workers in Q1 to Q4, particularly in small and medium-sized firms. After inflating, the upper end of the 95 percent confidence interval for Q1 to Q4 would be 0.87 and 0.92 in the case of the 2SLS regressions for firms of 10 or fewer or 30 or fewer employees, respectively, and would be 0.49, 0.64, and 0.94 in the case of the

[26] Of course, instrumental variables quantile regressions do not rely on a Wald estimate. In practice, however, 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.

median regressions in firms with 10 or fewer employees, 30 or fewer employees, or all firm sizes, respectively—all of which are below 1.

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 unexpected H-1Bs on the number of approved H-1B visas for applications received after the cap was reached (*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., but the majority of H-1B applications were for workers previously locating outside the U.S. (USCIS 2006, 2007). The ITT results again are policy-relevant effects of interest, regardless of whether these H-1Bs crowd out other visas.

If firms respond to an extra capped H-1B visa by reducing contracting work or outsourcing to other firms or countries—neither of which appears in our measure of employment at the firm itself—then by examining only employment at the firm, new H-1Bs will appear to be *less* substitutable with other potential employees than they actually are. Thus, it is all the more notable that we are able to rule out a coefficient on unexpected H-1Bs of one or greater. Fraud has also sometimes been alleged in the context of H-1Bs; this could lead to a larger coefficient on unexpected H-1Bs (if firms fraudulently obtain other types of visas for the workers who would have been H-1Bs if the firm had been awarded an H-1B), or to a smaller coefficient (if the firm responds to not receiving an H-1B by hiring a worker "off the books").

Effects on foreigners and non-foreigners

As described in detail in Appendix 4, Appendix Table 18 separately estimates the effect on employment of foreigners and non-foreigners. Foreigners constitute a majority of the workforce in our sample of firms, and we find that new H-1Bs crowd out employment of other foreigners at least to some extent. The point estimates suggest essentially no crowdout of U.S. natives/citizens. The confidence intervals rule out one-for-one crowdout of U.S. natives/citizens, but are compatible with a more moderate degree of crowdout of U.S. natives/citizens. We place these results in the Appendix because our two measures of the number of foreigners and non-foreigners are both imperfect—and, as the Appendix explains, one of these measures is liable to be somewhat biased toward finding crowdout of foreigners rather than non-foreigners—though the concordance of the results across two separate measures is reassuring.

Our goal is to examine the effect of additional H-1B visas specifically—a question of clear policy relevance. Some previous studies examine the effect of skilled immigrants on outcomes (e.g. Pekkala Kerr *et al*., forthcoming), but our imperfect measures of immigration status would hamper such an investigation here.

8. Effects on profits and payroll per employee

Table 8 shows the effect of unexpected H-1B visas on median firm payroll costs per employee during Years 0 to 3, calculated by dividing total firm payroll costs in a given year by the total number of employees at the firm in that year. It is possible that firms sponsoring H-1Bs could pay H-1Bs less relative to other comparable workers, for example if the frictions described earlier give sponsoring firms monopsony power. However, a reduction in average pay could appear not only if the firm pays the new H-1B less than an alternative worker, but also if the unexpected H-1B causes a reduction in average earnings of *other* employees at the firm. In firms with 10 or fewer, or 30 or fewer, employees, we find some evidence that the additional H-1B reduces median payroll costs per employee ($p<0.05$ in one estimate, and $p<0.10$ in two other estimates, of the four total). The point estimates suggest substantial decreases in payroll costs per employee in these firms (with larger and more significant estimates in the smaller size category), though the confidence intervals encompass much smaller effects. In the full sample of firms, an additional H-1B worker typically reflects only a small percentage of total employment and would be expected to influence payroll costs per employee little, and unsurprisingly we find no significant effect in these firms.[27] Appendix Table 19 shows insignificant impacts in later years, consistent with the hypothesis that in this later period the H-1B has typically left the firm and therefore no longer measurably reduces the firm's average pay.

[27] At other firm size thresholds, we typically find negative effects, though they unsurprisingly become increasingly attenuated at larger firm size thresholds. At other quantiles, we generally continue to find negative and often significant effects.

Table 9 examines the effect of an unexpected H-1B visa on the firm's profits in Years 0 to 3, using median regressions. The point estimate is positive across all the firm size cutoffs considered and is sometimes significant. The point estimates generally cluster around showing an increase in profits of five to ten thousand dollars per year (in \$2014).[28] The median regressions do not converge for many firm size cutoffs, including among firms of all sizes and for firm size thresholds over 200 employees; thus, the largest firm size cutoff we show is 200 employees or fewer. Across thresholds between 30 and 200 for which the regressions converged, the regressions generally also cluster around showing a positive effect of approximately five to ten thousand dollars per year. Overall, we find some evidence of a positive effect on profits, though it is not robustly significant. Profits regressions for later years did not converge.

Profits and payroll per employee are important outcomes, but we consider these results to be secondary because the results on profits and payroll per employee are less robust than others in the paper.[29]

Proxies for firm productivity are also of interest. Appendix Table 20 shows that the effects on revenue per employee, or total income per employee, are imprecise, which is again unsurprising given their large standard deviations.[30]

9. Conclusion

The effect of raising the H-1B visa cap is one of the centrally important U.S. immigration policy questions. We examine the marginal impact on a firm's outcomes of allowing extra capped H-1B visas to the firm, which is relevant information for those considering changing the number of H-1Bs granted by a modest amount. We find an insignificant effect of additional H-1B visas on patenting and the R&E credit, and across a variety of specifications the preponderance of evidence allows us to rule out more than modest effects, particularly in small and medium-sized firms. Parallel to these results, our

---

[28] Winsorized OLS regressions also tend to show positive point estimates. At other quantiles, many regressions did not converge, and those that did often showed imprecise results (though others showed results comparable to the median results).

[29] It is also possible that an unexpected H-1B lottery win affects a firm's competitors. We find no significant impact of unexpected H-1B lottery wins on any of the outcome variables among all other firms in that firm's 6-digit NAICS code, which is unsurprising given the large size of a six-digit industry.

[30] The effects on firm gross income, total firm payroll, or non-payroll costs are also imprecise. Ghosh, Mayda, and Ortega (2015) study effects of H-1Bs on productivity, firm size, profits, and other outcomes.

primary finding on employment is that additional H-1Bs at most increase total firm employment by a moderate amount. The preponderance of evidence indicates that H-1B workers at least partially crowd out other workers, with the estimates typically indicating substantial crowdout of other workers. This is consistent with Government Accountability Office (2011), in which firms reported that when they did not receive an H-1B, they generally filled the slot with another worker instead. It is notable that we find at most modest positive effects on patenting, R&E, and employment even among firms applying on the day the cap is reached, which 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 and intended salaries. Among firms of any size (including the largest firms), our results are illuminating in some cases—*e.g.* in the baseline we bound any increase in patenting or employment under a modest level—but in some others are imprecise.

Consistent with firm profit maximization, we find some evidence that extra H-1B visas increase median firm profits. We also find some evidence that extra H-1B visas lead to a decrease in median earnings per employee. If this reflects lower pay for H-1Bs than for alternative workers—which is not necessarily implied by our results, but is nonetheless consistent with our results—this would suggest frictions allowing such lower pay, such as firm labor market monopsony power and regulations restricting the free flow of workers across borders. Future research should try to investigate more directly whether H-1B workers' pay is consistent with prevailing wage regulations.

Overall, our results are more supportive of the narrative about the effects of H-1Bs on firms in which H-1Bs crowd out alternative workers, are paid less than the alternative workers whom they crowd out, and thus increase the firm's profits despite no measurable effect on innovation. Even though firms attest that hiring the H-1B does not adversely affect similarly employed workers, our results raise the possibility that in many cases firms could be employing H-1Bs instead of employing other workers.[31] Although we find little impact on measures of firms' quantity of innovation, further assessing impacts on measures related to productivity should be a priority for further research.

[31] 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.

Our results are consistent with the possibility that H-1B and non-H-1B workers are perfect substitutes. This is notable in light of frequent claims that H-1Bs have unique skills that cannot easily be obtained elsewhere. However, one cannot interpret our estimates as necessarily *implying* that H-1Bs are perfect substitutes with other technical workers. Our study focuses on estimating the causal impacts of additional H-1Bs, which could provide some of the building blocks for estimating parameters such as the elasticity of substitution between new H-1Bs and other workers in future work. However, such an estimate would be limited by having one instrument—unexpected wins—but multiple relevant parameters. The degree of crowdout of other workers should depend not only on the degree of substitutability or complementarity of additional H-1B and other workers (and/or labor and capital), but also on factors like possible frictions in matching firms with workers (*e.g.* search frictions).[32] If the firm faces frictions in finding a new employee that *limit* the degree of crowdout of other workers, it would be all the more notable that we find that a new H-1B worker *does* partially crowd out other workers, and that we cannot rule out that a new H-1B worker has no effect on total employment.

In several notable ways, our study is different from previous work on the effects of H-1Bs on outcomes in the receiving economy. First, ours is the only study that relies on random variation. Second, we examine the effects of H-1Bs given to a particular firm on that firm's outcomes (holding constant H-1Bs at other firms), but our empirical strategy does not estimate general equilibrium effects like impacts on employment, innovation, pay, or profits in the entire U.S., or in specific areas of the U.S. Third, some previous literature examines the effects of temporary visas in general, not specifically capped H-1B visas (though other literature is identified from changes in the H-1B cap, *e.g.* Kerr and Lincoln 2010). Fourth, our results are estimated from the FY2006 and 2007 lotteries, which may differ from other environments (*e.g.* Kerr and Lincoln 2010 exploit variation in the cap that also covers other time periods). At first pass, our results also apparently differ starkly from those in previous economics literature, as previous work has found very large positive effects of H-1Bs on patenting, and in some cases on

[32] Lewis (2011) studies the interaction of immigration with capital.

employment.[33] Though our paper focuses only on estimating the effects of additional capped H-1Bs on firm outcomes, future work could try to clarify to what extent the seeming divergence in results between our study and previous literature relates to any of the four factors above (or other factors). For example, large aggregate effects could potentially be reconciled with small firm-level effects through very large spillovers onto other workers, including very strongly increasing returns to scale. Finally, to address the effects of H-1Bs in other years, it could be helpful for those running the H-1B lotteries to begin regularly saving the data on lottery winners and losers.

[33] For example, Kerr and Lincoln (2010) find that a 10 percent increase in a city's H-1B population led to a 0.3 percent to 0.7 percent increase in total patenting for each standard deviation growth in "city dependency," a measure of H-1B applications per capita in each city. Given the standard deviation of city dependency in their sample, this would imply an increase in patenting at least 10 times as large as the maximum positive effect allowed by our 95% confidence interval. However, these estimates are not directly comparable to ours, as described above.

References

Acs, Zoltan, and David Audretsch. *Innovation and Small Firms* (1990). MIT Press.

Åslund, Olof, Per-Anders Edin, Peter Fredriksson, and Hans Grönqvist. "Peers, Neighborhoods and Immigrant Student Achievement: Evidence from a Placement Policy." *American Economic Journal: Applied Economics* 3.2 (2011): 67-95.

Azoulay, Pierre, Waverly Ding, and Toby Stuart. "The Determinants of Faculty Patenting Behavior: Demographics or Opportunities?" *Journal of Economic Behavior & Organizations*, 63.4 (2007): 599-623.

Bloom, Nicholas, Mark Schankerman, and John van Reenen. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica* 81.4 (2013): 1347-1393.

Borjas, George. "The Labor Demand Curve Is Downward-Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics* 118 (2003): 1335-374.

Borjas, George. "The Economics of Immigration." *Journal of Economic Literature* 32 (1994): 1667–1717.

Borjas, George, and Kirk Doran. "Cognitive Mobility: Native Responses to Supply Shocks in the Space of Ideas." *Journal of Labor Economics* (forthcoming 2015).

Borjas, George, and Kirk Doran. "The Collapse of the Soviet Union and the Productivity of American Mathematicians." *Quarterly Journal of Economics* 127.3 (2012): 1143-203.

Borjas, George, Richard Freeman, and Lawrence Katz. "How Much Do Immigration and Trade Affect Labor Market Outcomes?" *Brookings Papers on Economic Activity* (1997): 1-90.

Borjas, George, Jeffrey Grogger, and Gordon Hanson. "Comment: Substitution between Immigrants, Natives, and Skill Groups." *Journal of the European Economic Association* 10.1 (2012): 198-210.

Bound, John, Breno Braga, Joseph Golden, and Gaurav Khanna. "Recruitment of Foreigners in the Market for Computer Scientists in the U.S." Forthcoming, *Journal of Labor Economics*.

Burbidge, John, Lonnie Magee, and A. Leslie Robb. "Alternative Transformations to Handle Extreme Values of the Dependent Variable." *Journal of the American Statistical Association* 83 (1988): 123-127.

Card, David. "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review* 43.2 (1990): 245-57.

Card, David. "Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration." *Journal of Labor Economics* 19.1 (2001): 22-64.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. "How does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *The Quarterly Journal of Economics* 126.4 (2011): 1593-1660.

Doran, Kirk, Alexander Gelber, and Adam Isen. "The Effect of High-Skilled Immigration on Patenting and Employment: Evidence from H-1B Visa Lotteries." NBER Working Paper 20668.

Dustmann, Christian, Albrecht Glitz, and Tommaso Frattini. "The Labour Market Impact of Immigration." *Oxford Review of Economic Policy* 24.3 (2008): 477-94.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund. "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment." *The Quarterly Journal of Economics* 118.1 (2003): 329-57.

Foley, Fritz, and William Kerr. "Ethnic Innovation and U.S. Multinational Firm Activity." *Management Science* 59.7: 1529-1544.

Freeman, Richard. "People Flows in Globalization." *Journal of Economic Perspectives* 20.2 (2006): 145–70.

Friedberg, Rachel M. "The Impact of Mass Migration on the Israeli Labor Market." *The Quarterly Journal of Economics* 116.4 (2001): 1373-408.

Friedberg, Rachel, and Jennifer Hunt. "The Impact of Immigrants on Host Country Wages, Employment and Growth." *Journal of Economic Perspectives* 9.2 (1995): 23-44.

Gates, William H. "Testimony before the Committee on Science and Technology." U.S. House of Representatives, Washington D.C. 12 March. 2008. Address.

Gelber, Alexander M. "How Do 401(k)s Affect Saving? Evidence from Changes in 401(k) Eligibility." *American Economic Journal: Economic Policy* 3.4 (2011): 103-22.

Government Accountability Office. "H-1B Visa Program: Reforms Are Needed to Minimize the Risks and Costs of Current Program." Washington, D.C.: Government Printing Office (2011).

Ghosh, Anirban, Anna Maria Maybe, and Francesc Ortega. "The Impact of Skilled Foreign Workers on Firms: an Investigation of Public Traded U.S. Firms." CUNY Working Paper (2015).

Grogger, Jeff, and Gordon Hanson. "Attracting Talent: Location Choices of Foreign-Born Ph.D.s in the U.S." Forthcoming, *Journal of Labor Economics*.

Hall, Bronwyn, and Dietmar Harhoff. "Recent Research on the Economics of Patents." *Annual Review of Economics* 4.1 (2012): 541-565.

Hall, Bronwyn, and John Van Reenen. "How Effective are Fiscal Incentives for R&D? A Review of the Evidence." *Research Policy* 29 (2000): 449–469.

Hall, Bronwyn, Jacques Mairesse, and Pierre Mohnen. "Measuring the Returns to R&D." In Bronwyn Hall and Nathan Rosenberg, eds., *Handbook of the Economics of Innovation* (2010): 1033-1082. Amsterdam: Elsevier.

Hanson, Gordon. "The Economic Consequences of International Migration." *Annual Review of Economics* 1 (2009): 179-208.

Hira, Ron. "The H-1B and L-1 Visa programs: Out of Control", EPI Policy Paper (2010).

Hunt, Jennifer. "Which Immigrants Are Most Innovative and Entrepreneurial? Distinctions by Entry Visa." *Journal of Labor Economics* 29.3 (2011): 417-57.

Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. "How Much Does Immigration Boost Innovation." *American Economic Journal:Macroeconomics* 2.2 (2010): 31-56.

Jones, Benjamin. 2010. "Age and Great Invention." *Review of Economics and Statistics* 92(1): 1-14.

Kerr, William R. "U.S. High-Skilled Immigration, Innovation, and Entrepreneurship: Empirical Approaches and Evidence." HBS Working Paper (2013): 14-17.

Kerr, William R., and William F. Lincoln. "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention." *Journal of Labor Economics* 28.3 (2010): 473-508.

Levin, Sharon, and Paula Stephan. "Research Productivity over the Life Cycle: Evidence for Academic Scientists." *American Economic Review* 81.1 (1991): 114-132.

Lerner, Josh, and Amit Seru. "The use and misuse of patent data: Issues for corporate finance and beyond." Booth/Harvard Business School Working Paper.

Lewis, Ethan. "Immigration, Skill Mix, and Capital Skill Complementarity." *Quarterly Journal of Economics* 126.2 (2011): 1029-1069.

Lowell, B. Lindsay. "H-1B Temporary Workers: Estimating the Population." UCSD Center for Comparative Immigration Studies Working Paper No. 12 (2000).

Lubotsky, Darren. "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings." *Journal of Political Economy* 115.5 (2007): 820-867.

Matloff, Norman. "On the Need for Reform of the H-1B Non-immigrant Work Visa in Computer-Related Occupations." *University of Michigan Journal of Law Reform* 36.4 (2003): 815-914.

Miano, John. "H-1B Visa Numbers: No Relationship to Economic Need." *Center for Immigration Studies Policy Brief* (2008).

Moser, Petra, Alessandra Voena, and Fabian Waldinger. "German Jewish Émigrés and U.S. Invention." *American Economic Review* 104: 3222-3255.

Nagaoka, Sadao, Kazuyuki Motohashi, and Akira Goto. "Patent Statistics as an Innovation Indicator." In Bronwyn Hall and Nathan Rosenberg, eds., *Handbook of the Economics of Innovation* (2010): 1083-1127. Amsterdam: Elsevier.

National Foundation for American Policy. "H-1B visas and job creation." *Policy Brief*, Arlington, VA (2008).

North, David. "Estimating the Size of the H-1B Population in the U.S." *Center for Immigration Studies Memorandum* (2011).

Pekkala Kerr, Sari, and William R. Kerr. "Economic Impacts of Immigration: A Survey." *Finnish Economic Papers* 24.1 (2011): 1-32.

Pekkala Kerr, Sari, William R. Kerr, and William F. Lincoln. "Skilled Immigration and the Employment Structures of U.S. Firms." *Journal of Labor Economics* (forthcoming).

Pence, Karen. "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 (2006): 1-24.

Peri, Giovanni, Kevin Shih, and Chad Sparber. "STEM Workers, H-1B Visas, and Productivity in US Cities." *Journal of Labor Economics* (forthcoming).

Peri, Giovanni, Kevin Shih, and Chad Sparber. "The Effects of Foreign Skilled Workers on Natives: Evidence from the H-1B Visa Lottery." UC Davis Working Paper (2015).

Stuen, Eric, Ahmed Mushfiq Mobarak, and Keith E. Maskus. "Skilled Immigration and Innovation: Evidence from Enrolment Fluctuations in U.S. Doctoral Programs." *Economic Journal* 122 (2012): 1143-1176.

U.S. Customs and Immigration Services. "Characteristics of Specialty Worker Occupations." Washington, D.C.: U.S. Government Printing Office (2006, 2007).

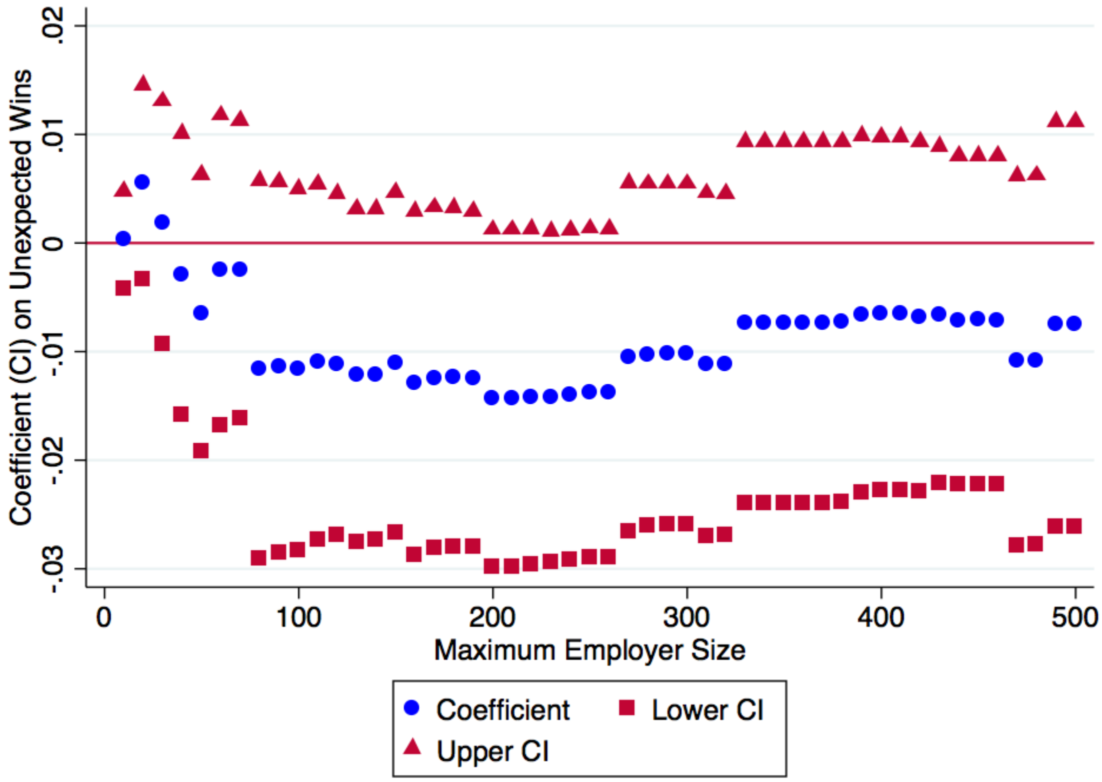
U.S. Customs and Immigration Services. "Change in H-1B Procedures Trims Weeks Off Final Selection Process." Web. http://www.uscis.gov/sites/default/files/files/pressrelease/H1Bfy08CapUpdate041907.pdf (accessed October 7, 2014).

U.S. Department of Labor. Office of Foreign Labor Certification Data. Web. http://www.foreignlaborcert.doleta.gov/performancedata.cfm#stat (accessed October 7, 2014).

U.S. Patent and Trade Office. "Performance and Accountability Report: Fiscal Year 2012." Washington, D.C.: U.S. Government Printing Office.

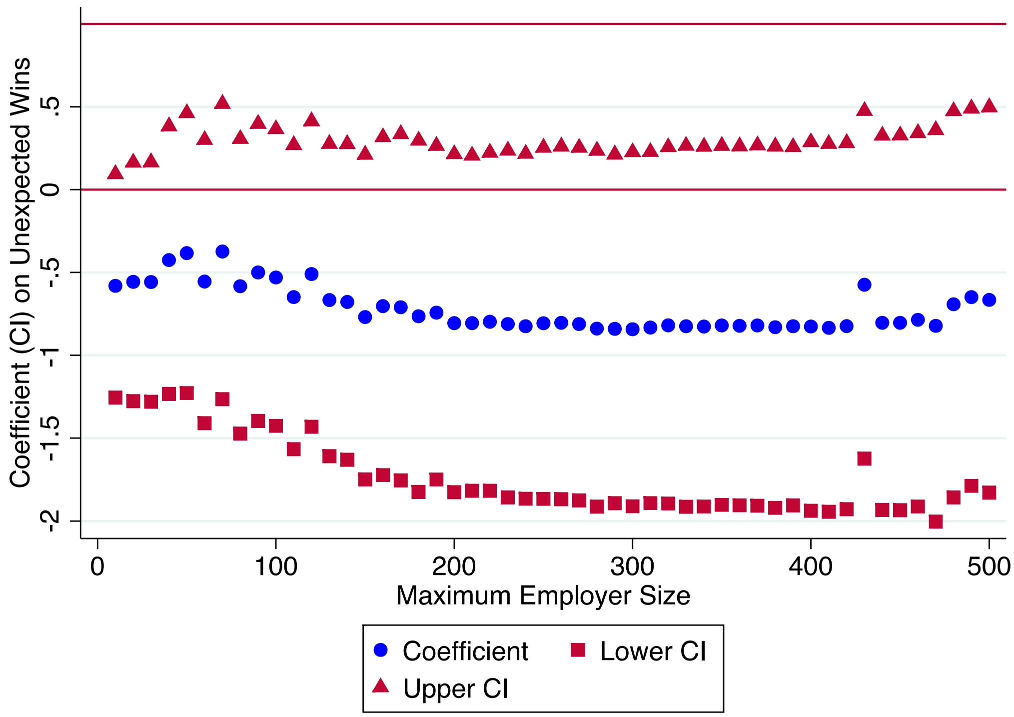
Yagan, Danny. "Moving to Opportunity? Migratory Insurance over the Great Recession." UC Berkeley Working Paper (2014).

Figure 1. *Effect of Unexpected H-1B Visas on Patents, by Employer Size*



Notes: The figure shows the coefficient and 95 percent confidence interval on unexpected H-1B visas when the dependent variable is the inverse hyperbolic sine 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). We show the coefficient for employers of each size range 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 patenting specification, in which we control for lagged number of patents and expected lottery wins. After multiplying by 100, the coefficient should be interpreted as the approximate percentage increase in firm patenting due to an unexpected H-1B visa lottery win.

Figure 2. *Effect of H-1B Visas on Total Firm Employment, by Employer Size*



Notes: The figure shows the coefficient and 95 percent confidence interval on unexpected lottery wins from median regressions in which the dependent variable is the total number of employees in a firm, 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 +1 on the *y*-axis corresponds to the case where hiring an extra H-1B visa worker leaves other employment unchanged (so that total employment would increase by exactly one). The horizontal line at 0 on the *y*-axis corresponds to the case where hiring an extra H-1B visa worker crowds out other workers one-for-one (so that total employment would increase by zero). We show the coefficient for employers of each size range 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 lagged employment and expected lottery wins.

Table 1. *Summary Statistics*

Variable	Mean (SD)	*n*
Number of patents (all)	4.52 (56.11)	3,050
Number of patents (≤30)	0.23 (8.59)	1,276
Number of patents (≤10)	0.23 (0.49)	749
IHS of patents (all)	0.15 (0.80)	3,050
IHS of patents (≤30)	0.017 (0.22)	1,276
IHS of patents (≤10)	0.010 (0.14)	749
Fraction patenting (all)	0.048 (0.21)	3,050
Fraction patenting (≤30)	0.010 (0.10)	1,276
Fraction patenting (≤10)	0.0075 (0.086)	749
IHS of R&E (all)	1.55 (4.74)	1,000
IHS of R&E (≤30)	0.15 (1.39)	470
IHS of R&E (≤10)	0.14 (1.22)	284
Fraction with R&E (all)	0.099 (0.30)	1,000
Fraction with R&E (≤30)	0.013 (0.11)	470
Fraction with R&E (≤10)	0.013 (0.11)	284
Number of employees (all)	1,877.84 (39,721.31)	2,281
Number of employees (≤30)	43.09 (1,904.34)	1,183
Number of employees in Q1-Q4 (≤10)	9.64 (55.63)	712
Median employees (all)	31	2,281
Median employees (≤30)	10	1,183
Median employees (≤10)	6	712
Winsorized emp. first diff. (all)	27.28 (92.39)	2,281
Winsorized emp. first diff. (≤30)	4.35 (9.43)	1,183
Winsorized emp. first diff. (≤10)	3.22 (6.84)	712
Median payroll per employee (all)	$49,331.89	2,191
Median payroll per employee (≤30)	$42,280.76	1,123
Median payroll per employee (≤10)	$38,656.64	636
Median firm profits (≤200)	$80,249.73	1,520
Median firm profits (≤30)	$43,300.70	1,033
Median firm profits (≤10)	$30,397.45	615
Fraction winning lottery		
2006 Regular	0.038	2,687
2006 ADE	0.17	306
2007 Regular	0.98	3,954
2007 ADE	0.55	296
Fraction in NAICS=54 (all)	56.43	3,050
Fraction in NAICS=54 (≤30)	65.60	1,276
Fraction in NAICS=54 (≤10)	64.62	749

Notes: The data are from IRS and USCIS administrative sources, and from the Patent Dataverse. "All" refers to the full sample of firms entering the lottery; "≤30" ("≤10") refers to those firms with 30 (10) or fewer employees in Year -1. 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 ≤200 employees; our regressions did not converge for larger thresholds. NAICS code 54 is professional, scientific, and technical services. "*n*" refers to the number of firm-lottery years in the sample (*i.e.* firms appearing in both lottery years count as two observations), except when reporting the fraction winning the lottery, where we report the number of applications entering the lottery. *n*'s vary across outcomes because the number of missing observations in the IRS data varies across outcomes; here and everywhere else, the results are similar when we restrict to the same sample across outcomes. 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 mean of a yearly dummy for taking the R&E 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. *Comparison of Applications on Day of Lottery to Other Applications*

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

Table 3. *Validity of the Randomized Design*

Dependent Variable	Coefficient (SE) on Unexpected Wins
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)
Number of patents in Year -2 (all)	-0.94 (1.34)
Number of patents in Year -2 (≤30)	0.037 (0.036)
Number of patents in Year -2 (≤10)	0.0029 (0.010)
IHS of patents in Year -2 (all)	0.0019 (0.019)
IHS of patents in Year -2 (≤30)	-0.013 (0.019)
IHS of patents in Year -2 (≤10)	-0.0028 (0.0044)
Patented in Year -2 (all)	-0.0024 (0.0089)
Patented in Year -2 (≤30)	-0.0077 (0.011)
Patented in Year -2 (≤10)	-0.0033 (0.0033)
IHS of R&E in Year -2 (all)	-0.30 (0.28)
IHS of R&E in Year -2 (≤30)	-0.0037 (0.015)
IHS of R&E in Year -2 (≤10)	-0.0040 (0.0034)
Fraction with R&E in Year -2 (all)	-0.019 (0.017)
Fraction with R&E in Year -2 (≤30)	-0.00029 (0.0013)
Fraction with R&E in Year -2 (≤10)	0.00 (0.00)
Employment in Year -2 (all, median)	0.50 (1.30)
Employment in Year -2 (≤30, median)	-0.55 (0.81)
Employment in Year -2 (≤10, median)	-0.31 (0.69)
Employment in Year -2 (all, winsorized first-difference)	0.082 (9.71)
Employment in Year -2 (≤30, winsorized first-difference)	0.56 (0.89)
Employment in Year -2 (≤10, winsorized first-difference)	-0.091 (0.57)
Payroll per employee in Year -2 (all, median)	91.01 (594.95)
Payroll per employee in Year -2 (≤30, median)	1,591.82 (1,519.61)
Payroll per employee in Year -2 (≤10, median)	1,645.07 (3,141.91)
Profits in Year -2 (≤200, median)	-6,268.96 (4,528.82)
Profits in Year -2 (≤30, median)	-8,027.92 (5,498.00)
Profits in Year -2 (≤10, median)	-20,306.35 (19,756.56)
Dummy for NAICS=54 (all)	0.007 (0.03)
Dummy for NAICS=54 (≤30)	-0.033 (0.043)
Dummy for NAICS=54 (≤10)	0.010 (0.058)

Notes: The table regresses placebo outcomes on unexpected H-1B lottery wins. We run OLS regressions for outcomes when our 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 we run 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 (in order of appearance): 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, by investigating Year -2 outcomes, we can 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 Year -1 outcomes are the dependent variables, and we control for Year -2 values of the variables, the regressions are insignificant except in one of 27 cases. We investigate the profits regressions in the sample with 200 or fewer employees because the regressions did not converge for the full sample. "Winsorized first-difference" means that the dependent variable is the first-difference of employment between Year -2 and Year -1, winsorized at the 5th and 95th percentiles. Standard errors are clustered by firm. Table 1 shows sample sizes. *** means p<1%; ** p<5%; and * p<10%.

Table 4. *Effect of H-1B Lottery Wins on Patenting*

	IHS of # patents		Patenting dummy	
	(1)	(2)	(3)	(4)
Panel A: ≤10 employees				
A) Years 0 to 8	0.00023	0.00026	-0.0010	-0.0010
	[-0.0046, 0.0050]	[-0.0042, 0.0047]	[-0.0042, 0.0022]	[-0.0041, 0.0020]
B) Years 0 to 3	-0.00033	-0.00015	-0.0023	-0.0023
	[-0.0090, 0.0084]	[-0.0082, 0.0079]	[-0.0083, 0.0037]	[-0.0081, 0.0035]
Panel B: ≤30 employees				
C) Years 0 to 8	0.0017	0.0018	-0.0029	-0.0028
	[-0.0096, 0.013]	[-0.0094, 0.013]	[-0.0095, 0.0038]	[-0.0094, 0.0038]
D) Years 0 to 3	-0.00053	-0.00030	-0.0047	-0.0046
	[-0.018, 0.017]	[-0.018, 0.017]	[-0.016, 0.0064]	[-0.016, 0.0065]
Panel C: All				
E) Years 0 to 8	-0.0087	-0.0089	-0.0014	-0.0012
	[-0.038, 0.020]	[-0.037, 0.019]	[-0.014, 0.011]	[-0.014, 0.011]
F) Years 0 to 3	-0.021	-0.021	-0.011	-0.010
	[-0.052, 0.010]	[-0.052, 0.010]	[-0.026, 0.0039]	[-0.025, 0.0041]
Prior patents	X	X	X	X
E[wins]		X		X

Notes: The table shows OLS regressions of the IHS of patents in each year over Years 0 to 8 (Rows A, C, and E), or over the duration of the H-1B visa in Years 0 to 3 (Rows B, D, and F), on unexpected H-1B lottery wins, defined as actual wins minus the expectation of wins conditional on number of applications and the probability each application wins. We pool and stack years. The table shows coefficients on unexpected H-1B visas, with 95 percent confidence intervals in brackets. The "prior patents" specifications control for the IHS of the total number of patents in Year -1 (for the IHS specifications), or for a dummy for patenting in Year -1 (for the patenting dummy specifications). The "prior patents, E[wins]" specifications control for patents 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 coefficients in the IHS specifications should be interpreted as the approximate percent effect on the number of patents. "Patenting dummy" refers to 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. See Table 1 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.

Table 5. *Effect of H-1B Lottery Wins on Research and Experimentation Credit*

	Amount of credit (IHS)		Claiming dummy	
	(1)	(2)	(3)	(4)
A) ≤10 employees	-0.13	-0.12	-0.012	-0.011
	[-0.30, 0.043]	[-0.27, 0.041]	[-0.027, 0.0043]	[-0.025, 0.0041]
B) ≤30 employees	-0.073	-0.065	-0.0069	-0.0061
	[-0.16, 0.018]	[-0.15, 0.018]	[-0.015, 0.0016]	[-0.014, 0.0016]
C) All	0.19	0.19	0.016	0.016
	[-0.33, 0.70]	[-0.33, 0.72]	[-0.018, 0.049]	[-0.018, 0.049]
Prior R&E credit	X	X	X	X
E[wins]		X		X

Notes: The table shows OLS regressions of the R&E credit over the duration of the H-1B visa (pooling and stacking Years 0 to 3), on unexpected H-1B lottery wins. The table shows coefficients on unexpected H-1B visas, with 95 percent confidence intervals in brackets. In Columns 1 and 2, the dependent variable is the IHS of the amount of the R&E credit claimed in each year over Years 0 to 3. In Columns 3 and 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. See Tables 1 and 4 for additional notes and sample sizes. *** refers to significance at the 1% level; ** at the 5% level; and * at the 10% level.

Table 6. *Effect of H-1B Lottery Wins on Employment in First Year*

	Median Regressions		Two-stage least squares	
	(1)	(2)	(3)	(4)
A) ≤10 employees	-0.53	-0.52	-0.54	-1.10
	[-1.18, 0.12]***	[-1.15, 0.11]***	[-1.95, 0.88]**	[-2.88, 0.68]**
B) ≤30 employees	-0.44	-0.36	-0.97	-1.26
	[-1.16, 0.28]***	[-1.09, 0.37]***	[-2.96, 1.01]*	[-3.25, 0.71]**
C) All	-1.27	-1.05	-20.37	-2.41
	[-3.08, 0.55]***	[-2.67, 0.57]**	[-230.99, 190.24]	[-17.76, 12.94]
Prior employment	X	X	X	X
E[wins]		X		X

Notes: The table shows point estimates and 95% confidence intervals on unexpected lottery wins. The first two columns show median regressions of total firm employment in Q1 to Q4, on unexpected lottery wins. The next two columns show 2SLS regressions where the dependent variable, the difference of total firm employment from the first quarter of Year -1 to the quarter in question from Q1 to Q4, has been winsorized at the 95th percentile. We pool and stack observations across quarters. The 5th and 95th 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 unexpected lottery wins and the endogenous variable is approved capped H-1B visas. The "prior employment" specifications control for employment from the first quarter of Year -1, and the "prior employment, E[wins]" specifications additionally control for the number of expected lottery wins. See Tables 1 and 4 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 1 corresponds to no crowd-out or crowd-in of other employment, and a coefficient of 0 corresponds to one-for-one-crowdout of other employment. None of the estimates is significantly different from 0 at any conventional significance level.

Table 7. *Effect of Unexpected Lottery Wins on Later Employment*

Outcome	(1) All	(2) ≤30 employees	(3) ≤10 employees
A) Q5-Q8	-2.03	-0.95	-0.99
	[-4.97, 0.90]**	[-2.29, 0.39]***	[-2.05, 0.065]***
n	2,213	1,142	682
B) Q9-Q12	-1.97	-1.57	-1.02
	[-5.46, 1.52]*	[-3.70, 0.56]**	[-2.28, 0.25]***
n	2,120	1,087	647
C) Q13-Q32	-3.24	-0.0096	0.92
	[-7.14, 0.67]**	[-2.26, 2.25]	[-1.31, 3.14]
n	2,048	1,045	618

Notes: The table shows the effect of unexpected lottery wins on employment in later time periods, reporting point estimates and 95 percent confidence intervals in square brackets for median regressions of employment on unexpected lottery wins. We pool and stack observations across quarters. All specifications control for employment in Year -1 and expected lottery wins, as in the baseline. *n*'s refer to the number of firms in each regression. See Table 6 for additional notes. Sample sizes fall in later years because fewer firms are still in business. In Q13 to Q32, the H-1B visa has expired and the worker has typically left the firm, so the test of a difference in the coefficient from 1 no longer indicates crowdout of other workers. *** shows estimates that are significantly different *from 1* at the 1% level; ** at the 5% level; * at the 10% level. None of the estimates is significantly different from zero at any conventional significance level.

Table 8. *Effect of Unexpected Lottery Wins on Payroll per Employee*

	(1) Fewer controls	(2) More controls
(A) ≤10 employees	-4,527.58	-4,860.54
	[-9,258.68, 203.52]*	[-9,552.97, -168.12]**
B) ≤30 employees	-2,618.66	-2,725.03
	[-6,200.56, 963.24]	[-5,976.60, 526.54]*
C) All firm sizes	26.64	80.21
	[-1,277.42, 1,330.69]	[-1,348.07, 1,508.50]
Prior payroll per employee	X	X
E[wins]		X

Note: The table shows median regressions of payroll costs per employee in Years 0 to 3 on unexpected H-1B visas and controls, pooling and stacking years. Years 0 to 3 cover the duration of the H-1B visa. The table shows coefficients and 95 percent confidence intervals on unexpected H-1B visas. 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. See Table 1 for sample sizes. Standard errors are clustered by firm. *** refers to $p<0.01$; ** to $p<0.05$; and * to $p<0.10$.

Table 9. *Effect of Unexpected Lottery Wins on Profits*

	(1) Fewer controls	(2) More controls
(A) ≤10 employees	8,163.43	6,518.156
	[-4,724.93, 21,051.79]	[-6,942.69, 19,979.00]
B) ≤30 employees	3,970.10	11,468.61
	[-6,583.254, 14,523.46]	[200.86, 22,736.37]**
C) ≤200 employees	11,538.41	2,526.67
	[-1,490.03, 24,566.86]*	[-32,168.54, 37,221.88]
Prior profits	X	X
E[wins]		X

Notes: The table shows median regressions of profits in Years 0 to 3 on unexpected H-1B visas and controls, pooling and stacking years. The table shows coefficients and 95% confidence intervals on unexpected H-1B visas. 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. We do not show the effect on median profits in Years 4 to 8 because it is unstable and often did not converge. Standard errors are clustered by firm. See Table 1 for sample sizes. *** refers to significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendices (for online publication)

Appendix 1. Description of matching procedure

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.

In the results that we report in the main tables, we exclude the 63 possible matches from the list of matched companies. In the Appendix, 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.

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.

Appendix 2. Description of Heterogeneity Results

Heterogeneity in patenting effects

We examine heterogeneity in the patenting effects across subsamples in Appendix Table 8. Row A examines the Regular H-1B lottery. The results are comparable to those in the full sample—with point estimates that cluster near zero, and the upper end of the 95 percent confidence interval ruling out more than a modest effect—which should not be surprising since 85.96 percent of the full sample participates in the Regular lottery. Row B examines the ADE lottery, where the confidence intervals also rule out more than a modest effect[34]

The effect on patenting is particularly relevant in professional, scientific, and technical services (NAICS code 54), since the bulk of patents occur in this industry. We find no evidence of an effect on patenting in this group, with confidence intervals that again rule out more than a modest effect. In firms outside NAICS code 54, the results are comparable.

Many H-1Bs are given for workers in firms like Infosys or Wipro that primarily offer outsourcing for temporary support services (often temporary technical support services). By contrast, other H-1Bs are given to companies like Intel or Google that do not specialize in such services. Although it is not possible to determine with certainty which visas fall in the broadly-defined "temporary support services" category, it is illuminating to investigate the effects in firms that likely specialize in such services. To probabilistically identify such firms, we first compiled a list of those firms among the largest 100 H-1B sponsors and that had "outsourcing services" or "IT support services" in the description of the company on its website. We found that these firms were in only

[34] When we investigate the effect separately in each year of the lottery (*i.e.* separating the FY2006 lotteries from the FY2007 lotteries), or separately in each of the four lotteries (FY2006 Regular, FY2006 ADE, FY2007 Regular, and FY2007 ADE), we again estimate insignificant effects in each year separately, with comparable point estimates to those in the full sample, though again with larger confidence intervals.

seven, six-digit NAICS categories.[35] We then ran our regressions only in firms in these industries, and separately ran the regressions only among firms in other industries. The point estimates and top end of the 95 percent confidence intervals are smaller in "temporary support services" industries, although the estimates are insignificant in both sets of industries (and insignificantly different across the two different samples).

Appendix Table 9 shows the coefficients on interactions of unexpected H-1B visas with continuous covariates. In principle, it is possible that the H-1B visa could tend to have more (or less) positive effects on firms that apply earlier for the visas. For example, the visas have the largest positive effects in such firms, motivating their earlier applications. In Appendix Table 9 Row A we interact the number of unexpected H-1B visas with the number of days taken to reach the cap in each lottery (which ranges across the four lotteries from 55 to 291). We find no significant interaction in Column 1. However, this evidence is merely suggestive: heterogeneity across the lotteries in the effect of H-1Bs visas that happens to be correlated with the time taken to reach the cap would confound our estimate of the interaction. In Row B Column 1 we show that the interaction of the IHS of prior patents (Year -1) with unexpected visa lottery wins is insignificant. Note that the estimates shown in Row A and Row B are from separate regressions. The interaction of unexpected visas with prior firm size is also insignificant.

Heterogeneity in employment effects

Appendix Table 17 investigates whether there is heterogeneity in the employment results across samples, using our baseline employment specification in Q1 to Q4 with median regressions and the more extensive set of controls. (Other specifications show similar results.) The point estimates 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. In fact, the point estimates are often positive and substantial in the case of the ADE lotteries, and in the case of non-scientific services—particularly when we examine firms of all sizes. The point estimates are negative in likely "temporary support services" employers but positive in other six-digit industries (though the estimates are insignificantly different across the two samples), and among "temporary support services" the coefficient estimate can be distinguished from unity in more firm size categories than in other industries. However, there are no significant differences across the different samples, including when we compare the 2006 and 2007 lotteries.

In Appendix Table 9 Column 2 shows that the estimated interaction of unexpected wins with the number of days taken to reach the cap is positive but insignificant. It also shows that the interaction of unexpected visas with the IHS of prior patents is extremely imprecise. The interaction of unexpected visas with prior firm size is also insignificant.[36]

[35] The NAICS codes are 541511, 541519, 541600, 541330, 519100, 423600, and 541512.

[36] More generally, there are many factors that theoretically could influence the size of the impacts, but we tend to find that the estimates are similar across groups.

Appendix 3. Further description of results when parallel specifications are run in patenting and employment contexts

It is worth additionally describing further results when we run other parallel specifications in the patenting and employment contexts. When the first-difference (or level) of the number of patents (or the IHS of patents) is the dependent variable and we winsorize at the 95^{th} (or 99^{th}) percentile, parallel to those in the employment context, our results are very similar to those shown in Appendix Table 3 (or Table 4) but are more precise and allow us to bound the maximum increase in patenting at a still lower level. These first differences are taken by subtracting patents in the pre-period (*i.e.* Year -1) from patents in Years 0 to 8.

When we run the two-stage least squares employment regressions but do not winsorize the dependent variable, the results are extremely imprecise among firms of all sizes or among firms with 30 or fewer employees in Year -1, which is unsurprising given the very large standard deviation of employment and long right tail. However, when we do not winsorize and run this specification among firms with 10 or fewer employees in Year -1, the top end of the 95 percent confidence interval is 0.31, and we are able to rule out a coefficient of 1 (p=0.015).

In sum, running parallel specifications in these two main contexts does not change any of our conclusions, except that our results are unsurprisingly imprecise when we examine employment in a two-stage least squares regression and do not winsorize. All of these results are available upon request.

Appendix 4. Estimating effects on employment of foreigners and non-foreigners

Measure 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. Although citizenship status is available through IRS data on W-2 forms, these data only have information on the individual's citizenship status *most recently reported* to the Social Security Administration (SSA), as opposed to always being measured in the year in question in our regressions (*e.g.* Year 0 or Year 1). Thus, one way to measure citizenship status is through this measure, which will probabilistically identify those who were citizens and non-citizens around the time of the lotteries (though with measurement error). We use W-2 data in this case (rather than form 941 data) because the W-2 data have this information on citizenship, but the form 941 data have no information on citizenship.

The data on past citizenship status is not directly available, which is a relevant limitation because a substantial fraction of H-1Bs go on to become permanent residents and in many cases citizens (Lowell 2000). At the same time, for many of those who go on to become permanent residents or citizens, the SSA data will *not* reflect the their updated citizenship status, for example because the Tax Identification Number under which they filed taxes as a non-citizen no longer applies once they become a citizen and gain a Social Security Number; thus, our measure of citizenship status is likely to estimate citizenship status at the time of being admitted to the U.S. with only modest error.

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 a high 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.[37]

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. Thus, 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 an immigrant, but those whose SSNs indicate that they applied early in life have a much smaller probability of being an immigrant. 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. Our results are robust to choosing other thresholds, as well as to assigning different cutoffs (*e.g.* 15 percent rather than 10 percent) in different geographic areas with different percentages of immigrants as identified in Census data (results available upon request).[38]

Estimated effect on employment of foreigners and non-foreigners

We estimate the effect on employment of foreigners vs. natives in Appendix Table 18. 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.[39] We investigate our baseline

[37] See http://www.ssa.gov/history/ssn/geocard.html
[38] Even if both were perfectly measured, citizenship at the time of the lotteries (or in the most recent IRS data) could be different than whether an individual is a native—namely, in those cases in which a non-native became a citizen prior to the time of the lotteries. Thus, there is no presumption that regressions with number of natives as the dependent variable should show the same results as regressions in which the dependent variable is the number of citizens at a later point in time.
[39] When the dependent variable is overall employment, the W-2 data show comparable results to the form 941 data. Of course, in interpreting the median regressions, we must recognize that the effects across

specification across the three employer size categories we investigate elsewhere, though our results hold robustly across other employer size thresholds and other specifications.

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 unexpected 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).

An important caveat to the results in Rows A and B is that because the IRS data measure most recent citizenship status rather than citizenship status at the time of application, these 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.

To address this ambiguity of interpretation, we also show results in Appendix Table 18 (rows C and D) where we probabilistically identify natives and non-natives using their SSNs as in Yagan (2014). 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.

separate regressions for foreigners and non-foreigners do not "add" to the median effect on overall employment.

Appendix Tables (for online publication)

Appendix Table 1. *First stage regressions*

Sample	Employment First Stage		Patenting First Stage	
	Coefficient (SE) on Unexpected Lottery Wins	First-stage F-statistic	Coefficient (SE) on Unexpected Lottery Wins	First-stage F-statistic
All	0.88 (0.029)***	935.14	0.87 (0.027)***	1053.65
≤30	0.89 (0.040)***	495.51	0.88 (0.042)***	435.14
≤10	0.88 (0.052)***	281.57	0.86 (0.059)***	214.04

The table shows the first stage regression of the number of approved H-1Bs on the number of unexpected wins. See Table 1 for other notes and sample sizes. *** denotes $p<0.01$; ** denotes $p<0.05$; * denotes $p<0.10$.

Appendix Table 2. *Effect of Unexpected Lottery Wins on Patenting, Sample Reweighted to Match Full Population*

	(1) Fewer controls	(2) More controls
(A) ≤10 employees	0.0025	0.0016
	[-0.0056, 0.011]	[-0.0024, 0.0056]
B) ≤30 employees	0.0027	0.0028
	[-0.0082, 0.014]	[-0.0080, 0.014]
C) All	-0.0017	-0.0017
	[-0.025, 0.022]	[-0.025, 0.022]
Prior profits	X	X
E[wins]		X

Notes: the specification is the same as the IHS specifications from Years 0 to 8 in Table 4, except that we weight each observation by weights reflecting the relative proportion of firm or worker characteristics among applications on the last day relative to applications for capped H-1B visas on any day (including those on the last day and those before the last day). Specifically, we run a probit in which the dependent variable is a dummy for whether a firm applies on a day *other* than the last day, and the independent variables are all firm and worker characteristics shown in Table 2. We then calculate the fitted values $\hat{p}$. Finally, we weight each firm by $1/(1 - \hat{p})$. See other notes to Table 4.

Appendix Table 3. *Effect of Unexpected H-1B Lottery Wins on the Number of Patents*

A) ≤10 employees	0.0024	0.0027
	[-0.0089, 0.014]	[-0.0077, 0.013]
B) ≤30 employees	0.0085	0.0088
	[-0.015, 0.033]	[-0.015, 0.033]
C) All firm sizes	0.43	0.32
	[-1.21, 2.08]	[-1.23, 1.88]
Prior patents	X	X
E[wins]		X

Notes: See notes to Table 4. The table runs the same specification as Table 4, except that in Appendix Table 3 the dependent variable is the *number* of patents in each year from Year 0 to 8. When we consider the number of patents in the sample of firms of all sizes, the results are extremely imprecise, which is unsurprising since the standard deviation of patents in this sample is so large, and since an extra H-1B worker represents only a small fraction of mean employment in the full sample of firms. The positive point estimate in this context is very sensitive to outliers; for example, when we winsorize the number of patents at the 99th percentile in the sample of all firms, we obtain negative point estimates, but the estimates are imprecise and insignificant. In the smaller firm size categories, the effects are far more precise. *** refers to significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 4. *Effects of Unexpected H-1B Lottery Wins on Patenting in Years 4-8*

	IHS of number of patents		Patenting dummy	
A) ≤10 employees	0.000041	0.000022	-0.000096	-0.00013
	[-0.0022, 0.0023]	[-0.0021, 0.0021]	[-0.0021, 0.0019]	[-0.0019, 0.0017]
B) ≤30 employees	0.0043	0.0044	-0.00098	-0.00094
	[-0.0054, 0.014]	[-0.0053, 0.014]	[-0.0055, 0.0036]	[-0.0055, 0.0036]
C) All firm sizes	-0.00081	-0.0017	0.0063	0.0062
	[-0.033, 0.031]	[-0.033, 0.029]	[-0.0095, 0.022]	[-0.0091, 0.021]
Prior patents	X	X	X	X
E[wins]		X		X

Notes: The table shows the effect of an extra unexpected H-1B visa on patent outcomes over the indicated years. The table is identical to Table 4, except that the dependent variable is the IHS of patents in each year over Years 4 to 8. See Tables 1 and 4 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 5. *Effect of Unexpected H-1B Visas on Patent Citations*

	(1) Fewer controls	(2) More controls
(A) ≤10 employees	-0.0059	-0.0057
	[-0.023, 0.011]	[-0.022, 0.010]
B) ≤30 employees	-0.0053	-0.0049
	[-0.032, 0.022]	[-0.032, 0.022]
C) All	-0.022	-0.025
	[-0.071, 0.028]	[-0.074, 0.023]
Prior profits	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. The most recent firm-level patent citation data in the Patent Dataverse run through 2010, so the last year covered by these data is Year 4 (for the FY2006 lotteries) or Year 5 (for the FY2007 lotteries). Thus, the dependent variable in these regressions is the IHS of patents in each year from Year 0 to 2010. Otherwise, the specification is the same as in the Table 4 IHS specifications. The results here and in all other patenting appendix tables are very similar when the dependent variable is the patenting dummy instead. 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. See other notes to Table 4. *** refers to significance at the 1% level; ** at the 5% level, and * at the 10% level.

Appendix Table 6. *Effect of Unexpected H-1B Lottery Wins on Patenting, using Alternative Matching Procedure*

	IHS of number of patents		Patenting dummy	
A) ≤10 employees	-0.011	-0.011	-0.012	-0.012
	[-0.029, 0.0065]	[-0.029, 0.0068]	[-0.024, 0.0000]*	[-0.024, 0.00027]*
B) ≤30 employees	-0.0090	-0.0088	-0.012	-0.011
	[-0.026, 0.0084]	[-0.026, 0.0085]	[-0.024, 0.00076]*	[-0.024, 0.00079]*
C) All firm sizes	-0.028	-0.027	-0.011	-0.010
	[-0.067, 0.011]	[-0.066, 0.011]	[-0.028, 0.0067]	[-0.027, 0.0072]
Prior patents	X	X	X	X
E[wins]		X		X

Notes: See notes to Table 4. The table is similar to Table 4, except in defining the firms that match between the USCIS data and the Patent Dataverse, Appendix Table 6 includes those firms that are "possible" matches (whereas Table 4 excludes those firms). The table examines patenting in each year from Year 0 to Year 8, as in the baseline. *** refers to significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 7. *Effect of Unexpected Lottery Wins on ln(1+Patents)*

	(1) Fewer controls	(2) More controls
(A) ≤10 employees	0.00020	0.00023
	[-0.0036, 0.0040]	[-0.033, 0.0037]
B) ≤30 employees	0.0016	0.0017
	[-0.0068, 0.010]	[-0.0067, 0.010]
C) All firm sizes	-0.0068	-0.0069
	[-0.030, 0.016]	[-0.030, 0.016]
Prior patents	X	X
E[wins]		X

Note: The table shows OLS regressions of ln(1+ number of patents) on unexpected H-1B lottery wins, measuring this outcome in each year from Year 0 to Year 8 and pooling and stacking the years. This is an alternative way of addressing the skewness of the outcome distribution while recognizing that the number of patents is often zero, and without resorting to the less-known inverse hyperbolic sine transformation—but at the cost of adding an arbitrary constant (*i.e.* 1). The table shows coefficients on unexpected H-1B visas, with 95 percent confidence intervals in brackets. The table shows that the results are similar to those when the dependent variable is the IHS of patents over Years 0 to 8. See other notes to Table 4. *** refers to $p<0.01$; ** $p<0.05$; and * $p<0.10$.

Appendix Table 8. *Effect of Unexpected H-1B Lottery Wins on Patenting in Subgroups*

	(1) ≤10 employees	(2) ≤30 employees	(3) All firm sizes
A) Regular	0.0017	0.0045	0.0070
	[-0.0040, 0.0074]	[-0.011, 0.020]	[-0.011, 0.025]
	{654}	{1,062}	{2,327}
B) ADE	-0.0038	0.00076	-0.031
	[-0.012, 0.0038]	[-0.0087, 0.010]	[-0.11, 0.046]
	{67}	{137}	{494}
C) Professional, sci., and tech. services	-0.0010	0.0021	-0.010
	[-0.0046, 0.0026]	[-0.012, 0.017]	[-0.041, 0.021]
	{459}	{762}	{1,486}
D) Industries other than professional, sci., and tech. services	0.0011	0.0018	-0.0087
	[-0.0057, 0.0080]	[-0.0075, 0.011]	[-0.066, 0.049]
	{261}	{432}	{1,273}
E) "Temporary support services" industries	-0.0015	0.0048	-0.010
	[-0.0057, 0.0028]	[-0.012, 0.021]	[-0.044, 0.024]
	{388}	{632}	{1,191}
F) Non-"temporary support services" industries	0.0014	-0.0015	-0.0051
	[-0.0042, 0.0070]	[-0.0085, 0.055]	[-0.056, 0.046]
	{333}	{565}	{1,572}

Notes: The table shows OLS regressions of the IHS of patents in each year from Year 0 to Year 8 on unexpected 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 6-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 4 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, and Rows E+F 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 8 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 9. *Interactions of Unexpected Visa Lottery Wins with Covariates*

Outcome:	(1) IHS of patents, Years 0 to 8	(2) Employment in Q1 to Q4
A) Interaction of unexpected visas with days to reach cap		
	0.023	0.038
	[-0.029, 0.074]	[-0.030, 0.11]
B) Interaction of unexpected visas with IHS of patents in Year -1		
	-0.018	-5.94
	[-0.044, 0.0069]	[-31.48, 19.58]

Notes: The table indicates that there is no significant difference in the effects on patenting or employment by time taken to reach the visa cap or by amount of patenting in the pre-period. In Column 1, the dependent variable is the IHS of the number of patents in each year from Year 0 to Year 8, and the specification is an OLS regression. The coefficient reported is the coefficient on the interaction. In Column 2, the dependent variable is the number of employees in Q1 through Q4 (pooled and stacked, with each quarter as a separate observation), and the specification is a median regression (again as in the baseline). In Row A, the main independent variables are the number of unexpected H-1B visas; the number of days taken to reach the visa cap in the year and lottery in question; and the interaction of these two variables. In Row B, the main independent variables are the number of unexpected H-1B visas; the IHS of total patents in Year -1; and the interaction of these two variables. The table shows coefficients and 95% confidence intervals on the interactions. All specifications additionally control for expected lottery wins, as well as patents in the pre-period (in Column 1) or employment in the pre-period (in Column 2) as in the baseline specifications. The time taken to reach the visa cap was 291 days in FY2006 Regular lottery, 131 days in the FY2006 ADE lottery, 116 days in the FY2007 Regular lottery, and 55 days in the FY2007 ADE lottery. When we allow the time taken to reach the cap to have a different impact in the two ADE lotteries together and the two Regular lotteries together, we also find no significant interaction in each set of lotteries taken together. Standard errors are clustered by firm. See Tables 4 and 6 for sample sizes and additional notes. *** refers to significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 10. *Effect of Unexpected Lottery Wins on Patents per Employee*

	(1) Fewer controls	(2) More controls
(A) ≤10 employees	-0.00028	0.000021
	[-0.0021, 0.0016]	[-0.0017, 0.0017]
B) ≤30 employees	0.00070	0.00075
	[-0.0021, 0.0035]	[-0.0020, 0.0035]
C) All	-0.00015	-0.00020
	[-0.0013, 0.00098]	[-0.0014, 0.00097]
Prior patents/employee	X	X
E[wins]		X

Notes: The table shows the effect of an extra unexpected H-1B visa on the number of patents per employee over Years 0 to 8, pooling and stacking years. In calculating the mean number of employees in a given quarter, when the number of employees is missing in a given quarter it does not count in the average number of employees from Years 0 to 8. The mean of the dependent variable among all firms is 0.0056; the mean among firms with 30 or fewer employees is 0.0078; and the mean among firms with 10 or fewer employees is 0.0049. See Tables 1 and 4 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 11. *Effect of H-1B Lottery Wins on Research and Experimentation Credit in Years 4 to 8*

	Amount of Credit (IHS)		Claiming dummy	
A) ≤10 employees	-0.44	-0.41	-0.039	-0.036
(*n*=396)	[-1.02, 0.14]	[-0.97, 0.14]	[-0.089, 0.012]	[-.085, 0.012]
B) ≤30 employees	-0.36	-0.35	-0.031	-0.029
(*n*=353)	[-0.73, 0.0063]*	[-0.70, 0.0097]*	[-0.061, -0.00046]**	[-0.059, -0.000079]**
C) All firm sizes	-0.098	-0.10	-0.0088	-0.0089
(*n*=770)	[-0.85, 0.65]	[-0.85, 0.65]	[-0.056, 0.038]	[-0.056, 0.038]
Prior R&E	X	X	X	X
E[wins]		X		X

Notes: The table shows OLS regressions of the R&E credit after the duration of the H-1B visa (*i.e.* Years 4 to 8), on unexpected H-1B lottery wins. The table shows coefficients on unexpected H-1B visas, with 95 percent confidence intervals in brackets. In Columns 1 and 2, the dependent variable is the IHS of the total amount of the R&E credit claimed in each year over Years 4 to 8. *n*'s refer to the number of firms in the regressions. In Columns 3 and 4, the dependent variable is a dummy variable for whether the firm claimed any R&E credit in each of the years from Years 4 to 8, so that the coefficient reflects the effect on the fraction of years claiming the R&E credit. See other notes to Table 5. *** refers to significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 12. *Effect of Unexpected Lottery Wins on Employment, Sample Reweighted to Match Full Population of Firms*

	(1) Fewer controls	(2) More controls
(A) ≤10 employees	0.00	-0.29
	[-0.98, 0.98]**	[-0.95, 0.37]***
B) ≤30 employees	-0.038	-0.098
	[-1.20, 1.12]*	[-0.53, 0.33]***
C) All	-0.32	-1.22
	[-2.68, 2.03]	[-3.03, 0.58]**
Prior profits	X	X
E[wins]		X

Notes: the specification is the same as the IHS specifications in Table 6, except that we weight each observation by weights reflecting the relative proportion of firm or worker characteristics among applications on the last day relative to applications for capped H-1B visas on any day (including those on the last day and those before the last day). Specifically, we run a probit in which the dependent variable is a dummy for whether a firm applies on a day *other* than the last day, and the independent variables are all firm and worker characteristics shown in Table 2. We then calculate the fitted values $\hat{p}$. Finally, we weight each firm by $1/(1 - \hat{p})$. See other notes to Table 6.

Appendix Table 13. *Employment regressions by quarter in Q1 to Q4*

	Median Regressions		Two-stage least squares	
	(1)	(2)	(3)	(4)
Panel A: ≤10 employees				
A) Q1	-0.00	-0.031	0.072	-0.15
	[-1.28, 1.28]	[-1.64, 1.58]	[-1.24, 1.39]	[-2.15, 1.86]
B) Q2	-0.00	-0.41	-0.80	-1.46
	[-0.68, 0.68]***	[-1.17, 0.36]***	[-2.34, 0.75]**	[-3.29, 0.36]***
C) Q3	-0.78	-0.53	-0.66	-1.33
	[-1.78, 0.23]***	[-1.42, 0.36]***	[-2.40, 1.08]*	[-3.47, 0.80]**
D) Q4	-0.76	-0.61	-0.90	-1.72
	[-2.05, 0.51]***	[-1.79, 0.57]***	[-3.12, 1.31]*	[-4.52, 1.08]*
Panel B: ≤30 employees				
E) Q1	-0.35	-0.32	-1.05	-1.31
	[-1.41, 0.72]***	[-1.38, 0.73]**	[-3.17, 1.06]*	[-3.47, 0.85]**
F) Q2	-0.22	-0.17	-0.73	-0.95
	[-1.08, 0.65]***	[-1.11, 0.78]**	[-2.57, 1.10]*	[-2.90, 1.00]*
G) Q3	-0.95	-0.76	-1.00	-1.33
	[-2.17, 0.27]***	[-1.83, 0.31]***	[-3.23, 1.23]*	[-3.62, 0.96]**
H) Q4	-0.53	-0.53	-0.92	-1.25
	[-1.82, 0.76]***	[-1.85, 0.79]**	[-3.51, 1.67]	[-3.99, 1.49]
Panel C: All				
I) Q1	-1.41	-1.67	-62.10	-9.40
	[-3.40, 0.58]***	[-3.89, 0.54]**	[-768.40, 644.19]	[-22.73, 3.92]
J) Q2	-1.35	-1.00	-17.32	-2.75
	[-3.72, 1.02]*	[-3.11, 1.12]*	[-180.09, 145.44]	[-18.09, 12.58]
K) Q3	-0.055	0.25	4.76	4.43
	[-3.15, 3.03]	[-2.33, 2.83]	[-72.71, 82.24]	[-15.97, 24.83]
L) Q4	1.36	-0.31	-13.70	0.04
	[-4.80, 2.07]	[-3.64, 3.01]	[-191.01, 163.60]	[-21.57, 21.64]
Prior employment	X	X	X	X
E[wins]		X		X

Notes: None of the estimates is significantly different from 0 at any conventional significance level. See other notes to Table 6. See Table 1 for sample sizes. *** denotes estimates that are significantly different *from 1* at the 1% level; ** at the 5% level; * at the 10% level.

Appendix Table 14. *Additional Employment Specifications for Employment in the First Year*

	(1) Winsorize at 99%	(2) IHS	(3) IHS of difference, winsorized at 99%	(4) IHS of level, winsorized at 99%	(5) First difference of employment, no controls
A) ≤10 employees	-1.86	-0.18	-0.18	-0.18	-0.53
	[-4.34, 0.62]**	[-0.43, 0.066]**	[-0.43, 0.067]**	[-0.42, 0.068]**	[-1.37, 0.31]***
B) ≤30 employees	-1.69	-0.16	-0.15	-0.16	-0.69
	[-4.55, 1.17]*	[-0.35, 0.035]*	[-0.34, 0.034]**	[-0.35, 0.037]**	[-1.68, 0.31]***
C) All	1.06	0.034	0.045	0.032	-1.07
	[-73.91, 76.03]	[-0.15, 0.22]	[-0.14, 0.23]	[-0.14, 0.21]	[-3.05, 0.92]**

Notes: Columns 1-4 of the table show the baseline two-stage least squares regressions of employment outcomes on approved H-1B visas, where unexpected lottery wins are the instrument for approved H-1B visas. (The corresponding ITT regressions show very similar results.) In Column 1, the dependent variable is the difference of employment from the first quarter of Year -1 to Q1, Q2, Q3, or Q4 (pooled), and winsorized at the 1st and 99th percentiles. The 1st and 99th percentiles of the first difference in employment are -5,559 and 2,430, respectively, in the full sample; are -20 and 62, respectively, among those with 30 or fewer employees; and are -10 and 53, respectively, among those with 10 or fewer employees. In Column 2, the dependent variable is the IHS of the difference in employment over the same periods. In Column 3, the dependent variable is the IHS of the difference in employment over the same periods, winsorized at the 99th percentile. In Column 4, the dependent variable is the IHS of the level of employment in Q1 through Q4 (pooled), winsorized at the 99th percentile, and the results are nearly identical to those in Column 3. All specifications in Columns 1, 2, 3, and 4 control for prior employment and the number of expected lottery wins, as in the baseline; the results are similar with other controls. In Column 5, we run median regressions (as in Table 6) and the dependent variable is the first difference of employment (from the first quarter of calendar Year -1 to a given quarter of Year 0, and pooling this measure from Q1 to Q4), but we do not include any controls. In all columns, we pool across Q1 to Q4, as in the baseline. None of the estimates is significantly different from 0 at any conventional significance level. In the case of these IHS specifications, before testing whether a coefficient is equal to 1, we transform the coefficient from the regression (which reflects the percentage increase in employment, rather than the increase in the absolute level of employment) by multiplying it by the mean level of employment. We then test whether this transformed coefficient is equal to 1. The test results reported above refer to this test. *** denotes estimates that are significantly different *from 1* at the 1% level; ** at the 5% level; * at the 10% level. See Tables 1 and 6 for other notes and sample sizes.

Appendix Table 15. *Effect of H-1B Visa on Being out of Business*

Panel A: ≤10 employees (*n*=719)		
A) Q1 to Q4	0.024	0.033
	[-0.016, 0.063]	[-0.022, 0.088]
B) Q1	0.016	0.023
	[-0.020, 0.052]	[-0.030, 0.077]
C) Q2	0.017	0.022
	[-0.033, 0.066]	[-0.051, 0.095]
D) Q3	0.032	0.046
	[-0.014, 0.079]	[-0.015, 0.11]
E) Q4	0.029	0.041
	[-0.017, 0.076]	[-0.022, 0.10]
Panel B: ≤30 employees (*n*=1,134)		
F) Q1 to Q4	0.010	0.012
	[-0.019, 0.040]	[-0.024, 0.047]
G) Q1	0.0033	0.0033
	[-0.028, 0.034]	[-0.034, 0.040]
H) Q2	0.0030	0.0029
	[-0.035, 0.041]	[-0.043, 0.049]
I) Q3	0.015	0.017
	[-0.020, 0.050]	[-0.023, 0.058]
J) Q4	0.020	0.023
	[-0.013, 0.052]	[-0.014, 0.060]
Panel C: All (*n*=2,292)		
L) Q1 to Q4	0.0050	0.0024
	[-0.068, 0.078]	[-0.014, 0.019]
M) Q1	-0.032	-0.0053
	[-0.39, 0.32]	[-0.022, 0.011]
O) Q2	-0.013	-0.0024
	[-0.13, 0.11]	[-0.024, 0.019]
P) Q3	-0.015	0.0054
	[-0.10, 0.13]	[-0.014, 0.025]
Q) Q4	0.037	-0.011
	[-0.21, 0.28]	[-0.0084, 0.031]
Prior employment	X	X
E[wins]		X

Notes: The table shows point estimates and 95% confidence intervals on unexpected lottery wins, from OLS (linear probability) regressions a dummy for whether the firm is "out of business" is regressed on unexpected lottery wins and controls. We define a firm as being "out of business" if it has either zero employees or is missing the number of employees. The results are similar with other definitions of being out of business. The "prior employment" specifications control for employment from the first quarter of Year -1, and the "prior employment, E[wins]" specifications additionally control for the number of expected lottery wins. None of the estimates is significantly different from 0 at any conventional significance level. Since the table measures the effect on whether the firm has employment in the U.S., these results also encompass effects on whether a firm chooses to locate in the U.S. "*n*" refers to the total number of firms in the regressions. See Tables 1 and 6 for other notes. *** denotes estimates that are significant at the 1% level; ** at the 5% level; * at the 10% level.

Appendix Table 16. *Effect of H-1B Visa on Being out of Business*

Panel A: ≤10 employees (*n*=719)		
A) Q5 to Q8	0.020	0.018
	[-0.088, 0.13]	[-0.090, 0.13]
B) Q8 to Q12	0.016	0.023
	[-0.020, 0.052]	[-0.030, 0.077]
C) Q13 to Q32	0.065	0.068
	[-0.041, 0.17]	[-0.039, 0.17]
Panel B: ≤30 employees (*n*=1,191)		
D) Q5 to Q8	-0.014	-0.013
	[-0.081, 0.054]	[-0.081, 0.054]
E) Q8 to Q12	-0.022	-0.022
	[-0.092, 0.048]	[-0.092, 0.047]
F) Q13 to Q32	0.023	0.024
	[-0.053, 0.099]	[-0.052, 0.10]
Panel C: All (*n*=2,289)		
G) Q5 to Q8	-0.00025	0.00092
	[-0.033, 0.033]	[-0.032, 0.034]
H) Q8 to Q12	-0.015	-0.012
	[-0.053, 0.024]	[-0.050, 0.026]
I) Q13 to Q32	-0.0097	-0.0079
	[-0.052, 0.033]	[-0.050, 0.034]
Prior employment	X	X
E[wins]		X

Notes: The table shows point estimates and 95% confidence intervals on unexpected lottery wins, from OLS (linear probability) regressions a dummy for whether the firm is "out of business" is regressed on unexpected lottery wins and controls. We define a firm as being "out of business" if it has zero employees or is missing number of employees. The results are similar with other definitions of being out of business. The "prior employment" specifications control for employment from the first quarter of Year -1, and the "prior employment, E[wins]" specifications additionally control for the number of expected lottery wins. None of the estimates is significantly different from 0 at any conventional significance level. Since the table measures the effect on whether the firm has employment in the U.S., these results also encompass effects on whether a firm chooses to locate in the U.S. "*n*" refers to the total number of firms in the regressions. See Tables 1 and 6 for other notes. *** denotes estimates that are significant at the 1% level; ** at the 5% level; * at the 10% level.

Appendix Table 17. *Effect of Unexpected Lottery Wins on Employment in Subgroups*

	(1) ≤10 employees	(2) ≤30 employees	(3) All firm sizes
A) Regular	-0.41	-0.59	-1.26
	[-1.10, 0.27]***	[-1.46, 0.28]***	[-3.33, 0.81]**
	{651}	{1,069}	{1,969}
B) ADE	-0.0000002	0.52	1.38
	[-1.36, 1.36]	[-1.51, 2.55]	[-5.63, 8.39]
	{67}	{134}	{400}
C) Professional, sci.,	-0.58	-0.72	-1.46
and tech. services	[-1.54, 0.39]***	[-1.92, 0.48]***	[-3.60, 0.67]**
	{456}	{759}	{1,275}
D) Industries other	0.36	0.65	1.16
than professional, sci.,	[-0.50, 1.22]	[-0.36, 1.65]	[-2.74, 5.05]
and tech. services	{257}	{426}	{1,015}
E) "Temporary	-1.56	-0.68	-1.54
support services"	[-5.70, 2.57]	[-2.09, 0.73]**	[-4.03, 0.95]**
industries	{384}	{628}	{4,738}
F) Non-"temporary	0.65	0.00	0.14
support services"	[-0.42, 1.72]	[-0.95, 0.95]**	[-2.46, 2.74]
industries	{330}	{560}	{1,265}

Notes: The table shows the effect of unexpected lottery wins on employment, displaying point estimates and 95% confidence intervals in [square brackets] for median regressions of employment in Q1-Q4 on unexpected 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. See other notes to Appendix Table 8. *** shows $p<0.01$ for the test of difference *from 1*; ** $p<0.05$; * $p<0.01$. None of the estimates is significantly different from zero.

Appendix Table 18. *Effect of Unexpected Lottery Wins on Employment of Foreigners and non-Foreigners*

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

Notes: The table shows the effect of unexpected lottery wins on employment of foreigners or non-foreigners, displaying point estimates of the coefficient on unexpected 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 contain information on 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 6 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 19. *Effect of Unexpected Lottery Wins on Payroll per Employee in Years 4-8*

	(1) Fewer controls	(2) More controls
(A) ≤10 employees	374.49	428.54
	[-1,214.22, 1,963.20]	[-1,263.43, 2,120.50]
B) ≤30 employees	-258.04	-1,325.10
	[-5,625.2, 5,109.12]	[-6,443.69, 3,793.48]
C) All	2,645.54	1,123.09
	[-658.12, 5,949.20]	[-7,018.17, 9,264.35]
Prior patents/employee	X	X
E[wins]		X

Notes: The table shows the effect of an extra unexpected H-1B visa on firms' payroll costs per employee over Years 4 to 8. The median of the dependent variable among all firms is 54,761.65; the median among firms with 30 or fewer employees is 49584.98; and the median among firms with 10 or fewer employees is 48,551.45. See Tables 1 and 8 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 20. *Effect of Unexpected Lottery Wins on Revenue or Total Income per Employee*

	(1) Revenue per Employee	(2) Total Income per Employee
(A) ≤10 employees	8,376.40	6,191.85
(*n*=615)	[-6,483.59, 23,236.40]	[-9,414.77, 21,798.48]
B) ≤30 employees	8,326.45	5,220.35
(*n*=1,033)	[-2,194.90, 18,845.80]	[-2,660.61, 13,101.31]
C) ≤200 employees	2,600.74	2,730.51
(*n*=1,520)	[-1,985.04, 7,186.51]	[-1,426.81, 6,887.82]

Notes: The table shows median regressions of revenue per employee (Column 1) or total income per employee (Column 2) in Years 0 to 3 on unexpected H-1B visas and controls, pooling and stacking the years. The table shows coefficients and 95% confidence intervals on unexpected H-1B visas. 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. Estimated effects in the shorter or longer term are comparable. Standard errors are clustered by firm. *n*'s refer to the number of firm-lottery years. *** refers to significance at the 1% level; ** at the 5% level; and * at the 10% level.