

The Impact of Immigration on Firms and Workers: Insights from the H-1B lottery and US Employer-Employee Data^{*}

Agostina Brinatti
University of Michigan

Mingyu Chen
IZA

Parag Mahajan
University of Delaware

Nicolas Morales
FRB Richmond

Kevin Shih
Queens College
(CUNY), IZA

March 10, 2023

^{*} Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2104. (CBDRB-FY22-P2104-R9930).

Acknowledgement & Disclaimer

This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation.

Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1582 and FSRDC Project Number 2104.

The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Board of Governors.

What are the labor market impacts of skilled immigrants?

- **Extensive literature on the US:** *Stephan & Levin (1999, 2001), Hunt and Gauthier-Loiselle (2010), Kerr and Lincoln (2010); Borjas & Doran (2012, 2015a, 2015b); Kerr et. al (2015a, 2015b), Peri, Shih & Sparber (2015a), Bound et al. (2015), Bound, Khanna & Morales (2018), Mayda et al. (2018, 2020), Morales (2022), Bernstein et al. (2022)*
 - Concentrated in STEM (Science, Technology, Engineering, Mathematics) occupations, responsible for large growth in STEM labor force
 - Contributors to innovation, which in turn fuels productivity growth
 - Since 1990, regulated by the H-1B Visa program, which allows for large role of employers in selection
 - Role of firms is crucial to understanding labor market effects
- However, major barriers to studying immigration at the firm-level:
 1. Endogenous selection of immigrants and firms
 2. Granular worker-firm level data by nativity is scarce, particularly in U.S.

This Paper

- **Research question:** how does lottery-induced variation in skilled immigration affect firms and workers?
- **Data**
 - USCIS/DOL data on H-1B visa applications in the FY 2008 and 2009 lotteries
 - Linked employer-employee data from the US Census Bureau
- **Research Design**
 - Construct a measure of firm success in the lotteries
 - Set up an DiD/event study approach: compare "lucky" and "unlucky" firms over time
- **Key, preliminary findings:** lottery-induced increases in foreign-born, college-educated workforce, no crowd-out of natives, increases in revenues and productivity

Existing Studies using H-1B Lotteries

- Clemens (2013): FYs 08-09 to identify wage gap bet. winners and losers $\approx 6\times$
- Peri, Shih & Sparber (2015b) FYs 08-09 lotteries and city-level data
- Doran, Gelber & Isen (2022): FYs 05-06 lotteries ($\sim 3\text{K}$ companies), 1 H-1B leads to 1.5 fewer natives, little evidence of positive effects on innovation; FY 08 robustness check (0.78 fewer natives)
- Dimmock et al. (2022) FYs 08, 09, 14, 15 to assess startups ($\sim 2.5\text{K}$ companies)

Existing Studies using H-1B Lotteries

- Clemens (2013): FYs 08-09 to identify wage gap bet. winners and losers $\approx 6\times$
- Peri, Shih & Sparber (2015b) FYs 08-09 lotteries and city-level data
- Doran, Gelber & Isen (2022): FYs 05-06 lotteries ($\sim 3\text{K}$ companies), 1 H-1B leads to 1.5 fewer natives, little evidence of positive effects on innovation; FY 08 robustness check (0.78 fewer natives)
- Dimmock et al. (2022) FYs 08, 09, 14, 15 to assess startups ($\sim 2.5\text{K}$ companies)
- We contribute by:
 - Matching first “full” H-1B lotteries FY 08/09 to US Census Worker-Firm data: $> 4\times$ as many firms as in 05-06 smaller lotteries, representative of user base and H-1B allocation since 2013
 - Careful attention to identification in lotteries where applications are not observed
 - Track *firms* (and their outcomes) over time and sub-groups, including hiring of H-1B (first-stage), and more/less “substitutable” natives
 - Follow individual workers over time and examine rich individual-level outcomes/transitions

Outline of Talk

1. The H-1B Visa Lotteries of FY 2008 & 2009
2. Measurement
3. Research Design/Identification
4. Preliminary Results
5. Ongoing Work

The H-1B Lotteries of FY 2008 and 2009

The H-1B Program

- Aimed at skilled workers in specialty occupations
- 3-year duration with possibility of renewal for +3 years
- New workers at for-profit employers are subject to a yearly cap
 - Regular cap: 65,000 per year
 - +20,000 Advanced Degree Exemption (ADE) – masters+ from a US institution
- Cap-exempt: Non-profits, renewals, change of status, Chile/Singapore

FYs 2008 and 2009: First Large-Scale Lotteries

	FY 2008	FY 2009
Regular Lottery		
Start of Filing	4/2/07	4/1/08
Final Receipt Date	4/3/07	4/7/08
Total Applications	123,480	143,000
Total Cap	65,000	65,000
<i>Fraction Winning</i>	0.53	0.45
ADE Lottery		
Start of Filing	4/2/07	4/1/08
Final Receipt Dates	N/A	4/7/08
Total Applications	N/A	31,200
Total Cap	20,000	20,000
<i>Fraction Winning</i>	N/A	0.64

- First-come first-served unless cap exceeded within 1st days of filing period (begins in April)
- Applications of lottery losers *were returned without processing.*
- In FY 08, no ADE lottery.
- In FY 09, ADE first participated in regular lottery, with losers then participating in the ADE lottery.

Measurement

Measuring Firm-Level Lottery Success

- Ideal measure of firm j 's success in a lottery (*I-129 is the official application form for H-1B visas*):

$$wr_{jt} \equiv \frac{\text{I-129 Lottery Wins}_{jt}}{\text{I-129 Lottery Applications}_{jt}}$$

- **Problem:** USCIS did not process applications of lottery losers, so I-129 Lottery Applications $_{jt}$ not observed
- **Solution:** Approximate measure of firm j 's success in a lottery:

$$\widehat{wr}_{jt} \equiv \frac{\text{I-129 Lottery Wins}_{jt}}{\text{LCA Lottery Applications}_{jt}}$$

H-1B Data

- **I-129 Lottery Wins** come from individual I-129 records obtained by FOIA from USCIS
 - *Old data*: Fuzzy education and company identifiers (firm name, address)
 - *New data*: Exact education level, company identifiers, and more...
- **LCA Lottery Applications** come from Labor Conditions Applications data available from Dept. of Labor
 - Job postings by employer/filing date, no fee to file, no identification of applications subject to lottery
 - Take filings from February-April, due to predating (Peri, Shih & Sparber 2015b)
- Construct win-rates
 - Fuzzy match I-129s/LCAs using name and address
 - Aggregate I-129 and LCAs according to lottery windows
 - Combine FY08 & 09 to create overall win-rate for each firm

Lottery summary statistics

Real vs predicted lotteries	CY 2007		CY 2008	
	Real	Predicted	Real	Predicted
Granted, lottery-subject, I-129 petitions	65,000	79,231	85,000	78,177
Lottery subject, I-129 applications	123,400	185,225	163,000	191,948
Aggregate win rate	0.527	0.428	0.521	0.407

- We underestimate the share of winning applications

Firm-level statistics	CY 2007	CY 2008	Combined 07/08
Mean firm-level predicted win rate	0.41	0.40	0.41
Share of firms that apply for one LCA	0.62	0.62	0.60

- 60% of firms only apply for one LCA

Research Design/Identification

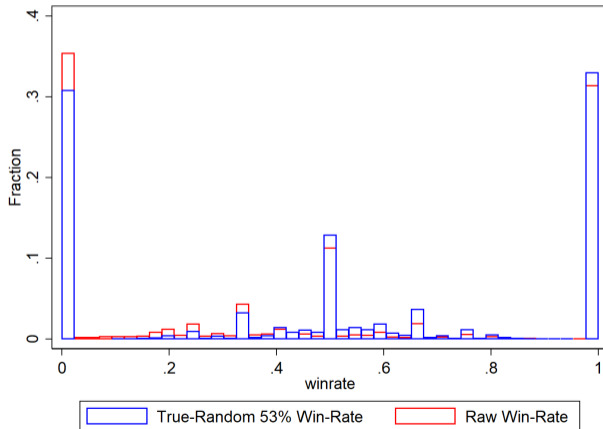
Adapting to Mismeasured Win Rates

- **Measurement Error** in win rates: $LCA_s \geq Applications$. Occurs under (at-least) 2 scenarios:
 - Firms always file some proportion of excess LCAs, a time-invariant characteristic perhaps correlated with other firm characteristics (e.g. size)
 - Shock leads firms to submit different proportion of LCA applications usual (e.g. Q1-Q2)
- **Difference-in-differences** nets out this time-invariant measurement error out between more and less successful firms
- **Threat to DD**: Time-varying shocks prior to lottery affect win-rates and potentially outcomes
- **Event study** assesses if differences between more and less successful firms are apparent in the periods leading up to the lotteries

Win-rate Diagnostics: Comparing Distributions

- Compare our “raw” winrate to a “truly random” winrate (53%)

Random vs. Raw



Research Design

$$\log(y_{jt} + 1) = \sum_{\tau \neq 2007} \beta_{\tau} [\underbrace{\widehat{wr}_j}_{\text{Win Rate}} \times \mathbb{1}(\tau = t)] + \underbrace{\Gamma X_{jt}}_{\text{Controls}} + \underbrace{\alpha_j}_{\text{Firm FE}} + \underbrace{\alpha_{k(j),t}}_{\text{Industry-time FE}} + \varepsilon_{jt}$$

- **ID assumption:** parallel trends in absence of different win rates
- α_j takes care of issues relating to imperfect lottery measure
- $X_{jt} = \log(\text{March 12 Employment}_{2007}) \times \mathbb{1}(\tau = t)$
- Compare similar firms with different lottery success rates
 1. Assess research design ($\beta_{\tau}, \tau < 2007$)
 2. Trace out effect of lottery success over time ($\beta_{\tau}, \tau > 2007$)

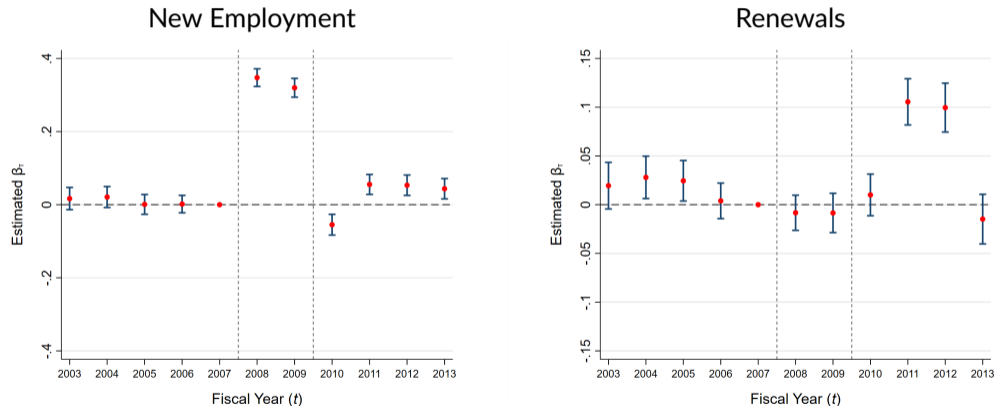
Firm-Level US Census Bureau Data

- Longitudinal Business Dynamics (LBD)
 - Near-universe of US private sector establishments
 - Key variables for today: revenues, employment
- The Longitudinal Employer-Household Dynamics (LEHD)
 - Employer-employee matched data for the universe of firms and workers
 - 25 states
 - Key variables for today: age, country of birth, education of each worker
- Link to H-1B data via fuzzy string match on name/address

Preliminary Results

Testing the “First Stage” in the H-1B Data

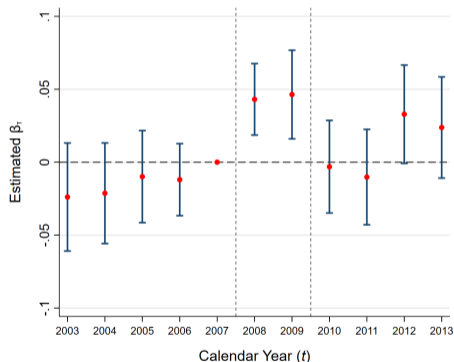
Figure: Log Successful I-129 Applications (USCIS Data)



Notes: Number of observations is 116,000. Number of firms is 14,500. Standard errors clustered at the firm level.

Testing the “First Stage” in the Census Data

Figure: Log Employment of “Likely H-1B” Immigrants

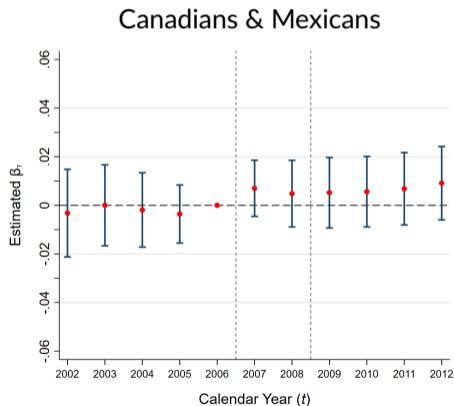
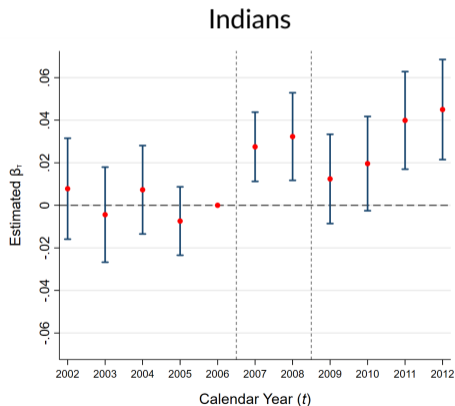


- Using LEHD demographic variables
- “Likely H-1B” immigrant: 25-40 year-old, foreign-born, college graduate with less than 3 years of tenure at the firm.

Notes: Number of observations is 116,000. Number of firms is 14,500. Standard errors clustered at the firm level.

Changes in “H-1B-Like” Employment by Origin

- Indians account for 50-70% of H-1B employment
- Canadian/Mexicans (4-6% of H-1B) have TN visa as an alternative
- Winning firms have a persistent increase in Indian, “H-1B-like” workers (networks?)



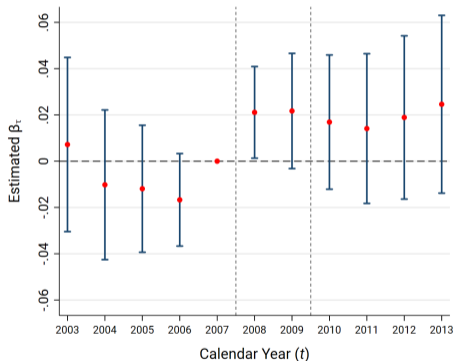
Notes: Number of observations is 116,000. Number of firms is 14,500. Standard errors clustered at the firm level. Outcomes measured using the LEHD.

Firm Performance and Composition

- Impacts on firm performance
 - Answer using outcomes from LEHD/LBD
 - Total employment
 - Revenues
 - Revenues per worker
- Impacts on employment opportunities for native workers
 - Answer by measuring firm employee composition in LEHD
 - Closest substitutes: *young* (25-40 y.o.), *low-tenure* (<3 years at firm), college graduates
 - All other native college graduates

Total Employment

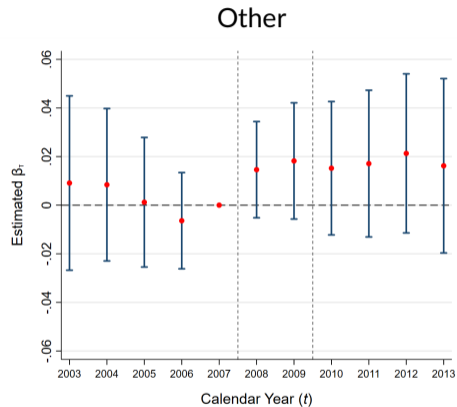
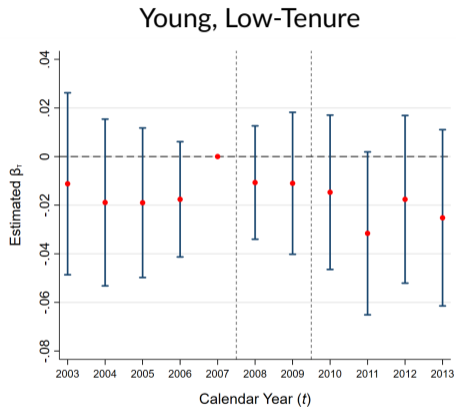
Figure: Log Total Employment



- Employment increases in years after lottery
- Precision is an issue
- Each successful lottery I-129 applications increases firm employment by 1.23 workers by 2009

Notes: Number of observations is 116,000. Number of firms is 14,500. Standard errors clustered at the firm level.

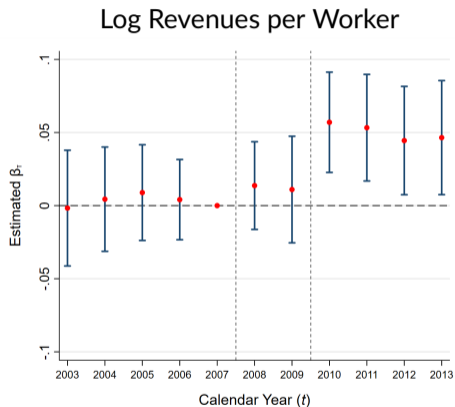
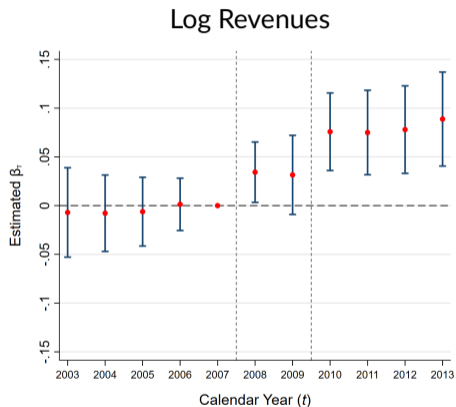
Employment of Native College Graduates



Notes: Number of observations is 116,000. Number of firms is 14,500. Standard errors clustered at the firm level. Outcomes measured using the LEHD.

Non-College Workers

Persistent Increases in Revenues, Labor Productivity



Notes: Number of observations is 116,000. Number of firms is 14,500. Standard errors clustered at the firm level. Outcomes measured using the LBD.

Ongoing Work

Ongoing Work

1. Identification

- Balancing tests
- Randomization inference
- Check Q1-Q2 pretrend
- Win-rate corrections

2. Firm-level Analysis

- Firm Entry/Exit
- Employment, Wages, Revenues, Productivity, Profit
- Separating lotteries, ADE applications
- Heterogeneity across firms (e.g. large, small, by industry etc.)

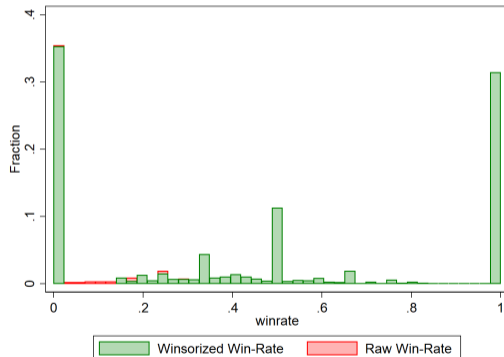
3. Individual-level Analysis

- Individual level regressions following workers over time
- Examine displacement, job switching, earnings
- Heterogeneity across workers (immigrants, natives, young, old, high/low earners, etc)

Win-rate Diagnostics: Bounding Extreme Cases

- Extra mass in left tail of raw win-rate distribution
- LCAs are an over-count of true applications
- We can bound extreme cases:
 - Given measured applications and known win-rate (53%)
 - Binomial distribution tells us if observed win-rate falls beyond 99th percentile
 - Bound cases above the 99th percentile to the 99th percentile (Winsorize applications)

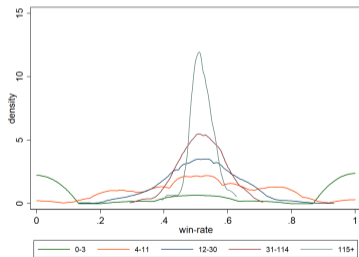
Raw vs. Winsorized



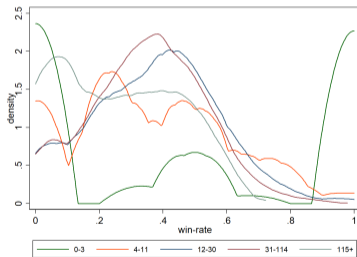
Win-rate Diagnostics: Consistent with randomness?

- Simple diagnostic: variation in win-rates should decline with applications, remain centered on mean

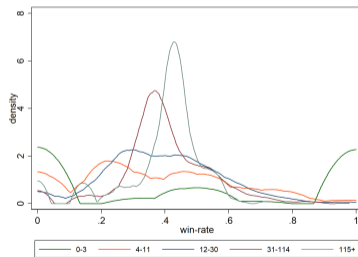
“Random”



Raw



Winsorized



Conclusion

- We examine the effect of skilled immigration on workers and firms using the H-1B lotteries of FYs 2008 and 2009
- Link H-1B lotteries to U.S. Census Bureau worker-firm matched data
- Preliminary results indicate employment increases, primarily through H-1B hiring, little crowd-out of other workers, improvements in revenue
- Lots of additional work underway
 - New, improved H-1B data
 - Improvements to our win rate measure
 - Heterogeneity across firms
 - Individual-level regressions
 - Firm-level patent data