CESifo WORKING PAPERS

9874
2022

July 2022

Immigration and Business Dynamics: Evidence from U.S. Firms

Parag Mahajan

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9874*

Immigration and Business Dynamics: Evidence from U.S. Firms

Abstract

Prior literature on the economic impact of immigration has largely ignored changes to the composition of labor demand. In contrast, this paper uses a comprehensive collection of survey and administrative data to show that heterogeneous establishment entry and exit drive immigrant-induced job creation and a rightward shift of the productivity distribution in U.S. local industries. High-productivity establishments are more likely to enter and less likely to exit in high immigration environments, whereas low-productivity establishments are more likely to exit. These dynamics result in productivity growth. A general equilibrium model proposes a mechanism that ties immigrant workers to high-productivity firms and shows how accounting for changes to the employer distribution can yield substantially larger estimates of immigrant-generated economic surplus than canonical models of labor demand.

JEL-Codes: J230, J610, L110, F220.

Keywords: immigration, business dynamics, productivity, firm heterogeneity.

Parag Mahajan
Department of Economics
University of Delaware / Newark / DE / USA
paragma@udel.edu

July 29, 2022
I would like to thank John Bound, Dean Yang, Jeffrey Smith, Hoyt Bleakley, and Jagadeesh Sivadasan for invaluable advice and support. I would also like to thank Emek Basker, Ariel Binder, Kirill Borusyak, Laurien Gilbert, Steven Haider, Peter Hull, Brian Kovak, Andrei Levchenko, Joan Llull, Ben Lipsius, Nicolas Morales, Gianluca Orefice, Pia Orrenius, Sebastian Ottinger, Dhiren Patki, Max Risch, and seminar participants at the University of Michigan, U.S. Census Bureau CES, University of Wisconsin, Colby College, University of Delaware, RAND Corporation, Federal Reserve Board, University of Virginia, FDIC, Federal Reserve Bank of Dallas, University of Colorado—Denver, Wesleyan University, U.S. Treasury OTA, Michigan State University, Baruch College, MEA Annual Conference, SOLE Annual Meetings, APPAM Fall Research Conference, SEA Annual Meetings, the NBER Immigrants and the U.S. Economy conference, and the CESifo Venice Summer Institute for helpful comments and suggestions. Finally, I would like to thank Joseph Ballegeer and J. Clint Carter for help and support in accessing data and disclosing results and Jarcy Zee for data visualization assistance. This is a greatly revised version of Mahajan (2021). 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 2105. (CBDRB-FY22-P2105-R9761).

1 Introduction

By 2030, immigration will overtake natural increase as the primary driver of population growth in the United States.[1] This far-reaching demographic change will translate into a workforce that increasingly relies on the foreign-born, magnifying the need for a complete understanding of how they are absorbed into labor markets and ultimately shape economies. Recent advances to data and theory have dramatically expanded our insight into the role of the firm in these processes, with a particular focus on the form and choice of production technique.[2] Nonetheless, most of this literature has either implicitly or explicitly restricted its attention to representative firm models of production that do not feature differences across firms in input use or total factor productivity. Moreover, empirical work has largely focused on non-U.S. settings and has not emphasized business entry, business exit, or subsequent changes to the employer distribution as important mediators of the economic impact of immigration.

In contrast, broader study of the U.S. economy finds that business entry and exit are crucial determinants of job creation and productivity growth, particularly when entry is accompanied by the exit of less productive businesses (e.g., Bartelsman and Doms, 2000; Foster et al., 2008; Haltiwanger et al., 2013).[3] Further, in the U.S., immigrant entrepreneurs and employees are uniquely important to business entry (Kerr and Kerr, 2016), and enforcement policies that expel immigrant populations generate business exit (Ayromloo et al., 2020). Motivated by these facts, I recast the economic impact of immigration in the world's largest immigrant destination through the lens of heterogeneous business dynamics. In doing so, I introduce a new channel through which immigrant workers impact economic efficiency: by changing the distribution of operating firms.

This paper ultimately finds that heterogeneous business dynamics drive the U.S. immigrant absorption process and its economic impact. I integrate several confidential sources of demographic and business data from the U.S. Census Bureau to draw this conclusion. I study immigrant worker inflows into a local industry—defined as a pairing between one of 722 commuting zones and 41 industry groups—over the time period spanning 2000 to 2015, measured using demographic data that includes all survey responses to the 2000 Decennial Census Long Form and 2005–2017 American Community Surveys.[4] I test how these immigrant inflows affect establishment entry and exit dynamics, which are primarily measured using the Longitudinal Business Database, an administrative data set with near-complete coverage of the U.S. private sector. To resolve endogeneity concerns endemic to the study of immigration on economic outcomes, I develop a new shift-share instrument that uses international migration data from thousands of origin-destination pairs to isolate

[1] Natural increase is defined as births minus deaths of natives. See Vespa et al. (2020).

[2] See, e.g., Lewis (2005, 2012); Clemens et al. (2018); Peri (2012); Lewis (2011); Peri and Sparber (2009); Mitaritonna et al. (2017); Arellano-Bover and San (2020); Brinatti and Morales (2021); Orefice and Peri (2020); Beerli et al. (2021)

[3] A vast majority of firms in the U.S. are single-unit. However, because several important employers in the U.S. are multi-unit, it is important to distinguish between a firm and an establishment. I use the term "business" to encompass firms and establishments more generally.

[4] This paper will use the term workers to encompass both the self-employed and employees.

exogenous migration pushes to the U.S.

Using this setup, I present three core empirical results. I first characterize the relationship between immigrant worker inflows, establishment dynamics, and immigrant absorption. Immigrant inflows increase establishment counts within local industries, and this effect is roughly equally driven by increased establishment entry and reduced establishment exit. Furthermore, these extensive margin responses account for nearly all of immigrant-induced job creation, leaving a minimal role for growth from incumbent, continuing establishments. Thus, the first core result of this paper is that the extensive margin of labor demand drives immigrant absorption in U.S. labor markets.

Reductions in establishment exit prompt concerns that immigrant inflows may stunt creative destruction. Shifting our attention to the productivity consequences of immigration, I therefore study the exit decisions of over 4.7 million individual establishments from 2000 to 2015. Contrary to these concerns, I find that immigrant worker inflows *increase* the likelihood of exit by establishments whose parent firms are in the lowest quintile of the productivity distribution. Meanwhile, immigrant worker inflows substantially reduce the probability that establishments from high productivity firms exit the market. A complementary analysis of all 9.4 million entering establishments over the study period finds that high-immigration environments spawn entry by establishments that are more likely to be long-lived and high-grossing. Put together, the second—and most important—core empirical result of the paper is that immigrant worker inflows benefit high-productivity firms at the expense of low-productivity firms on the extensive margin, consistent with creative destruction.

The third and final set of analyses corroborate that these dynamics shift the employer productivity distribution to the right and ultimately increase productivity. I show that immigrant-induced increases in establishment counts are heavily concentrated at the top of the productivity distribution, with the top quintile alone accounting for a near-majority. I then show that immigrant worker inflows lead to increases in three proxies for productivity at the local industry level: average earnings, revenues, and revenues per worker. Thus, the third core empirical result of the paper is that immigrant worker inflows lead to higher aggregate productivity by altering the employer distribution.

Such changes to the employer distribution are not built into canonical models of immigration and the labor market. In light of my empirical results, I reevaluate the aggregate economic impacts of immigration in general equilibrium within a novel modeling framework that ties immigrant workers to high-productivity firms. The key mechanism is that firms must pay fixed recruiting costs to access immigrant workers. Firms that pay these fixed costs lower their variable costs by assigning immigrant and native workers to tasks for which they have a comparative advantage—a notion embodied by a finite elasticity of substitution across workers of different nativity. Only larger, more productive firms find it profitable to invest in hiring immigrants because they spread the fixed recruiting cost over many units of output. With increased immigration, these higher-productivity, immigrant-hiring firms see larger reductions in labor costs than their lower-productivity counterparts, and resulting price competition drives the lowest-productivity firms out of the market. The effect of immigration on native incomes—the "immigration surplus"—can be written entirely as a function of changes to the employer distribution. Unlike canonical, representative firm models of labor demand, immigrant-induced changes to the composition of firms generate first-order productivity gains in this model.

1.1 Contribution to Literature

This paper is unique in presenting a comprehensive analysis of how increases in immigrant labor supply impact establishment dynamics in the U.S. and elucidating the consequences of this relationship for immigrant absorption and productivity. To my knowledge, it is the first to harness both confidential demographic and economic data from the Census Bureau to study the impact of immigration on the U.S. economy. These data allow for breadth and detail that help this paper make several contributions to the empirical literature.

While previous and concurrent research has studied the impact of immigrant inflows on establishment counts in the U.S. (Olney, 2013; Orrenius et al., 2020), this paper directly links this relationship to job creation and productivity. Like Olney (2013), I find that immigrant inflows increase establishment counts, and like Orrenius et al. (2020), I find that this effect is driven partially by reduced establishment exit. Relative to these works, I introduce a new identification strategy and both a finer-grained level of study and more comprehensive coverage of the U.S. economy—including establishment-level analyses. More importantly, by studying the heterogeneity in employer responses to immigration, I am able to show that increased business entry and reduced business exit are not just an additional consequence of immigrant inflows, but a singular mechanism through which the impacts of immigration manifest.

This paper is motivated by and contributes to literature from advanced economies other than the U.S., which has studied establishment-level responses to immigration but has not focused on the interaction between extensive margin establishment dynamics, on one hand, and productivity and job creation, on the other. Mitaritonna et al. (2017), for example, show that establishment exit declines in response to immigration in France, but find that this decline is not stratified by initial productivity. While the results are different, the analysis in Section 3.2.1 of this paper is motivated by theirs. Beerli et al. (2021) find that establishment entry increases in response to increased availability of foreign workers in Switzerland, but do not find an effect on reduced establishment exit. The differences between these results and those found in this paper broach the possibility that the extent to which heterogeneous entry and exit drive the impact of immigration is unique to the U.S.

In that vein, this paper builds on previous work linking immigration to technological change and productivity increases throughout U.S. history (e.g., Peri, 2012; Lewis, 2012; Clemens et al., 2018; Khanna and Lee, 2018; Sequeira et al., 2019). Of particular note are Burchardi et al. (2020), who find that immigrant inflows increase county-level measures of patenting and earnings in recent decades, and Peri (2012), who analyzes the effect of immigration on total factor productivity (TFP) in U.S. states from 1960–2006 and finds that a positive effect is mediated by task specialization across immigrants and natives. This paper's unique data allow me to present a novel, complementary channel—changes to the employer distribution—through which immigrant inflows increase economic efficiency. Indeed, both increased patenting and task specialization may interact with changes to the employer distribution in generating increased productivity. Motivated by these considerations, I am the first since Peri (2012) to go beyond earnings and provide estimates of the effect of immigrant worker inflows on revenue and revenue per worker (labor productivity) growth; and, unlike Peri (2012), I can do so at the more granular, local industry level. Of additional note from the U.S. literature is Ayromloo et al. (2020), who find that state enforcement of e-Verify laws leads to exits by larger firms, consistent with the importance of heterogeneous exit to my results.

A deep literature studies the questions of whether and how immigrants are absorbed into local economies, questions most directly addressed in Section 3.1. Motivated by Lewis (2005) and Lewis (2012), this paper is specifically focused on how production—specifically, the firms that are in the market for immigrant labor—respond to immigrant-induced increases in labor supply within an industry. This contrasts with other important immigrant absorbing mechanisms—including changes in cross-industry output mix (e.g., Gonzalez and Ortega, 2011; Burstein et al., 2020), increases in consumer demand (e.g., Hong and McLaren, 2015), and respondent changes in labor supply (e.g., Monras, 2020).[5] I isolate these responses in production by looking within local industries while controlling for commuting-zone-wide fixed effects. Dustmann and Glitz (2015) were the first to broach business entry and exit as important immigrant absorption mechanisms on the production side of the economy, finding that it accounts for 15 percent of immigrant-induced net job creation in Germany's tradable sector. Their decomposition exercise motivates the decomposition presented in Section 3.1. I am the first to conduct this decomposition in the U.S. context, and my results suggest that extensive margin responses account for a substantially larger proportion of immigrant-induced job creation in the U.S. compared to Germany.

The empirical results in this paper are also related to but set apart from recent literature on entrepreneurship. It builds on work identifying the link between population growth and business entry in the U.S. (Hopenhayn et al., 2018; Karahan et al., 2019) by showing that immigrant workers are particularly active in changing extensive margin firm decisions. This comports with previous literature showing that immigrants are more likely to work at new firms, both as owners and employees (Kerr and Kerr, 2016). However, despite important recent work on the particular importance of immigrant entrepreneurship to the U.S. economy (Kerr and Kerr, 2016, 2018; Azoulay et al., 2022), I find that only 23 percent of establishment entry generated by immigrant worker inflows comes from immigrant entrepreneurs, whereas 56 percent comes from publicly-held firms. This is another novel result that reinforces a major theme of this paper: ties between immigrant *employees* and high-productivity firms drive the dynamics found here.

[5] Important recent works comparing channels of adjustment include Dustmann and Glitz (2015) and Monras (2021).

An overarching empirical contribution of this paper comes from its approach to identification. Motivated by Autor et al. (2013) and Llull (2017), I build a shift-share instrument that replaces the more standard use of immigrant inflows from origin countries with emigrant outflows from origin countries to non-U.S. destinations in the shift component. This innovation, coupled with a rich set of controls and fixed effects in my estimating equations, makes tangible improvements to the plausibility and usability of this type of identifying variation in the face of recent work on shift-share instrumentation (Jaeger et al., 2018; Borusyak et al., 2021; Adao et al., 2019). Furthermore, this approach identifies the effect of immigrant inflows that are primarily comprised of "low-skilled" workers. This paper therefore complements a bevy of recent research that has emphasized "high-skilled" immigration by instead focusing on a set of workers that are more typically part of the average U.S. immigrant worker inflow and less obviously connected to productivity through factors like patenting and innovations. Orefice and Peri (2020) independently use a similar strategy for immigration to French regions.

Section 4 develops a theoretical model in which immigrant-hiring firms are positively selected on productivity. The fixed cost mechanism that produces this selection is motivated by Bustos (2011), who introduces endogenous technological change to the Melitz (2003) framework by allowing a subset of firms to pay a higher fixed cost to access a better production technology. Tying immigrant workers to higher productivity firms introduces a new channel through which immigrants increase productivity and lower prices: by shifting the firm productivity distribution rightward. Immigrant-induced price decreases that increase consumer welfare have been studied by Cortes (2008) (through a wage reduction channel) and both di Giovanni et al. (2014) and Hong and McLaren (2015) (through an increased variety channel). The productivity distribution channel is a novel addition to this theoretical literature that ties immigrant-induced technological change (e.g., Clemens et al., 2018) to firm heterogeneity. Independently, Brinatti and Morales (2021) also develop a model that incorporates worker heterogeneity by nativity into a model of firm heterogeneity with immigrant recruiting costs and find that this setup increases the immigration surplus. Unlike their work, this paper is focused on the extensive margin and the firm distribution rather than the decisions of continuing firms on the intensive margin.

The rest of this paper is organized as follows: Section 2 describes the U.S. Census data and shift-share identification approach used in subsequent analyses. Section 3.1 quantifies the positive relationship between immigrant worker inflows and establishment entry and exit and culminates by showing how these relationships drive immigrant absorption. Section 3.2 analyzes heterogeneous establishment shut-down decisions in response to immigrant worker inflows, then analyzes entrants over that same time period. Section 3.3 shows how these dynamics alter the productivity distribution and ultimately increase productivity growth. Section 3.4 probes additional heterogeneity, including the role of immigrant entrepreneurship. Section 4 incorporates the new insights into the production process generated by the empirical results into a parsimonious theoretical model and reassesses the welfare impact of U.S. immigration in general equilibrium. Section 5 concludes.

2 Data and Identification

2.1 Data

The analyses presented in Section 3 are facilitated most importantly by access to confidential data from the U.S. Census Bureau's Longitudinal Business Database (LBD). The LBD is an establishment level panel dataset constructed from administrative tax records for each U.S. non-farm, employee-hiring, private-sector establishment. Establishments are assigned unique, consistent identifiers that can be linked over time to create a true panel. The LBD also contains unique firm identifiers, which allows me to link establishments to their parent firms. A majority (≈75%) of establishments can be matched to parent-firm-level revenue information from the Census Bureau's BRFIRM_REV dataset starting in 1997 (see Haltiwanger et al., 2019). For large, representative samples in 2002 and 2017, the establishment identifiers can also be linked to establishment-level revenue information from the 2002 Survey of Business Owners (SBO) and 2018 Annual Business Survey (ABS).[6] Each of these data sets covers roughly 33% of all LBD establishments in the survey year. The 2018 ABS additionally contains information on ownership nativity of privately-owned businesses that will be used in Section 3.4.1.[7]

In order to study the effect of immigrant presence on outcomes constructed from the LBD, I also exploit restricted-access U.S. Census Bureau demographic data from the 1980, 1990, and 2000 Long-Form Decennial Censuses and the 2005 through 2017 American Community Surveys (ACS).[8] The demographic data allow for unusually precise measures of immigrant inflows, not just into geographies, but into relatively detailed industry groups within U.S. counties and by country of origin. These elements are important to the identification strategy presented in Section 2.2. I limit the sample to employed individuals (both self-employed and employees) so that I can assign workers an industry.

The empirical analyses in this paper are based on immigrant exposure in commuting zone-industry group pairings—local industries—over time.[9] I study the 722 commuting zones in the contiguous United States and the 41 industry groups seen in Table A1. This results in coverage of 29,602 local industries per time period in each estimated model. Appendix Section A provides additional details on the data and sample construction, including Section A.4 on industry group construction.

[6] The 2018 ABS covers the year 2017.

[7] Note that this paper's focus on the time period 2000–2015—despite the fact that the LBD stretches back to 1976—is necessitated by lack of revenue data prior to 1997. Revenues form the basis of several key analyses in Sections 3.2 and 3.3. I also find it useful to leave out pre-period data for use in balance checks, as seen in Section 2.2.

[8] Most importantly, I use the 2000 Long-Form Decennial Census responses to measure immigrant presence in 2000 and the 2013–2017 ACS to measure immigrant presence in 2015 (taken as an average to increase underlying sample size). However, in Section 3.2.1, I also use the 2005 and 2008–2012 ACS to measure immigrant presence in 2005 and 2010, respectively. And in Section 2.2, I use the 1980 and 1990 Long-Form Decennial Census responses to measure pre-period outcome variables for use in an instrument balance test.

[9] Commuting zones are groupings of counties meant to mimic local labor markets. Crosswalks from counties to commuting zones are downloaded from David Dorn's Data Page, as used in Autor and Dorn (2013).

2.2 Identification Strategy

To facilitate discussion of the identification strategy, I first present the primary specification used in Section 3:[10]

$$\Delta y_{gk} = \beta[\Delta I_{gk}] + \alpha_g + \alpha_{d(g),k} + \Gamma X_{gk} + \varepsilon_{gk} \tag{1}$$

where g indexes a commuting zone, $d(g)$ indexes the Census Division that contains commuting zone g, k indexes an industry group, and Δ represents the 15-year long difference within a local industry gk between 2000 and 2015.[11] Δy_{gk} is an outcome related to business dynamics in local industry gk; for example, the log change in the establishment count between 2000 and 2015. The independent variable of interest, ΔI_{gk}, is the change in immigrant worker stock in gk between 2000 and 2015, divided by initial (2000) workforce size in gk—a relative immigration shock to labor supply.

The granularity of the data allows for a rich set of fixed effects and control variables. α_g removes any effects immigrant inflows have at the commuting zone level as a whole. Under the premise that immigrants do not solely demand goods in the industry in which they work, α_g then insulates β from being primarily identified by changes in consumption patterns that can result from immigration. This premise is strengthened by the fact that we compare across 41 industry groups within a commuting zone.[12] Removing the impact of consumer demand helps align the identification strategy with the goal of this paper, which is to focus on the responses of firms that are in the market for labor represented by ΔI_{gk}.[13] $\alpha_{d(g),k}$ restricts comparisons across commuting zones to those in the same division and industry, which eliminates the influence of regional shocks to an industry from contaminating estimates. X_{gk} is a vector of control variables that includes initial (2000) immigrant share of the local industry workforce, initial college share, initial workforce size, a Bartik control for predicted employment growth between 2000 and 2015, the proportion of initial national employment in industry group k accounted for by gk, shift-share controls for sending-country shocks to imports, exports, and population, the sum of shares from the shift-share instrument described in 2.2.1, and a control for differences in measured employment growth across the LBD and the U.S. Census Bureau demographic data. Additional details on the inclusion and construction of these control variables can be seen in Section A.6.

OLS estimates from Equation (1)—even with its rich set of controls and fixed effects—are subject to a source of bias shared by much of the immigration literature: immigrant workers are

[10] Minor deviations from Equation (1) are necessitated in Section 3.2 to accommodate establishment-level analyses.

[11] From this point forward, use of the term "division" will refer to a Census Division: East, West, South, or Midwest.

[12] For example, an inflow of immigrants into the "Hospitals" industry group can generate an increase in economic activity in other nontradable industry groups because the new immigrant workers in the "Hospitals" industry group also consume goods and services locally. By including α_g and thus inducing comparison across industry groups within a given commuting zone and decade, β measures the immigrant-induced increase in economic activity in the "Hospitals" industry group above and beyond what other industry groups experienced due to this consumer demand effect.

[13] As opposed to the responses of firms to immigrant-induced increases in consumer demand.

often attracted by employment opportunities,[14] which confounds their effect on economic outcomes. Isolating exogenous variation that pushes immigrants into local industries substantially strengthens our case for a causal interpretation of β. To this end, I turn to a shift-share instrumental variables approach.

2.2.1 Emigrants IV

The primary instrumental variable used in this paper takes the following form:

$$\Delta z_{gk}^{M} = \sum_{o} \underbrace{\pi_{og}\rho_{ok}}_{\text{shares}} \left[\underbrace{\Delta \log(M_{o,-US})}_{\text{shifts}} \right].$$

π_{og} is the proportion of 2000 origin o immigrants living in commuting zone g, ρ_{ok} is the proportion of 2000 origin o immigrants working in industry group k, and $\Delta \log(M_{o,-US})$ is the log change in the number of emigrants from origin o living in one of 18 *non-US* OECD destinations between 2000 and 2015, measured from the United Nations Population Division's (UNDP) International Migration Stock 2019.[15]

Relative to a standard shift-share immigration instrument (e.g. Card, 2001), the primary innovation of this approach is to use *non-U.S. emigration* instead of *U.S. immigration* from origin country o in the shift component. By replacing a measure of origin-specific immigration to the U.S. with $\Delta \log(M_{o,-US})$, I retain the full suite of factors that are *pushing* individuals in origin country o to emigrate—both to the U.S. and to other destinations—while discarding any factors that are specifically *pulling* immigrants into the U.S.—including and most importantly, local-industry-specific labor demand.

Previous empirical literature has implicitly recognized the value of reducing endogeneity in the shift component of a shift-share instrument. In a well-known, recent example, Autor et al. (2013) instrument for Chinese import exposure in U.S. local labor markets by interacting measures of initial industry concentration in commuting zones (shares) with Chinese, industry-level *export* growth to eight advanced, *non-U.S.* economies (shifts). Within the immigration literature, Llull (2017) uses a suite of origin-specific migration push factors—conflict, natural disasters, changes in per capita income, and changes to political regimes—as shifts in shift-share instruments. In both Autor et al. (2013) and Llull (2017), the use of shifts that originate from abroad is an attempt to purge a shift-share instrument of latent economic trends specific to any of the local labor markets being studied. Relative to the specific push factors used in Llull (2017), Δz_{gk}^{M} retains *all* push factors and therefore helps with instrument strength.

[14] While this is true of worker movements more generally, Cadena and Kovak (2016) show that immigrant workers can be more economically motivated in their *internal* migration decisions than their native counterparts. This idea also underlies recent work in Abramitzky and Boustan (2022).

[15] The 18 destination countries are Australia, Austria, Canada, Chile, Denmark, Finland, France, Germany, Greece, Ireland, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom.

Borusyak et al. (2021) and Jaeger et al. (2018) provide some explicit theoretical backing for using more plausibly exogenous shifts. Borusyak et al. (2021) show that when shifts represent a set of relatively disbursed and uncorrelated shocks across origin countries, their quasi-random assignment conditional on shares can overcome endogeneity in the share component of the instrument.[16,17] Meanwhile, Jaeger et al. (2018) show that aggregate immigrant inflows into the U.S. are highly serially correlated in their origin country composition. This generates a potentially severe bias that confounds short- and long-run responses to immigration. To the extent that they are more plausibly exogenous and less serially correlated at the origin-country-level than aggregate U.S. immigrant inflows, non-U.S. emigrant outflows are also less likely to generate biased coefficient estimates $\hat{\beta}$.

Shares π_{og} and ρ_{ok} apportion emigration shocks into commuting zones and industry groups, respectively, in the spirit of Card (2001). π_{og} is a standard measure of the origin o immigration network in commuting zone g, meant to leverage immigrant preferences for co-location with compatriots. ρ_{ok} is a measure of origin-specific comparative advantage that leverages the tendency of immigrant workers from particular countries to specialize in certain industries.[18]

2.2.2 Assessing Conditional Exogeneity of the Emigrants IV

The identifying assumption underlying consistent two-stage-least-squares (2SLS) estimation of Equation (1) with Δz^M_{gk} as the instrumental variable is that local industries with highest exposure to shifts $\Delta \log(M_{o,-US})$ through π_{og} and ρ_{ok} do not have systematically different potential outcomes than local industries with lower exposure to these shifts, conditional on fixed effects and controls. I take a simple approach to assessing this assumption, described in the reduced form regression below:

$$\Delta y^{\text{Norm}}_{gk} = \phi[\Delta z^M_{gk}] + \alpha_g + \alpha_{d(g),k} + \Gamma \tilde{X}_{gk} + \varepsilon_{gk} \tag{2}$$

where $\Delta y^{\text{Norm}}_{gk}$ is a normalized version of a given outcome. Outcomes include growth in workforce size, payroll, immigrant workforce size, the establishment count, employment, and average earnings, each measured during (true outcomes) and prior (balance outcomes) to the study period.[19] Outcomes measured during the study period may be rightfully affected by the immigration shock represented by Δz^M_{gk}, but correlations between Δz^M_{gk} and pre-period outcomes indicate likely viola-

[16] Since this identification strategy is conditional on shares, the sum of shares is always included in X_{gk}.

[17] Their identifying assumption contrasts with Goldsmith-Pinkham et al. (2020), who focus on exogeneity in the share component as a sufficient condition for instrument validity.

[18] Card (2001) also examines immigrant inflows into groupings *within* local areas, but his groupings are occupation-based instead of industry-based. In addition, the analogous measure to ρ_{ok} in Card (2001) is not measured in the base year, but rather is the proportion of national immigrants from origin o *during* the study period that ended up in a given occupation. I follow guidance in Borusyak et al. (2021) to measure shares exclusively prior to the shock.

[19] Workforce growth and growth in the immigrant workforce are measured from Census Bureau demographic data. Payroll, employment, and average earnings growth are measured from the LBD. Beyond being measured from a different data source, "workforce" includes the self-employed and other employees that are enumerated in the Census demographic data but not as employees in the LBD data. "Employment"—measured in the LBD—only reflects employment at LBD-covered establishments.

tions of the identifying assumption. All control variables described in Section 2.2 are included in $\tilde{X}_{gk}$, except for the Bartik labor demand control for predicted employment growth, which I leave out for use as an additional balance test.

Figure 1 presents the result of this exercise and offers support for conditional exogeneity of Δz_{gk}^M. Estimated coefficients $\hat{\phi}$ are not statistically significant and hover around zero for growth rates of the five pre-period outcome variables in the 1980s and 1990s (maroon squares). Of particular note are the lack of effects on immigrant workforce growth in the 1980s and 1990s, which mitigate concerns that Δz_{gk}^M picks up long-run responses to immigration shocks from previous time periods (Jaeger et al., 2018). $\hat{\phi}$ is also small and insignificant when the Bartik control for predicted employment growth *during* the study period is the outcome, providing evidence that Δz_{gk}^M is a true "supply-push" that does not simultaneously shift demand (black circle).[20] The true outcomes at the top of the figure (blue diamonds) offer a preview of some basic empirical results and reassurance that the pre-period results are not just the result of excessive noise. All show statistically significant, positive effects, with the smallest point estimate higher than the largest upper bound of the 95% confidence intervals from the balance outcomes.

2.2.3 Educational Composition of Immigrant Inflows

Figure 2 benchmarks the differences in educational attainment associated with immigrant inflows relative to receiving native workforces in the local industries studied throughout this paper. The bars reflecting immigrant inflows are obtained by setting $\Delta y_{gk} = \Delta I_{gk}^e$ in Equation (1), where ΔI_{gk}^e represents net inflows of immigrant workers with educational attainment level e, such that $\sum_e \Delta I_{gk}^e = \Delta I_{gk}$.[21]

Figure 2 contains two key findings. First, comparing the "Δz^M Pushed Immigrant Inflow" bar to the "2000 Receiving Native Workforce" bar reveals that inflows of immigrant workers tilt the composition of the workforce towards less-educated workers. That is, the results in Section 3 are primarily driven by "low-skilled" immigrant workers. On the education dimension, then, standard immigration surplus arguments should apply.[22] Second, comparing the "Δz^M Pushed Immigrant Inflow" bar to the "OLS Immigrant Inflow" bar reveals that Δz_{gk}^M tends to push immigrants of slightly lower, but largely similar educational attainment into the U.S. relative to the typical immigrant inflow during the study period. This alleviates concerns that the preferred 2SLS approach identifies a less policy-relevant estimand.

[20] The Bartik control variable is given by $\text{Bartik}_{gk} = \sum_{k'} \frac{E_{gk',2000}}{E_{gk,2000}} \times \Delta \log(E_{gk'})$, where k' indexes a 6-digit NAICS industry and E stands for employment. It is motivated by the seminal work in Bartik (1991). See Section A.6 for more details.

[21] Exploiting the adding-up property of linear regression, estimating Equation (1) for mutually exclusive and exhaustive educational groupings decomposes how many workers of each educational attainment category are brought in by each immigrant, on average.

[22] See Borjas (1999).

2.2.4 First Stage Strength, Inference, and Native Displacement

Table 1 provides important, additional context around this paper's identification strategy along with a first set of results. Column (1) demonstrates a strong first stage, consistent with the notion that non-U.S. emigrant outflows reflect a bevy of factors that are also relevant in pushing immigrants to the U.S. Throughout the table, conventional standard errors—robust to heteroskedasticity—are presented in parentheses. Where applicable, Borusyak et al. (2021) "exposure-robust" standard errors are presented in square brackets.[23] Under regularity assumptions, these standard errors are robust to the inferential concerns specific to shift share instruments broached by Adao et al. (2019).[24] In Column (1), implied first-stage F statistics are 93.6 under conventional standard errors and 101.1 under the Borusyak et al. (2021) standard errors. The similarity across inference procedures reflects the ability of the rich fixed effect and control structure in Equation (1) to prevent excessive correlation across local industries with similar shares from undermining inference.[25]

Columns (2) and (3) estimate Equation (1) using the net change in a local industry's *native* workforce over 2000–2015 divided by initial workforce size as the outcome variable. Given that ΔI_{gk} is the net change in a local industry's *immigrant* workforce over 2000–2015 divided by initial workforce size and that both variables are measured using the same Census demographic data sources, β simply measures the number of additional native workers per immigrant, with $\beta = 0$ indicating no native displacement, on net.[26] $\beta \neq 0$ would prompt concerns that general equilibrium channels may seriously undermine the interpretation of results presented in Sections 3.1 through 3.4, whereby low-immigration local industries are experiencing spillover effects due to incoming natives who were displaced in high-immigration local industries or due to outgoing natives who are drawn in by high-immigration local industries.

Column (2) estimates Equation (1) using OLS and shows that immigrant worker inflows are correlated with growth in the native workforce. While there is a plausible causal interpretation of this result, it more likely reflects the standard endogeneity bias that arises because immigrant workers are attracted to areas with growing employment opportunities. Column (3) shows that estimating Equation (1) using 2SLS with Δz^{M}_{gk} as the instrumental variable likely eliminates this

[23]These are calculated using `ssaggregate.do` in `Stata` (Borusyak et al., 2018).

[24]These standard errors are obtained by "transforming" Equation (1) into an equivalent, origin-o level regression, where standard errors can be clustered at the level of the shock. Because regional factors may impact global migration pushes, I cluster these "exposure-robust" standard errors at the UN region level. There are 18 UN regions represented: Central America, Caribbean, South America, Northern Europe, Western Europe, Southern Europe, Eastern Europe, Northern Africa, Western Africa, Middle Africa, Eastern Africa, Southern Africa, Western Asia, Central Asia, Southern Asia, Eastern Asia, Polynesia, and Southeastern Asia. Borusyak et al. (2021) show that "exposure-robust" standard errors are conservative under the assumption that X_{gk} can be written as a function of unobserved o-level shifters, p_o: $X_{gk} = \sum_o \text{Immigrants}_{og,2000} \times \text{Immigrants}_{ok,2000} \times p_o + \omega_{gk}$ where ω_{gk} is white noise.

[25]This will be seen again in various results from Sections 3.1 and 3.3.2. Because "exposure-robust" standard errors proposed by Adao et al. (2019) and Borusyak et al. (2021) do not readily translate to the more granular analyses presented in Section 3.2, I report and prefer conventional standard errors throughout the paper, whereas Borusyak et al. (2021) are reported for basic local-industry-level analyses in order to alleviate the concerns broached by Adao et al. (2019).

[26]This specification and outcome have a deep tradition in immigration economics. See Peri and Sparber (2010) for more. Full native displacement would mean $\beta = -1$.

bias, with $\hat{\beta} \approx 0$.[27] While this alleviates first order concerns regarding general equilibrium spillovers, an important caveat is that *net* zero inflows of native workers does not preclude compositional changes. I nonetheless take the results from Table 1 as evidence that Δz_{gk}^M is a strong instrument that corrects for a first-order endogeneity bias and pushes in a set of immigrant workers that are roughly fully absorbed into the local industries that they enter.

2.2.5 Additional Vetting of the Emigrants IV

Appendix Section B presents several additional checks and analyses that further vet and characterize Δz_{gk}^M. This includes a direct check of the double-instrumentation procedure advocated by Jaeger et al. (2018), a comparison between Δz_{gk}^M and a more standard shift-share immigration instrument, more on the composition of immigrant inflows pushed in by Δz_{gk}^M, and a representative example from the U.S. housing bubble of how Δz_{gk}^M corrects for endogeneity. All told, Appendix Section B finds that Δz_{gk}^M holds up to a battery of checks and further justifies its use as the primary instrumental variable in Section 3.

3 Empirical Results

This section presents the three core empirical results of the paper using the identification strategy outlined above. Section 3.1 focuses on the role of establishment entry and exit in immigrant absorption. Section 3.2 uncovers creative destruction in response to immigrant worker inflows using granular, establishment-level analyses. Section 3.3 finds that immigrant worker inflows ultimately lead to economic activity heavily concentrated at the top of the productivity distribution and an increase in aggregate productivity. Section 3.4 complements these core findings with additional heterogeneity analyses.

3.1 Entry, Exit, and Immigrant Absorption

3.1.1 Decomposing Growth Rates

Throughout this section, I utilize Davis-Haltiwanger-Schuh (DHS) growth rates in the establishment count and employment for a local industry over the time period 2000–2015 as outcome variables in Equation (1):

$$\Delta y_{gk} = \frac{Y_{gk,2015} - Y_{gk,2000}}{(Y_{gk,2015} + Y_{gk,2000})/2} \equiv \frac{\Delta Y_{gk}}{\text{Denom}_{gk}^Y} \approx \Delta \log(Y_{gk})$$

[27]Note that this result does not necessarily contradict those found in Hong and McLaren (2015). Their commuting-zone-level study incorporates a job multiplier that arises due to consumer demand, which I largely control away using the fixed effect α_g.

where Y_{gkt} is either the LBD-measured establishment count or LBD-measured employment in local industry gk in year t.

DHS growth rates closely approximate log changes but easily allow for decomposition analysis. Specifically, I denote $\mathcal{E}_{gk}$ as the set of establishments that were not active in 2000 but were active in 2015 in local industry gk ($\mathcal{E}$ for entrants), $\mathcal{X}_{gk}$ as the set of establishments that were active as of 2000 but that were no longer active as of 2015 ($\mathcal{X}$ for exiters)[28], and $\mathcal{C}_{gk}$ as the set of establishments that were active in gk in both 2000 and 2015 ($\mathcal{C}$ for continuers). Then, letting e index an establishment, we can split the numerator ΔY_{gk} for each DHS growth rate using the following decompositions:

$$\Delta Y_{gk} = \Delta \text{Estab Count}_{gk} = \underbrace{\sum_e \mathbf{1}\{e \in \mathcal{E}_{gk}\}}_{\text{Entry}} - \underbrace{\sum_e \mathbf{1}\{e \in \mathcal{X}_{gk}\}}_{\text{Reduced Exit}}$$

$$\Delta Y_{gk} = \Delta \text{Emp}_{gk} = \underbrace{\sum_{e \in \mathcal{E}_{gk}} \text{Emp}_{e,2015}}_{\text{Entry}} - \underbrace{\sum_{e \in \mathcal{X}_{gk}} \text{Emp}_{e,2000}}_{\text{Reduced Exit}} + \underbrace{\sum_{e \in \mathcal{C}_{gk}} \Delta \text{Emp}_e}_{\text{Incumbent Growth}}$$

Dividing each of these components by Denom^{Y}_{gk} generates a set of outcomes that decompose the overall growth in establishment count and employment into their component parts that stem from entrants, exiters, and continuers.

3.1.2 Results

Figure 3 presents results from the setup described above and contains several novel results on U.S. immigration. Starting on the left, 2SLS estimates indicate that a one percent shock to a local industry's workforce due to immigration results in a 0.9 percent increase in the establishment count (black dot and surrounding 95% confidence bands), with 56 percent of this effect coming from establishment entry and 44 percent of this effect coming from reductions in establishment exit.[29] These results comport with and extend a sparse extant literature on immigration and establishment dynamics in the U.S. economy. Olney (2013) and Orrenius et al. (2020) also find that immigrant inflows increase establishment presence, with Orrenius et al. (2020) also finding that a one percent relative immigration shock leads to a 0.9 percent increase in the establishment count.

Much of the rest of the empirical results can be thought of as flowing out of the left bar in Figure 3—moving beyond the current literature by generating a more complete understanding of *how* entry and exit factor into job creation, alter a local industry's productivity distribution and ultimately lead to productivity growth. This starts with the right bar of Figure 3. A one percent

[28] For the remainder of the paper, I will use "exit" and "inactive" synonymously, where inactive means no payroll or employment. Note that this means that "exit" is not necessarily an absorbing state in this paper. Empirically, true (absorbing) exit and "inactive" often coincide.

[29] See Table C2 for the underlying coefficients and standard errors of the establishment count growth rate decomposition.

increase in workforce size due to immigration increases LBD-covered employment by 0.76 percent.[30]

While this headline number is not surprising, the decomposition presents a striking picture. Not only do immigrant worker inflows generate the increases in establishment entry and reductions in establishment exit seen in the left bar of Figure 3, it is precisely through these channels that immigrants induce job creation in U.S. local industries. In fact, I cannot reject the null hypotheses that employment growth at continuing, incumbent establishments plays no role whatsoever in immigrant-induced job creation (see Table C3). Meanwhile, increases in establishment entry and reudctions in establishment exit account for 50 percent and 44 percent of immigrant-induced net job creation, respectively. In other words, the immigrant absorption process—the process through which enough jobs are created to keeps pace with immigrant inflows—appears to occur entirely on the extensive margin of labor demand. It is also worth noting that this result may be quite specific to the U.S.: when Dustmann and Glitz (2015) performed a similar decomposition exercise among tradable firms in Germany, they found that firm entry and exit explained only 15 percent of immigrant-induced net job creation.

3.2 Establishment Exit and Creative Destruction

Section 3.1 found that increased establishment entry and reduced establishment exit play roughly equal roles in generating increased establishment presence and employment growth in U.S. local industries. While this delivers new insights into the immigrant absorption process, the role of reduced exit also broaches concerns that immigration may stunt the creative destruction process associated with productivity growth. In this section and in Section 3.3, I shift the focus of the empirical analysis towards the productivity consequences of immigrant worker inflows in U.S. local industries. I start by conducting granular analyses of exit and entry in Sections 3.2.1 and 3.2.2, respectively, taking full advantage of the LBD establishment panel.

3.2.1 The Exit Margin

The analyses in this subsection utilize a fully balanced panel of LBD establishments that were active as of 2000 and that can be linked to their parent firm's revenue information in 2000. I follow these 4.7 million establishments every five years through 2015: $t \in \{2000, 2005, 2010, 2015\}$. To start, I re-confirm the relationship between immigrant worker inflows and establishment exit in this establishment-level panel using the following specification:

$$\mathbf{1}\{\text{Inactive}_{et}\} = \beta[I_{g(e),k(e),t}] + \alpha_e + \alpha_{gt} + \alpha_{d(g),kt} + \Gamma X_{gkt} + \varepsilon_{et} \tag{3}$$

[30]Note that this number does not preclude full absorption of immigrant workers into the local industry—in fact, full absorption cannot be rejected, as shown in Table 1. Rather, the less than one-for-one elasticity found here reflects the fact that ΔI_{gk} includes self-employed individuals and employees in the non-profit and public sectors, many of whom are not enumerated by the LBD.

where $I_{g(e),k(e),t}$ is now the stock of immigrant workers in establishment e's local industry $(g(e), k(e))$, divided by initial (2000) workforce size in the local industry.[31] Equation (3) is the direct analog to Equation (1) but broken down to the establishment level. Of particular note is the inclusion of an establishment fixed effect, α_e, which subsumes the first-difference employed in Equation (1). Outcome $\mathbf{1}\{\text{Inactive}\}_{et}$ is an indicator for whether an establishment is no longer active, with both zero employment and payroll. Because the immigrant exposure variable is now in levels, I employ a levels version of the instrumental variable:

$$z_{gkt}^{M} = \sum_{o} \pi_{og}\rho_{ok}\big[\log(M_{ot,-US})\big]$$

where I take advantage of the UNDP International Migration Stock data's quinquennial frequency to measure $\log(M_{ot,-US})$.[32]

Table 2 displays the first stage, OLS, and 2SLS estimates from Equation (3). These results corroborate results from Section 3.1 and provide a benchmark against similar estimates from non-U.S. settings. Column (1) indicates a reassuringly similar first stage from the establishment panel regression as that found in Table 1 from the local-industry-level regression, with an implied first stage F statistic of 114.4. Both OLS and 2SLS estimates show a negative effect of increased immigrant worker presence on establishment exit. The preferred, 2SLS estimate finds that a one percent relative immigration shock reduces the probability of establishment by 0.21 percentage points. This estimate is about one-third of the size of analogous estimates from France contained in Mitaritonna et al. (2017). One reason for this discrepancy may be the findings regarding increased exit among lower productivity establishments in the U.S., described below.

Specifically, to address concerns of stunted creative destruction that may arise as a byproduct of reduced establishment exit, I modify Equation (3) as follows:

$$\mathbf{1}\{\text{Inactive}_{et}\} = \sum_{q=1}^{5} \beta_q[I_{g(e),k(e),t} \times \mathbf{1}\{q(f(e)) = q\}] + \alpha_e + \alpha_{gt} + \alpha_{d(g),kt} + \Gamma X_{gkt} + \varepsilon_{et} \tag{4}$$

where $f(e)$ indexes establishment e's parent firm and $q(f(e))$ indexes $f(e)$'s quintile in the initial (2000) productivity distribution. In this model, β_q represents the effect of increased exposure to immigrant workers on the probability of exit for an establishment whose parent firm is in the qth quintile of the productivity distribution.

My preferred proxy for initial productivity is based on firm-level revenues at the start of the analysis period (2000), obtained by linking LBD establishments to their parent firms in the BR-FIRM_REV data set.[33] I assign firms to quintiles by ranking them according to their log revenues

[31]Control variables are the same as in Equation (1), but also modified to be in levels, where appropriate. See Section A.6 for details.

[32]This every-five-year frequency allows me to track the relationship between immigrant workers and exit more closely than in Section 3.1 but also precludes an annual analysis.

[33]Recall that establishment-level revenues are only available for a sample of establishments in the SBO and ABS.

in 2000 within 6-digit NAICS code by age group bins.[34] Section A.5 discusses the motivations, strengths, and weaknesses of this measure in detail. To summarize, ranking within detailed industry compares firms with similar input requirements and ranking within age bins compares firms with similar footholds in the market. Removing the influence of input mix and name recognition brings us closer to a proxy for firm-level total factor productivity. I also present results based on revenues-*per-worker* quintiles alongside the revenue-based results.[35,36]

Given that it is stratified by firm productivity, Equation (4) tests whether firm heterogeneity modifies the impact of immigration by asking who stays and who goes in response to immigrant worker inflows. A reasonable prior—that explains the reductions in exit found above—might be that immigrant workers are simply pushing down labor costs in local industries that they enter. In this case, we would expect to see muted heterogeneity, with β_q of similar magnitude across q. We may even expect $\beta_1 < ... < \beta_5 < 0$ in this case, with establishments from marginal firms—who are initially at high hazard risk—being subsidized away from exit. Similar effects may arise from immigrant-induced increases in consumer demand, if they are not adequately swept out by the commuting-zone-by-year fixed effect, α_{gt}.

Figure 4 plots 2SLS estimates of $\hat{\beta}_q$ from Equation (4) and presents a strikingly different picture. Regardless of our productivity proxy, establishments whose parent firms are in the lowest productivity quintile are *more* likely to exit with increased exposure to immigrant workers. $\hat{\beta}_q$ declines monotonically, with establishments whose parent firms are in the top two quintiles driving the overall reduction in exit found in Table 2.

The stark heterogeneity found in Figure 4 suggests that inflows of immigrant workers raise the productivity bar that firms need to clear in order to continue operating establishments. In many models of firm heterogeneity, including the one presented in Section 4, raising this bar also raises aggregate productivity. Indeed, Section 3.3.2 finds such an increase. Figure 4 also presents a strong signal that immigrant workers are specifically tied to higher productivity firms, a relationship that also has profound implications for the welfare impacts of immigration that are probed in Section 4. At the very least, these results suggest that firm heterogeneity is critical to the understanding of immigration. Finally, as seen in Figure 4 and Appendix Section C.3, these results are robust to alternate ways of defining firm productivity.

[34] Firm age groups are 0-1, 2-4, 5-9, 10-19, and 20+, and firm age is determined by the age of a firm's oldest establishment. Firm-level NAICS codes are provided in the BRFIRM_REV data set. Rankings are weighted by inverse probability weights given in the BRFIRM_REV data set in order to generate quintile cutoffs. These are inverse probability weights that are designed to make the subset of firms that have revenue information representative of the national firm distribution in 2000.

[35] The revenue-based measure is preferred here because the model presented in Section 4 contains a one-for-one relationship between revenues and TFP at the firm level. Meanwhile, in the model, revenues per worker and TFP are independent. However, because revenues per worker is a more common proxy for TFP in the literature, I present results using this proxy alongside the revenue-based proxy.

[36] In the Appendix, I also present results in which I base productivity quintiles on rankings within local industries rather than national 6-digit NAICS by age bins. Results are robust to all of these alternate choices.

3.2.2 Entrants in High Immigration Environments

Section 3.2.1 showed that reductions in exit found in 3.1 do not benefit low-productivity businesses; in fact, they mask *increased* exit by these marginal businesses. In other words, the "destruction" in "creative destruction" does unfold in response to immigrant worker inflows in U.S. local industries. This section asks whether the "creative" part manifests by studying how entrants in high immigration environments differ from entrants in low immigration environments, beyond being more numerous.

I approach this question using the following specification, which takes the 9.4 million LBD-covered establishments that were born between 2001–2015 and tracks them for up to 18 years into their existence:

$$y_{et} = \sum_a \beta_a[\Delta I_{g(e),k(e)} \times \mathbf{1}\{a(e,t) = a\}] + \alpha_{a(e,t)} + \alpha_t + \alpha_g + \alpha_{d(g),k} + \Gamma X_{gk} + \varepsilon_{et} \tag{5}$$

where $a(e,t)$ is the age of establishment e in year t, binned into three year intervals.[37] I return to the exposure variable from (1) and Section 3.1 and thus the instrumental variable Δz^M_{gk}.[38] Retaining the focus on productivity, y_{et} is one of three outcomes: $\mathbf{1}\{\text{Active}_{et}\}$, $\log(\text{Revenues}_{f(e),t})$, or $\log(\text{Revenues per Worker}_{f(e),t})$.[39] Note that the revenue-based outcomes are measured conditional on $\mathbf{1}\{\text{Active}_{et}\} = 1$ and on observing firm-level revenues.

In this model, β_a represents the impact of a higher immigration environment on one of these productivity proxies when an establishment is a years old. That is, it traces out the life cycle of an establishment that is born in a higher immigration environment and help adds context to the results on entry found in Section 3.1. For example, while we know there is increased entry from Figure 3, new establishments may not contribute materially to economic activity if they are "subsistence establishments" that primarily exist as an alternative to low-wage work for their founders (see, e.g. Schoar, 2010). Equation (5) sheds light on whether establishments that enter during high immigration periods are viable and productive in the long run.

Figure 5 presents 2SLS estimates of β_a from Equation (5) across our three productivity proxies. Across all three measures, but most clearly for survival and revenue-based scale, establishments that enter in high immigration environments appear more productive and to come from more productive parent firms. Moreover, these advantages are long-lived, lasting through the first 18 years of an establishment's life; and, in the case of firm-level log revenues, increasing over time.[40] Note that for revenue-based outcomes, the "control group" is a highly selected, successful group by age 17

[37] This binning helps keep the 1st Stage F statistically at a reasonable value. Equation (5) requires six instruments instead of the 18 required by interactions between all ages and ΔI_{gk}.

[38] More precisely, I instrument $\sum_a[\Delta I_{g(e),k(e)} \times \mathbf{1}\{a(e,t) = a\}]$ with $\sum_a[\Delta z_{g(e),k(e)} \times \mathbf{1}\{a(e,t) = a\}]$.

[39] $\mathbf{1}\{\text{Active}_{et}\}$ is a measure of productivity insofar as we expect more productive firms to be more likely to be profitable and therefore survive (see, e.g., Syverson, 2011), while revenues and revenues per worker more directly measure productivity conditional on survival but are measured at the firm level and not available for all firms.

[40] Note that a changing set of cohorts identifies each coefficient in Equation (5). For example, all cohorts identify β_a for a = 0-2 years old, but only establishments that were born between 2001 and 2003 identify β_a for a = 15-17 years old.

since β_{17} measures the difference in outcomes across 17 year-old *surviving* establishments that were born in high versus low immigration environments. The durability of the effects found in Figure 5 are striking because they find an effect above and beyond this selection. Section C.4 provides supplemental results.

Thus, while establishments from less productive parent firms are culled from the market (Section 3.2.1), they are also replaced by establishments from high-productivity firms that are long-lived (here). Further evidence for the high productivity levels of entrants comes from Section 3.4.1, which finds that entry is driven by publicly-held firms. In sum, the empirical analyses in this section provide direct, establishment level evidence that immigrant worker inflows lead to the kind of creative destruction associated with productivity growth. Next, I study the impact of these inflows on the productivity distribution as a whole and on productivity growth directly.

3.3 The Employer Distribution and Productivity Growth

Section 3.2 drilled down to the establishment level, fully exploiting the granularity of the LBD, to comprehensively characterize how entrants and potential exiters respond to immigrant worker inflows. In this section, I return to the local industry level and detail how these establishment-level dynamics alter the productivity distribution and, ultimately, lead to productivity growth.

3.3.1 The Employer Productivity Distribution

I return to Equation (1) to assess how immigrant worker inflows affect the employer productivity distribution as a whole:

$$\Delta y_{gk} = \beta[\Delta I_{gk}] + \alpha_g + \alpha_{d(g),k} + \Gamma X_{gk} + \varepsilon_{gk} \tag{1}$$

I do so by decomposing the DHS growth rate in the establishment count based on where establishments' parent firms fall in the productivity distribution. That is, I now estimate Equation (1) with outcome variables that reflect the change to establishment presence in each productivity quintile:

$$\Delta y_{gk} = \frac{\Delta \text{Rev Estabs}^q_{gk}}{\text{Denom}^{\text{Rev Estabs}}_{gk}}$$

for $q \in \{1, ..., 5\}$. Rev Estabs$_{gkt}$ represents the count of establishments whose parent firms can be assigned revenue information in year t in local industry gk.[41] Rev Estabs$^q_{gkt}$ represents the count of

[41]These counts are generated by adding up the weights given in the BRFIRM_REV revenue data set across establishments within gk in t. That is, if we denote $\mathcal{A}_{gkt}$ as the set of active establishments in local industry gk in year t, Rev Estabs$_{gkt} = \sum_e pw_{f(e)} \mathbf{1}\{e \in \mathcal{A}_{gkt}\}$, where $pw_{f(e)}$ is parent firm $f(e)$'s weight in the BRFIRM_REV data set. These are inverse probability weights that are designed to make the subset of firms that have revenue information representative of the national firm distribution in a given year. Note that Denom$^{\text{Rev Estabs}}_{gk} = \frac{\text{Rev Estabs}_{gk,2015} + \text{Rev Estabs}_{gk,2000}}{2}$. Then, Rev Estabs$^q_{gkt} = \sum_e pw_{f(e)} \mathbf{1}\{e \in \mathcal{A}_{gkt}\} \mathbf{1}\{q(f(e)) = q\}$.

those establishments whose parent firms are in quintile q based on the same 2000 firm productivity distributions used in Section 3.2.1. Then, $\Delta\text{Rev Estabs}^{q}_{gk} = \text{Rev Estabs}^{q}_{gk,2015} - \text{Rev Estabs}^{q}_{gk,2000}$, the change in establishment count in productivity quintile q between 2000 and 2015, and estimated coefficients $\hat{\beta}$ represent the contribution of quintile q to the impact of immigrant worker inflows on the establishment count growth rate.

Figure 6 presents the results. Across measures, it finds that immigrant-induced increases in the establishment count are heavily concentrated at the top of the productivity distribution, with roughly half of the total effect explained by the top productivity quintile alone. This represents a substantive change in the composition of firms that operate in a local industry and one that implies a rightward shift of the distribution. Figure 6 is a succinct illustration of the overarching novelty of this paper, which is to center changes to the employer distribution in the economic analysis of immigration. It also confirms that the granular, establishment-level results found in Section 3.2 generate substantive changes to this distribution.

3.3.2 Aggregate Productivity Growth

Analyses in Sections 3.2 and Section 3.3.1 deliver detailed insight into how immigrant worker inflows lead to establishment dynamics that are consistent with productivity growth in U.S. local industries. Here, I complete the picture by directly showing that proxies for productivity increase in response to these inflows at the local industry level. I do so by estimating Equation (1) using three local-industry-level outcomes: $\Delta\log(\text{Mean Earnings}_{gk})$, $\Delta\log(\text{Revenues}_{gk})$, and $\Delta\log(\text{Revenues per Worker}_{gk})$.

While $\Delta\log(\text{Mean Earnings}_{gk})$ is readily measured from the LBD by summing payroll and employment across establishments in gk and taking the quotient for $t = \{2000, 2015\}$, establishment-level revenues are not measured in the LBD.[42] Thus, to construct local industry revenues and revenues per worker, I rely on approximately 1-in-3 samples of U.S. private sector establishments enumerated in the 2002 SBO and 2018 ABS (which covers 2017). Each of these data sets contain establishment-level revenues and survey weights that can be used to aggregate revenues across surveyed establishments to a measure of local industry revenues.[43] For outcomes $\Delta\log(\text{Revenues}_{gk})$ and $\Delta\log(\text{Revenues per Worker}_{gk})$, I therefore study the effect of immigrant inflows between 2000 and 2015 on sample-measured revenue and revenue-per-worker growth between 2002 and 2017. These caveats notwithstanding, to my knowledge, I am the first to provide causal estimates of the effect of immigrant worker inflows on revenue and labor productivity growth in U.S. local industries.[44]

[42]Firm-level revenues cannot be used to measure local-industry-level because multi-unit firms play a large role in overall revenue growth in the U.S.

[43]Specifically, if we once again denote $\mathcal{A}_{gkt}$ as the set of active establishments in local industry gk in year t, $\text{Revenues}_{gkt} = \sum_e \text{Revenues}_{et} \times \text{SBO/ABS Weight}_{f(e)} \times \mathbf{1}\{e \in \mathcal{A}_{gkt}\}$, where $\text{SBO/ABS Weight}_{f(e)} = 0$ if an LBD establishment's parent firm is not enumerated in the SBO/ABS. Revenues per worker are obtained by dividing Revenues_{gkt} by LBD-measured employment in local industry gk in t.

[44]While many studies proxy for labor productivity using mean earnings, revenues per worker can differ from average earnings when product markets or labor markets are not perfectly competitive. Results in Table 3 provide a clear

Table 3 presents the 2SLS estimates from Equation (1) using these outcomes. A clear picture emerges, with immigrant worker inflows into local industries increasing aggregate productivity. A 1% relative shock to the workforce due to immigration increases average earnings by 0.4%, output by 1.9% and labor productivity by 1.2%. The divergence between earnings and labor productivity is consistent with an emerging literature on rent-sharing in the U.S. economy (see, e.g. Kline et al., 2019).[45] These results provide a concise takeaway finding that summarizes the impact of the detailed dynamics studied above. The establishment entry and exit decisions that undergird immigrant absorption and generate creative destruction ultimately lead to economically significant increases in productivity.

3.4 Additional Heterogeneity

3.4.1 Does Immigrant Entrepreneurship Drive Entry?

Given a bevy of recent literature on the importance of immigrant business owners to the U.S. economy (see, e.g., Azoulay et al., 2022; Kerr and Kerr, 2016, 2018, among others), it is natural to wonder how much immigrant entrepreneurship is contributing to the effects found above. I explore this question using outcomes derived from the 2018 ABS in Equation (1).[46] Specifically, I estimate Equation (1) with the following outcome variables that reflect the contribution of group o entrepreneurship to the entry effects found in Section 3.1:

$$\Delta y_{gk} = \frac{[\text{Group } o \text{ Owned Establishments Born After 2000}]_{gk,2017}}{\text{Denom}^{\text{Estabs}}_{gk,2000,2017}}$$

$$\Delta y_{gk} = \frac{[\text{Employment at Group } o \text{ Owned Establishments Born After 2000}]_{gk,2017}}{\text{Denom}^{\text{Emp}}_{gk,2000,2017}}$$

where o is one of four groups: 1) privately-held, immigrant-owned, 2) privately-held, native-owned, 3) publicly-held, and 4) unknown (not answered on ABS survey).[47] Immigrant-owned firms are defined as those for whom at least 50% of the ownership stake is held by individuals who were not

demonstration.

[45] Kline et al. (2019) find that the elasticity of earnings to revenues is 0.35. The implied elasticity here is 0.32. Note that this strongly argues against competitive labor markets, which are nonetheless assumed in Section 4 for simplicity. Incorporating rent sharing into the economic analysis of immigration is an important area for future research that is outside of the scope of this paper.

[46] Beyond the specific interest in *new* firms that are started by the *new* immigrants represented in ΔI_{gk}, recall that the 2002 SBO does not contain ownership nativity information. So, I cannot fully measure the importance of immigrant entrepreneurs to the reduced exit found above.

[47] To be more precise, denote $\mathcal{E}^{\text{ABS}}_{gk}$ as the set of establishments that were enumerated in the 2018 ABS who were born after 2000 and operating in local industry gk in 2017. Then, $[\text{Group } o \text{ Owned Establishments Born After 2000}]_{gk,2017} = \sum_e \mathbf{1}\{e \in \mathcal{E}^{\text{ABS}}_{gk}\} \times [\text{ABS Weight}_e] \times \mathbf{1}\{o(f(e))_{2017} = o\}$, where $o(f(e))$ indedex the ownership group of parent firm $f(e)$. Similarly, $[\text{Employment at Group } o \text{ Owned Establishments Born After 2000}]_{gk,2017} = \sum_{e \in \mathcal{E}^{\text{ABS}}_{gk}} [\text{Emp}_{e,2017}] \times [\text{ABS Weight}_e] \times \mathbf{1}\{o(f(e))_{2017} = o\}$. Note that $\text{Denom}^{\text{Estabs}}_{gk,2000,2017} = \frac{\text{Estab Count}_{gk,2017} + \text{Estab Count}_{gk,2000}}{2}$ and $\text{Denom}^{\text{Emp}}_{gk,2000,2017} = \frac{\text{Emp}_{gk,2017} + \text{Emp}_{gk,2000}}{2}$, both measured from the full-count 2000 and 2017 LBD.

born in the U.S.

Several important points emerge from Figure 7, which presents he results of this exercise. Several points emerge. First, there is an increase in new immigrant entrepreneurship and in employment growth at new, immigrant-owned firms, consistent with the literature alluded to above. Further, this increase comes at the cost of native entrepreneurship. Prospective native entrepreneurs appear to be vulnerable to the dynamics found in this paper. Whether or not these prospective entrepreneurs end up in better wage-work opportunities or are frozen out of economic opportunity altogether is an important avenue for future research that is beyond the scope of this paper. I also note that the reduction in native entrepreneurship provides additional reassurance that the results found in this paper do not just reflect reverse causation from increased economic activity to immigrant inflows.

Nonetheless, these immigrant-native dynamics pale in comparison to the influx of establishments and employment at publicly-held firms. 56% of establishment entry and 88% of employment at entrants are accounted for by publicly-held firms. The analogous numbers are 23% and 15% for privately-held, immigrant-owned firms. This is a striking, new result to the literature on U.S. immigration that directly ties immigrant workers to large businesses and better-places the importance of immigrant entrepreneurship in context. Links between immigrant *employees* and higher productivity firms most likely drive the results presented in this paper and the overall impact of immigration on the U.S. economy.

3.4.2 Heterogeneity by Industry

Appendix Section C.7 presents additional heterogeneity by industry, based on tradability of and the average education level of a worker in a given industry. I find that tradable industries experience more growth in establishment count, employment and payroll than nontradable industries, consistent with Burstein et al. (2020). I also find that effects are bigger, but less precisely estimated in industries that hire workers with higher educational attainment, on average. This could be the result of larger complementarities between the "low-skilled" immigrant inflows and "higher skilled" incumbents in these industries.

4 A Synthesizing Model

Section 3 presented several novel, partial equilibrium results on how increases in the supply of immigrant labor change the composition of labor demand, and how these changes in composition ultimately lead to increased productivity. In this section, I use a model to widen our focus. I ask how much our estimates of the economic impact of U.S. immigration in general equilibrium change when we incorporate the ties between immigrant workers and high-productivity firms implied by my partial equilibrium results.

4.1 Overview

The model has three main ingredients. The first is firm heterogeneity in productivity. In light of the large, heterogeneous responses found in Section 3, accounting for firm heterogeneity in the economic impact of immigration becomes paramount. I do this using a simple Melitz (2003) framework. The second ingredient is imperfect substitutability across immigrant and native workers. Imperfect substitutability allows me to track the specific effect of immigrant workers on the aggregate economy and can arise because immigrant and native workers have comparative advantages in different occupational tasks (e.g., Peri and Sparber, 2009).

Alone, melding firm heterogeneity with imperfect substitutiability across immigrant and native workers does not capture the specific linkages between higher productivity firms and immigrant employees that are implied by Section 3. I thus add an additional heterogeneity to my theoretical framework as the final ingredient. In the spirit of Bustos (2011), I recast the choice to hire immigrants as one of costly endogenous technological adoption. Firms must pay additional fixed operating costs to recruit from the pool of foreign-born workers. Those that do obtain lower per unit labor costs thanks to the task specialization embodied by the imperfect elasticity of substitution. Due to a spreading effect, only larger, more productive firms find it profitable to invest in hiring immigrants to obtain these lower variable costs, and this ties higher productivity firms to immigrant workers. Below, I re-analyze the economic impact of immigration within this new modeling environment.

4.2 Evidence of Immigrant Ties to High-Productivity Firms

Before formalizing the model, I first discuss the plausibility that there are fixed costs to immigrant recruiting and they result in ties between immigrant workers and high productivity firms. While the results presented in Section 3 strongly imply these ties by showing that inflows of immigrant workers benefit high-productivity firms more than low-productivity firms, they do not demonstrate them *directly*. Further, the fixed cost investment is not directly observable in the data.

Nonetheless, there are several forms of fixed costs that are associated with hiring immigrant workers. These costs include hiring translators and liaisons to be able to enter into immigrant job search networks and direct search for immigrants,[48] hiring lawyers to work on visa issues (see Section 4.2.1), paying enforcement costs (in expectation) when hiring undocumented immigrants, and discovery costs required to assess labor pools from foreign countries. Note that these costs only need to be associated with lumpy investments in order to generate ties between immigrant workers and larger firms. Hiring a single translator enables a firm to recruit multiple immigrant workers. Small firms that only want to hire one additional worker would not engage in such an investment,

[48]See e.g., this Center for American Progress report about Tyson Fresh Meats and its willingness to hire translators, liaisons, and chaplains in order to utilize "low-skilled" immigrant labor: https://www.americanprogress.org/issues/immigration/reports/2018/11/14/460894/proactive-and-patient/.

whereas large firms may.

Related literature from Europe also provides evidence that firms that employ immigrant workers tend to be more productive. Mitaritonna et al. (2017) find that manufacturing firms in France that employ immigrant workers are more than 11 percent more productive than those that employ zero immigrants. Brinatti and Morales (2021) find that larger firms in Germany have larger immigrant shares in their wage bills.[49] Next, I provide evidence from the U.S. by examining firm usage of the H-1B and H-2B visa programs.

4.2.1 The H-1B and H-2B Visa Programs

Starting in 2002, publicly-available sources have tracked firm-level applications for temporary labor certifications from the Department of Labor under the H-1B and H-2B visa programs—two employer-based visas that facilitate the hiring of foreign-born workers with a bachelor's degree (H-1B) and foreign-born workers without a bachelor's degree (H-2B), respectively.[50,51] I use a fuzzy matching procedure to link these data for 2002-2017 to the LBD, described further in Appendix Section D.1. I then use this matched data set to assess whether firms that are more intensive in the use of these programs are more productive than their counterparts.

Beyond data availability, there are two reasons why examining H-1B and H-2B visa use is particularly compelling in this setting. First, use of these programs is an archetypal example of how hiring immigrant workers can be associated with fixed costs. Obtaining a worker through each program is a complex process that almost always involves external help from a law firm or an internal apparatus that has been built specifically to handle visa applications.[52] Second, these programs make up a quantitatively important component of immigrant worker inflows. For example, the ratio of new H-1B approvals to the number of foreign-born workers with a bachelor's degree who entered the U.S. was roughly 0.36 in 2015. In the same year, the ratio of new H-2B approvals to the number of foreign-born workers without a bachelor's degree who entered the U.S. was 0.22.[53] Thus, analyzing usage of these programs helps link the model to data for a set of firms and workers that are non-trivial to the overall impact of immigration on the U.S. economy. Notably, it also helps assess whether the link between immigrant-use and productivity is extant for both high- and low-educated labor.

Figure 8 presents the key results of this exercise, plotting usage of each program by productivity quintiles. As above, productivity quintiles are based on ranks of 2000 log revenues and 2000 log

[49] Firm size is a one-for-one correlate with productivity in the model presented above.

[50] These data are obtained from the FLC Data Center for 2002–2007 and the Department of Labor for 2008–2017.

[51] Note that I use the terms "foreign-born worker" and "immigrant" interchangeably. This differs from the official definition of an immigrant, which is a foreign-born individual with permanent legal residence or citizenship. H-1B and H-2B workers do not fit under this formal definition.

[52] See this link for a description of the H-1B visa process from the Society for Human Resource Management (SHRM). See this link for a vivid illustration of the complexity in the H-2B visa application process in Bier (2021).

[53] Calculations from the 2016–2018 public-use ACS (denominator) and the USCIS H-1B and H-2B Data Hubs.

revenues per worker national 6-digit NAICS code by age group bins. Usage is then measured by the percent of firms in a given productivity quintile that apply for at least one worker under a given visa program and year, averaged across the years 2002–2017. Appendix Figure D1 supplements Figure 8 by analyzing the number of visa applications as a fraction of a given firm's workforce.

Several pieces of validating evidence emerge from Figure 8. First and most importantly, there is a clear gradient between productivity and usage of each visa program. Across productivity measure and visa type, usage is highest in the top productivity quintile, and differences between the top quintile and other quintiles are always statistically significant at the 1% level. Importantly, this is true for *both* visa programs: while the H-1B is generally used by a wider swath of firms, the distribution of usage across productivity is strikingly similar between the H-1B and H-2B program. High productivity firms are more likely to access both low- and high-educated foreign-born workers through visas.

The noticeable non-linearity in the revenues-based productivity measure also validates the key precept of the model, which is that scale is a critical determinant of immigrant intensity. Figure 8 shows that the largest firms in a given industry-by-age-bin are far more likely to utilize each visa program, while removing scale from our productivity measure makes this relationship more linear. Figure D1 of the appendix further shows that, despite being larger, more productive firms file more applications *per worker*. All told, these findings provide compelling evidence that larger, more productive firms are much more likely to overcome hurdles in hiring foreign-born workers, consistent with the existence of fixed costs in such hiring.

4.3 Model Setup

4.3.1 Worker-Consumers

Individuals are consumer-employees of type $i \in \{I, N\}$, with I representing the foreign-born and N representing the native-born. The mass of each labor type in the economy is fixed and employees supply their labor inelastically—the primary comparative static will increase immigrant mass I. Consumers have constant elasticity of substitution (CES) preferences with elasticity of substitution μ across products produced by different firms ($\mu > 1$). I allow the standard taste for variety in these models to be optional using a parameter $\eta \in \{0, 1\}$. If $\eta = 1$, consumers have a taste for variety, which generates external scale effects. When $\eta = 0$, we shut down this channel from market size to welfare and focus on the productivity distribution (see, e.g., Egger and Kreickemeier, 2009).[54]

[54] See Appendix Section E.1 for more details on the utility function that generates this optional taste for variety.

4.3.2 Firms

The market structure is monopolistic competition, and each firm indexes a product.[55] An endogenous mass of potential entrepreneurs pay an entry cost to take a draw of total factor productivity z from a known Pareto distribution with shape parameter ϕ ($\phi > \mu - 1$) and minimum value m.

Once they draw their productivity, entrepreneurs decide whether or not to produce and which production technology to use. Firm production technologies are given by:

$$Q_j(z) = zL_j(z)$$
$$L_j(z) = \left(aI(z)^{\frac{\sigma_j - 1}{\sigma_j}} + N(z)^{\frac{\sigma_j - 1}{\sigma_j}}\right)^{\frac{\sigma_j}{\sigma_j - 1}}$$

where $Q_j(z)$ is total production of a firm whose owner draws productivity z and produces with technology type $j \in \{0, 1\}$. $L_j(z)$ is a CES aggregator of immigrant and native labor employed by the firm—the only two factors or production. σ_j is the elasticity of substitution between immigrant and native workers under production technology j ($\sigma_j > 1$).[56] Labor markets are perfectly competitive.

The first key assumption of the model is:

$$\sigma_1 < \sigma_0 \to \infty$$

Under the basic condition that $a < w_I$, where w_I is the relative immigrant wage, this assumption means that firms that utilize production technology $j = 0$ only hire native workers. This is an important assumption because it adds both tractability and realism to the model: native-only firms are an important feature of advanced economies. In both Germany and France, they comprise roughly 40% of all firms (Mitaritonna et al., 2017; Brinatti and Morales, 2021). $j = 1$ firms, meanwhile, hire immigrant and native workers and separate them into different tasks, as implied by $\sigma_1 < \infty$.

The cost function is given by

$$\left(\frac{c_j}{z}\right) Q_j(z) + (\kappa_f + \mathbf{1}\{j = 1\}\kappa_I)$$

where

$$c_j \equiv \left(a^{\sigma_j} {w_I}^{1-\sigma_j} + {w_N}^{1-\sigma_j}\right)^{\frac{1}{1-\sigma_j}}$$

[55] There is no distinction between a firm and an establishment in the model. Each firm/establishment represents an additional variety.

[56] Consensus estimates of this all-important parameter are generally $\sigma_1 \in [5, 20]$ (see, e.g., Ottaviano and Peri, 2012). The present model is more likely to be at the lower end of this range because the production function does not include additional nests for worker characteristics like education and experience. Thus, σ_1 can be thought of as a reduced form parameter that incorporates *all* the ways in which immigrant and native workers differ from each other on average.

and $w_N = 1$ is the normalized native wage. κ_f is a fixed operating cost that all firms face, and κ_I is the additional fixed cost borne by immigrant-hiring firms in order to access the better, labor-cost saving production technology ($\sigma_1 > 1 \Rightarrow c_1 < c_0$). Given CES consumer preferences, firms charge a constant mark-up over their marginal costs:

$$p_j(z) = \left(\frac{\mu}{\mu - 1}\right)\left(\frac{c_j}{z}\right)$$

4.4 Technology Choice in Equilibrium

Let $\pi_j(z)$ index profits for the firm with productivity z producing with technology type j. Following the logic of Bustos (2011), there exists a cutoff, z_1^* at which producers are indifferent between the two technologies:

$$\pi_0(z_1^*) \equiv \pi_1(z_1^*)$$

Entrepreneurs only stay in the market if they are profitable. This defines the usual cutoff productivity for type 0 firms:

$$\pi_0(z_0^*) \equiv 0$$

Figure 9 presents an illustration of the equilibrium technology choice in this model, featuring these two key cutoffs.[57] It plots profitability against productivity under both technologies. Bold lines indicate the production choices of operating firms. All firms would like to produce with the lower-variable-cost technology, $j = 1$, as illustrated by the dashed blue line, which removes the fixed immigrant recruiting cost κ_I from the profit function. However, the existence of these fixed costs preclude smaller, less productive firms from investing in the $j = 1$ technology. This is due to scale. Firms with higher productivity draws produce more output and require a larger workforce. These larger firms find it profitable to pay the augmented fixed cost in order to reduce their variable costs because this fixed cost is spread across more employees. The reductions in variable costs come from the imperfect substitutability across immigrant and native workers: the ability to separate immigrant and native workers into tasks for which they have a comparative advantage increases technical efficiency (Peri and Sparber, 2009). Meanwhile, firms with lower productivity draws do not produce at a scale that justifies paying a larger fixed cost, because variable costs are relatively less important to their cost structure.

At z_0^*, these smaller, less-productive firms are indifferent between operating and exiting the market because they make zero profits. While a lot of the intuition in the model revolves around z_1^*, it is only through its effect on z_0^* that z_1^* affects aggregate welfare. In other words, z_0^* is the

[57] The equilibrium depicted in Figure 9 requires the additional assumption that $\left(\frac{\kappa_I}{\kappa_f}\right) > \left((c_1)^{1-\mu} - 1\right)$, but this condition is always satisfied under plausible calibrations (see Appendix Sections E.2.1 and E.3).

key object of interest in this model. It defines the minimum of the realized employer productivity distribution; and in so doing, defines aggregate productivity. When z_0^* rises, the entire employer distribution is shifted to the right and marginal, least productive firms exit the market.

4.5 New Insights into the Economic Impact of Immigration

Closing the model requires four additional components, described further in Appendix Section E: a free entry condition, the aggregate price level (P), equating income and expenditure, and labor market clearing. With these fully specified, this model delivers substantive new insights into the general equilibrium impact of immigration. These revolve around the profitability cutoff, z_0^*, and its impact on the immigration surplus.

4.5.1 The Profitability Cutoff, z_0^*

Appendix Section E.2.2 shows that under the setup described above, z_0^* is given by

$$(z_0^*)^\phi = \underbrace{\theta}_{\text{Endogenous}} \underbrace{m^\phi \left(\frac{\kappa_f}{\kappa_e}\right)\left(\frac{\mu - 1}{\phi - (\mu - 1)}\right)}_{\text{Exogenous Parameters}} \tag{6}$$

where

$$\theta \equiv 1 + \left(\frac{z_1^*}{z_0^*}\right)^{-\phi}\left(\frac{\kappa_I}{\kappa_f}\right)$$

The profitability cutoff z_0^* is constant except for θ, an endogenous variable that sets this model apart from more standard models of firm heterogeneity (e.g., Melitz, 2003). It introduces the notion that entry and exit decisions for marginal, $j = 0$ firms depend on inframarginal, $j = 1$ firms, through their ability to steal away market share when their production costs go down. This notion generates the key proposition from this model, which finds a direct effect of immigration on total factor productivity:

Proposition 4.1. *If* $\frac{dw_I}{dI} < 0$, *then* $\frac{dz_0^*}{dI} > 0$

Proof. See Appendix Section E.2.4

$\frac{dw_I}{dI} < 0$ is perhaps the most consistent empirical regularity in the study of immigration (see, e.g., Ottaviano and Peri, 2012), and it holds under all plausible calibrations used here (see Appendix Section E.3). Under this basic condition, the dynamics in this model are as follows. Falling immigrant wages reduce per unit variable costs for $j = 1$ firms. These $j = 1$, immigrant-hiring firms then pass savings on to consumers by charging lower prices, as dictated by the pricing rule. $j = 0$, native-only firms now have to compete with these lower prices without having experienced reductions in their variable costs, since they do not hire immigrant workers. As a result, the productivity bar that they need to cross in order to remain profitable, z_0^*, rises.

Meanwhile, a rising z_0^* represents a rightward shift of the productivity distribution that delivers an increase in aggregate TFP, which is a weighted average from this distribution. That is, in this model, immigration increases aggregate TFP *through its effect on the employer distribution.* The partial equilibrium, empirical analog to a rising z_0^* can be seen in Section 3.2.1, in which establishments at the low end of the productivity distribution exit the market in response to immigration, consistent with a rising bar for zero-profitability. When z_0^* rises, entrants are also more productive, which is consistent with results from Section 3.2.2.

4.5.2 The Immigration Surplus

How much does the rise in z_0^* matter to our assessment of the economic impact of immigration? Appendix Section E.2.3 generates a simple expression for the "immigration surplus"—the effect of an increase in immigration on log real native incomes:

$$\frac{d\log\left({}^{w_N}/{}_P\right)}{dI} = \underbrace{\left(\frac{\eta}{\mu-1}\right)\frac{d\log(F)}{dI}}_{\text{Gains from Variety}} + \underbrace{\left(1+\frac{\phi}{\mu-1}\right)\frac{d\log(z_0^*)}{dI}}_{\text{Gains in Efficiency}} \tag{7}$$

where F is the mass of operating firms in the economy. Equation (7) expresses the immigration surplus *entirely* as a function of features of the employer distribution. Its mass imparts gains from variety, insofar as consumers value them (i.e., if $\eta = 1$). Meanwhile, changes to the employer distribution's support impart gains in efficiency that are the focus of this paper: a rightward shift of the productivity distribution that results from a rising z_0^*.

In Figure 10, I compare immigration surplus estimates generated from Equation (7) to those from a standard, representative firm model of production under a series of simulations. A striking picture emerges: accounting for heterogeneous effects immigration has on the composition of operating firms can more than double our estimates of the immigration surplus. These augmented, first-order gains comes through the aforementioned, direct effect of immigration on TFP that goes through the employer distribution; specifically, through z_0^*. In short, the welfare impact of immigration is substantially amplified when we account for the fact that it is higher productivity firms that engage in the task separation embodied by $\sigma_1 < \infty$. Details regarding these simulations are provided in Section E.3.

5 Conclusion

This paper uses several novel empirical results and a new model to reorient the literature on the economic impact of immigration around the employer distribution. Increased establishment entry and reduced establishment exit account for nearly all of immigrant-induced job creation over a 15 year period in which immigration was a defining feature of demographic change in the U.S. Contrary to what we would expect if these effects were driven solely by consumer demand or uniformly lower labor costs across firms, immigrant inflows cull establishments from lower productivity firms and draw in new establishments from higher productivity firms. Consequently, immigrant worker inflows increase the number of operating establishments in the right tail of the productivity distribution and increase productivity. Ties between immigrant workers and higher productivity firms imply a substantially larger "immigration surplus" than we would estimate with standard, representative firm models of the product market. In fact, responses by business entry and exit—and subsequent changes to firm composition—represent a plausible channel through which the immigration surplus is first-order.

Future work can corroborate and add nuance to this paper's overarching conclusion that employers and their heterogeneity drive immigration's benefits to the U.S. economy. Employer-employee linked data can test the implication that more productive firms hire more immigrant workers, beyond the visa programs studied in Section 4.2.1. Detailed data on prices can test the model's implication that the welfare impacts of immigration largely go through a negative correlation between firm productivity and output prices. Finally, results in Section 3.3.2 imply that accounting for labor market imperfections can add realism to our baseline models of immigration. Understanding whether or not these imperfections alter the conclusions of this paper is an important next step.

References

Abramitzky, R. and L. Boustan, *Streets of Gold: America's Untold Story of Immigrant Success*, PublicAffairs, 2022.

Adao, Rodrigo, Michal Kolesar, and Eduardo Morales, "Shift-Share Designs: Theory and Inference," *The Quarterly Journal of Economics*, 08 2019, *134* (4), 1949–2010.

Arellano-Bover, Jaime and Shmuel San, "The Role of Firms in the Assimilation of Immigrants," Technical Report 2020.

Autor, David H. and David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, August 2013, *103* (5), 1553–97.

__ , __ , and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, October 2013, *103* (6), 2121–68.

Ayromloo, Shalise, Benjamin Feigenberg, and Darren Lubotsky, "States Taking the Reins? Employment Verification Requirements and Local Labor Market Outcomes," NBER Working Papers 26676, National Bureau of Economic Research, Inc January 2020.

Azoulay, Pierre, Benjamin Jones, J. Daniel Kim, and Javier Miranda, "Immigration and Entrepreneurship in the United States," *American Economic Review: Insights*, September 2022, (27778).

Bartelsman, Eric J. and Mark Doms, "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, September 2000, *38* (3), 569–594.

Bartik, Timothy J., *Who Benefits from State and Local Economic Development Policies?*, Kalamazoo, Mich.: W.E. Upjohn Institute for Employment Research, 1991.

Beerli, Andreas, Jan Ruffner, Michael Siegenthaler, and Giovanni Peri, "The Abolition of Immigration Restrictions and the Performance of Firms and Workers: Evidence from Switzerland," *American Economic Review*, 2021.

Bier, David J., "H-2B Visas: The Complex Process for Nonagricultural Employers to Hire Guest Workers," Technical Report, Cato Institute 2021.

Borjas, George J., "The economic analysis of immigration," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, May 1999, chapter 28, pp. 1697–1760.

__ , "The labor supply of undocumented immigrants," *Labour Economics*, 2017, *46*, 1–13.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "SSAGGREGATE: Stata module to create shock-level aggregates for shift-share IV," Statistical Software Components, Boston College Department of Economics September 2018.

__ , __ , and __ , "Quasi-Experimental Shift-Share Research Designs," *The Review of Economic Studies*, 06 2021, *89* (1), 181–213.

Brinatti, Agostina and Nicolas Morales, "Firm Heterogeneity and the Impact of Immigration: Evidence from German Establishments," 2021. Working Paper.

Burchardi, Konrad B, Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J Terry, "Immigration, Innovation, and Growth," Working Paper 27075, National Bureau of Economic Research May 2020.

Burstein, Ariel, Gordon Hanson, Lin Tian, and Jonathan Vogel, "Tradability and the Labor-Market Impact of Immigration: Theory and Evidence From the United States," *Econometrica*, 2020, *88* (3), 1071–1112.

Bustos, Paula, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, February 2011, *101* (1), 304–40.

Cadena, Brian C. and Brian K. Kovak, "Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession," *American Economic Journal: Applied Economics*, 2016.

Card, David, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, *19* (1), pp. 22–64.

Clemens, Michael A., Ethan G. Lewis, and Hannah M. Postel, "Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion," *American Economic Review*, June 2018, *108* (6), 1468–87.

Cortes, Patricia, "The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data," *Journal of Political Economy*, 2008, *116* (3), 381–422.

di Giovanni, Julian, Andrei A. Levchenko, and Francesc Ortega, "A Global View of Cross-Border Migration," *Journal of the European Economic Association*, 2014, *13* (1), 168–202.

Dustmann, Christian and Albrecht Glitz, "How Do Industries and Firms Respond to Changes in Local Labor Supply?," *Journal of Labor Economics*, 2015, *33* (3), 711–750.

Egger, Hartmut and Udo Kreickemeier, "Firm Heterogeneity and the Labor Market Effects of Trade Liberalization," *International Economic Review*, 2009, *50* (1), 187–216.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, March 2008, *98* (1), 394–425.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik Instruments: What, When, Why, and How," *American Economic Review*, August 2020, *110* (8), 2586–2624.

Gonzalez, Libertad and Francesc Ortega, "How do very open economies adjust to large immigration flows? Evidence from Spanish regions," *Labour Economics*, January 2011, *18* (1), 57–70.

Haltiwanger, John, Ron Jarmin, Robert Kulick, Javier Miranda, Veronika Penciakova, and Cristina Tello-Trillo, "Firm-level Revenue Dataset," CES Technical Notes Series 19-02, Center for Economic Studies, U.S. Census Bureau August 2019.

__ , Ron S. Jarmin, and Javier Miranda, "Who Creates Jobs? Small versus Large versus Young," *The Review of Economics and Statistics*, 2013, *95* (2), 347–361.

Hong, Gihoon and John McLaren, "Are Immigrants a Shot in the Arm for the Local Economy?," Working Paper 21123, National Bureau of Economic Research April 2015.

Hopenhayn, Hugo, Julian Neira, and Rish Singhania, "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," Working Paper 25382, National Bureau of Economic Research December 2018.

Jaeger, David A, Joakim Ruist, and Jan Stuhler, "Shift-Share Instruments and the Impact of Immigration," Working Paper 24285, National Bureau of Economic Research February 2018.

Karahan, Fatih, Benjamin Pugsley, and Aysegul Sahin, "Demographic Origins of the Startup Deficit," Working Paper 25874, National Bureau of Economic Research May 2019.

Kerr, Sari Pekkala and William R. Kerr, "Immigrant Entrepreneurship," Working Paper 22385, National Bureau of Economic Research July 2016.

__ and __ , "Immigrant Entrepreneurship in America: Evidence from the Survey of Business Owners 2007 & 2012," Working Paper 24494, National Bureau of Economic Research April 2018.

Khanna, Gaurav and Munseob Lee, "High-Skill Immigration, Innovation, and Creative Destruction," Working Paper 24824, National Bureau of Economic Research July 2018.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who Profits from Patents? Rent-Sharing at Innovative Firms*," *The Quarterly Journal of Economics*, 03 2019, *134* (3), 1343–1404.

Lewis, Ethan G., "Immigration, skill mix, and the choice of technique," Working Papers 05-8, Federal Reserve Bank of Philadelphia 2005.

— , "Immigrant-Native Substitutability: The Role of Language Ability," Working Paper 17609, National Bureau of Economic Research November 2011.

— , "Immigration and Production Technology," Working Paper 18310, National Bureau of Economic Research August 2012.

Llull, Joan, "The effect of immigration on wages: exploiting exogenous variation at the national level," *Journal of Human Resources*, 2017, pp. 0315–7032R2.

Mahajan, Parag, "Immigration and Local Business Dynamics: Evidence from U.S. Firms," CES Working Paper 21-8, U.S. Census Bureau 2021.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Mitaritonna, Cristina, Gianluca Orefice, and Giovanni Peri, "Immigrants and firms' outcomes: Evidence from France," *European Economic Review*, 2017, *96*, 62 – 82.

Monras, Joan, "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis," *Journal of Political Economy*, 2020, *128* (8), 3017–3089.

— , "Local Adjustment to Immigrant-Driven Labor Supply Shocks," *Journal of Human Capital*, 2021, *15* (1), 204–235.

Olney, William W., "Immigration and Firm Expansion," *Journal of Regional Science*, 2013, *53* (1), 142–157.

Orefice, Gianluca and Giovanni Peri, "Immigration and Worker-Firm Matching," Working Paper 26860, National Bureau of Economic Research March 2020.

Orrenius, Pia M., Madeline Zavodny, and Alexander Abraham, "The Effect of Immigration on Business Dynamics and Employment," IZA Discussion Papers 13014, Institute of Labor Economics (IZA) February 2020.

Ottaviano, Gianmarco I. P. and Giovanni Peri, "Rethinking the Effect of Immigration on Wages," *Journal of the European Economic Association*, 2012, *10* (1), 152–197.

Peri, Giovanni, "The Effect Of Immigration On Productivity: Evidence From U.S. States," *The Review of Economics and Statistics*, 2012, *94* (1), 348–358.

— **and Chad Sparber**, "Task Specialization, Immigration, and Wages," *American Economic Journal: Applied Economics*, 2009, *1* (3).

— **and** — , "Assessing Inherent Model Bias: An Application to Native Displacement in Response to Immigration," Working Paper 16332, National Bureau of Economic Research September 2010.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "IPUMS USA: Version 9.0 [dataset]," 2019.

Schoar, Antoinette, "The Divide between Subsistence and Transformational Entrepreneurship," in "Innovation Policy and the Economy, Volume 10," University of Chicago Press, February 2010, pp. 57–81.

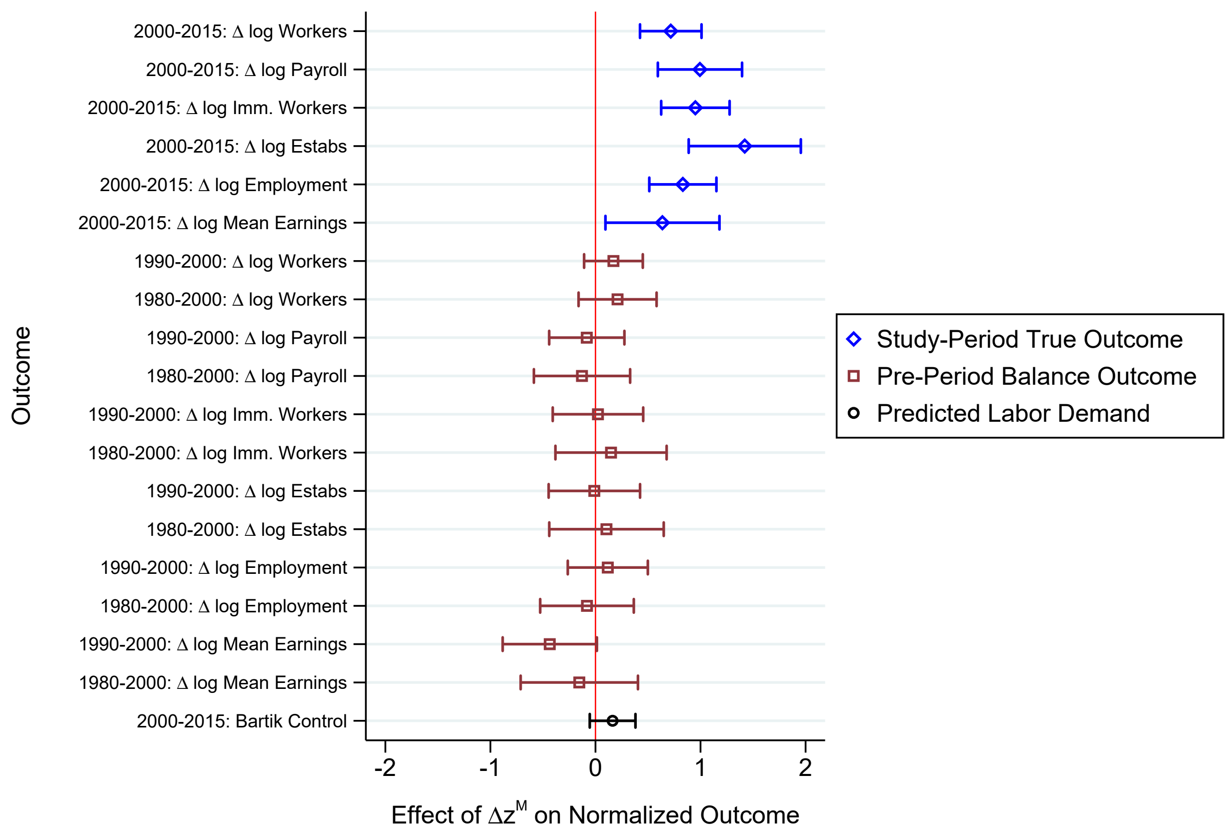
Sequeira, Sandra, Nathan Nunn, and Nancy Qian, "Immigrants and the Making of America," *The Review of Economic Studies*, 03 2019, *87* (1), 382–419.

Syverson, Chad, "What Determines Productivity?," *Journal of Economic Literature*, June 2011, *49* (2), 326–65.

Vespa, Jonathan, David M. Armstrong, and Lauren Medina, "PS 25-1144: Demographic Turning Points for the United States: Population Projections for 2020 to 2060," Technical Report, U.S. Census Bureau 2020.

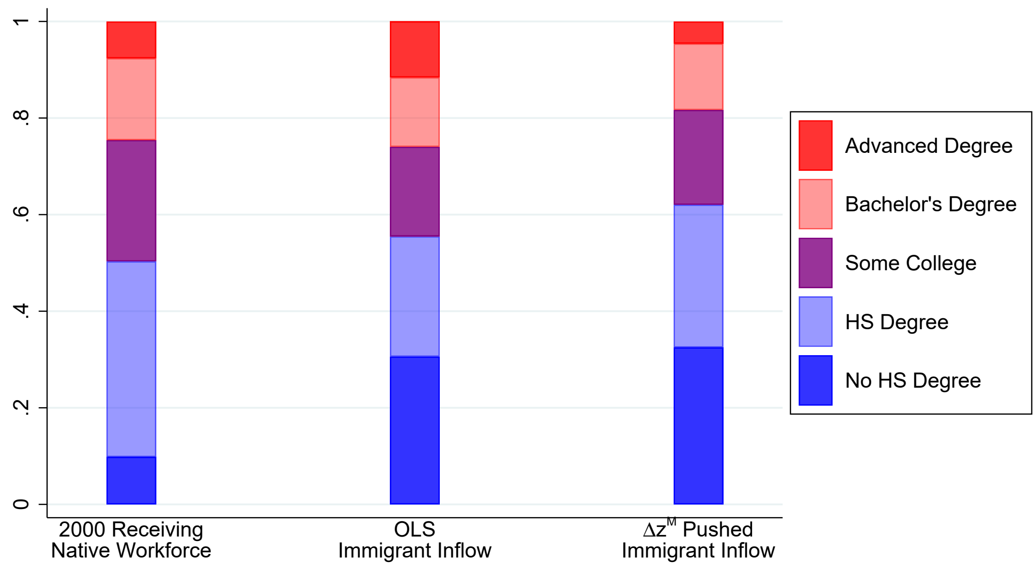
Figures and Tables

Figure 1: Assessing Conditional Exogeneity of Instrument



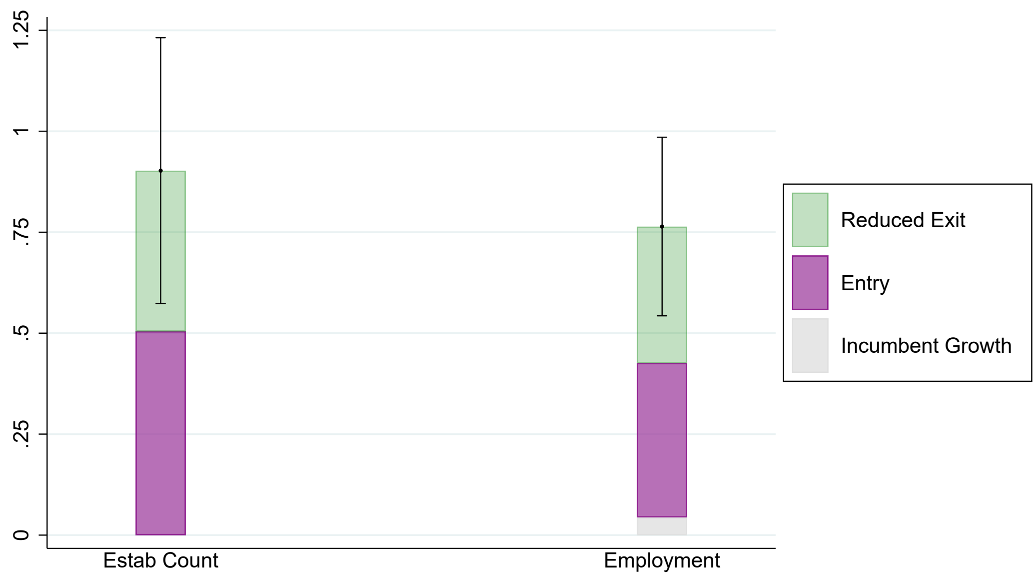
Notes: Each plotted coefficient $\hat{\phi}$ comes from Equation (2). All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. Capped spikes indicate 95% confidence intervals, derived from conventional standard errors that are robust to heteroskedasticity. All specifications include the set of control variables described in Section 2.2 except for the Bartik labor demand control.

Figure 2: Skill Composition of Immigrant Inflows and Receiving Native Workforce



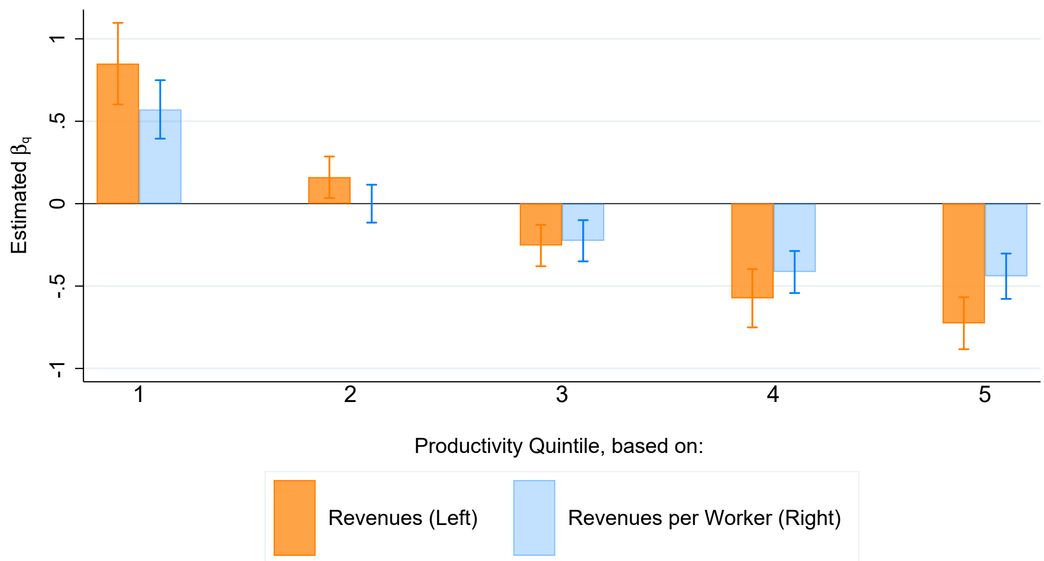
Notes: "2000 Receiving Workforce" (left bar) is constructed from IPUMS-USA data (Ruggles et al., 2019), where each component represents the proportion of native workers in 2000 in a given education grouping in 2000. Each component of the "Inflow" bars (middle and right) are obtained by estimating Equation (1) with ΔI^{e}_{gk} as the outcome, for the mutually exclusive and exhaustive education levels e shown in the legend. ΔI^{e}_{gk} are net inflows of immigrant workers with education level e into local industry gk between 2000 and 2015, divided by initial (2000) workforce size. The resulting estimates add up to one and illustrate the educational make-up of immigrant inflows into the U.S. based on whether Equation (1) is estimated using OLS (middle bar) or 2SLS with Δz^{M}_{gk} as the instrumental variable (right bar). All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2.

Figure 3: The Effect of Immigrant Worker Inflows on Establishment Counts and Employment



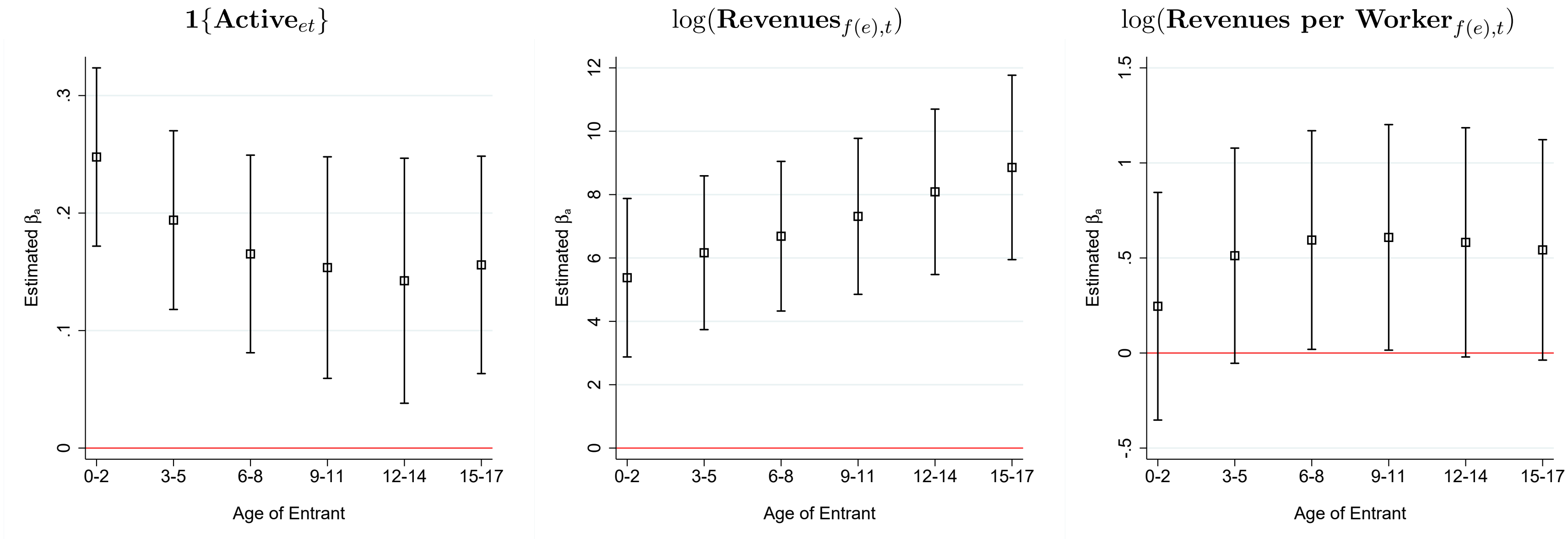
Notes: See Equation (1) for specification, estimated using 2SLS with instrumental variable Δz^M_{gk}. Each bar adds up to the estimated effect of a 1% increase in a local industry's workforce due to immigration on the growth rate in establishment count or employment in that local industry. These total effects are also plotted using black points, with capped spikes representing conventional, heteroskedasticity-robust 95% confidence intervals. See Table C2 for underlying coefficients and standard errors for each component of the establishment count decomposition and Table C3 for underlying coefficients and standard errors for each component of the employment decomposition. Tables C2 and C3 also compare results from the DHS growth rate outcome to the log change outcome. See Figure C1 for corresponding OLS results. Table C1 presents the totals effects as log changes and also computes corresponding Borusyak et al. (2021) standard errors. All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2.

Figure 4: Immigrant Workers and Establishment Exit, by Initial Productivity



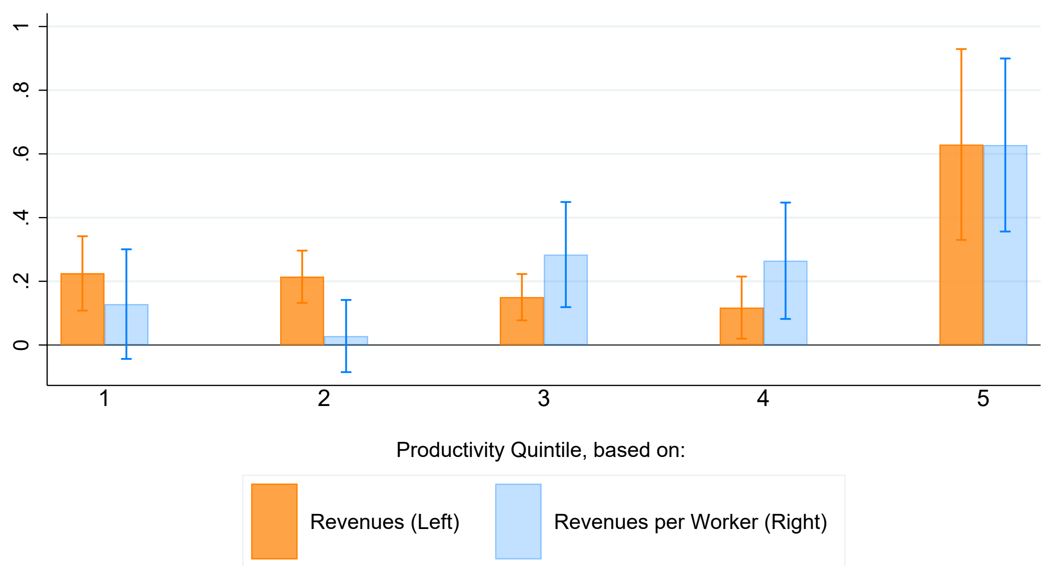
Notes: See Equation (4) for specification, estimated using 2SLS with instrumental variable z^M_{gkt}. Each coefficient $\hat{\beta}_q$ plotted as a bar represents the effect of a one percent increase in a local industry's workforce due to immigration on the probability that an establishment in the given quintile has zero payroll and employment. Equation (4) is estimated once using revenue-based quintiles (orange bars) and once using revenues-per-worker-based quintiles (blue bars). Each specification covers 4.7 million establishments that were operating (had positive payroll or employment) as of 2000, followed every five years until 2015. Establishments are split into productivity quintiles based on their parent firm's national rank in either log revenues (orange bars) or log revenues per worker (blue bars) within 6-digit NAICS codes and age bins in 2000. Capped spikes indicate conventional 95% confidence intervals, with standard errors clustered at the local industry level. The 1st Stage F statistic for the model estimated using revenues-based quintiles and the model estimated using revenues-per-worker are 25.65 and 26.02, respectively. See Figure C2 for analogous OLS results. See Figure C3 for analogous 2SLS results when firms are assigned productivity quintiles based on their rankings within the local industries that their establishments operate, rather than national 6-digit-NAICS-by-age bins. See Figure C4 for OLS results with this alternate ranking. All specifications include the control variables described in Section 2.2.

Figure 5: Immigrant Worker Inflows and Entrant Productivity Characteristics over their Life Cycle



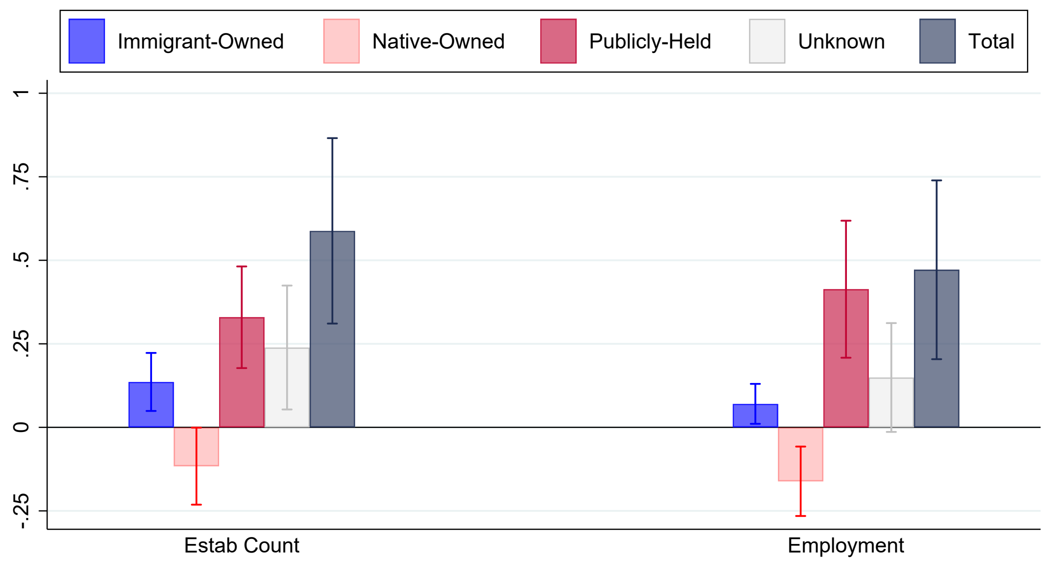
Notes: See Equation (5) for specification, estimated using 2SLS with instrumental variable $\sum_a [\Delta z_{g(e),k(e)} \times \mathbf{1}\{a(e,t) = a\}]$. Specification with outcome variable $\mathbf{1}\{\text{Active}_{et}\}$ is estimated on 9.4 million establishments and has a conventional First Stage F statistic of 14.21 (standard errors clustered at the local industry level). Specifications with revenue-based outcome variables are conditional on $\mathbf{1}\{\text{Active}_{et}\} = 1$, estimated on 6.7 million establishments, and have a conventional First Stage F statistic of 12.83. Conventional 95% confidence intervals shown in capped spikes. See C5 for corresponding OLS results. See C4 for average effects, not split by age. All specifications include the control variables described in Section 2.2.

Figure 6: Immigrant Worker Inflows and the Productivity Distribution



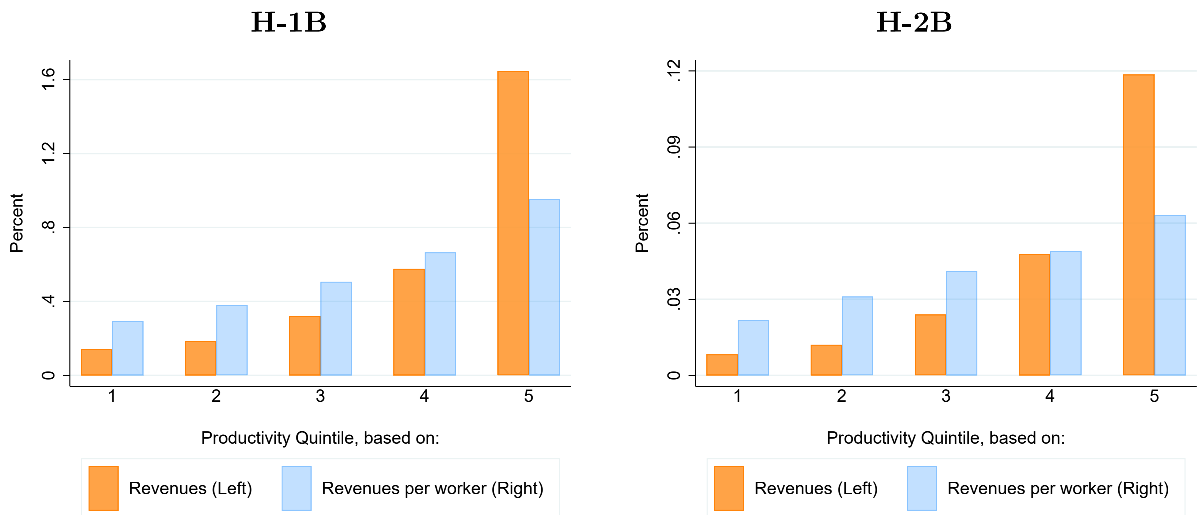
Notes: See Equation (1) for specification, estimated using 2SLS with instrumental variables Δz^{M}_{gk}. Each estimated coefficient (presented as a bar) plots the contribution of a given productivity quintile to the overall growth rate in establishment count in local industry gk. Establishments are assigned quintiles based on their parent firm's revenues (orange) or revenues per worker (blue). Qunitile cutoffs for (real) revenues and revenues per worker are determined by firm rankings within 6-digit NAICS code and age bins in 2000. Capped spikes indicate conventional 95% confidence intervals, with standard errors clustered at the local industry level. See Figure C6 for analogous OLS results. See Figure C7 for analogous 2SLS results when firms are assigned productivity quintiles based on their rankings within their local industries rather than national 6-digit-NAICS-by-age bins. See Figure C8 for OLS results with this alternate ranking. All specifications include the control variables described in Section 2.2.

Figure 7: The Role of Entry by Ownership Nativity and Type



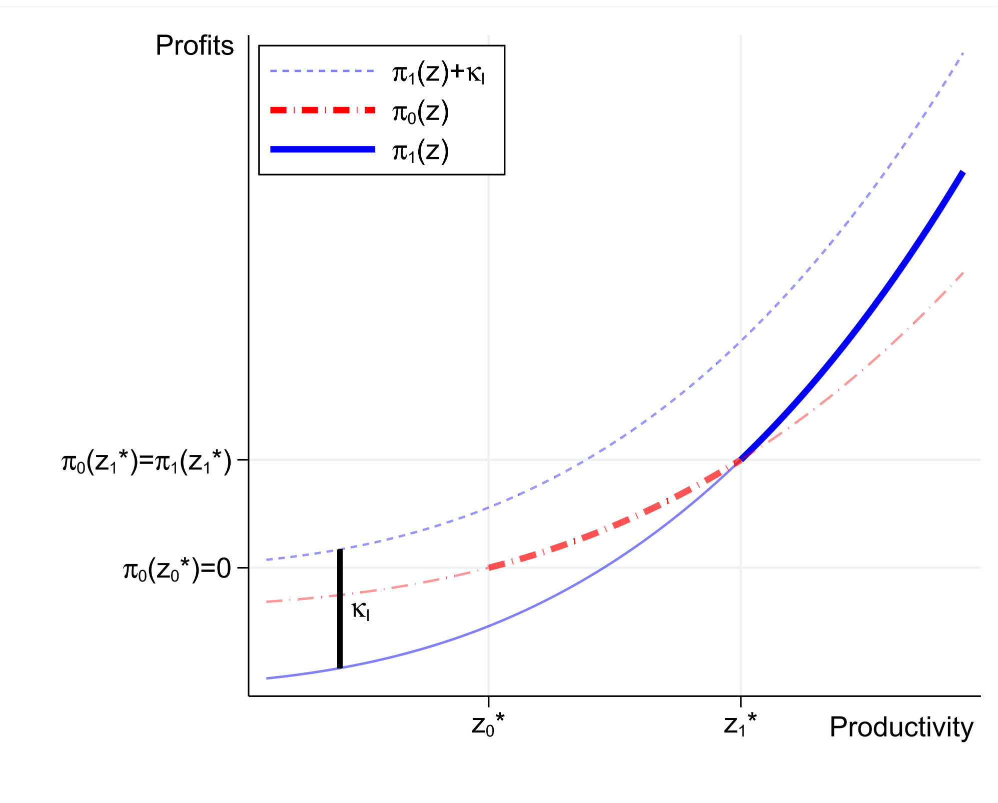
Notes: See Equation (1) for specification, estimated using 2SLS with instrumental variable Δz_{gk}^{M}. The rightmost, dark navy bars represent the contribution of entry to local industry growth rates in establishment counts and employment over the time period 2000–2017, as measured using the 2000 LBD (full-count) and 2018 ABS. The left four bars add up to this overall contribution of entry in each case. Each of these bars represent the role of a given firm ownership group (entrepreneur type) in generating the overall entry effect. The four groups are: privately-held, at least 50% immigrant-owned (immigrant-owned), privately held, less than 50% immigrant-owned (native-owned), publicly-held, and unknown. Conventional 95% confidence intervals shown in capped spikes. See Figure C9 for corresponding OLS results. All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2.

Figure 8: Usage of Employer Visa Programs by Productivity Quintile



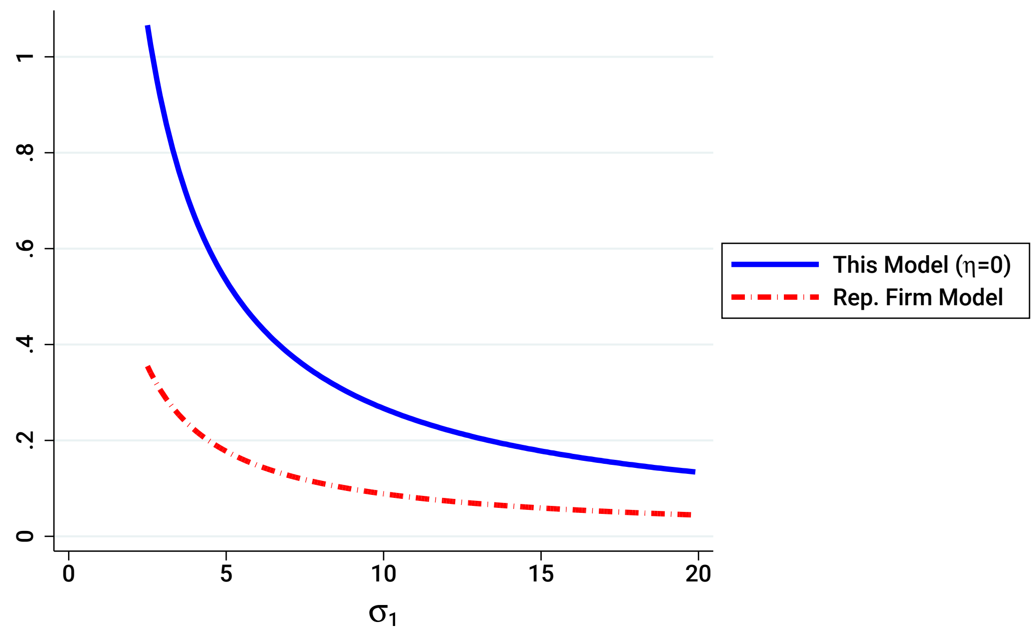
Notes: Each bar plots the percentage of firms in a given productivity quintile that applied for a given visa program in a given year, averaged across the time period 2002–2017. For example, more than 1.6 percent of firms in the top quintile of the revenue-based productivity distribution applied for an H-1B worker, on average (darker orange bar in 5th Productivity Quintile of left panel). Meanwhile, less than 0.2 percent of first in the bottom quintile of the revenue-based productivity distribution applied for an H-1B worker (darker orange bar in 1st Productivity Quintile of left panel). See Section 4.2.1 for more details on each visa program and D.1 for more details on the matching procedure used to link firms that use each visa program to measures of productivity. Within each visa program and productivity measure, the percent of firms using the program is smaller in the 1st, 2nd, 3rd, and 4th quintile compared to the 5th quintile—these differences are statistically significant at the 1% level in all cases. See Figure D1 for an analogous figure using applications per worker.

Figure 9: Technology Choice in Equilibrium



Notes: This figure plots the profit function for a firm given its entrepreneur's productivity draw and under each production technology $j \in \{0, 1\}$. κ_I is the fixed cost of recruiting immigrant workers. Thick lines indicate profits of firms that choose to operate in the market. The thick lines' colors and patterns indicate their technology choice, with the blue, solid line indicating immigrant-hiring firms and the red, dash-dotted line indicating native-only firms. The blue, short-dashed line indicates profits under $j = 1$ production if there were no fixed cost of recruiting immigrant workers. See Appendix Section E.2.1 for additional details.

Figure 10: Percent Increase in Native Incomes from a 1% Immigration Shock



Notes: This figure plots the simulated "immigration surplus" from the model described in Section 4 against that generated by a standard, representative firm model of labor demand. η is an indicator for whether consumers value variety. σ_1 is the elasticity of substitution across immigrant and native workers in production among firms that hire both. See Appendix Section E.3 for additional details on the simulations and additional simulation results.

Table 1: Immigrant Worker Inflows and Native Displacement

	Outcome:		
	Immigrant Worker Inflows (ΔI_{gk})	Native Worker Inflows	
	(1)	(2)	(3)
Emigrants Instrument (Δz^M_{gk})	0.517***		
	(0.053)		
	[0.051]		
Immigrant Worker Inflows (ΔI_{gk})		0.474***	0.036
		(0.050)	(0.138)
			[0.178]
Estimation	OLS: 1st Stage	OLS	2SLS

Notes: See Equation (1) for specification. Conventional standard errors—robust to heteroskedasticity—are presented in parentheses. Where applicable, Borusyak et al. (2021) standard errors—from equivalent origin o level regressions in which standard errors are clustered at the UN region level—are presented in square brackets. All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. Outcomes are both divided by initial local industry workforce. In Columns (2) and (3), coefficient estimates can be interpreted as the number of additional native workers per immigrant in a local industry over the time period 2000–2015, where $\beta = 0$ indicates no native displacement, on net. All specifications include the control variables described in Section 2.2.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Immigrant Workers and Establishment Exit

	Outcome:		
	Immigrant Workers (I_{gkt})	$\mathbf{1}\{\text{Inactive}_{et}\}$	
	(1)	(2)	(3)
Emigrants Instrument (z^M_{gkt})	0.460***		
	(0.043)		
Immigrant Workers (I_{gkt})		-0.031***	-0.209***
		(0.004)	(0.052)
Estimation	OLS: 1st Stage	OLS	2SLS

Notes: See Equation (3) for specification. Each specification estimated on 4.7 million establishments, all of which were active as of 2000. Every establishment is observed four times, in $t \in \{2000, 2005, 2010, 2015\}$. Immigrant Workers, I_{gkt} is divided by 2000 local industry workforce size. Standard errors clustered at the local industry level. All specifications include the control variables described in Section 2.2.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Immigrant Worker Inflows and Productivity Growth

	Outcome: $\Delta \log$ of		
	Mean Earnings	Revenues	Revenues per Worker
Immigrant Worker Inflows (ΔI_{gk})	0.367**	1.940***	1.147***
	(0.168)	(0.315)	(0.291)
	[0.125]	[0.165]	[0.191]
Outcome Data Source	LBD	SBO/ABS	SBO/ABS
Outcome Long Difference Span	2000–2015	2002–2017	2002–2017

Notes: See Equation (1) for specification, estimated using 2SLS with instrumental variable Δz^M_{gk}. Conventional standard errors—robust to heteroskedasticity—are presented in parentheses. Borusyak et al. (2021) standard errors—from equivalent origin o level regressions in which standard errors are clustered at the UN region level—are presented in square brackets. Table C1 provides corresponding OLS results. Outcomes with Data Source "LBD" are measured using the Longitudinal Business Database, which is a full-count panel of establishments with near-complete coverage of the U.S. private sector. Outcomes with Data Source "SBO/ABS" are measured using repeated cross-sections from the 2002 Survey of Business Owners and 2017 Annual Business Survey, which are each approximately 10% representative samples of U.S. firms. See Section 2.2 for additional details on data sources. All models are estimated using 722 CZ × 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (based on conventional standard errors).

Appendix

A Data

A.1 Business Data

The U.S. Census Bureau's Longitudinal Business Database (LBD) is constructed from administrative tax records for each U.S. non-farm, employee-hiring, private-sector establishment. In this paper, I use the newly revised version of the LBD (see Chow et al., 2021, for details).

Several important components of the revised LBD aid in my analysis. First, it comes with consistent, 2017-based NAICS codes over time, based on the methodology developed by Fort and Klimek (2018). This allows me to construct local-industry-level measures accurately going back to the 1980s based on the classifications generated in Table A1.[58] Second, I utilize the restriction `bds_tab==‘‘1’’` to restrict observation to the universe of establishments that were used in the construction of the Business Dynamics Statistics (BDS). This enables a consistent way of eliminating outlier observations and increases comparability with publicly-available data. For the exit analyses in Section 3.2.1, this means I restrict the sample to establishments that were active and with `bds_tab==‘‘1’’` in 2000. For the entrants analyses in Section 3.2.1, this means I restrict the sample to establishments that had `bds_tab==‘‘1’’` for at least one year of observation. Finally, the LBD contains unique parent firm identifiers that can be linked to firm-level revenue information from the BRFIRM_REV data set for approximately 75% of firms in 1997–2018.

As described in the main text, I also utilize the 2002 Survey of Business Owners (2002) and 2018 Annual Business Survey (ABS). Both of these data sets have establishment level revenues for a sample of establishments, along with weights that make the sample representative. The 2002 SBO covered roughly 33% of firms in the U.S. private sector in 2002, while the 2018 ABS also covered roughly 33% of firms in the U.S. private sector in 2017. The 2018 ABS also asks about birthplace of owners with the three highest ownership stakes in a given firm. If at least 50% of a given firm is owned by individuals not born in the U.S., it is coded as an immigrant-owned firm in 2017. If less than 50% is accounted for by foreign-born individuals, it is coded as a native-owned firm. For several firms, this question is unanswered, so I separately enumerate "unknown," privately-held firms. A limitation of the 2002 SBO from the perspective of this paper is that it did not ask about ownership nativity.

A.2 Demographic Data

Immigrant exposure variables along with several outcome and control variables are measured using restricted-access U.S. Census Bureau demographic data from the 1980, 1990, and 2000 Long-Form Decennial Censuses and the 2005 through 2019 American Community Surveys (ACS). The underlying sample of respondents that I use to construct these measures consists of employed workers (self-employed or employees) who can be assigned a country of origin, who reside in the contiguous United States (including Washington, D.C., but excluding Alaska and Hawaii), and who work in an industry group from Table A1. Immigrant workers are defined as those who indicated that they are either a naturalized citizen or a non-citizen. All other workers are defined as native. Population estimates are generated by summing over survey weights from the underlying sample.

[58] LBD coverage begins in 1976.

The ACS is a yearly survey that contains smaller underlying sample sizes than the Decennial Census Long Form. Given the level of detail in the unit of analysis—commuting zone by industry group—I average across ACS years to increase the underlying sample size. Specifically, for all years other than 2005, measures from the ACS are averaged over five years to eliminate noise and keep underlying sample sizes similar to measures from the Decennial Censuses. For example, estimates of immigrant presence in a local industry in 2010 is the average of immigrant presence in that local industry over the period 2008-2012. This cannot be done for 2005 because the 2001-2004 ACS were experimental products that were not representative at sub-state levels.

A.3 Emigration Data

I use the United Nations Population Division's (UNPD) International Migration Stock 2019 to construct $\Delta z^{\mathrm{M}}_{gk}$ and z^{M}_{gkt}, the instrumental variables for Section 3. These data contain total emigration stocks in origin-destination pairs for several counties, including the U.S. Stocks are mostly obtained from destination country microdata, and are sometimes imputed, as described here.

In order to focus on migration destinations that are similar to the U.S., I sum emigrants in non-U.S. destinations for 18 OECD member nations: Australia, Austria, Canada, Chile, Denmark, Finland, France, Germany, Greece, Ireland, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. Other than the U.S., Luxembourg is dropped from the destination list because it is combined with Belgium as an origin country in the data used to construct the control variables for trade exposure (described in Appendix Section A.6). All destination countries are dropped as potential origins. In particular, this removes many Europe-to-Europe moves that started to occur after the introduction of the Eurozone in 1999. Results are robust to dropping Canada as a destination country, prompted by concerns of destination substitution between the U.S. and Canada (available upon request).

Well over 100 origin countries are covered.[59] This covers the set of countries that are in the UNPD data, match to the WITS trade data, and match to a consistent geography in the Census data. These generally represent the largest migration origin countries, but some aggregations are made in order to account for changing boundaries over time during the study period and aggregations that come with the Census and UNPD data.

A.4 Industry Classifications and Summary Statistics

Because industry classifications differ both across Census years and between the Census and North American Industrial Classification System (NAICS) codes contained in the LBD, constructing the industry groups involved multiple steps. I first use the 1990 Decennial Census industry codes as a bridge between different Census industry classification systems, as is done in IPUMS-USA (Ruggles et al., 2019).[60] I then construct a crosswalk between the 1990 Decennial Census industry codes and 3-digit 2017 NAICS codes, which available for all years in the revised LBD. In some cases, the 1990 industry classification corresponds to more than one 3-digit NAICS code, and in some cases a 3-digit NAICS code corresponds to more than one 1990 industry classification. The industry groups I use therefore generally represent the smallest possible mutually-exclusive sets of industry classifications.

[59] The exact number and set of countries can be disclosed if necessary, upon request.

[60] Crosswalks provided by IPUMS-USA between the 1990 and other Census year classifications, as well as between the 1990 Census industry classifications and NAICS codes, were crucial to this process.

For example, 1990 Census Industry classification code 132 is "Knitting mills" and corresponds to NAICS Codes 315 and 313. However, NAICS code 313 also covers "Yarn, thread, and fabric mills," which is 1990 Census Industry classification code 142. Additionally, NAICS code 315 also includes manufacturing of "Apparel and accessories other than knitting." Manufacturing of apparel and accessories, knitting mills, and yarn, thread, and fabric mills are therefore all covered in the same industry grouping in my analysis.

Some additional aggregations are made to ensure that industry groups do not vary excessively in size. The Agriculture (NAICS 11) and Public (NAICS 92) sectors along with the Postal Service (NAICS 491), Fund, Trusts, and Private Households (NAICS 814) are dropped from the analysis due to relatively less reliable coverage in the LBD. The final set of industry groups can be seen in Table A1.

Note that starting in 2000, the Census Bureau began basing industry codes in its demographic data on NAICS codes. Prior to 2000, the Census Bureau had been basing these codes on the Standard Industrial Classification (SIC) system. Constructing a uniform set of industry codes across the Census and LBD and across different Census waves is thus particularly important for pre-trend testing conducted in Figure 1.

Table A1: Industry Groups

Industry Group	1990 Census Codes	2017 NAICS Codes	Worker Educ. Designation	Tradability Designation
Mining	40, 41, 42, 50	21	High-School Equivalent	Tradable
Construction	60	23	High-School Equivalent	Non-Tradable
Management of companies	710	55, 523, 525	College Equivalent	Tradable
Utilities	422, 450, 451, 470–472	22, 486, 562	High-School Equivalent	Non-Tradable
Manufacturing – Food	100, 101, 102, 110 , 111, 112, 120, 121, 122, 130, 610	311–312	High-School Equivalent	Tradable
Manufacturing – Clothing	132, 140, 142, 150, 151, 152, 220, 221, 222	313–316	High-School Equivalent	Tradable
Manufacturing – Wood & Furniture +	160–162, 231, 232, 241, 242, 250–252, 261, 262	321, 322, 327, 337	High-School Equivalent	Tradable
Manufacturing – Plastics +	180–182, 190–192, 200, 201, 210–212	324–326	College Equivalent	Tradable
Manufacturing – Metals & Machinery	270–272, 280–282, 290–292, 300, 301, 310–312, 320, 321, 331, 332, 380	331–333	High-School Equivalent	Tradable
Manufacturing – Electrical & Household	322, 340–342, 350, 371, 372, 381, 390, 391	334, 335, 339	College Equivalent	Tradable
Manufacturing – Transportation	351, 352, 360–362, 370	336	High-School Equivalent	Tradable
Printing & Publishing	171, 172	323, 511	College-Equivalent	Tradable
Wholesale Trade – Durable	500, 501, 502, 510–512, 521, 530–532	423	College Equivalent	Tradable
Wholesale Trade – Nondurable	540–542, 550–552, 560–562	424, 425	High-School Equivalent	Tradable
Retail Trade – Vehicles	612, 620, 622	441	High-School Equivalent	Non-Tradable
Retail Trade – Household Durables	580–582, 631–633	442–444	College Equivalent	Non-Tradable
Retail Trade – Food & Gas	601, 602, 611, 621, 650	445, 447	High-School Equivalent	Non-Tradable
Retail Trade – Misc.	590, 640, 642, 651, 652, 661, 662, 681, 682	446, 451, 453	College Equivalent	Non-Tradable
Retail Trade – Apparel	623, 630, 660	448	High-School Equivalent	Non-Tradable
Retail Trade – Dept. & Variety Stores	591, 592, 600	452	High-School Equivalent	Non-Tradable
Retail Trade – Fuel, Catalog, Vending	663, 670–672	454	High-School Equivalent	Non-Tradable
Misc. Transportation	400, 401, 402, 420, 421	481–483, 485	High-School Equivalent	Tradable
Trucking	410	484, 492	High-School Equivalent	Non-Tradable
Warehousing & Storage	411	493	High-School Equivalent	Tradable
Non-Telephone Communication	440, 852	515, 519	College Equivalent	Non-Tradable
Telecomm & Data Processing	441, 442, 732	517, 518, 533	College Equivalent	Non-Tradable
Savings Institutions	700–702	521, 522	College Equivalent	Non-Tradable
Insurance	711	524	College Equivalent	Tradable
Real Estate	712	531	College Equivalent	Non-Tradable
Professional Services	12, 721, 741, 841, 882, 890–893	541, 711	College Equivalent	Tradable
Admin. & Support Services	20, 432, 722, 731, 740	487, 488, 561	College Equivalent	Non-Tradable
Health Services excl. Hospitals	812, 820–822, 830, 840	621	College Equivalent	Non-Tradable
Hospitals	831	622	College Equivalent	Non-Tradable
Nursing & Residential Care Facilities	832, 870	623	High-School Equivalent	Non-Tradable
Social Services	861–863	624	College Equivalent	Non-Tradable
Entertainment Services	742, 800–802, 810, 872	512, 532, 712, 713	College Equivalent	Non-Tradable
Lodging	762, 770	721	High-School Equivalent	Tradable
Eating & Drinking Places	641	722	High-School Equivalent	Non-Tradable
Repair Services	750–752, 760, 782, 790	811	High-School Equivalent	Non-Tradable
Personal Services	771, 772, 780, 781, 791	812	High-School Equivalent	Non-Tradable
Unions & Religious Organizations	873, 880, 881	813	College Equivalent	Non-Tradable

Table A2: Summary Statistics (Publicly-Available Data)

Industry Group	2000 Workforce Mean	2000–2015 ΔI_{gkt}: Immigrant Inflows per Initial Worker Mean	Std Dev	2000–2015 Δ log Employment Mean	Std Dev
Admin. & Support Services	64,297	0.20	0.13	-0.03	0.27
Construction	82,964	0.12	0.11	-0.19	0.21
Eating & Drinking Places	67,257	0.14	0.12	0.31	0.16
Entertainment Services	42,171	0.06	0.06	0.11	0.33
Health Services excl. Hospitals	57,308	0.13	0.12	0.41	0.18
Hospitals	61,998	0.09	0.08	0.13	0.21
Insurance	33,377	0.04	0.04	0.01	0.24
Lodging	18,232	0.12	0.12	0.16	0.31
Management of Companies	48,474	0.04	0.06	0.09	0.30
Manufacturing, Clothing	28,351	-0.12	0.17	-1.33	0.63
Manufacturing, Electrical & Household	59,469	-0.01	0.06	-0.58	0.45
Manufacturing, Food	12,263	0.11	0.12	0.00	0.41
Manufacturing, Metals & Machinery	35,745	0.01	0.06	-0.29	0.31
Manufacturing, Plastics +	28,051	0.04	0.07	-0.12	0.36
Manufacturing, Transportation	63,382	0.03	0.06	-0.52	0.63
Manufacturing, Wood & Furniture +	14,764	0.00	0.05	-0.49	0.29
Mining	7,299	0.13	0.20	0.56	0.79
Non-Telephone Communication	13,879	0.05	0.12	-0.49	0.60
Nursing & Residential Care Facilities	17,125	0.12	0.14	0.28	0.19
Personal Services	28,164	0.16	0.11	0.06	0.22
Printing & Publishing	27,855	-0.02	0.05	-0.22	0.33
Professional Services	150,852	0.06	0.04	0.23	0.20
Real Estate	33,757	0.08	0.07	0.17	0.22
Repair Services	21,943	0.06	0.08	-0.07	0.20
Retail Trade – Household Durables	28,895	0.01	0.04	-0.07	0.16
Retail Trade, Apparel	17,939	0.08	0.09	0.20	0.25
Retail Trade, Dept. & Variety Stores	19,760	0.06	0.08	0.13	0.17
Retail Trade, Food & Gas	31,410	0.07	0.06	0.02	0.18
Retail Trade, Fuel, Catalog, Vending	6,858	0.09	0.13	0.04	0.52
Retail Trade, Misc.	33,288	0.05	0.05	0.03	0.15
Retail Trade, Vehicles	14,985	0.05	0.05	0.02	0.14
Savings Institutions	40,740	0.05	0.05	-0.02	0.24
Social Services	17,834	0.05	0.06	0.43	0.26
Telecomm & Data Processing	56,454	0.15	0.12	-0.16	0.32
Transportation	38,121	0.11	0.09	-0.07	0.43
Trucking	19,472	0.11	0.12	0.01	0.33
Unions & Religious Organizations	23,732	0.03	0.04	0.06	0.24
Utilities	12,199	0.06	0.07	0.08	0.28
Warehousing & Storage	3,637	0.31	0.39	1.77	0.87
Wholesale Trade, Durable	34,370	0.00	0.04	1.02	0.21
Wholesale Trade, Nondurable	24,394	0.05	0.07	-0.66	0.24
Mean Across Industries	50,297	0.08	0.11	0.01	0.45

Notes: Data obtained from IPUMS-USA (Ruggles et al., 2019) and Eckert et al. (2020) version of County Business Patterns. All statistics weighted by 2000 workforce size of local industry.

A.5 Proxying for Productivity

A combination of the modeling framework in Section 4 and previous literature—particularly, Foster et al. (2008)—motivate the use of revenues as my primary proxy for productivity, the adjustments I make to align it more closely with total factor, "physical" productivity (TFP), and the use of revenues per worker as an alternative measure in every set of results. First, Foster et al. (2008) find that revenue-per-input measures of productivity correlate strongly with measures of physical productivity in industries where these concepts can be separated cleanly. However, in many simplified models with firm heterogeneity and competitive labor markets—like that presented in Section 4—revenues per worker are independent of TFP at the firm level, while revenues and TFP move one-for-one. Thus, I always present results with both revenues-per-worker-based and revenues-based productivity proxies.

Second, the measures of revenue-based total factor productivity in Foster et al. (2008) account for capital, materials, and energy inputs along with labor. My primary measures of productivity thus rank firms within detailed industries—where input requirements are more likely to be similar—to mute differences in non-labor input use from driving differences in my productivity measure. Ranking within detailed industry also removes cross-industry market-power differentials from driving revenue per worker differences. Finally, Foster et al. (2008) also find that divergence between revenue-based and physical measures of productivity—even within detailed industries—often occurs because of demand shocks that reflect market foothold. For example, older firms that produce the same product as younger firms may generate more demand because of non-quality-related factors like name recognition. This motivates ranking firms within age groups in my primary measures of productivity.

A summary of the construction of my productivity ranking-based measures is:

1. Rank the set firms operating in 2000 within national 6-digit-NAICS-code-by-age bins according to log real revenues measured in the BRFIRM_REV data set.
2. Construct (4) quintile cutoffs within each bin, weighting firms by the inverse probability weights of inclusion in the revenue data sets in the construction of these cutoffs.
3. For Section 3.2.1, assign the set of establishments operating as of 2000 whose parent firms have revenue information in 2000 a productivity quintile based on their log real revenues in 2000 the cutoffs from 2.
4. For Section 3.3.1, assign the set of establishments operating as of 2000 whose parent firms have revenue information in 2000 a productivity quintile based on their 2000 log real revenues and the cutoffs from 2. Assign the set of establishments operating as of 2015 whose parent firms have revenues information in 2015 a productivity quintile based on their 2015 log real revenues and the cutoffs from 2.
5. Repeat Steps 1.-4. based on log real revenues per worker instead of log real revenues.
6. Repeat Steps 1-5. based on rankings within local industries (commuting-zone-by-industry bins) instead of national 6-digit-NAICS-code-by-age bins for robustness checks. See Figures C3, C3, C7, and C8 for the relevant results under these alternate rankings.

While the ability to conduct these detailed rankings is unique to my data, a large literature starting with Klette and Griliches (1996) and including Foster et al. (2008) imply that this measure still has clear limitations as a proxy for true physical productivity, TFP. Because I do not observe firm-level prices or non-labor inputs, I ultimately cannot directly eliminate the influence of these factors in my measure of productivity.[61] To view results as reflective of differences in physical productivity across firms, the key assumption can be stated as: ranking within 6-digit NAICS code and age groups removes enough influence from idiosyncratic demand and non-labor input use across

[61] Indeed, Foster et al. (2008) find that these factors are important even within very detailed industries.

firms such that remaining variation in revenues per worker primarily reflects differences in physical productivity. The plausibility of this assumption is strengthened by robustness of results to using alternate measures that also correlate with physical productivity.

As additional motivation, I also note the prima facie interest understanding how immigrant worker inflows affect firms along the revenue per worker and revenue distributions. At best, these also measure the impact of immigrant worker inflows on the total factor productivity distribution. At worst, they measure important features of the firms and establishments that operate in local economies.

A.6 Construction of Control Variables

A.6.1 Predetermined Local Industry Controls

Unless otherwise specified, all models include controls for "start-of-period" (2000) college share, and immigrant share—the proportion of employed workers that have a college degree or are foreign-born in local industry gk in 2000, respectively. They also include the basic market size control $\text{Workers}_{gk,2000}$. Finally, all models include a control for the proportion of the national workforce in industry k that is accounted for by commuting zone g in 2000:

$$\text{Ind Share}_{gk} = \frac{\text{Workers}_{gk,2000}}{\sum_g \text{Workers}_{gk,2000}}$$

All of these variables are measured from the 2000 Decennial Census Long Form. When estimating Equations (3) and (4), these variables are interacted with time fixed effects.

A.6.2 Bartik Labor Demand Control

The structure of the control variable for labor demand mimics the instrumental variable for labor demand proposed by Bartik (1991). It is included because this paper seeks to isolate labor demand *responses* to labor supply shocks from immigration as opposed to labor demand shocks that are generated by the same factor that induces the immigration labor supply shocks. The control variable takes advantage of the fact that the LBD data contains consistent 6-digit NAICS codes over time (see Section A and Fort and Klimek, 2018), whereas my industry groupings are aggregations of 3-digit NAICS codes. Specifically, letting k' denote a 6-digit NAICS code and k denoting an industry group as usual,

$$\Delta\text{Bartik}_{gk} = \sum_{k' \in k} \left[\left(\frac{\text{Emp}_{gk',2000}}{\text{Emp}_{gk,2000}} \right) \times \Delta \log(\text{Emp}_{k'}) \right]$$

where $\Delta \log(\text{Emp}_{k'})$ is LBD-measured national employment growth in 6-digit NAICS code k'. That is, the growth rate in national employment in 6-digit NAICS code between 2000 and 2015 is projected into local industries based on the proportion of that local industry's workforce accounted for by that 6-digit NAICS industry in 2000.

ΔBartik_{gk} is used in the estimation of Equations (1), (5), and (8). When estimating Equations

(3) and (4), I use the levels version:

$$\text{Bartik}_{gkt} = \sum_{k' \in k} \left[\left(\frac{\text{Emp}_{gk',2000}}{\text{Emp}_{gk,2000}} \right) \times \log(\text{Emp}_{k't}) \right]$$

for $t \in \{2000, 2005, 2010, 2015\}$.

A.6.3 Shift-Share Controls

I also control for origin-country shocks that may correlate with out-migratory pressure and therefore confound the relationship between the instrument and local industry outcomes in areas and industries with large shares from particular origin countries. As advocated by Borusyak et al. (2021), I construct these controls using the same shares as my instrumental variable.

The most important of these control variables account for exposure to trade. Immigrant worker ties to origin countries can create international trade linkages (see, e.g., Parsons and Vézina, 2018). Local industries heavily exposed to workers from a given origin country may therefore experience both increased immigration and reduced trade due to the same (economic) shock in that origin country. To account for this, I utilize data from the World Bank's World Integrated Trade Solution, which contain data on real trade flows from sending countries to and from the U.S. I use these data to construct shift-share control variables for trade exposure:

$$\Delta\text{Trade Exposure}_{gk}^{[\text{Flow}]} = \sum_{o} \pi_{og}\rho_{ok} \times \text{Prop. Traded}_{k,2000} \times \Delta \log([\text{Flow}]_o)$$

where $[\text{Flow}] \in \{\text{Imports}, \text{Exports}\}$ measured in real USD, and $\Delta \log([\text{Flow}]_o)$ represents the growth rate over the period 2000–2015. Prop. Traded_{kt} is the proportion of industry k's workforce in 2000 that is employed in a "Traded" 6-digit NAICS code industry according to the Porter classification system (Porter, 2003). $\Delta\text{Trade Exposure}_{gk}^{[\text{Flow}]}$ is used in the estimation of Equations (1), (5), and (8). When estimating Equations (3) and (4), I use the levels version:

$$\text{Trade Exposure}_{gkt}^{[\text{Flow}]} = \sum_{o} \pi_{og}\rho_{ok} \times \text{Prop. Traded}_{k,2000} \times \log([\text{Flow}]_{ot})$$

for $t \in \{2000, 2005, 2010, 2015\}$.

In order to isolate emigration shocks from general population growth, I also include the following shift-share control:

$$\text{Population Exposure}_{gk} = \sum_{o} \pi_{og}\rho_{ok} \times \text{Population}_{o,2000}$$

This variable is interacted with year fixed effects when estimating Equations (3) and (4).

Finally, as proposed by Borusyak et al. (2021), I include the sum of shares as as control variable in order to isolate shifts in my identifying variation. In "shift-share" form, the sum of shares is $\sum_0 \text{Immigrants}_{og,2000} \times \text{Immigrants}_{ok,2000}$ since the denominators of π_{og} and ρ_{ok} are not based on either g or k and are therefore technically part of the shift. In the main text, I describe the instrument using π_{og} and ρ_{ok} to retain consistency with the vast tradition in immigration economics that uses similar shares to apportioning shifts into local areas, occupations, and industries.

A.6.4 LBD-CEN Measurement Difference Control

Most of the analyses in this paper rely on outcomes measured at the gk level from the LBD (because they cannot be measured in the demographic data) and an immigrant exposure variable measured from Census Bureau demographic data (because they cannot be measured in the LBD). A key precept of these analyses is that these two data sources can each be used to measure outcomes at the gk level, and the creation of the industry groups described in Section A.4 paid particular attention to this concern.

Nonetheless, in order to alleviate any remnant concerns, reduce noise, and align the two data sources completely, I include an additional control variable based on measurement of a variable that can be measured in both data sources: employment growth. Let CEN Emp_{gkt} be the count of workers who are classified as private-sector, wage-earning employees in local industry gk at time t.[62] I then construct as a control variable

$$\text{Meas Control}_{gk} = \frac{\text{CEN Emp}_{gk,2015} - \text{CEN Emp}_{gk,2000}}{\text{Denom}_{gk}^{\text{CEN Emp}}} - \frac{\text{LBD Emp}_{gk,2015} - \text{LBD Emp}_{gk,2000}}{\text{Denom}_{gk}^{\text{LBD Emp}}}.$$

That is, Meas Control_{gk} is the difference in the measured employment growth rates across the two data sources when they are most closely aligned in terms of their underlying composition. Reassuringly, the instrument is not statistically significantly correlated with this measure, and it primarily serves to tighten standard errors.

[62] Recall that for $t = 2000$, this variable will be measured from the 2000 Decennial Census Long Form, and for $t = 2015$, it represents an average across the 2013–2017 ACS. I can isolate private-sector, wage-earning employees thanks to the Census Bureau demographic data's class of worker variables.

B Additional Instrument Vetting Using Publicly-Available Data

In this section, I conduct further vetting of the instrumental variable Δz^M_{gk}. In order to avoid excessive disclosure avoidance review of the confidential data, I use publicly-available data sources. Most notably, I use IPUMS-USA data to construct a version of Δz^M_{gk} and ΔI_{gk} (Ruggles et al., 2019).

B.1 Country-Level Pushes

To begin, I show that shocks to non-U.S., OECD emigration predicts immigration to the U.S. at the origin country level. This check ensures that the shifts that I rely on for identification are meaningful and not just noise being distributed into local industries by shares π_{og} and ρ_{ok}. Specifically, I estimate the following regression models:

$$\log(I_{ot}) = \tau \log(M_{o,-US,t}) + \alpha_o + \alpha_t + \Gamma X_{ot} + \varepsilon_{ot}$$

where I_{ot} is the stock of immigrants from origin o in the U.S. at time $t \in \{2000, 2005, 2010, 2015\}$ and $M_{o,-US,t}$ is the stock of emigrants from origin o living in one of 18 non-U.S. OECD countries at time t.

Table B1 presents the results of this exercise confirms that $\log(M_{o,-US,t})$ have strong predictive power on $\log(I_{ot})$, regardless of an increasingly strict set of controls. At least some portion of the first-stage strength found in the main text is coming from the shift component of the instrument, which is more plausibly exogenous. I also note that the strength of these results very likely understates the relationship between $\log(M_{o,-US})$ and $\log(I_{ot})$ because $\log(I_{ot})$ is measured with substantially less precision in publicly-available data. See Mahajan and Yang (2020) for more details on the benefits of measuring country-level migration using the restricted-access Census Bureau data. This will be relevant in Section B.2, when a version of Δz^M_{gk} constructed from publicly-available data has less first-stage strength than found in the results presented in the main text.

Table B1 also presents results in which I replace $\log(I_{ot})$ with I_{ot} and $\log(M_{o,-US,t})$ with $M_{o,-US,t}$, for comparison. The lack of precision found here motivate the use of log shifts instead of level shifts, although the latter is more common in the prior literature. Both instrument strength and identification are much more likely to rely on shares when using levels, given their lack of predictive power.

Table B1: Non-U.S. OECD Emigration and U.S. Immigration

	Outcome: $\log(I_{ot})$			Outcome: I_{ot}		
$\log(M_{o,-US,t})$	0.967***	0.921***	0.993***			
	(0.315)	(0.315)	(0.353)			
	[0.472]	[0.472]	[0.497]			
$M_{o,-US,t}$				0.139**	0.026	0.072*
				(0.069)	(0.030)	(0.039)
				[0.093]	[0.040]	[0.056]
2000 Population Control		✓	✓		✓	✓
Trade Controls		✓	✓		✓	✓
UN Region by Year FE			✓			✓

Notes: Standard errors clustered at the origin country o level presented in parentheses. Standard errors clustered at the UN Region level presented in square brackets. 2000 Population interacted with Year FE. Trade controls are the log of imports from origin o in the U.S. and exports from the U.S. to origin o in year t. All models are estimated using 147 countries $\times$ 4 Years = 588 Country-Years.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (based on origin country clustering).

B.2 Comparison Across Data Sources and Instruments

In this section, I present a set of results that helps assess two key questions:

1. How comparable are results generated from publicly-available data to results from restricted-access data, presented in the main text?

2. How does Δz_{gk}^{M} compare to a more "standard" instrumental variable for immigrant inflows into local industries?

Answering the first question requires an outcome variable that is used in the main analysis and, preferably, constructed using the same source data as the LBD in a consistent manner over time. Recent work by Eckert et al. (2020) produce such a variable: County Business Patterns (CBP) employment measured using consistent, Fort and Klimek (2018) NAICS codes over time at the county level. The CBP uses the same source data as the LBD (the Census Bureau's Business Registrar). I thus aggregate the Eckert et al. (2020) from the county-by-6-digit-NAICS code level to the local industry level used in this paper's analysis and use this daa to construct CBP employment growth in local industry gk over the time periods 2000–2015, $\Delta \log(\text{CBP Emp})$.

To answer the second question, I construct a version of z_{gk}^{M} from IPUMS-USA that uses the aforementioned 90 origin countries from Section B.1. I also construct a more "standard" version of the instrumental variable, which takes changes to the aggregate immigrant stock in the U.S. and distributes into local industries=:

$$\Delta z_{gk}^{\text{Std}} = \frac{1}{\text{Workers}_{gk,2000}} \sum_{o} \pi_{ok} \tilde{\rho}_{ok} \Delta I_o$$

ΔI_o is the change in the stock of immigrants from origin o living in the U.S. between 2000 and 2015 and where $\tilde{\rho}_{ok}$ is the proportion of those immigrants that when to industry k:

$$\tilde{\rho}_{ok} \equiv \frac{\Delta I_{ok}}{\Delta I_o}$$

See Equation (10) in Card, 2001 for the analogous instrument into city-occupation groups instead of CZ-industry groups. In order to directly address concerns broached by Jaeger et al. (2018) regarding confounding short- and long-run responses to immigration, I also construct:

$$\Delta I^M_{gk,2000} = \frac{\text{Immigrants}_{gk,2000} - \text{Immigrants}_{gk,1990}}{\text{Workers}_{1990}}$$

$$\Delta z^M_{gk,2000} = \sum_o \pi_{og,1990}\rho_{ok,1990}\left[\log(M_{o,-US,2000}) - \log(M_{o,-US,1990})\right]$$

$$\Delta z^{\text{Std}}_{gk,2000} = \frac{1}{\text{Workers}_{gk,1990}} \sum_o \pi_{ok,1990}\tilde{\rho}_{ok,1990}\left[I_{o,2000} - I_{o,1990})\right]$$

Table B2 presents several important results. Columns (1) and (2) are directly comparable to the left-most column in Table C3. They demonstrate a remarkable similarity in results across those generated by restricted-access data (Table C3) and those generated by publicly-available data (Table B2). The 2SLS-estimated local industry effect of a 1% relative immigration on employment growth 0.734% in the restricted-access data and 0.795% in the publicly-available data. This similarity reassures us that the results presented in the rest of Table B2 are meaningful and that the Eckert et al. (2020) data is useful for this kind of detailed analysis.

One key difference between the results from restricted-access and publicly-available data, however, is in the strength of the instrument. Likely due to increased precision in ΔI_{gk}, the ability to include additional origin countries, and the ability to better-measure π_{og} and ρ_{ok}, the instrument is about four times as strong in the restricted-access data. The similar ultimate 2SLS effect estimates $\hat{\beta}$ support this measurement-error based explanation.

The rest of Table B2 focuses on comparing Δz^M_{gk} with $\Delta z^{\text{Std}}_{gk}$. Column (3) shows that the standard instrument is considerably stronger, but ultimately yields an effect estimate that is essentially identical to the OLS estimate. While this is theoretically possible if various sources of bias are cancelling out, it is also consistent with the more plausible explanation that the increased first-stage strength and resulting increase in precision come at the cost of increased bias. This is perhaps most clearly illustrated in Columns (4)-(6), which present basic balance tests on 1980–2000 employment growth using normalized versions of each source of variation for ease of comparison. While both ΔI^M_{gk} and Δz^M_{gk} are not statistically significantly correlated with pre-period employment growth, there is a highly significant, strong, and negative relationship between $\Delta z^{\text{Std}}_{gk}$ and this outcome. There is thus strong evidence that reversion to the mean may be driving some of the estimate in Column (3).

Column (7)-(9) implement the double-instrumentation strategy proposed in Jaeger et al. (2018) to address concerns that results found for the time period 2000–2015 actually reflect longer-run responses to shocks that occurred prior to 2000. In the context of Jaeger et al. (2018), one potential benefit of using non-U.S. emigrant outflows instead of U.S. immigrant inflows is that it is more likely to change over time in terms of origin country composition, reducing the serial correlation in the instrument that generates this bias. Column (8) indicates that Δz^M_{gk} is indeed quite resilient to the data demanding double-instrumentation procedure, showing an *increase* in the 1st Stage F Statistic and a nearly identical point estimate of interest (compared to Column 2).

Column (9) indicates that $\Delta z_{gk}^{\text{Std}}$ is less-so but adequately robust to double-instrumentation, with the 1st Stage F Statistic declining by a factor of 7, but still at a reasonable value of 20.5. In addition, the point estimate of interest in Column (9) is also nearly identical to that found in Column (3). However, there is an important caveat to this result. When I use a consistent denominator across $\Delta z_{gk}^{\text{Std}}$ and $\Delta z_{gk,2000}^{\text{Std}}$, the 1st Stage F Statistic drops below 1 and results become unstable and insignificant. That is, if I replace $\frac{1}{\text{Workers}_{gk,2000}}$ with $\frac{1}{\text{Workers}_{gk,1990}}$ in $\Delta z_{gk}^{\text{Std}}$, the standard instrument is no longer robust to double-instrumentation. This indicates that much of the remaining 1st Stage strength in Column (9) is due to correlation between the denominators of the instruments and endogenous variables rather than the numerators—a "blunt instruments" issue (Bazzi and Clemens, 2013) that is not present in Δz_{gk}^{M}.

All told, I draw the following conclusions:

1. For aggregate variables, extant publicly-available data can roughly match results from restricted-access data. However, very few of the analyses in this paper can be replicated with publicly-available data.

2. Δz_{gk}^{M} outperforms a more standard instrument along several dimensions because it is more likely to isolate exogenous, contemporaneous immigration pushes, conditional on controls.

Table B2: Immigrant Worker Inflows and Employment Growth in Publicly-Available Data

	Outcome: $\Delta \log(\text{CBP Emp}_{gk})$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ΔI_{gk}	1.021***	0.795***	1.028***				1.056***	0.767***	1.043***
	(0.045)	(0.162)	(0.122)				(0.043)	(0.138)	(0.134)
$\Delta I_{gk,2000}$							-0.311***	-0.056	0.100
							(0.044)	(0.139)	(0.166)
Normalized ΔI_{gk}				-0.025					
				(0.018)					
Normalized Δz_{gk}^{M}					0.001				
					(0.003)				
Normalized $\Delta z_{gk}^{\text{Std.}}$						-0.308***			
						(0.062)			
Outcome Long Difference Span	2000–2015	2000–2015	2000–2015	1980–2000	1980–2000	1980–2000	2000–2015	2000–2015	2000–2015
Estimation	OLS	2SLS	2SLS	OLS	OLS	OLS	OLS	2SLS	2SLS
Instrument(s)	—	Δz_{gk}^{M}	$\Delta z_{gk}^{\text{Std}}$	—	—	—	—	$\begin{pmatrix} \Delta z_{gk}^{M} \\ \Delta z_{gk,2000}^{M} \end{pmatrix}$	$\begin{pmatrix} \Delta z_{gk}^{\text{Std}} \\ \Delta z_{gk,2000}^{\text{Std}} \end{pmatrix}$
Conventional 1st Stage F Statistic		24.0	146.9					30.0	20.5

Notes: See Equation (1) for specification. Conventional standard errors—robust to heteroskedasticity—are presented in parentheses. All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2, constructed from publicly-available data.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (based on conventional standard errors).

B.3 More on the Composition of Δz_{gk}^{M} Pushed Immigrants

Using publicly-available data, this section estimates Equation (1) with outcome variables $\Delta I_{gk}^{\text{Char.}}$ for several sets of mutually exclusive and exhaustive characteristics (Char.). The idea is to expand on Section 2.2.3 in analyzing the characteristics of inflows pushed in by the instrumental variable, Δz_{gk}^{M}.

B.3.1 Origin Country and its Importance to Conditional Exogeneity

The first set of characteristics is origin country. Figure B1 plots estimated coefficients $\hat{\beta}$ from Equation (1) with ΔI_{gk}^{o} as the outcome variable across origin countries o, such that $\sum_o \Delta I_{gk}^{o} = \Delta I_{gk}$, in map form. Each coefficient plotted represents the number of immigrants from origin country o brought in per immigrant worker. Estimating Equation (1) using OLS gives us a sense of the "average" immigrant inflow, whereas estimating Equation (1) using 2SLS and gives us a sense of the immigrant inflow used as identifying variation. Figure B1 is most clearly an illustration of how Δz_{gk}^{M} reduces the concentration of inflows from Mexico, instead bringing in a set of inflows from a more balanced set of origin countries that are less susceptible to U.S. pull factors. Concerns that this alters the skill distribution of inflows are alleviated by Figure 2 from the main text.

A simple and relevant example that helps illustrate why it is important to reduce the influence of Mexican inflows comes from the housing bubble that crested during this paper's study period—in the mid-2000s—largely in the South and West of the U.S. The housing bubble created a large labor demand shock for construction workers in the South and West Census Regions of the U.S., and induced immigrant workers from Mexico to fill this demand. Between 2000 and 2005, 66% of immigrant inflows into the Construction industry, and 39% of overall workforce growth in the Construction industry were explained by net inflows of *Mexican* workers alone. These numbers rise to 85% and 50% when considering all of Central America. These are the precise kind of inflow an instrumental variable should not use for identification of β in Equation (1) because it confounds immigration that was induced by labor demand with the response of labor demand to immigration.

Δz_{gk}^{M} exploits the fact that there is no reason to believe that the U.S. housing bubble would cause large outflows of Mexican (or other Central American) emigrants to non-U.S. OECD countries. Thus, measured emigration shocks during the study period are much less likely to reflect the pull of labor demand from the Construction industry in U.S. southwest commuting zones. Furthermore, by measuring $\rho_{0,\text{Construction}}$ in 2000, Δz_{gk}^{M} eliminates the influence national shifts towards the construction industry. Note that $\Delta z_{gk}^{\text{Std}}$ does not have these features. Initial shares of Mexican workers, $\pi_{\text{Mexico},g}$, are also high in South and West commuting zones, as is the *changing* share of Mexican immigrants working in the construction sector, $\tilde{\rho}_{\text{Mexico},k}$.

In order to summarize the ultimate result of these corrections, I utilize the Housing Price Index (HPI) developed by Bogin et al. (2016) as an outcome in the following specification:

$$\Delta HPI_{gt} = \xi \left[\Delta v_{k=\text{CONS},gt}\right] + \Gamma X_{k=\text{CONS},gt} + \alpha_g + \alpha_{d(g),t} + \varepsilon_{gt}$$

where I study five-year changes in the construction industry across commuting zones in order to track the boom and bust of the housing bubble during my study period ($t \in \{2005, 2010, 2015\}$). Specifically, ΔHPI_{gt} is the change in the housing price index over a five year interval and $\Delta v_{k=\text{CONS},gt} \in \{\text{Normalized } \Delta I_{k=\text{CONS},gt}, \text{Normalized } \Delta z_{k=\text{CONS},gt}^{M}, \text{Normalized } \Delta z_{k=\text{CONS},gt}^{\text{Std}}\}$ are the 5-year differenced, normalized versions of the endogenous exposure variable, the emigrants instrument, and the standard instrument, respectively.

Results from estimating this equation are presented in Table B3. Conditional on controls, Δz_{gk}^{M} is not correlated with housing price changes, whereas both the endogenous variable and the alternate, standard instrument are (though the standard instrument does substantially reduce the magnitude of the correlation). This representative example demonstrates the importance of bringing Mexico's influence on the instrument in line with other origin countries and strengthens the notion that Δz_{gk}^{M} is also making appropriate corrections when confronted with similar labor demand shocks in other industries.

Table B3: Immigrant Worker Inflows and Housing Prices in the Construction Industry

	Outcome: ΔHPI_{gt}		
	(1)	(2)	(3)
Normalized $\Delta I_{k=\text{CONS},gt}$	23.32***		
	(3.33)		
Normalized $\Delta z_{k=\text{CONS},gt}^{M}$		0.71	
		(0.83)	
Normalized $\Delta z_{k=\text{CONS},gt}^{\text{Std}}$			2.24**
			(1.06)
Commuting Zones	649	649	649

Notes: Conventional standard errors—clustered at the commuting zone level—are presented in parentheses. All models are estimated using 649 CZ $\times$ 3 Years = 1,947 CZ-Years, weighted by the initial (2000) workforce size in the Construction industry. The Housing Price Index (HPI) is only available for 649 commuting zones. It is the county-level HPI averaged to the commuting zone level, weighted by county population. All independent variables are normalized for ease of comparison. All specifications include the control variables described in Section 2.2, constructed from publicly-available data.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (based on conventional standard errors).

B.3.2 Undocumented Immigrants

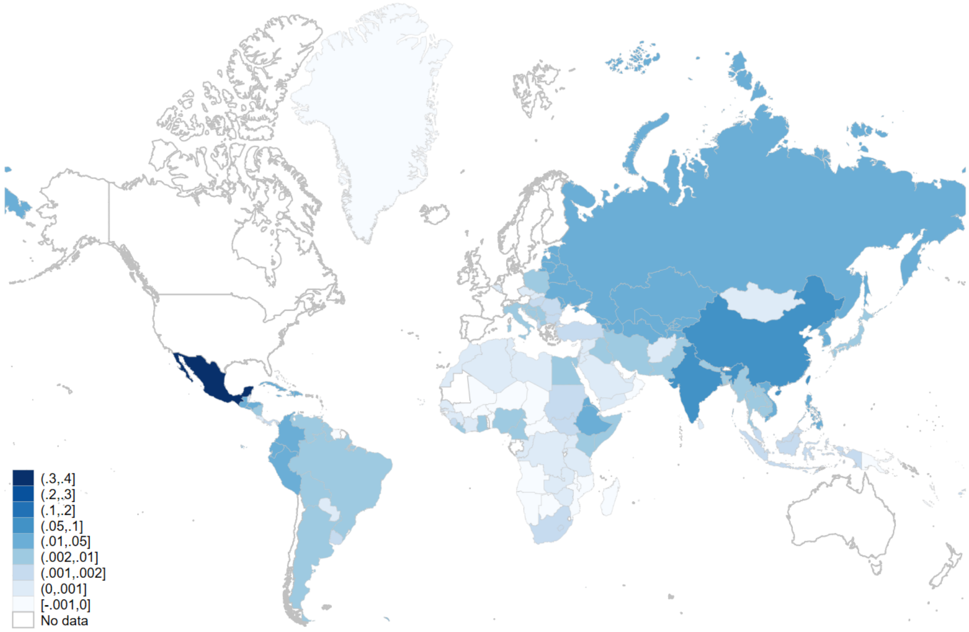
How much does the reduced influence of Mexico in the identifying variation impact the role that undocumented immigrants play in generating results? I employ two methods of measuring the undocumented workforce from the recent literature to approach this question.

The first adapts the Pew Center methodology described in Borjas (2017) to the IPUMS-USA 2000 Census and ACS demographic files. It mimics the "residual" method of identifying undocumented immigrants in survey data. It starts with the assumption that all non-citizen immigrants are undocumented, then systematically recodes them as legal based on specific characteristics. Specifically, non-citizen immigrants are considered legal if they meet any of the following criteria:

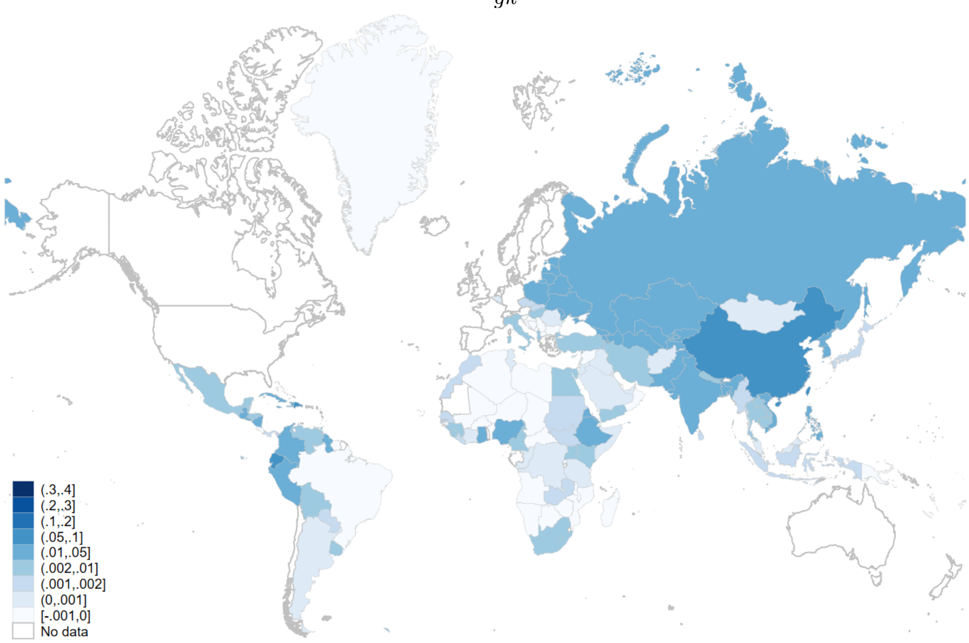
1. Arrived in the U.S. before 1980
2. Receiving social security, Supplemental Security Income (SSI), or welfare benefits
3. Participating in the armed forces

Figure B1: Immigrant Inflows by Origin Country

Panel A: OLS



Panel B: Δ^M_{gk} Pushed



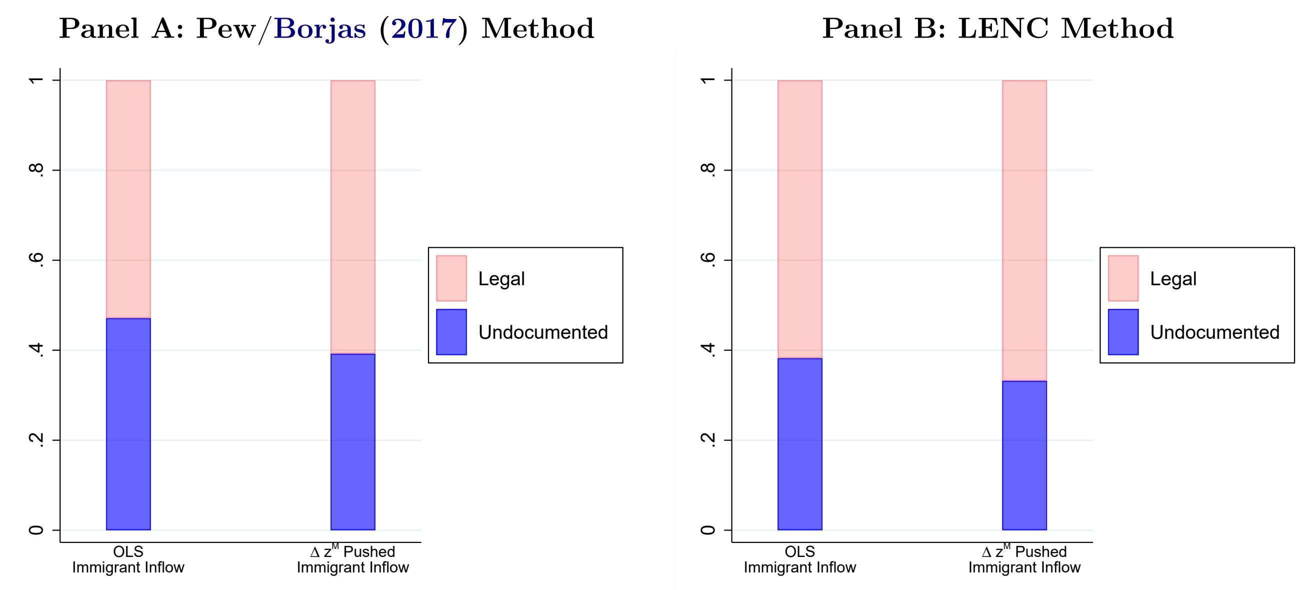
Notes: Each plotted coefficient is obtained by estimating Equation (1) with ΔI^o_{gk} as the outcome, for the mutually exclusive and exhaustive origins o shown in the map. ΔI^o_{gk} are net inflows of immigrant workers from origin o into local industry gk between 2000 and 2015. The resulting estimates add up to one and illustrate the origin country make-up of immigrant inflows into the U.S. based on whether Equation (1) is estimated using OLS (Panel A) or 2SLS with Δz^M_{gk} as the instrumental variable (Panel B). All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2.

4. Work in the public sector
5. Born in Cuba
6. Working in licensed occupations: healthcare professionals, lawyers, judges, teachers, accountants, psychologists, pilots, air traffic controllers, architects, and engineers
7. Spouse is either a citizen or meets any of the criteria above

The second methodology simply considers all non-citizens with at most a high school degree as undocumented. Similar strategies have been used in previous research (see, e.g., Bohn et al., 2014; East et al., forthcoming).

Figure B2 plots estimated coefficients $\hat{\beta}$ from Equation (1) with ΔI_{gk}^{ℓ} as the outcome variable across legal status ℓ, such that $\sum_{\ell} \Delta I_{gk}^{\ell} = \Delta I_{gk}$, where $\ell \in \{\text{Undocumented}, \text{Legal}\}$. Results indicate that a sizeable component–between 30% and 40%—of the inflows represented in ΔI_{gk} are likely to come from undocumented workers. This reflects the overall importance of undocumented workers to immigration in the U.S., particularly when it is primarily "low-skilled," as is the case here. Reassuringly, the legal-status composition of inflows pushed by the instrument are similar to the "average inflow" represented by the OLS bars in each panel, despite the change to the origin country composition of inflows. This is an important check because undocumented workers may lead to specific labor market dynamics that are different for legal immigrant workers (see, e.g., Peri and Chassamboulli, 2015; Borjas, 2017; East et al., forthcoming; Mahajan, 2017; Albert, 2021).

Figure B2: Legal Status Composition of Immigrant Inflows



Notes: Each plotted coefficient is obtained by estimating Equation (1) with ΔI_{gk}^{ℓ} as the outcome, for the mutually exclusive and exhaustive groups $\ell \in \{\text{Undocumented}, \text{Legal}\}$. ΔI_{gk}^{ℓ} are net inflows of immigrant workers of legal status ℓ into local industry gk between 2000 and 2015. The resulting estimates add up to one and illustrate the legal status make-up of immigrant inflows into the U.S. based on whether Equation (1) is estimated using OLS (left bar within panel) or 2SLS with Δz_{gk}^{M} as the instrumental variable (right bar within panel). Panels A and B use different methods of identifying undocumented workers in IPUMS-USA data, described in the text. LENC stands for less-educated non-citizen. All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2.

B.3.3 Task Content of Occupation and English Language Ability

A key reason why immigrant workers may be imperfect substitutes for native workers (i.e., $\sigma_1 < \infty$ in Section 4) is because they bring different skills and abilities with them and therefore have comparative advantage in different tasks (Peri and Sparber, 2009; Lewis, 2011). In order to assess whether these are plausible channels to immigrant-native imperfect substitutability, I also assess the composition of immigrant inflows relative to the receiving population in terms of the task content of occupations and English language speaking ability.

In order to assess whether immigrant workers are more likely to be in occupations that require manual or routine tasks, I use the task designations produced for IPUMS-USA `occ1990` in Autor and Dorn (2013), available here. Autor and Dorn (2013) calculate the number of manual, routine, and abstract tasks associated with each `occ1990`. I calculate the number of immigrants working on a given task type as follows. Let T^{m}_{occ} denote the number of manual tasks associated with `occ1990` code *occ*. T^{r}_{occ} denote the number of routine tasks associated with *occ* and T^{a}_{occ} denote the number of abstract tasks associated with *occ*. Then,

$$I^{task}_{gkt} = \sum_{occ} \frac{T^{task}_{occ} I^{occ}_{gkt}}{(T^{m}_{occ} + T^{r}_{occ} + T^{a}_{occ})}$$

where $task \in \{m, r, a\}$ and I^{occ}_{gkt} is the number of immigrant workers in occupation *occ* and local industry gk at time t. I can then construct ΔI^{task}_{gk} over the period 2000–2015, where *task* is again a set of mutually exclusive and exhaustive groupings and ΔI^{task}_{gk} is used as an outcome variable in Equation (1). The same procedure is used to construct $N^{task}_{gk,2000}$, where N stands for native workers.

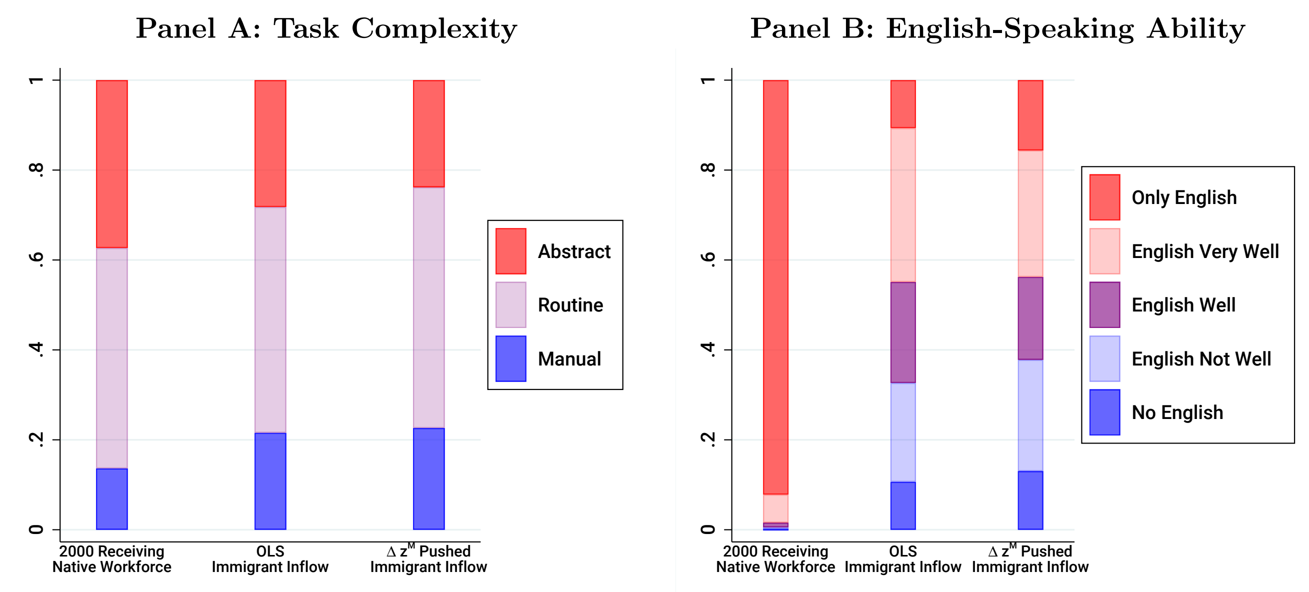
A simpler and related decomposition from the perspective of immigrant-native substitutability is English language ability. Here, I simply study use Δz^{spk}_{gk} as the outcome variable in Equation (1), where *spk* covers five mutually exclusive and exhaustive categories based on the IPUMS-USA variable `speakeng`: No English, Speaks English, Not Well, Speaks English, Well, Speaks English, Very Well, Only Speaks English.

The resulting decompositions are presented in Figure B3. Across both OLS- and IV-pushed inflows, immigrant workers are substantially more likely to perform manual and routine tasks on the job, indicating a strong case for imperfect substitutability (Panel A). A primary reason for this is potentially reflected in Panel B, which shows that more than one-third of immigrant worker inflows are accounted for by those who either speak no English or do not speak it well.

B.3.4 Age and Gender

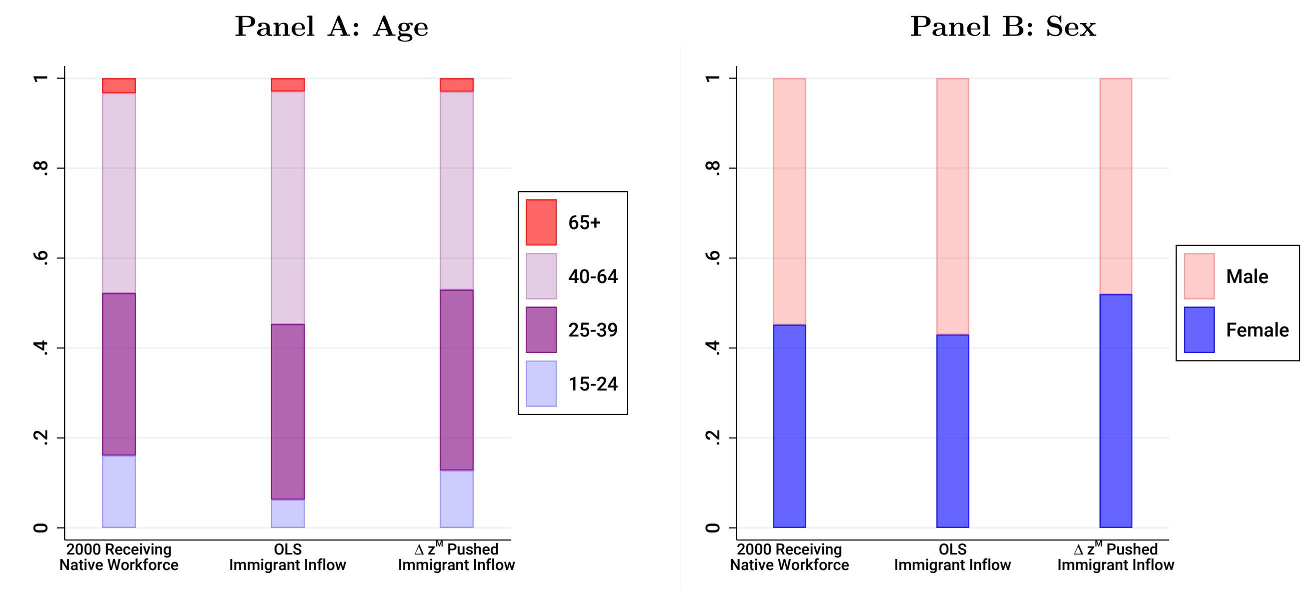
Finally, Figure B4 uses the same decomposition method studies how immigrant inflows compare to the receiving population in terms of age and gender. It finds that inflows pushed in by Δz^{M}_{gk} are slightly younger and more gender balanced than the "average inflow." Furthermore, the IV-pushed distributions look more similar to the receiving native population than the OLS-pushed distributions, indicating slightly less scope for imperfect substitutability than is associated with the "average inflow" on these characteristics.

Figure B3: Occupational and English Language Ability Composition of Immigrant Inflows



Notes: Each plotted coefficient is obtained by estimating Equation (1) with $\Delta I_{gk}^{\text{Char.}}$ as the outcome, for mutually exclusive and exhaustive groups. In Panel A, these mutually exclusive and exhaustive groups are $task \in \{m, r, a\}$. In Panel B, these mutually exclusive and exhaustive groups are based on the IPUMS-USA variable `speakeng`: No English, Speaks English, Not Well, Speaks English, Well, Speaks English, Very Well, Only Speaks English. The resulting estimates add up to one and illustrate the legal status make-up of immigrant inflows into the U.S. based on whether Equation (1) is estimated using OLS (left bar within panel) or 2SLS with Δz_{gk}^{M} as the instrumental variable (right bar within panel). All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2.

Figure B4: Age and Sex Composition of Immigrant Inflows



Notes: Each plotted coefficient is obtained by estimating Equation (1) with $\Delta I_{gk}^{\text{Char.}}$ as the outcome, for mutually exclusive and exhaustive groups. In Panel A, these mutually exclusive and exhaustive groups are $age \in \{15-24, 25-39, 40-64, 65+\}$. In Panel B, these mutually exclusive and exhaustive groups are $sex \in \{\text{Male}, Female\}$. The resulting estimates add up to one and illustrate the legal status make-up of immigrant inflows into the U.S. based on whether Equation (1) is estimated using OLS (left bar within panel) or 2SLS with Δz_{gk}^{M} as the instrumental variable (right bar within panel). All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2.

C Supplemental Figures and Tables

C.1 Immigrant Worker Inflows and Aggregate Local Industry Outcomes

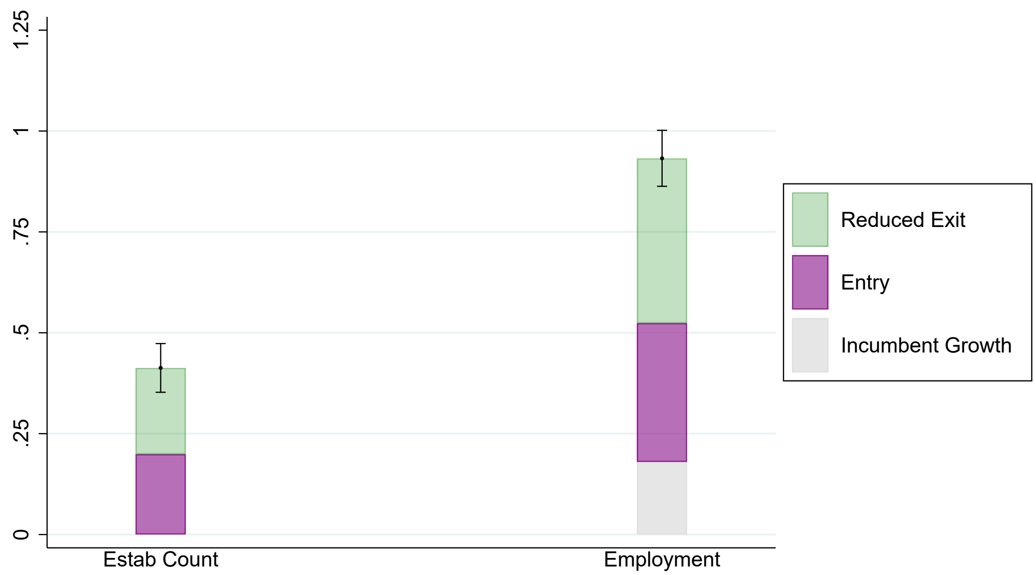
Table C1: The Effect of Immigrant Worker Inflows on Local Industry Aggregates

	Outcome: Δ log of					
	Average Earnings	Employment	Estab Count	Payroll	Revenues	Revenues per Worker
Panel A: OLS						
Immigrant Worker Inflows (ΔI_{gk})	-0.010	1.018***	0.423***	1.008***	1.050***	0.167***
	(0.030)	(0.041)	(0.032)	(0.052)	(0.079)	(0.060)
Panel B: 2SLS (Δz_{gk}^{M})						
Immigrant Worker Inflows (ΔI_{gk})	0.367**	0.734***	0.921***	1.101***	1.940***	1.147***
	(0.168)	(0.118)	(0.171)	(0.197)	(0.315)	(0.291)
	[0.125]	[0.115]	[0.336]	[0.165]	[0.165]	[0.191]
Outcome Data Source	LBD	LBD	LBD	LBD	SBO/ABS	SBO/ABS
Outcome Long Difference Span	2000–2015	2000–2015	2000–2015	2000–2015	2002–2017	2002–2017

Notes: See Equation (1) for specification. Conventional standard errors—robust to heteroskedasticity—are presented in parentheses. Where applicable, Borusyak et al. (2021) standard errors—from equivalent origin o level regressions in which standard errors are clustered at the UN region level—are presented in square brackets. All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2. Outcomes with Data Source "LBD" are measured using the Longitudinal Business Database, which is a full-count panel of establishments with near-complete coverage of the U.S. private sector. Outcomes with Data Source "SBO/ABS" are measured using repeated cross-sections from the 2002 Survey of Business Owners and 2017 Annual Business Survey, which are each approximately 10% representative samples of U.S. firms. See Section 2.2 for additional details on data sources.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (based on conventional standard errors).

C.2 Estab. Count and Employment Growth Decompositions (Section 3.1)

Figure C1: The Effect of Immigrant Worker Inflows on Establishment Counts and Employment (OLS Results)



Notes: See Equation (1) for specification, estimated using OLS. All models are estimated using 722 CZ × 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. Each bar adds up to the estimated effect of a 1% increase in a local industry's workforce on the growth rate in establishment count or employment in that local industry. These total effects are also plotted using black points, with capped spikes representing conventional, heteroskedasticity-robust 95% confidence intervals around them. See Table C2 for underlying coefficients and standard errors for each component of the establishment count decomposition and Table C3 for underlying coefficients and standard errors for each component of the employment decomposition. See Figure 3 for corresponding IV results. All specifications include the control variables described in Section 2.2.

Table C2: Decomposing the Effect of Immigrant Worker Inflows on Local Industry Establishment Count Growth

	Δ log Estab Count	DHS Estab Count Growth Rate	Contribution to DHS Growth Rate:	
			Entry	Reduced Exit
Panel A: OLS				
Immigrant Worker Inflows (ΔI_{gk})	0.423***	0.413***	0.199***	0.214***
	(0.032)	(0.031)	(0.023)	(0.016)
Panel B: 2SLS (Δz_{gk}^{M})				
Immigrant Worker Inflows (ΔI_{gk})	0.921***	0.903***	0.505***	0.398***
	(0.171)	(0.168)	(0.133)	(0.081)

Notes: See Equation (1) for specification. Conventional standard errors—robust to heteroskedasticity—are presented in parentheses. All models are estimated using 722 CZ × 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2. Figure 3 provides a visual illustration of this table.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

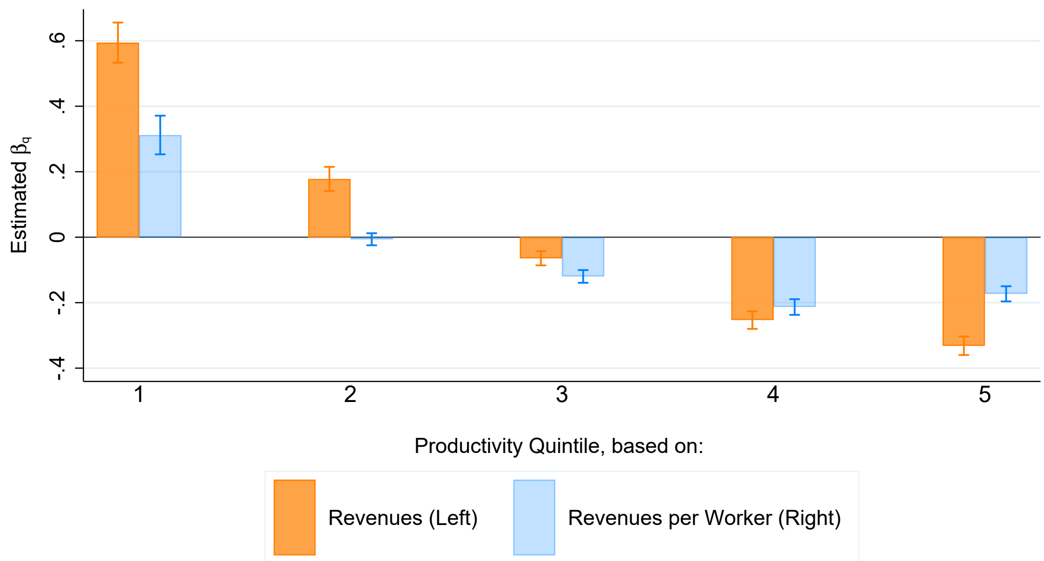
Table C3: Decomposing the Effect of Immigrant Worker Inflows on Local Industry Employment Growth

	Δ log Employment	DHS Employment Growth Rate	Contribution to DHS Growth Rate:		
			Entry	Reduced Exit	Incumbent Growth
Panel A: OLS					
Immigrant Worker Inflows (ΔI_{gk})	1.018***	0.932***	0.344***	0.409***	0.180***
	(0.041)	(0.035)	(0.024)	(0.024)	(0.023)
Panel B: 2SLS (Δz^M_{gk})					
Immigrant Worker Inflows (ΔI_{gk})	0.734***	0.764***	0.382***	0.338***	0.044
	(0.118)	(0.113)	(0.126)	(0.122)	(0.102)

Notes: See Equation (1) for specification. Conventional standard errors—robust to heteroskedasticity—are presented in parentheses. All models are estimated using 722 CZ $\times$ 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2. Figure 3 provides a visual illustration of this table.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

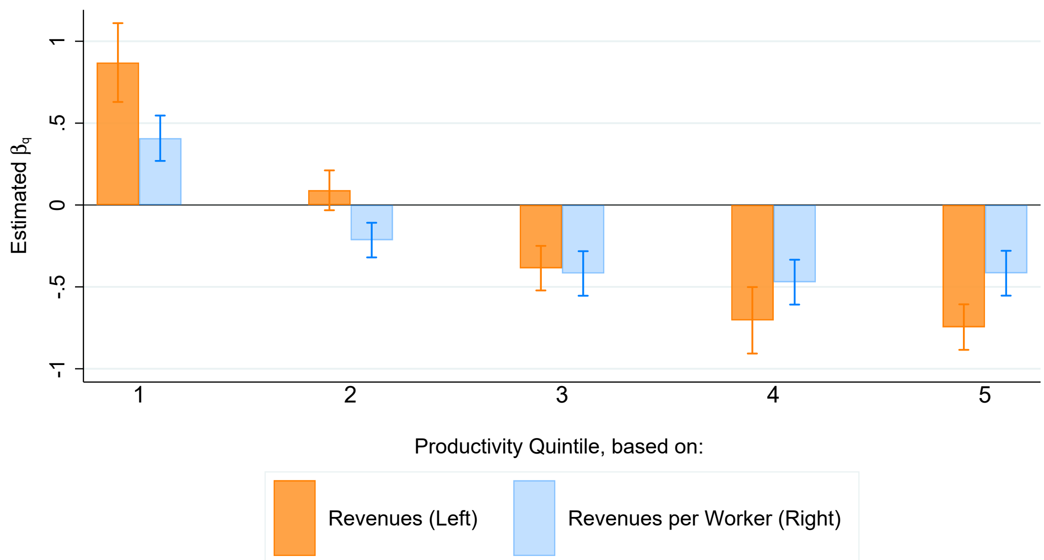
C.3 Heterogeneity at the Exit Margin (Section 3.2.1)

Figure C2: Immigrant Workers and Establishment Exit, by Initial Productivity (OLS Results)



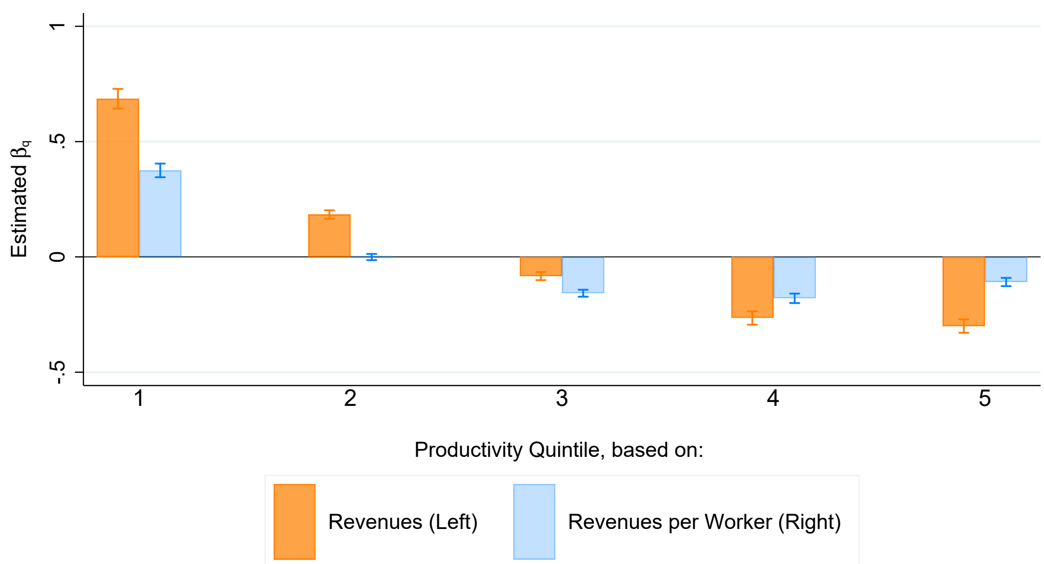
Notes: See Equation (4) for specification. Each coefficient $\hat{\beta}_q$ plotted as a bar represents the effect of a one percent increase in a local industry's workforce due to immigration on the probability that an establishment has zero payroll and employment. Each specification covers 4.7 million establishments that were operating (had positive payroll or employment) as of 2000, followed every five years until 2015. Establishments are split into productivity quintiles based on their parent firm's national rank in either log revenues (orange bars) or log revenues per worker (blue bars) within 6-digit NAICS codes and age bins in 2000. Capped spikes indicate conventional 95% confidence intervals, with standard errors clustered at the local industry level. The 1st Stage F statistic for the model estimated using revenue-based quintiles is 25.65, and it is 26.02 for the model estimated using revenues-per-worker-based quintiles. All specifications include the control variables described in Section 2.2. See Figure 4 for analogous 2SLS results. See Figure C3 for analogous 2SLS results when firms are assigned productivity quintiles based on their rankings within their local industries rather than national 6-digit-NAICS-by-age bins. See Figure C4 for OLS results with this alternate ranking.

Figure C3: Immigrant Workers and Establishment Exit, by Initial Productivity (IV Results, Alternate Ranking)



Notes: See Equation (4) for specification, estimated using 2SLS with instrumental variable z^M_{gkt}. Each coefficient $\hat{\beta}_q$ plotted as a bar represents the effect of a one percent increase in a local industry's workforce due to immigration on the probability that an establishment has zero payroll and employment. Each specification covers 4.7 million establishments that were operating (had positive payroll or employment) as of 2000, followed every five years until 2015. Establishments are split into productivity quintiles based on their parent firm's national rank in either log revenues (orange bars) or log revenues per worker (blue bars) within 6-digit NAICS codes and age bins in 2000. Capped spikes indicate conventional 95% confidence intervals, with standard errors clustered at the local industry level. The 1st Stage F statistic for the model estimated using revenue-based quintiles is 25.65, and it is 26.02 for the model estimated using revenues-per-worker-based quintiles. All specifications include the control variables described in Section 2.2. See Figure C4 for analogous OLS results. See Figure C3 for analogous 2SLS results when firms are assigned productivity quintiles based on their rankings within national 6-digit-NAICS-by-age bins. See Figure C2 for OLS results with this alternate ranking.

Figure C4: Immigrant Workers and Establishment Exit, by Initial Productivity (OLS Results, Alternate Ranking)



Notes: See Equation (4) for specification. Each coefficient $\hat{\beta}_q$ plotted as a bar represents the effect of a one percent increase in a local industry's workforce due to immigration on the probability that an establishment has zero payroll and employment. Each specification covers 4.7 million establishments that were operating (had positive payroll or employment) as of 2000, followed every five years until 2015. Establishments are split into productivity quintiles based on their parent firm's national rank in either log revenues (orange bars) or log revenues per worker (blue bars) within 6-digit NAICS codes and age bins in 2000. Capped spikes indicate conventional 95% confidence intervals, with standard errors clustered at the local industry level. The 1st Stage F statistic for the model estimated using revenue-based quintiles is 25.65, and it is 26.02 for the model estimated using revenues-per-worker-based quintiles. All specifications include the control variables described in Section 2.2. See Figure C3 for analogous 2SLS results. See Figure C3 for analogous 2SLS results when firms are assigned productivity quintiles based on their rankings within national 6-digit-NAICS-by-age bins. See Figure C3 for OLS results with this alternate ranking.

C.4 Entrants in High Immigration Environments (Section 3.2.2)

I first present results from the following specification, estimated on all entering establishments from 2001-2015, as with Equation (5), but which does not stratify effects by age.

$$y_{et} = \beta[\Delta I_{g(e),k(e)}] + \alpha_{a(e,t)} + \alpha_t + \alpha_g + \alpha_{d(g),k} + \Gamma X_{gk} + \varepsilon_{et} \tag{8}$$

This adds power and provides an overall assessment of entrants in high immigration environments.

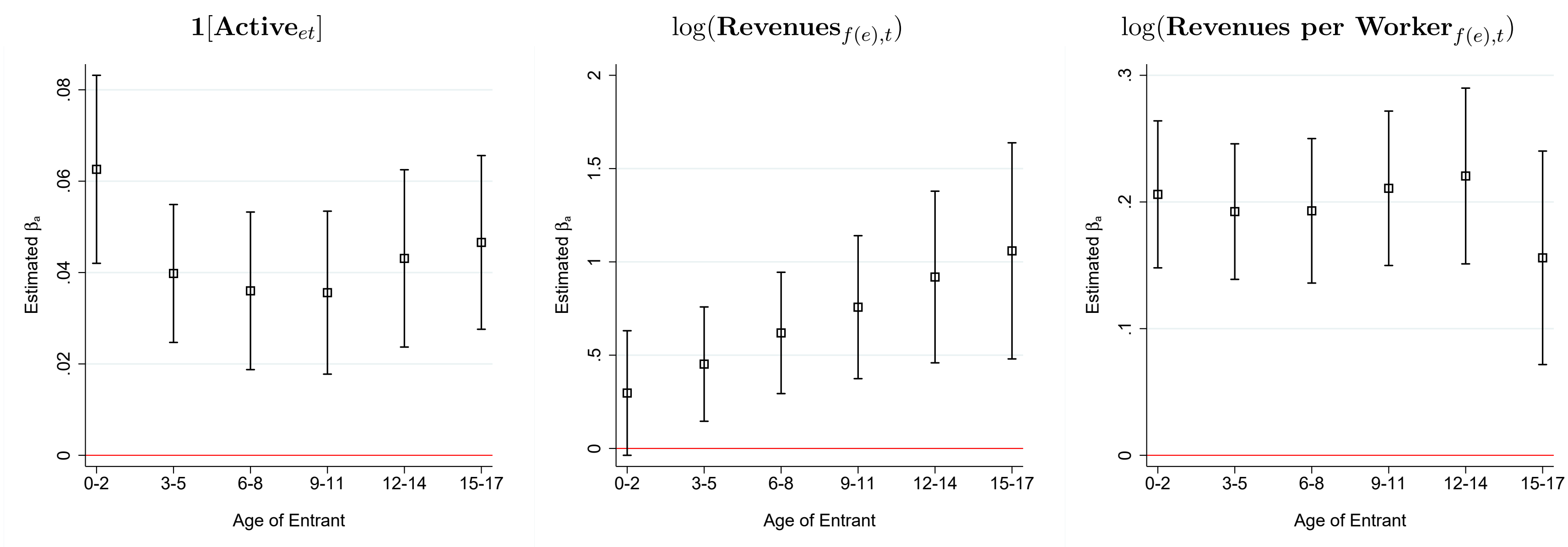
Table C4: Immigrant Worker Inflows and Entrant Productivity Characteristics

	Outcome:		
	$\mathbf{1}$[Active]	log(Revenues)	log(Revenues per Worker)
Panel A: OLS			
Immigrant Worker Inflows (ΔI_{gk})	0.046***	0.462***	0.201***
	(0.007)	(0.156)	(0.027)
Panel B: 2SLS (Δz^M_{gk})			
Immigrant Worker Inflows (ΔI_{gk})	0.195***	6.060***	0.410
	(0.037)	(1.217)	(0.295)
Establishments (millions)	9.4	6.7	6.7
Outcome Level	Establishment	Firm	Firm
Conventional First Stage F Statistic	85.26	76.88	76.88

Notes: See Equation (8) for specification. Standard errors clustered at the local industry level. All specifications include the control variables described in Section 2.2.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

I next present the OLS-estimated results that correspond to Figure 5.

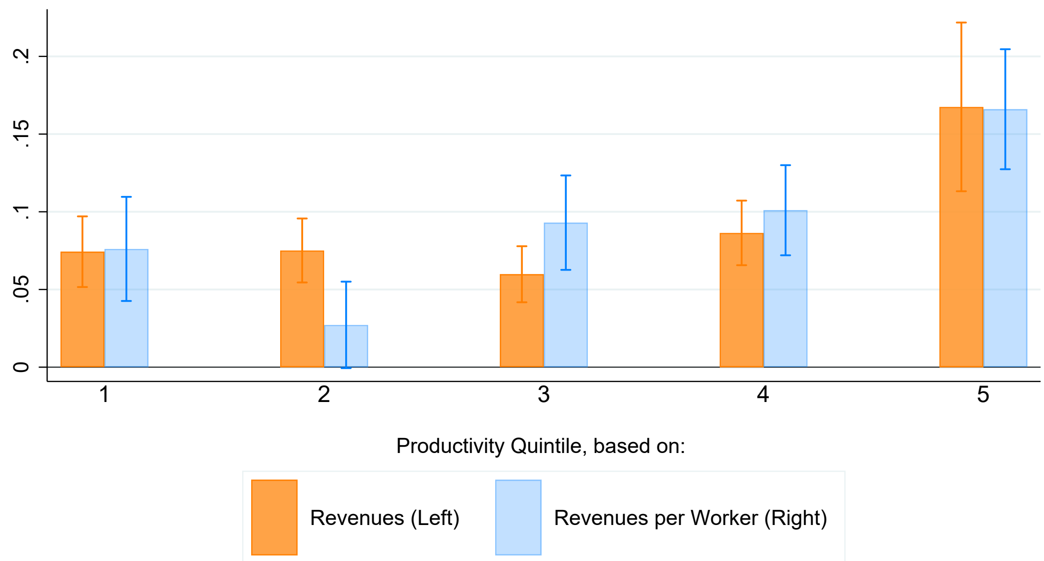
Figure C5: Immigrant Worker Inflows and Entrant Characteristics over their Life Cycle (OLS Results)



Notes: See Equation (5) for specification. Specification with outcome variable $\mathbf{1}[\text{Active}_{et}]$ is estimated on 9.4 million establishments and has a conventional First Stage F statistic of 14.21 (standard errors clustered at the local industry level). Specifications with revenue-based outcome variables are conditional on $\mathbf{1}[\text{Active}_{et}] = 1$, estimated on 6.7 million establishments, and have a conventional First Stage F statistic of 12.83. All specifications include the control variables described in Section 2.2. See 5 for corresponding IV results. See C4 for average effects, not split by age.

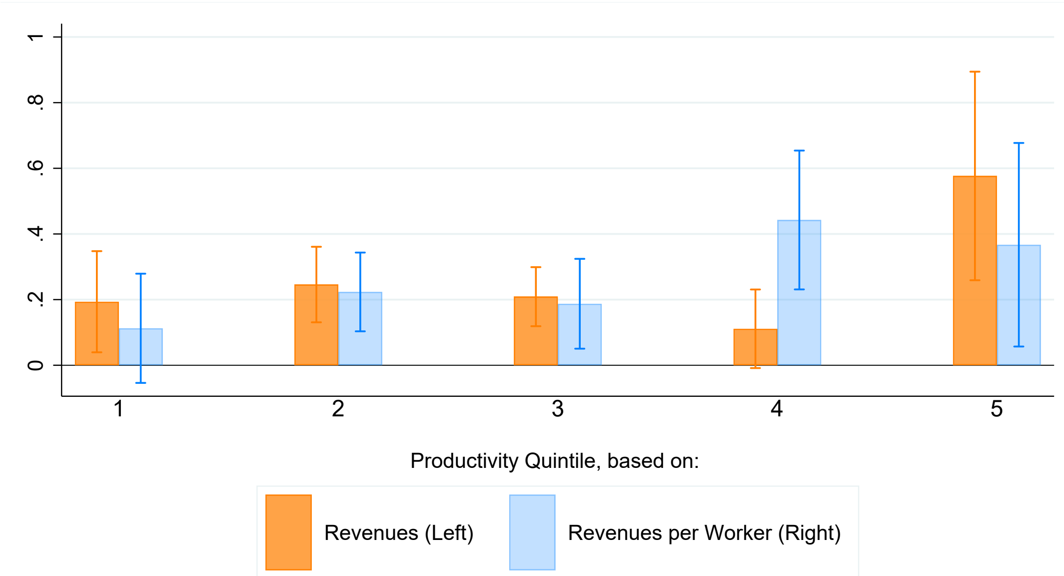
C.5 The Employer Productivity Distribution (Section 3.3.1)

Figure C6: Immigrant Worker Inflows and the Productivity Distribution (OLS Results)



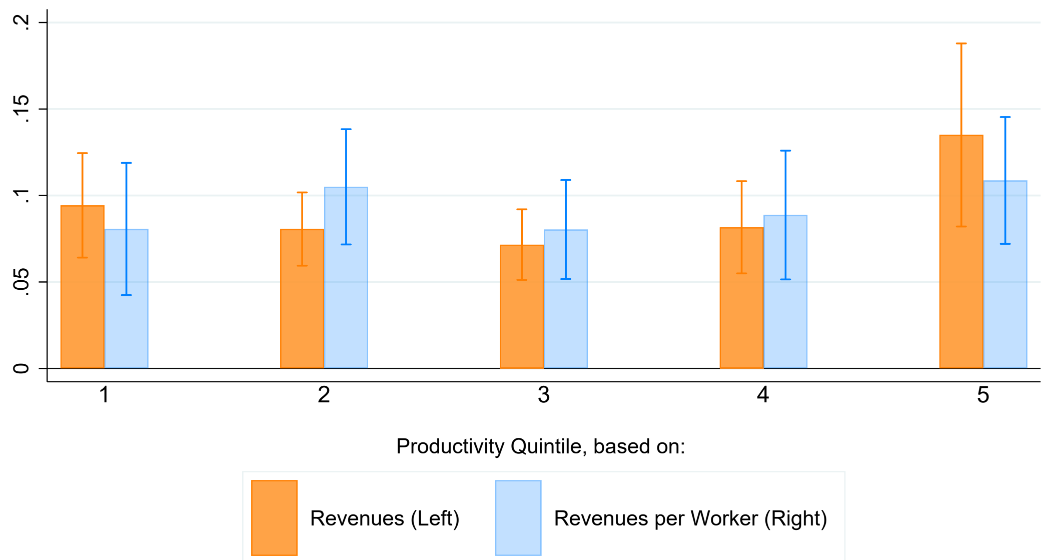
Notes: See Equation (1) for specification. Each estimated coefficient (presented as a bar) plots the contribution of a given productivity quintile to the overall growth rate in establishment count in local industry *gk*. Establishments are assigned quintiles based on their parent firm's revenues (orange) or revenues per worker (blue). Qunitile cutoffs for (real) revenues and revenues per worker are determined by firm rankings within 6-digit NAICS code and age bins in 2000. Capped spikes indicate conventional 95% confidence intervals, with standard errors clustered at the local industry level. All specifications include the control variables described in Section 2.2. See Figure 6 for analogous 2SLS results. See Figure C7 for analogous 2SLS results when firms are assigned productivity quintiles based on their rankings within their local industries rather than national 6-digit-NAICS-by-age bins. See Figure C8 for OLS results with this alternate ranking.

Figure C7: Immigrant Worker Inflows and the Productivity Distribution (IV Results, Alternate Ranking)



Notes: See Equation (1) for specification, estimated using 2SLS with instrumental variable Δz_{gk}^{M}. Each estimated coefficient (presented as a bar) plots the contribution of a given productivity quintile to the overall growth rate in establishment count in local industry gk. Establishments are assigned quintiles based on their parent firm's revenues (orange) or revenues per worker (blue). Qunitile cutoffs for (real) revenues and revenues per worker are determined by firm rankings within 6-digit NAICS code and age bins in 2000. Capped spikes indicate conventional 95% confidence intervals, with standard errors clustered at the local industry level. All specifications include the control variables described in Section 2.2. See Figure C8 for analogous OLS results. See Figure 6 for analogous 2SLS results when firms are assigned productivity quintiles based on their rankings within national 6-digit-NAICS-code-by-age bins rather than national 6-digit-NAICS-by-age bins. See Figure C6 for OLS results with this alternate ranking.

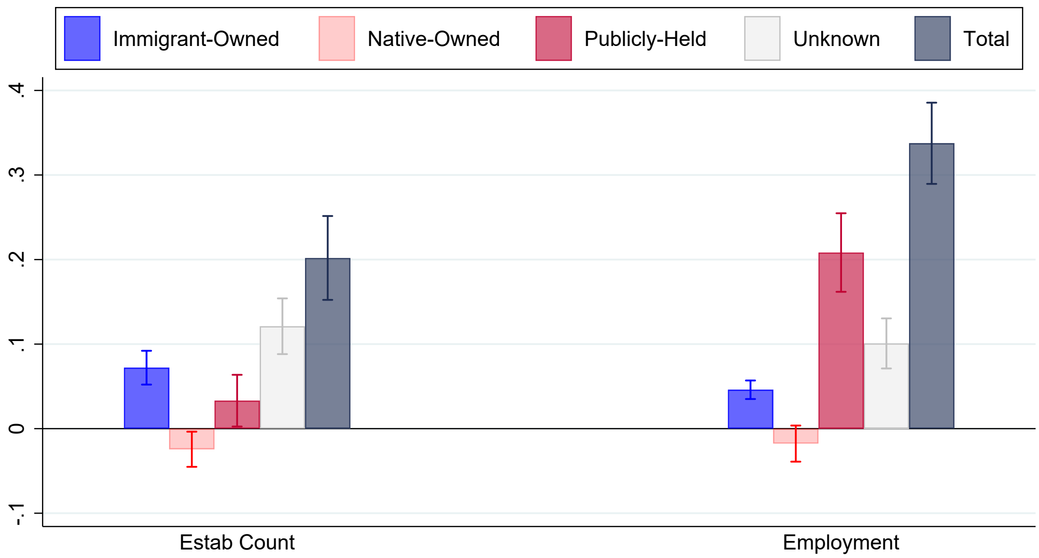
Figure C8: Immigrant Worker Inflows and the Productivity Distribution (OLS Results, Alternate Ranking)



Notes: See Equation (1) for specification. Each estimated coefficient (presented as a bar) plots the contribution of a given productivity quintile to the overall growth rate in establishment count in local industry gk. Establishments are assigned quintiles based on their parent firm's revenues (orange) or revenues per worker (blue). Qunitile cutoffs for (real) revenues and revenues per worker are determined by firm rankings within 6-digit NAICS code and age bins in 2000. Capped spikes indicate conventional 95% confidence intervals, with standard errors clustered at the local industry level. All specifications include the control variables described in Section 2.2. See Figure C7 for analogous 2SLS results. See Figure 6 for analogous 2SLS results when firms are assigned productivity quintiles based on their rankings within national 6-digit-NAICS-code-by-age bins rather than national 6-digit-NAICS-by-age bins. See Figure C6 for OLS results with this alternate ranking.

C.6 Entreprenuership by Nativity and Type (Section 3.4.1)

Figure C9: The Role of Entry by Ownership Nativity and Type (OLS Results)



Notes: See Equation (1) for specification, estimated using OLS. The rightmost, dark navy bars represent the contribution of entry to local industry growth rates in establishment counts and employment over the time period 2000–2017, as measured using the 2000 LBD (full-count) and 2018 ABS (≈16% sample). The left four bars add up to this overall contribution of entry in each case. Each of these bars represent the role of a given firm ownership group (entrepreneur type) in generating the overall entry effect. The four groups are: privately-held, at least 50% immigrant-owned (immigrant-owned), privately held, less than 50% immigrant-owned (native-owned), publicly-held, and unknown. All models are estimated using 722 CZ × 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2. See Figure 7 for corresponding IV results.

C.7 Heterogeneity by Industry

In order to construct Table C5, I split industries based on tradability and average educational attainment of workers (see Table A1 for final designations). I generate a tradability designation by aggregating 2000 traded and non-traded employment within each industry group based on the Porter classification system for 6-digit NAICS codes (Porter, 2003). Each industry group is then designated as tradable if more than 50 percent of its employment was in a tradable 6-digit NAICS code in 1980, and vice versa. Comparing Panel A to Panel B, I find that the impacts of immigrant worker inflows on local industry aggregates are generally larger in tradable industries, consistent with (Burstein et al., 2020).

I also designate industry groups based on whether they tend to hire higher- or lower-educated workers. Similar to Doms et al. (2010), I do this by assigning industry groups with below the median share (across industry groups) of college equivalent workers in 1980 the "high-school-equivalent-hiring" designation and industry groups with above the median share the "college-equivalent-hiring" designation. Comparing Panels C to D, I find that effects are generally larger in industries that effects are larger in industries that tend to hire "higher skilled" workers. This comports with the classical immigration surplus argument (see, e.g. Borjas, 1999), in which inflows of primarily "low-skilled" (recall Figure 2) immigrants are more complementary with incumbent populations that are more educated.

Table C5: Heterogeneity in Key Local Industry Results by Industry Groupings

	Outcome: Δ log of			
	Estab Count	Employment	Payroll	Revenues
Panel A: Tradable (1st Stage F=17.36)				
Immigrant Worker Inflows (ΔI_{gk})	1.454***	1.233***	1.949***	1.739
	(0.393)	(0.380)	(0.712)	(1.138)
Panel B: Nontradable (1st Stage F=39.96)				
Immigrant Worker Inflows (ΔI_{gk})	1.017***	0.912***	1.562***	1.851***
	(0.328)	(0.191)	(0.367)	(0.514)
Panel C: High School Equiv. Hiring (1st Stage F=69.85)				
Immigrant Worker Inflows (ΔI_{gk})	0.714***	0.511***	0.753***	1.284***
	(0.191)	(0.144)	(0.203)	(0.344)
Panel D: College Equiv. Hiring (1st Stage F=10.54)				
Immigrant Worker Inflows (ΔI_{gk})	0.592**	0.941***	1.260**	2.018***
	(0.260)	(0.241)	(0.561)	(0.760)
Outcome Data Source	LBD	LBD	LBD	SBO/ABS
Outcome Long Difference Span	2000–2015	2000–2015	2000–2015	2002–2017

Notes: See Equation (1) for specification, estimated using 2SLS with instrumental variable Δz^M_{gk}. Conventional standard errors—robust to heteroskedasticity—are presented in parentheses. All models are estimated using 722 CZ × 41 Industry Groups = 29,602 Local Industries, weighted by their initial (2000) workforce size. All specifications include the control variables described in Section 2.2. Outcomes with Data Source "LBD" are measured using the Longitudinal Business Database, which is a full-count panel of establishments with near-complete coverage of the U.S. private sector. Outcomes with Data Source "SBO/ABS" are measured using repeated cross-sections from the 2002 Survey of Business Owners and 2017 Annual Business Survey, which are each approximately 10% representative samples of U.S. firms. See Section 2.2 for additional details on data sources.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D Supplemental Information on Employer Visa Analysis

D.1 Matching Visa Data to the LBD

Here, I describe the process of matching employer-level information on initial applications for a temporary worker to the Department of Labor under the H-1B and H-2B visa programs to the establishment-level Census Bureau data that I use for the primary analyses in this paper. Though I describe broad contours here, code will be made publicly-available upon publication or by request. Match rates have not yet been reviewed for disclosure by the Census Bureau, but compare favorably to match rates between Compustat firms and H-1B visa applications in the literature (see, e.g., Mayda et al., 2018). To my knowledge, I am the first to attempt this match for the H-2B visa data.

As stated in the text, the visa data comes from FLC Data Center for 2002–2007 and the Department of Labor for 2008–2017. The key variables contained in the visa data are: employer name, employer state, and employer city, and employer ZIP code. Employer ZIP code is missing for H-2B applications from 2002–2005. Using pre-processing commands described in Wasi and Flaaen (2015b) along with some additional corrections of common mistakes, I clean the names of employers. I then collapse the dataset to the name-state-city-ZIP level.

On the Census side, I link the LBD to the Business Registrar (CBPBR) using unique, within-year establishment identifiers. The Business Registrar also contains name, state, city, and ZIP information for employers. Notably, it includes two name fields and both mailing and physical address for the establishment. Because visa applications are filled out by employers, they may use either the physical or mailing address on their form. I therefore reshape the LBD-CBPBR dataset to have a unique observation for each employer's address. I perform the same pre-processing commands and collapse to the `lbdnum`-name-state-city-ZIP level. `lbdnum` is the longitudinal, unique, establishment-level identifier that enables all of the analyses in this paper.

The match proceeds in 6 steps, looping over states (implicitly requiring a match on state), using the `reclink2` command (Wasi and Flaaen, 2015a):

1. Exact matching on all four variables.

2. Exact match on ZIP, fuzzy match on employer name1 and city, with more emphasis on name

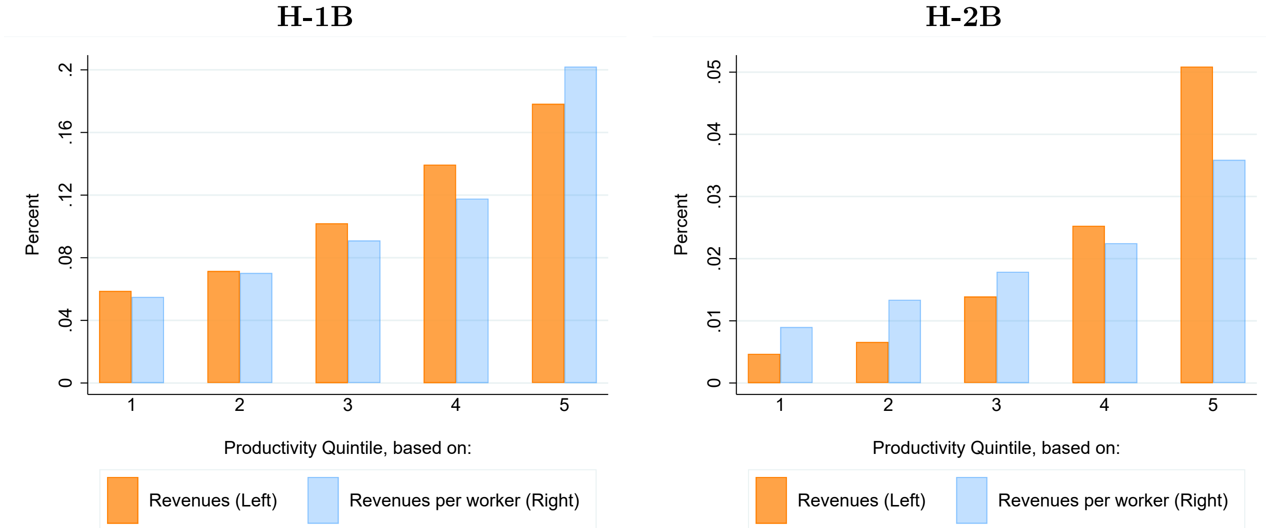
3. Exact match on ZIP, fuzzy match on employer name1 and city, with slightly less emphasis on name and a higher match score requirement

4. Fuzzy match on ZIP, employer name1, and city, with an even higher match score requirement.

5. Repeat Steps 2.-4. with employer name2

D.2 Results on Applications as a Share of Workforce

Figure D1: Employer Visa Applications as a Percent of Workforce, by Productivity



Notes: Each bar plots the average number of visa applications per worker in a given productivity quintile for a given visa program in a given year, averaged across the time period 2002–2017. For example, for firms in the top quintile of the revenue-per-worker based productivity distribution, the H-1B applications as a percent of workforce is 0.2 percent, on average (lighter blue bar in 5th Productivity Quintile of left panel). Meanwhile, for firms in the bottom quintile of the revenue-based productivity distribution, applications as a percent of workforce is roughly 0.05 percent (lighter blue bar in 1st Productivity Quintile of left panel), despite the fact that these firms have smaller workforces. See Section 4.2.1 for more details on each visa program and D.1 for more details on the matching procedure used to link firms that use each visa program to measures of productivity. Within each visa program and productivity measure, applications as a percent of workforce is smaller in the 1st, 2nd, 3rd, and 4th quintile compared to the 5th quintile—these differences are statistically significant at the 1% level in all cases.

E Additional Model Details and Results

For ease of reading, I list the key model equations here:

$$\text{[Pricing Rule]} \qquad p_j(z) = \left(\frac{\mu}{\mu-1}\right)\left(\frac{c_j}{z}\right) \tag{E1}$$

$$\text{[Technology Switching Point]} \qquad \pi_0(z_1^*) \equiv \pi_1(z_1^*) \tag{E2}$$

$$\text{[Zero-Profit Cutoff]} \qquad \pi_0(z_0^*) \equiv 0 \tag{E3}$$

$$\text{[Free Entry]} \qquad \mathbb{E}[\pi(z)] = \kappa_e \tag{E4}$$

$$\text{[Price Level]} \qquad P^{1-\mu} \equiv n_e \left[\int_{z_0{}^*}^{z_1{}^*} p_0(z)^{1-\mu} g(z)dz + \int_{z_1{}^*}^{\infty} p_1(z)^{1-\mu} g(z)dz\right] \tag{E5}$$

$$\text{[Income = Expenditure]} \qquad w_I I + w_N N = Y \tag{E6}$$

$$\text{[Labor Market Clearing]} \qquad I = n_e \left[\int_{z_0{}^*}^{z_1{}^*} I_0(z)g(z)dz + \int_{z_1{}^*}^{\infty} I_1(z)g(z)dz\right] \tag{E7}$$

Some additional notes regarding the equations that are not provided in the main text:

- In (E4), κ_e is a sunk (entry) cost potential entrepreneurs pay in order to take productivity draws. When average profits are high enough, entrepreneurs enter until expected profits equal this cost.
- In (E5) and (E7), $g(z)$ represents the Pareto PDF with shape parameter ϕ and minimum value m.
- In (E6), Y is total consumer spending
- In (E7), $I_j(z)$ represents immigrant hiring by a firm whose owner draws productivity z and chooses to produce with technology $j \in \{0, 1\}$.

E.1 Consumer Utility, Product Demand, and Firm Profits

A representative consumer has the utility function

$$\mathcal{U} = \left[F^{\frac{\eta-1}{\mu}} \int_0^F Q(f)^{\frac{\mu-1}{\mu}} df\right]^{\frac{\mu}{\mu-1}}$$

where F is the mass of operating firms in the economy and f indexes an individual firm from the consumer's perspective. Utility is maximized subject to (E6). So, for a firm whose owner draws productivity z and chooses technology j, we have the following expression for product demand:

$$Q_j(z) = YF^{\eta-1}P^{\mu-1}p_j(z)^{-\mu} \tag{E8}$$

Firm profits are

$$\pi_j(z) = p_j(z)Q_j(z) - \left(\frac{c_j}{z}\right)Q_j(z) - (\kappa_f + \mathbf{1}\{j=1\}\kappa_I)$$

Plugging in product demand, we get a key additional equation:

$$\text{[Firm Profits]}\ \pi_j(z) = p_j(z)^{1-\mu} Y F^{\eta-1} P^{\mu-1} \left(\frac{1}{\mu}\right) - \kappa_f - \mathbf{1}\{j=1\}\kappa_I \tag{E9}$$

E.2 Derivations and Proofs

E.2.1 Equilibrium Ordering of z_0^* and z_1^*

Here, I show that the equilibrium is correctly ordered under basic assumptions. That is, that $\pi_0(z) > \pi_1(z)$ for $z \in [z_0^*, z_1^*)$, $\pi_1(z) \geq \pi_0(z)$ for $z \in [z_1^*, \infty)$, and $z_1^* > z_0^*$.

Plugging (E3) into (E9), and (E1),

$$(z_0^*)^{\mu-1}(c_0)^{1-\mu}\left(\frac{\mu}{\mu-1}\right)^{1-\mu} Y F^{\eta-1} P^{\mu-1}\left(\frac{1}{\mu}\right) = \kappa_f \tag{E10}$$

Using (E2), (E9), and (E1),

$$(z_1^*)^{\mu-1}\left(\frac{\mu}{\mu-1}\right)^{1-\mu} Y F^{\eta-1} P^{\mu-1}\left(\frac{1}{\mu}\right)\left((c_1)^{1-\mu} - (c_0)^{1-\mu}\right) = \kappa_I \tag{E11}$$

Dividing (E11) by (E10),

$$\left(\frac{z_1^*}{z_0^*}\right)^{\mu-1} = \left(\frac{\kappa_I}{\kappa_f}\right)\left(\left(\frac{c_0}{c_1}\right)^{\mu-1} - 1\right)^{-1} \tag{E12}$$

So,

$$z_1^* > z_0^* \Leftrightarrow \left(\frac{\kappa_I}{\kappa_f}\right) > \left((c_1)^{1-\mu} - 1\right)$$

since $c_0 = 1$ under our normalization $w_N = 1$. Because $c_1 \equiv \left(a^{\sigma_j} {w_I}^{1-\sigma_j} + 1\right)^{\frac{1}{1-\sigma}}$ and $\sigma_1 > 1$, $c_1 < 1$. So, this condition is not trivial. However, it is always satisfied under plausible calibrations (see Section E.3).

Next, from (E9), (E1), and (E11)

$$\begin{aligned}
\pi_1(z) \geq \pi_0(z) &\Leftrightarrow p_1(z)^{1-\mu} Y F^{\eta-1} P^{\mu-1}\left(\frac{1}{\mu}\right) - \kappa_f - \kappa_I \geq p_0(z)^{1-\mu} Y F^{\eta-1} P^{\mu-1}\left(\frac{1}{\mu}\right) - \kappa_f \\
&\Leftrightarrow z^{\mu-1}\left(\frac{\mu}{\mu-1}\right)^{1-\mu} Y F^{\eta-1} P^{\mu-1}\left(\frac{1}{\mu}\right)\left((c_1)^{1-\mu} - (c_0)^{1-\mu}\right) \geq \kappa_I \\
&\Leftrightarrow z^{\mu-1} \geq (z_1^*)^{\mu-1} \\
&\Leftrightarrow z \geq z_1^*
\end{aligned}$$

So, firms do indeed switch to $j = 1$ technology when $z > z_1^*$ and use $j = 0$ technology otherwise. A similar exercise shows that $j = 0$ firms are profitable as long as $z > z_0^*$. So, we have the desired ordering and the equilibrium depicted in Figure 9.

E.2.2 Derivation of Zero-Profitability Cutoff, z_0^*: Equation (6)

Combining (E1) with (E4) tells us that the free entry condition is

$$Y F^{\eta-1} P^{\mu-1} \left(\frac{1}{\mu}\right) \underbrace{\left[\int_{z_0{}^*}^{z_1{}^*} p_0(z)^{1-\mu} g(z)dz + \int_{z_1{}^*}^{\infty} p_1(z)^{1-\mu} g(z)dz\right]}_{\equiv\left(\frac{1}{n_e}\right)P^{1-\mu}} = \int_{z_0^*}^{z_1^*} \kappa_f g(z)dz + \int_{z_1^*}^{\infty} (\kappa_f + \kappa_I) g(z)dz + \kappa_e$$

So, we have

$$Y\left(\frac{1}{\mu}\right)\left(\frac{1}{n_e}\right) = m^{\phi} \left(z_0^*\right)^{-\phi} \kappa_f \underbrace{\left(1 + \left(\frac{z_1^*}{z_0^*}\right)^{-\phi}\left(\frac{\kappa_I}{\kappa_f}\right)\right)}_{\equiv\theta} + \kappa_e$$

Defining $\theta \equiv \left(1 + \left(\frac{z_1^*}{z_0^*}\right)^{-\phi}\left(\frac{\kappa_I}{\kappa_f}\right)\right)$, we get

$$\frac{Y}{n_e\mu} = m^{\phi} \left(z_0^*\right)^{-\phi} \kappa_f \theta + \kappa_e \tag{E13}$$

Plugging (E1) into (E5) and solving, we get

$$P^{1-\mu} = \theta n_e F^{\eta-1} \left(\frac{\mu}{\mu-1}\right)^{1-\mu} \left(\frac{\phi m^{\phi}}{\phi - (\mu-1)}\right) (c_0)^{1-\mu} (z_0^*)^{-(\phi-(\mu-1))} \tag{E14}$$

Plugging (E14) into (E10), we get

$$\left(\frac{Y}{n_e\mu}\right) (z_0^*)^{\phi} = \theta\kappa_f \left(\frac{\phi m^{\phi}}{\phi - (\mu-1)}\right) \tag{E15}$$

And, plugging (E13) into (E15), we finally get

$$(z_0^*)^{\phi} = \theta m^{\phi} \left(\frac{\kappa_f}{\kappa_e}\right)\left(\frac{\mu-1}{\phi-(\mu-1)}\right)$$

which is Equation (6).

E.2.3 The Immigration Surplus: Equation (7)

Since $w_N = 1$, $\frac{d\log(w_N/P)}{dI} = -\frac{d\log(P)}{dI}$. so we focus on the price level P when thinking about the immigration surplus.

The mass of firms producing in the economy is given by

$$F = n_e \int_{z_0^*}^{\infty} g(z)dz \tag{E16}$$

$$= n_e m^{\phi} \left(z_0^*\right)^{-\phi} \tag{E17}$$

Plugging (E17), $c_0 = 1$, and (6) into (E14), we get

$$P^{1-\mu} = AF^{\eta}\left(z_0^*\right)^{\phi+\mu-1} \tag{E18}$$

where $A \equiv \left(\frac{\mu}{\mu-1}\right)^{1-\mu}\left(\frac{\phi}{\mu-1}\right)\left(\frac{\kappa_e}{\kappa_f}\right)m^{-\phi}$ is a constant that is a function of fixed, exogenous parameters. So,

$$\begin{aligned}\frac{d\log\left(w_N/P\right)}{dI} &= -\frac{d\log(P)}{dI} \\ &= \left(\frac{\eta}{\mu-1}\right)\frac{d\log(F)}{dI} + \left(1+\frac{\phi}{\mu-1}\right)\frac{d\log(z_0^*)}{dI}\end{aligned}$$

which is Equation (7).

E.2.4 Proposition 4.1: If $\frac{dw_I}{dI} < 0$, then $\frac{dz_0^*}{dI} > 0$

Proof. By Equation (6), $\frac{dz_0^*}{dI} > 0 \Leftrightarrow \frac{d\theta}{dI} > 0$. So, since $\frac{d\theta}{dI} = \frac{d\theta}{dc_1}\frac{dc_1}{dw_I}\frac{dw_I}{dI}$, we want to show that $\frac{d\theta}{dc_1}\frac{dc_1}{dwI} < 0$.

First, note that

$$\frac{dc_1}{dw_I} = (c_1)^{\sigma_1}a^{\sigma_1}w_I^{-\sigma_1} > 0$$

So, $\frac{dc_1}{dw_I} > 0$ and we now only need to show that $\frac{d\theta}{dc_1} < 0$.

Let $R_z \equiv \frac{z_1^*}{z_0^*}$. In Section E.2.1, we showed that $R_z > 1$. Then, Equation (E12) tells us

$$\begin{aligned}R_z^{\mu-1} &= \left(\frac{\kappa_I}{\kappa_f}\right)\left(c_1^{1-\mu}-1\right)^{-1} \\ \Rightarrow R_z^{\mu-2}\left(\frac{dR_z}{dc_1}\right) &= \left(\frac{\kappa_I}{\kappa_f}\right)\left(\underbrace{c_1^{1-\mu}}_{>1}-1\right)^{-2}c_1^{-\mu} \\ \Rightarrow \frac{dR_z}{dc_1} &> 0\end{aligned}$$

Meanwhile,

$$\begin{aligned}\theta &\equiv 1+\left(\frac{\kappa_I}{\kappa_f}\right)R_z^{-\phi} \\ \Rightarrow \frac{d\theta}{dc_1} &= -\left(\frac{\kappa_I}{\kappa_f}\right)\phi R_z^{-\phi-1}\underbrace{\frac{dR_z}{dc_1}}_{>0 \text{ from above}} \\ \Rightarrow \frac{d\theta}{dc_1} &< 0\end{aligned}$$

$\square$

E.2.5 Labor Market Equilibrium

To complete the model, we need to solve for relative immigrant wages, w_I. The supply of immigrant labor is exogenous and given by I. Since only $j = 1$ firms hire immigrant workers, (E7) can be simplified to:

$$I = n_e \left[\int_{z_1{}^*}^{\infty} I_1(z) g(z) dz \right]$$

So, we need to solve for $I_1(z)$ to determine w_I.

An individual firm whose owner draws productivity z and produces with technology j has the following first order conditions for each labor type:

$$a \left(\frac{I_1(z)}{N_1(z)} \right)^{-\frac{1}{\sigma_1}} = w_I$$

Using the definition of c_1, we can then re-write overall labor demand as a function of immigrant labor demand:

$$L_1(z) = I_1(z) a^{-\sigma_1} w_I^{\sigma_1} c_1^{-\sigma_1}$$

Since $L_1(z) = \frac{Q_1(z)}{z}$, using (E8) we have

$$I_1(z) = \left(\frac{\mu}{\mu - 1} \right)^{-1} a^{\sigma_1} w_I^{-\sigma_1} c_1^{\sigma_1 - 1} Y F^{\eta - 1} P^{\mu - 1} p_1(z)^{1-\mu}$$

Integrating, plugging in (E14), and plugging in the definition of θ, we arrive at our wage determination (labor market clearing) condition:

$$I = \left(\frac{\mu}{\mu - 1} \right)^{-1} a^{\sigma_1} Y c_1^{\sigma_1 - \mu} w_I^{-\sigma_1} \left(c_1^{1-\mu} - 1 \right)^{-1} \left(\frac{\theta - 1}{\theta} \right) \tag{E19}$$

E.3 Simulation Details

I simulate the model under different values of σ_1 and I, holding N fixed. The results revolve around Equation (E19), since the dynamics of the model flow out of $\frac{dw_I}{dI} < 0$, as seen in Proposition 4.1. To determine immigrant wages, we only need to solve the following relations:

$$I = \left(\frac{\mu}{\mu - 1} \right)^{-1} a^{\sigma_1} Y c_1^{\sigma_1 - \mu} w_I^{-\sigma_1} \left(c_1^{1-\mu} - 1 \right)^{-1} \left(\frac{\theta - 1}{\theta} \right) \tag{S1}$$

$$Y = w_I I + N \tag{S2}$$

$$c_1 = \left(a^{\sigma_1} w_I{}^{1-\sigma_1} + 1 \right)^{\frac{1}{1-\sigma_1}} \tag{S3}$$

$$\theta = \left(1 + (R_z)^{-\phi} \left(\frac{\kappa_I}{\kappa_f} \right) \right) \tag{S4}$$

$$R_z^{\mu - 1} = \left(\frac{\kappa_I}{\kappa_f} \right) \left(c_1^{1-\mu} - 1 \right)^{-1} \tag{S5}$$

once again using the normalization $w_N = 1 \Rightarrow c_0 = 1$.

E.3.1 Calibration

Solving Equations (S1)–(S5) require calibrations for a, μ, κ_f, κ_I, and ϕ. Using results from these simulations to further solve for key items of interest—like z_0^* and P—further require calibrations for m and κ_e. These key calibrations and their sources are given in Table E1.

Three of these parameters are calibrated independently. I use estimates of the average U.S. markup in 2000 from De Loecker et al. (2020) Figure 1 to calibrate μ. Axtell (2001) shows that the U.S. firm distribution closely follows a Power Law, and di Giovanni and Levchenko (2013) show that the Power Law parameter is an estimate of $\frac{\phi}{\mu-1}$ in Melitz frameworks. I therefore use the Power Law parameter estimates in Axtell (2001) to calibriate ϕ upon calibrating μ. m does not alter the dynamics of the model, but I use it to match realistic income levels.

Four key parameters are jointly calibrated within each simulation. I target relative productivity of immigrant workers at $j = 1$ firms—a—using the relative immigrant wage in 2000 from IPUMS-USA Census data. I target the fixed operating cost common to all firms using the overall firm mass in the economy divided by the labor force in 2000, tabluated in the Business Dynamics Statistics. I target the additional fixed cost $j = 1$ firms pay in order to hire immigrant workers using an estimate of the proportion of firms that hire immigrants from France, since this moment is not observable in my data and has not been publicly disclosed from employer-employee linked data in the U.S. Finally, I target the sunk entry cost entrepreneurs need to pay in order to discover their productivity level using the proportion of firms that started in 2000 and survived until 2015, from the Bureau of Labor Statistics' Business Employment Dynamics. The joint calibration step plugs values of μ, ϕ, m, I, N, σ_1, and w_I into Equations (S1)–(S5) and the following model relations:

$$\begin{aligned}
\frac{F}{I+N} &= Y\left(\frac{\phi-(\mu-1)}{\phi\mu}\right)\left(\frac{1}{\kappa_f}\right)\left(\frac{1}{\theta}\right) \\
\frac{F_1}{F} &= R_z^{-\phi} \\
\frac{F}{n_e} &= \left(\frac{\kappa_e}{\kappa_f}\right)\left(\frac{1}{\theta}\right)\left(\frac{\phi-(\mu-1)}{\phi\mu}\right)
\end{aligned}$$

I probe sensitivity of these calibrations in Section E.3.5.

Table E1: Calibration

Parameter	Range $\sigma_1 \in [2.5, 20]$	Value $\sigma_1 = 5$	Target Moment	Source
Panel A: Individually Calibrated				
μ	3.22	3.22	2000 U.S. Markup $\approx 45\%$	De Loecker et al. (2020)
ϕ	3.42	3.42	$\frac{\phi}{\mu-1} = 1.06$	Axtell (2001)
m	5.25	5.25	$w_N \approx \$25,000$	2000 Census IPUMS
Panel B: Jointly Calibrated				
a	[0.51,0.81]	0.65	$\frac{w_I}{w_N} = 0.87, w_N = 1$	2000 Census IPUMS
κ_f	[0.24, 0.33]	0.29	$\frac{F}{I+N} = 0.04$	Business Dynamics Statistics
κ_I	[0.02,0.16]	0.07	$\frac{F_1}{F} = 0.6$	Mitaritonna et al. (2017)
κ_e	0.75	0.75	$\frac{F}{n_e} = 0.26$	BLS Business Employment Dynamics

E.3.2 Model Dynamics

Figure E1 presents the results of simulations with $\sigma_1 = 5$ and $\eta = 0$ and I ranging from 0.05 to 0.2 in order to illustrate the dynamics of the model when immigration increases. The parameter values from this simulation can be seen in the third column of Table E1.

The dynamics are as described in the text: when immigrant wages fall, this lowers $j = 1$ labor costs but does not lower $j = 0$ labor costs. Through the pricing rule, $p_1 = \left(\frac{\mu}{\mu-1}\right)\left(\frac{c_1}{z}\right)$, $j = 1$ firms pass along these savings to consumers and compete the market away from $j = 0$ firms. This increases the productivity level $j = 0$ firms need in order to stay profitable. But, this increase in z_0^* raises aggregate productivity because it represents a rightward shift in the employer productivity distribution, and aggregate TFP is a weighted average from this distribution. Aggregate TFP increases translate to lower prices, so real native incomes rise substantially, and immigrant incomes do not fall as sharply (see Section E.3.4 for more on the latter point).

E.3.3 The Representative Firm Model for Comparison

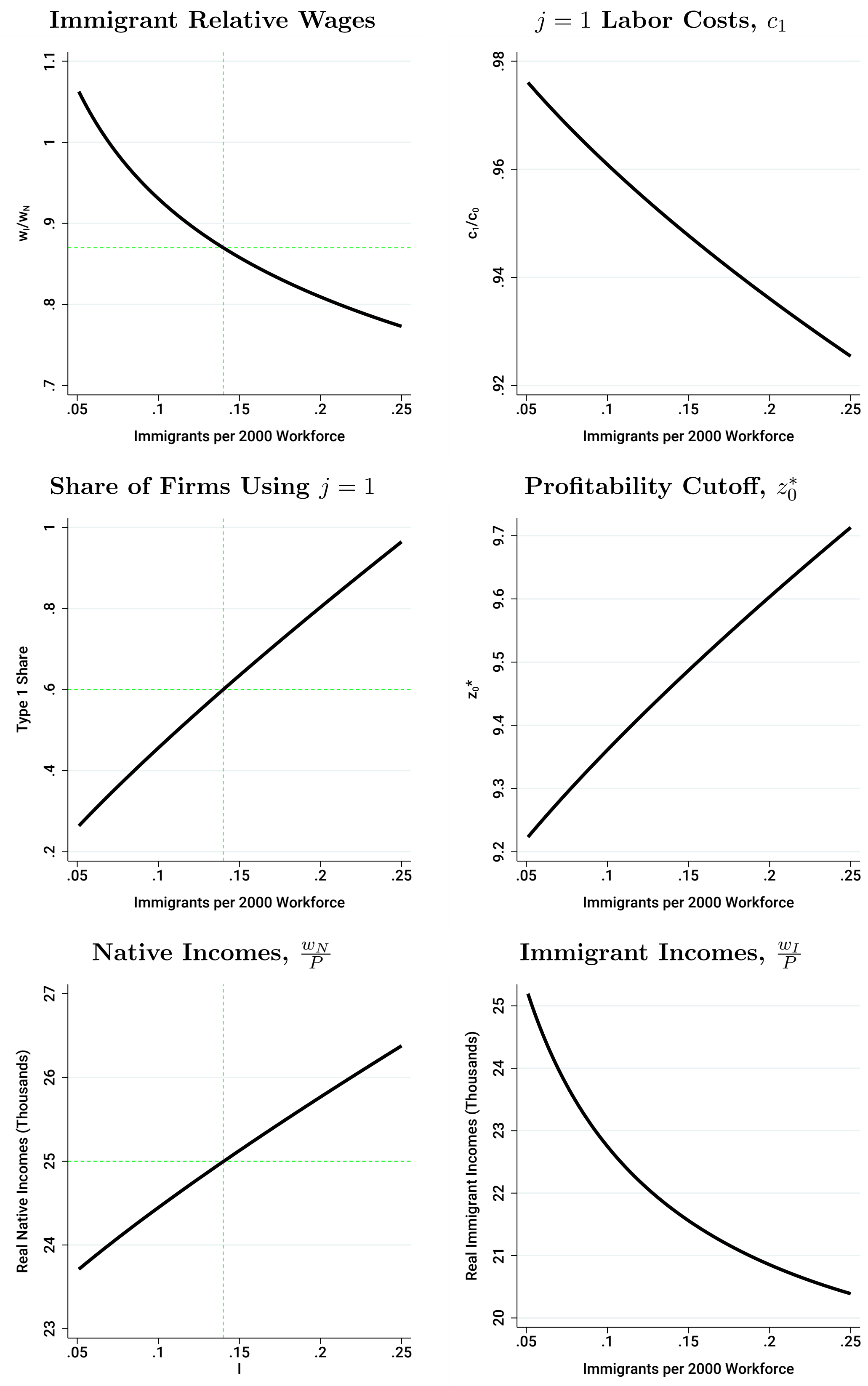
I benchmark results from the present model against the kind of "standard" representative firm model that has dominated the economic analysis of immigration. In this model, production for the representative firm is given by

$$Q = z\left(aI^{\frac{\sigma_1-1}{\sigma_1}} + N^{\frac{\sigma_1-1}{\sigma_1}}\right)^{\frac{\sigma_1}{\sigma_1-1}} = zL$$

where σ_1 is once again the elasticity of substitution across immigrant and native workers, but one that is shared by all firms in a perfectly competitive economy. Let prices be the numeraire. Under perfectly competitive labor markets, the representative firm's first order conditions are:

$$w_I = zL^{\frac{1}{\sigma_1}} aI^{-\frac{1}{\sigma_1}}$$
$$w_N = zL^{\frac{1}{\sigma_1}} N^{-\frac{1}{\sigma_1}}$$

Figure E1: Model Dynamics from Example Simulation with $\sigma_1 = 5$ and $\eta = 0$



Notes: Dashed green lines indicated targeted moments.

For a given value of σ_1 and once we set $I = 0.14$ and $N = 0.86$, we can easily calibrate a within each simulation. Then, up to z, w_I and w_N are fully determined. The key difference between the representative model and the model presented in Section 4 is that TFP, z, does not change in response to immigration in this model. So, when assessing the immigration surplus in this model—$\frac{d\log(w_N)}{dI}$—the choice of z is irrelevant:

$$\frac{d\log(w_N)}{dI} = \left(\frac{1}{\sigma_1}\right)\frac{d\log(L)}{dI}$$

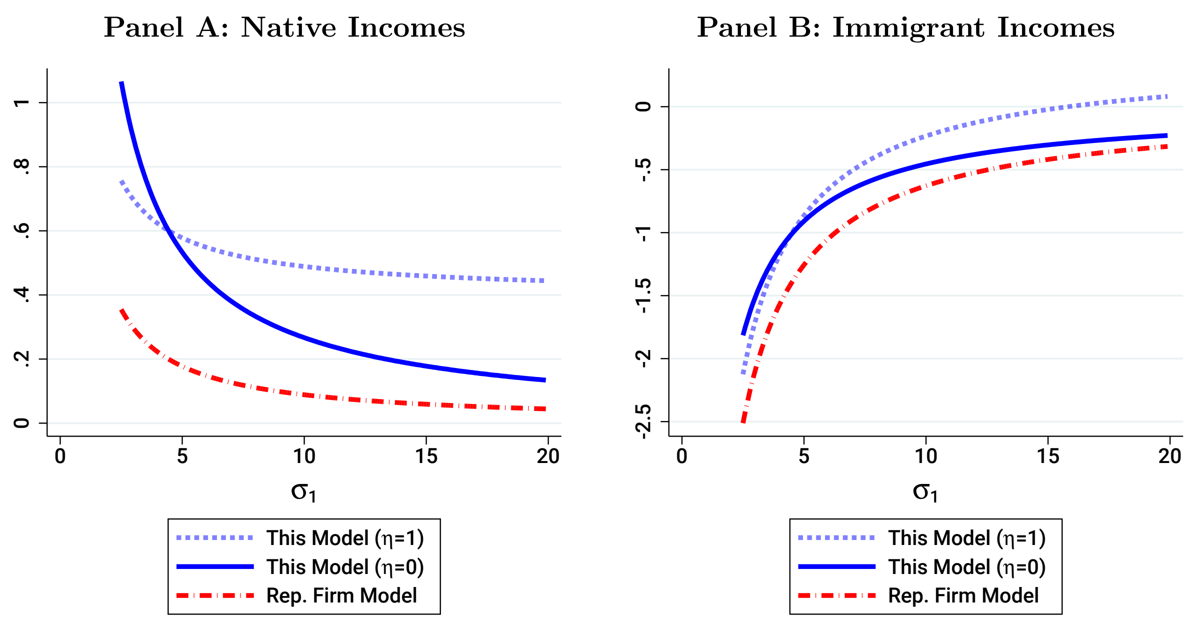
since the native stock is held fixed.

E.3.4 More on Figure 10 and the Impact on Immigrants

According to IPUMS-USA Census data, immigrants made up 14 percent of the workforce in 2000. I therefore estimate changes to incomes generated by small increases in I at $I = 0.14$.[63] In all simulations, our three key assumptions are met: $w_I > a$, $\frac{dw_I}{dI} < 0$, and $\left(\frac{\kappa_I}{\kappa_f}\right) > (c_1)^{1-\mu} - 1$.

Figure E2 focuses on the main results that emerge from this model. Its left panel replicates Figure 10, which plots estimates of the "immigration surplus"—$\frac{d\log(w_N/P)}{dI}$—calculated for small changes of I at $I = 0.14$ across different values of σ_1. It also adds the $\eta = 1$ case, which demonstrates that, insofar as they are valued, gains from can further amplify the immigration surplus (see Hong and McLaren, 2015). Panel B examines $\frac{d\log(w_I/P)}{dI}$, not shown in the main text. It finds that the aggregate productivity growth that stems from the rightward shift of the employer productivity distribution ($\frac{dz_0^*}{dI} > 0$) also buffers incumbent immigrant workers from some of the negative effects on income that stem from increased immigration.

Figure E2: Percent Increase in Incomes from a 1% Immigration Shock

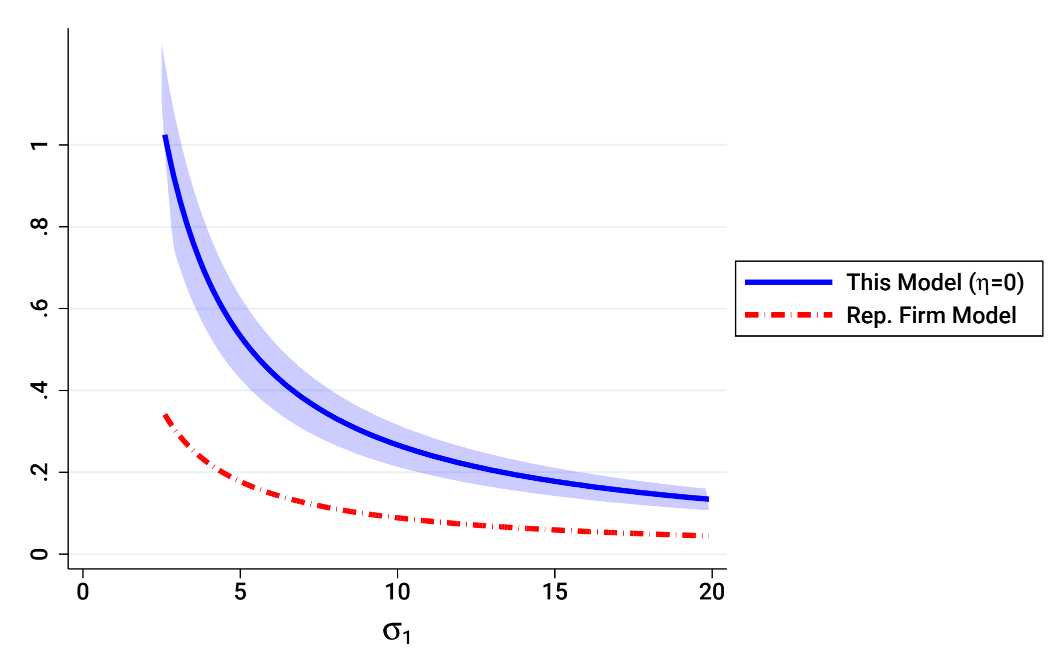


[63] In practice, from increases in I from 0.14 to 0.141.

E.3.5 Sensitivity

I probe sensitivity to two key calibrations whose data target moments do not have a strong consensus: μ and κ_e. Estimates of μ vary greatly across industry, time, and markup-estimation methodology. It is also unclear that the realized business survival rate in the data is an adequate approximation for $\frac{F}{n_e}$, given that entrepreneurs that draw low productivities do not literally start a business in the present model. I therefore conduct the same simulations as above, but that allow $\mu \in [2.5, 6]$ and $\frac{F}{n_e} \in [0.05, 0.85]$ using loops over each relevant target moment. I then reproduce Figure 10 with blue shaded areas around the immigration surplus estimate results with $\eta = 0$—our key item of interest—indicating the range of estimates across these simulations. These results can be seen in Figure E3. I find no evidence that these calibrations are driving the model results. For example, at $\sigma_1 = 5$, the lower bound of model-implied immigration surplus estimates are still more than twice the size of those from the benchmark representative firm model.

Figure E3: Percent Increase in Native Incomes from a 1% Immigration Shock—Sensitivity



Appendix References

Albert, Christoph, "The Labor Market Impact of Immigration: Job Creation versus Job Competition," *American Economic Journal: Macroeconomics*, January 2021, *13* (1), 35–78.

Autor, David H. and David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, August 2013, *103* (5), 1553–97.

Axtell, Robert L., "Zipf Distribution of U.S. Firm Sizes," *Science*, 2001, *293* (5536), 1818–1820.

Bartik, Timothy, *Who Benefits from State and Local Economic Development Policies?*, Upjohn Institute, 1991.

Bazzi, Samuel and Michael A. Clemens, "Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth," *American Economic Journal: Macroeconomics*, April 2013, *5* (2), 152–86.

Bogin, Alexander, William Doerner, and William Larson, "Local House Price Dynamics: New Indices and Stylized Facts," FHFA Staff Working Papers 16-01, Federal Housing Finance Agency 2016.

Bohn, Sarah, Magnus Lofstrom, and Steven Raphael, "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population," *Review of Economics and Statistics*, 2014, *96* (2).

Borjas, George J., "The Economic Analysis of Immigration," *Handbook of Labor Economics*, 1999, *3.*

— , "The labor supply of undocumented immigrants," *Labour Economics*, 2017, *46*, 1–13.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-Experimental Shift-Share Research Designs," *The Review of Economic Studies*, 06 2021, *89* (1), 181–213.

Burstein, Ariel, Gordon Hanson, Lin Tian, and Jonathan Vogel, "Tradability and the Labor-Market Impact of Immigration: Theory and Evidence From the United States," *Econometrica*, 2020, *88* (3), 1071–1112.

Card, David, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, *19* (1), pp. 22–64.

Chow, Melissa C, Teresa C Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elisabeth Ruth Perlman, Martha Stinson, and T. Kirk White, "Redesigning the Longitudinal Business Database," Working Paper 28839, National Bureau of Economic Research May 2021.

di Giovanni, Julian and Andrei Levchenko, "Firm entry, trade, and welfare in Zipf's world," *Journal of International Economics*, 2013, *89* (2), 283–296.

Doms, Mark, Ethan Lewis, and Alicia Robb, "Local labor force education, new business characteristics, and firm performance," *Journal of Urban Economics*, 2010, *67* (1), 61 – 77. Special Issue: Cities and Entrepreneurship.

East, Chloe N., Annie Laurie Hines, Philip Luck, Hani Mansour, and Andrea Velasquez, "The Labor Market Effects of Immigration Enforcement," *Journal of Labor Economics*, forthcoming.

Eckert, Fabian, Teresa C Fort, Peter K Schott, and Natalie J Yang, "Imputing Missing Values in the US Census Bureau's County Business Patterns," Working Paper 26632, National Bureau of Economic Research January 2020.

Fort, Teresa C. and Shawn D. Klimek, "The Effects of Industry Classification Changes on US Employment Composition," Working Papers 18-28, Center for Economic Studies, U.S. Census Bureau June 2018.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, March 2008, *98* (1), 394–425.

Hong, Gihoon and John McLaren, "Are Immigrants a Shot in the Arm for the Local Economy?," Working Paper 21123, National Bureau of Economic Research April 2015.

Jaeger, David A, Joakim Ruist, and Jan Stuhler, "Shift-Share Instruments and the Impact of Immigration," Working Paper 24285, National Bureau of Economic Research February 2018.

Klette, Tor Jakob and Zvi Griliches, "The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous," *Journal of Applied Econometrics*, 1996, *11* (4), 343–361.

Lewis, Ethan G., "Immigrant-Native Substitutability: The Role of Language Ability," Working Paper 17609, National Bureau of Economic Research November 2011.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications*," *The Quarterly Journal of Economics*, 01 2020, *135* (2), 561–644.

Mahajan, Parag, "Undocumented Immigration and Labor Market Fluidity: Evidence in the Context of Search Theory," 2017.

— **and Dean Yang**, "Taken by Storm: Hurricanes, Migrant Networks, and U.S. Immigration," *American Economic Journal: Applied Economics*, 2020.

Mayda, Anna Maria, Francesc Ortega, Giovanni Peri, Kevin Shih, and Chad Sparber, "New Data and Facts on H-1B Workers across Firms," in "The Roles of Immigrants and Foreign Students in US Science, Innovation, and Entrepreneurship," University of Chicago Press, October 2018, pp. 99–121.

Mitaritonna, Cristina, Gianluca Orefice, and Giovanni Peri, "Immigrants and firms' outcomes: Evidence from France," *European Economic Review*, 2017, *96*, 62 – 82.

Parsons, Christopher and Pierre-Louis Vézina, "Migrant Networks and Trade: The Vietnamese Boat People as a Natural Experiment," *The Economic Journal*, 01 2018, *128* (612), F210–F234.

Peri, Giovanni and Andri Chassamboulli, "The Labor Market Effects of Reducing the Number of Illegal Immigrants," *The Review of Economic Dynamics*, 2015, *18* (4).

— **and Chad Sparber**, "Task Specialization, Immigration, and Wages," *American Economic Journal: Applied Economics*, 2009, *1* (3).

Porter, Michael, "The Economic Performance of Regions," *Regional Studies*, 2003, *37* (6-7), 549–578.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "IPUMS USA: Version 9.0 [dataset]," 2019.

Wasi, N. and A. Flaaen, "Record linkage using Stata: Preprocessing, linking, and reviewing utilities," *Stata Journal*, 2015, *15* (3), 672–697(26).

Wasi, Nada and Aaron Flaaen, "Record Linkage Using Stata: Preprocessing, Linking, and Reviewing Utilities," *The Stata Journal*, 2015, *15* (3), 672–697.