The Labor Hub

Job Market Paper Series No. 5

Advisor: Todd Elder

DISCOUNTED DEGREES: EMPLOYER FRICTIONS AND THE MISALLOCATION OF HIGH-SKILLED IMMIGRANTS

Willy Chen

Discounted Degrees: Employer Frictions and the Misallocation of High-Skilled Immigrants

Willy Chen*

October 2026†

Abstract

Immigrant over-education in the United States is well documented, but its mechanism is not. I develop a signaling model in which an employer friction, together with rules that bar paying an immigrant less than a native in the same job, moves the adjustment from the wage to the job: immigrants keep their degrees and take jobs that require less education, by less at higher degrees. Using a new LinkedIn panel that follows college graduates across borders, I find that at their first US job, immigrant master's holders with a US degree and with only foreign degrees are 4.1 and 8.5 percentage points more likely to be over-educated than comparable US-educated workers. The gap is smaller for doctorate holders, and positive for those with a US degree at employers subject to the H-1B cap; it is largest, 16.0 points, for foreign-trained professionals such as physicians and lawyers, who must relicense. Over a decade, the occupational wage penalty specific to foreign degrees halves for doctorate and professional-degree holders but narrows by only a fifth for master's holders. The 2008 STEM Optional Practical Training extension, which lengthened work authorization without changing training, raised US career starts among 2008–2011 STEM master's graduates by 3.7 percentage points, and stayers moved into their field and into higher-paying occupations, mostly at employers subject to the H-1B lottery. The friction has more than one source, and part of it is a cost that policy can remove, not a productivity gap.

Keywords: Occupational Mismatch, Over-Education, Signaling, High-skilled Immigration
JEL: J15, J24, J31, J61, D82

*Michigan State University: willyc@msu.edu

†Acknowledgements: I thank Todd Elder, Alfonso Flores-Lagunes, Raoul Minetti, Stephen Woodbury, and Hanzhe Zhang for their continuing inspiration, advice, and support of this project. All errors are my own.

1 Introduction

Occupational mismatch, where a worker's schooling diverges substantially from the norm for their occupation, has been widely documented in the US labor market among immigrant workers (Bohn, 2010; Chiswick and Miller, 2009, 2013; Rendon and Cuecuecha, 2010). It is not confined to workers educated abroad. In the 2010–2019 American Community Survey, 42 percent of non-citizens who likely earned their highest degree in the United States work in occupations where their most common degree exceeds that of natives, against 13 percent where it falls short (Figure 16). In 2019, 21 percent of college-educated immigrants in the United States, about 2 million workers, were unemployed or in jobs requiring no more than a high school diploma, against 16 percent of US-born graduates; the rate was 24 percent among immigrants educated abroad and 17 percent among those educated in the United States (Batalova and Fix, 2021).

Extant research largely emphasizes the role of over-education and the associated wage penalties, yet the mechanisms behind it remain proposed rather than tested. This paper formalizes a leading mechanism by extending the canonical education signaling framework (Cho and Kreps, 1987; Mailath, 1987; Spence, 1978) to a hiring cost that decreases an immigrant's net productivity to the firm. In equilibrium, immigrants keep their degrees and sort into jobs that require less education. The question is therefore not whether immigrants are mismatched, which the literature has established, but how the friction is distributed across entry pathway, education attainment, and arrival cohort.

Introducing immigrant workers alters the classic signaling game through hiring costs: sponsorship reduces an immigrant's net productivity relative to an otherwise identical native worker. Crucially, firms observe which workers are foreign but cannot pay them less within an occupation due to prevailing-wage attestation and posted occupational scales. A firm free to discount an immigrant's wage would simply shift their schedule downward and leave every education level available to them; barred from doing so, it excludes them from the occupation instead.

Visa sponsorship is the prominent cost, but the friction is general: it stands equally for taste-based or statistical discrimination against foreign credentials and for language or cultural barriers. Because its size and persistence can differ by entry pathway, I separate Came-to-Study (CtS) workers, who arrived to study and completed a terminal post-secondary US degree, from Came-to-Work (CtW) workers, who entered the US labor market with only foreign credentials, and compare both with fully US-educated (FUS) workers, whose post-secondary degrees were all earned at US institutions.

The framework does not deliver an estimate of the friction, because the endpoints of the wage schedule are not observed. I instead measure the adjusted gap against FUS workers, who stand in for the native benchmark, and interpret a negative gap as consistent with a positive friction; its

size is informative about the friction only if the wage schedule is locally linear and selection is on observables. I place workers on the wage schedule using occupation-level wages and revealed-preference rankings of occupations and firms rather than individual pay (Section 3). The estimates describe immigrants who take US jobs; jobs taken abroad, which the panel records, lie off the US schedule, and whether the estimates over- or understate the friction faced by all entrants depends on how outside options correlate with it.

I report three findings on how displacement is distributed and a fourth showing that the friction can be moved. First, over-education at entry shrinks with education, as the model predicts. Relative to FUS workers with the same degree, institution tier, and field, CtS master's holders are 4.1 percentage points more likely to be over-educated at their first US job and CtW master's holders are 8.5 points more likely; pooled across employers, the doctorate gap is -0.1 points for CtS and -2.8 for CtW. The pool mixes in cap-exempt universities, which account for 37 to 40 percent of doctorate first jobs and where the H-1B lottery does not apply. At cap-subject employers, CtS doctorate holders are 2.0 percentage points more likely to be over-educated, about half the master's gap of 3.8, as the model predicts; CtW doctorate holders, at -2.1, are the exception, consistent with positive selection at entry and with their concentration in postdoctoral positions, which require a doctorate but pay little. Foreign-trained professional-degree holders, such as physicians and lawyers, have the largest gap, 16.0 points, which is consistent with a second friction after arrival: their credentials must be relicensed before they can be used in the occupations they trained for.

Second, over-education does not map directly into lower-paying occupations. CtS master's holders hold occupations that pay 11.9 log points more at the median than those of FUS master's holders, 5.6 points of it from the tasks of those occupations; CtW master's holders hold occupations that pay 7.8 log points less, and 10.1 less conditional on worker, job, and employer characteristics. These are differences in where workers sort, not in pay within an occupation.

Third, the part of the gap tied to foreign credentials fades for doctorate and professional-degree holders but not for master's holders. Workers in both pathways need sponsorship, but only CtW workers hold credentials that US employers cannot readily evaluate, so the distance between the two pathways isolates what a foreign credential adds. Among ten-year stayers, CtW workers start 12.8 to 17.2 log points behind CtS workers with the same degree in unexplained occupational median wages. Over the next decade the distance halves for doctorate and professional-degree holders, to 7 to 8 log points, but narrows only from 12.8 to 10.4 for master's holders. For professional-degree holders, the raw gap to FUS workers narrows less, from 9.4 to 6.3 log points. Both the incidence and the persistence of displacement therefore depend on how workers enter the US labor market and on the credentials they bring.

Fourth, as the model predicts, the friction responds to work-authorization policy. In April 2008, the US extended Optional Practical Training (OPT) for F-1 graduates in STEM fields from 12 to 29

months, delaying the point at which an employer must sponsor an H-1B visa; non-STEM graduates, whose friction was unchanged, are the comparison group. Because the rule changed the length of work authorization and not graduates' training, a pure productivity gap should not have moved placement. It did. For the 2008–2011 cohorts, the share of STEM master's graduates taking a first US job rose 3.7 percentage points relative to non-STEM CtS graduates, and by 2.6 points already for cohorts who enrolled before the rule; the increase grows to 6.2 points for the 2012–2015 cohorts. Five years out, STEM master's graduates still in the US were 7.6 percentage points less likely to work outside their field and held occupations paying 8 log points more; over-education did not change. Retention changed little, by -0.1 percentage points for these cohorts and 1.0 over all post-2008 cohorts at year five, so marginal stayers are at most one in eight of year-five stayers, too few to explain these gains. The gains are concentrated among graduates whose first employer was subject to the H-1B cap. Doctorate estimates are similar, but their pre-period coefficients are less stable, so I treat them as suggestive.

This paper contributes to three strands of literature, reviewed in Section 1.1. First, it formalizes the screening account of immigrant over-education (Chiswick and Miller, 2009): when wage floors bind, a group-specific friction acts on the job rather than on the wage (Lang and Manove, 2011), and the convexity of the wage schedule makes the over-education gap shrink with education. Second, because the panel records where every degree was earned (Friedberg, 2000) and follows fixed cohorts for ten years, it separates the part of the gap that fades from the part that persists. Third, where work on high-skilled visa policy measures effects on firms and on who stays, this paper measures whether graduates' jobs match their education. The distinction matters for welfare: if the mismatch reflects lower productivity, under-placement is efficient; if it is a cost created by visa rules, it is a misallocation of talent in the sense of Hsieh et al. (2019). A rule that changed only the length of work authorization moved both entry and placement, so at least part of the friction is of the second kind; Appendix H sets out the welfare loss it implies and why its size depends on what the friction represents.

Section 2 presents the model, Section 3 the data and occupation measures, and Section 4 the empirical strategy. Sections 5 to 7 report the friction at entry, over the first decade, and under the STEM-OPT extension. Section 8 concludes, and Appendix H discusses welfare.

1.1 Related Literature

The paper relates to three strands of literature: the mechanisms of over-education, the recognition of foreign credentials, and high-skilled immigration policy in the United States.

1.1.1 Immigrant Over-Education: Mechanisms and Consequences

Over-education is common and it is costly. Across 296 estimates from 39 countries, McGuinness et al. (2018) report an average over-education incidence of 21.5-25.9% and, across 74 estimates, an average wage penalty of 13.6%. They distinguish vertical mismatch, where attained schooling exceeds or falls short of what the job requires, from horizontal mismatch between a worker's field of study and occupation; I measure both. A recent survey names imperfect information about worker skill as one of the two primary sources of mismatch, alongside search frictions (Coraggio et al., 2026); the friction studied in this paper is of the first kind.

Over-educated workers earn less than adequately matched workers with the same schooling (Duncan and Hoffman, 1981; Verdugo and Verdugo, 1989). Whether the penalty persists matters as much as its size. Over-educated workers in the PSID move up more often, consistent with a stepping stone (Sicherman, 1991), whereas among Flemish school-leavers accepting such a job lowers the transition rate into an adequate one by 51 to 98 percent (Baert et al., 2013). For immigrants, a stepping stone predicts that the gap closes with US experience; a trap predicts that it does not.

Following work on the general labor force (Hartog, 2000; Kiker et al., 1997; Verdugo and Verdugo, 1989), studies of immigrants document larger and more persistent penalties (Chiswick and Miller, 2009, 2013; Li and Lu, 2023; Lu and Hou, 2020; Mattoo et al., 2008; Warman and Worswick, 2015), in the United States and in Europe (Aleksynska and Tritah, 2013; Nieto et al., 2013; Visintin et al., 2015; Weber et al., 2024). Cassidy and Gaulke (2024) report a 56 percent increase in the US immigrant over-education penalty between 1993 and 2019, and Lu and Li (2021) attribute a substantial share of nativity sorting gaps to educational mismatch. Using the 2000 Census, Chiswick and Miller (2008) find that immigrants and natives in occupations matched to their schooling earn the same return to it; immigrants' lower overall return comes mainly from under-education, which they link to favorable selection, whereas they link over-education to the imperfect transferability of foreign skills and leave open why so many immigrants work in occupations whose typical schooling is well below their own. Chiswick and Miller (2009) find that immigrants are more often over-educated the more work experience they acquired before migrating, which they attribute to US employers' uncertainty about the value of foreign experience, and less often over-educated the longer they have lived in the United States. They add a screening hypothesis to the search, human capital, and assignment accounts reviewed by Hartog (2000), but state it verbally rather than model it, and interpret their evidence as favorable selection that is realized only once immigrants overcome the imperfect transferability of their foreign skills. Whether over-education reflects unobserved skill gaps is also unsettled. Among UK graduates it partly does: graduates who are over-educated and dissatisfied with the match have a much lower skill endowment than matched graduates and a pay penalty of 22 to 26 percent, against 5 to 11 percent for those whose skills resemble matched graduates' (Chevalier, 2003; Chevalier and Lindley, 2009). In the US

PIAAC, however, first-generation immigrants are more often over-educated than natives, but over-educated workers score above average in literacy and numeracy, and numeracy does not predict the odds of mismatch once other characteristics are held fixed (Pivovarova and Powers, 2022). The immigrant-specific work establishes the size and persistence of the penalty without identifying why mismatched matches arise.

The closest precedent for a group-specific friction in a signaling model is Lang and Manove (2011): employers observe Black workers' productivity less precisely, so Black workers acquire more education than white workers of similar ability. In their model, competitive firms pay each worker their expected productivity given education and the noisy signal, so the friction operates entirely through the wage, and there are no occupations. In the context of immigrant workers, work visa prevailing-wage rules keep a firm from discounting an immigrant's wage within an occupation, so the friction cannot act on the wage and instead moves the worker into a job below their education. The terminal-degree pattern in Section 5, with a gap that is largest for master's holders and smaller for doctorate holders, is the evidence this strand has lacked.

1.1.2 Foreign Credentials, Assimilation, and Employer Learning

Where a credential was earned shapes what it is worth. Foreign schooling and experience are valued below their domestic equivalents in Israel (Friedberg, 2000); in the United States, immigrants with US schooling earn more and earn higher returns on their foreign schooling (Bratsberg and Ragan, 2002); and in a Toronto resume audit, foreign experience or Indian, Pakistani, Chinese, or Greek names lowered callbacks, and listing language fluency, multinational-firm experience, or education from highly selective schools did not close the gap (Oreopoulos, 2011). Immigrants who entered on student or temporary work visas outperform natives, largely through education and field (Hunt, 2011). The came-to-study/came-to-work comparison in this paper follows the same logic but asks how each group's job compares with its own education attainment over the career.

Language and cultural adaptation are the other leading account: immigrant earnings rise with English proficiency, especially in occupations that require it (Chiswick and Miller, 2010), and longer US residence reduces over-education (Chiswick and Miller, 2009). Cross-sectional evidence indexed by years since arrival, however, confounds assimilation with changes in cohort quality (Borjas, 1985), and selective emigration overstates assimilation among those who remain (Lubotsky, 2007). The assimilation estimates below condition on graduation cohort and restrict to workers who stay in the United States for ten years; they therefore describe stayers.

The task content of immigrants' jobs tells a related story. Male immigrants to Canada move from source-country occupations that require high cognitive and low manual skills into first Canadian jobs that require the opposite; the gap is small for immigrants fluent in English or French and narrows with time in Canada for the rest (Imai et al., 2019). In the United States, immigrants' jobs

require fewer analytical and interactive tasks and more manual tasks than natives' jobs, and the gap widened across arrival cohorts as occupational assimilation slowed, least so for immigrants with graduate degrees (Cassidy, 2019). Both studies point to language proficiency as a main source of the gap, a transferability deficit that should be small for fluent, highly educated immigrants and should narrow with time in the host country. The workers in this paper are highly educated, and many hold US degrees, so a gap that persists among them over ten years is hard to attribute to that deficit alone.

Employer learning offers a distinct account: wages come to depend less on schooling and more on productivity that employers learn over time (Altonji and Pierret, 2001), and how much there is to learn depends on the credential: for US college graduates, ability appears to be observed nearly perfectly at entry, while for high school graduates it is revealed gradually (Arcidiacono et al., 2010). Applied to immigrants, the gap should close fastest for the group whose credentials employers evaluate least accurately at entry. Comparing immigrants with foreign and with US degrees isolates that group: both need sponsorship, but only the first hold credentials that US employers cannot readily evaluate, so learning predicts that their path converges to that of immigrants with US degrees. Section 6 finds that it does for doctorate holders, whose distance halves over ten years, but barely for master's holders, whose distance narrows by a fifth. Professional-degree holders converge as fast as doctorate holders, but the share of them working in licensed occupations does not rise, so relicensing is unlikely to be the reason. Because the outcome is the median wage of the occupation held, it changes only when a worker changes occupation. The paths therefore measure occupational mobility. Employer learning can produce that mobility, as employers who learn a worker's productivity move them into better-paid occupations, but so can job search, the accumulation of US-specific skills, or the end of sponsorship once a worker obtains permanent residence.

1.1.3 High-Skilled Immigration Policy

Research on US high-skilled visa policy has focused on firms and natives. H-1B admissions raised immigrant science and engineering employment and patenting, with limited effects on natives (Kerr and Lincoln, 2010), and an additional H-1B visa crowds out about 1.5 other workers at the firm (Doran et al., 2022). The 2008 STEM-OPT extension raised the likelihood that international STEM graduates initially stay in the United States (Beine et al., 2023; Demirci, 2019) and foreign students' choice of STEM majors (Amuedo-Dorantes et al., 2019; Khoo, 2025); Guo et al. (2024) and Kwon et al. (2026) study its effects on natives. None of this work measures how well the jobs these graduates take match their education. Section 7 does, and interprets the extension as a fall in the hiring friction.

2 Conceptual Framework

The screening hypothesis for mismatch has been proposed by the broader educational mismatch literature (Chiswick and Miller, 2009; Leuven and Oosterbeek, 2011). However, to the best of my knowledge, no formal model has been introduced to formalize the main cause of immigrant over-education using a game of asymmetric information. Signaling has been used to explain over-education in the general population: when ability is imperfectly correlated with schooling costs, pooling equilibria can arise in which workers are over-educated (Ordine and Rose, 2009), and a worker's job assignment can itself signal ability to outside employers (Waldman, 1984). Neither model introduces a friction that applies to one group of workers and not another. In this section, I extend the canonical signaling model characterized with a unique separating equilibrium under regularity conditions (Cho and Kreps, 1987; Mailath, 1987; Spence, 1978). In this extension, immigrant workers are treated as additional players in the game after a separating equilibrium among natives has formed.

2.1 Separating Equilibrium of Natives

The classic signaling models of Spence (1978) and Mailath (1987) consider a labor market where there are two otherwise identical firms that compete for workers. There are a measure 1 of workers that can be hired.

- Each worker has a private innate productivity type $\theta \in \Theta \subset \mathbb{R}_+$.
- Each worker chooses an observable education $e \in \mathcal{E} \subset \mathbb{R}_+$.
- The cost of education attainment is represented by the function $C(\theta, e)$ such that $\frac{\partial}{\partial \theta} C < 0$, $\frac{\partial}{\partial e} C > 0$, $\frac{\partial^2}{\partial e^2} C > 0$, and $\frac{\partial^2}{\partial \theta \partial e} C < 0$.
- The production of a worker in a firm is represented by the function $F(\theta, e)$ such that $\frac{\partial}{\partial \theta} F, \frac{\partial}{\partial e} F, \frac{\partial^2}{\partial \theta \partial e} F > 0, \forall \theta \in \Theta$.[1]
- Firms simultaneously offer workers a wage W after observing e.
- The worker's utility function is $U(\theta, e, W) = W - C(\theta, e)$.
- The firm's profit function is $\pi(\theta, e, W) = F(\theta, e) - W$.

[1] The original model in Spence (1978) assumes that signals are costly and unproductive, i.e., $\frac{\partial}{\partial e} F(\theta, e) = 0$. However, this simple deviation from the original does not alter the result of the existence of separating equilibria (Mas-Colell et al., 1995).

Suppose that firms cannot perfectly observe each worker's type. Based on a worker's education level e, the firm makes a guess $\widehat{\Theta}(e) \subseteq \Theta$ on the worker's type and offers them $W = E\left[F(\theta, e) \mid \theta \in \widehat{\Theta}(e)\right]$. Let $\sigma : \Theta \to \mathcal{E}$ denote a worker's strategy for education attainment.

Definition 1 (Incentive Compatibility)**.** *The strategy σ is incentive compatible if*

$$\sigma(\theta) \in \underset{e}{argmax}\ U\left(\theta, e, W(e)\right),\ \forall \theta \in \Theta \tag{1}$$

While Spence (1978) shows that there exist multiple equilibria, both pooling and separating, Cho and Kreps (1987) shows that, with reasonable belief updates, only the most efficient separating equilibrium exists. In this separating equilibrium, the signaling function is a one-to-one function that satisfies incentive compatibility. Since σ is a one-to-one function, the competing firms can perfectly infer a worker's type θ based on their education. As such, the market permits a unique equilibrium wage schedule that is a one-to-one function mapping education level to a wage equal to the worker's implied productivity, i.e., $W(e) = F(\sigma^{-1}(e), e)$. Let $\bar{e} \equiv \sigma(\bar{\theta})$ denote the education the top type chooses in this equilibrium, so that a type-$\bar{\theta}$ native holds the contract $W(\bar{e})$.

Assumption 1 (Reasonable Beliefs of Education Signal)**.** *In equilibrium, a type-θ_w worker can only be inferred as θ_w if they choose education $\sigma(\theta_w) = e_w > \max\{\sigma(\theta) \mid \theta < \theta_w\}$, where σ is the incentive compatible separating equilibrium strategy that survives the Cho-Kreps intuitive criterion.*

Mailath (1987) proves that a continuous separating equilibrium exists given the environment laid out at the beginning of this section. The relevant theorems are provided in Appendix B and the graphical derivation of the equilibrium wage schedule is provided in Appendix B.1. Given that immigrant status such as foreign education and employment sponsorship needs are directly observable, I impose that a separating equilibrium exists for native workers, and extend the model to introduce immigrant workers as a new group of workers that adapt to the existing natives equilibrium wage schedule.

2.2 Extending the Model with Immigrant Workers

Suppose that the most efficient separating equilibrium exists among natives and that immigrants draw types from the same distribution Θ. A firm that hires an immigrant bears a hiring friction $d > 0$, so an immigrant of type θ with education e is worth $F(\theta, e) - d$ to the firm, against $F(\theta, e)$ for an otherwise identical native; hiring them into the job that pays $W(e)$ is profitable only if $F(\theta, e) - d \geq W(e)$. For intuition, let F be linear in e for each type, as drawn in Figure 1, and let $\sup \Theta = \bar{\theta}$ and $\inf \Theta = \underline{\theta}$. I characterize the contract of the top type $\bar{\theta}$ and then compare it with

that of a lower type. The analysis is a conceptual framework for the form displacement takes, not a model of the full distribution of immigrant placement.

The friction d is a generic employer-side friction: anything that lowers the net value of a foreign worker to the firm relative to an otherwise identical native. Its most directly observable component is visa sponsorship. Work authorization for high-skilled immigrants is rationed through capped programs and time-limited pathways, and sponsorship fees, legal costs, and lottery risk are closer to a fixed cost per hire than to a share of output, which is why d enters in levels. Lottery evidence makes the cost visible: Doran et al. (2022) compare firms that won and lost H-1B lotteries and find that an additional visa crowds out other hires, so sponsored workers are costly and imperfectly substitutable. Uncertain work authorization also lowers the firm's return to training the worker. The friction also covers discrimination, either taste-based (Becker, 1957) or statistical, when a foreign credential is a noisier signal of productivity (Altonji and Pierret, 2001; Friedberg, 2000). Finally, the gap in net value can reflect a genuine difference in the returns to foreign training, for instance from language barriers (Imai et al., 2019). Placement patterns alone cannot separate such a productivity difference from a pure cost; Section 7 separates them in part with a change in work authorization that moved the cost without changing training. The components also differ in persistence: statistical discrimination should attenuate as employers learn a worker's productivity, whereas taste-based discrimination and sponsorship costs persist until the worker no longer needs sponsorship, which Section 6 cannot separate with occupational-wage paths alone.

Assumption 2 (Stable Occupation-Level Posted Wages)**.** *The contract $W(e)$ is tied to the occupation rather than to the individual worker: a firm offering a position with equilibrium contract $W(e)$ pays $W(e)$ to whomever it hires and cannot adjust that wage downward for an immigrant.*

For H-1B workers the assumption is a legal requirement. An employer that sponsors an H-1B worker must attest in a Labor Condition Application that it will pay at least the higher of the prevailing wage for the occupation and area and the wage it pays other workers with similar experience and qualifications.[2] The certified offers are consistent with the rule binding: only 2 percent of sponsored offers in the occupations immigrant master's holders enter fall below the occupation's tenth wage percentile, against 10 percent for the occupation's own workforce (Section 5.4).[3] Workers in their first job on Optional Practical Training are not yet covered by this rule, but an employer that intends to keep them must sponsor them, so the prevailing wage binds from the H-1B petition onward and the employer anticipates it when it hires. For an OPT worker, the friction is the expected

[2]The attestation is tied to the position described in the petition, so an employer cannot certify a worker for one job and assign them to another. A material change in duties, worksite, or pay requires an amended petition, and the Fraud Detection and National Security Directorate of US Citizenship and Immigration Services verifies, through unannounced site visits, that the worker's actual duties, location, hours, and salary match the petition.

[3]Consistent with pay set by the job, immigrants and natives in occupations matched to their schooling earn the same return to it (Chiswick and Miller, 2008).

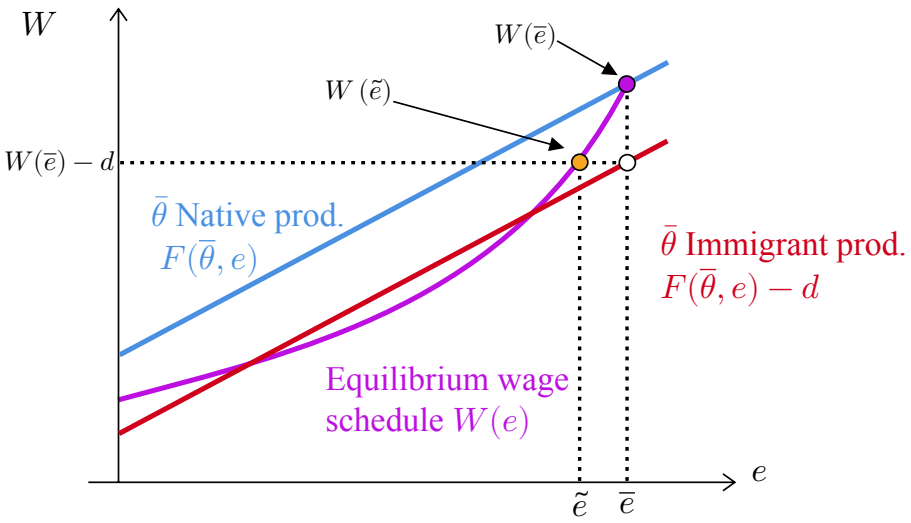
cost of that petition, borne at hiring, and the STEM-OPT extension studied in Section 7 lowers it by lengthening the period before the petition must be filed.[4] If the friction cannot be passed on through the wage, it must be absorbed by the job the worker is offered, which is what the rest of this section models.

Under Assumption 2, immigrants choose among the contracts on the natives' equilibrium wage schedule. By Assumption 1, a type-$\bar{\theta}$ immigrant must still choose $\bar{e}$ to be identified as $\bar{\theta}$, and has no reason to choose more. Their net value to the firm is then $F(\bar{\theta}, \bar{e}) - d = W(\bar{e}) - d < W(\bar{e})$, so no firm hires them into the job $W(\bar{e})$. Under Bertrand competition, firms bid for them until the job offered is worth their net value. The equilibrium contract is the under-placement contract

$$W(\tilde{e}) = W(\bar{e}) - d, \tag{2}$$

a job on the natives' schedule that is typically held by natives with education $\tilde{e} < \bar{e}$ (Figure 1).

Figure 1: Under-Placement of a Type-$\bar{\theta}$ Immigrant



Notes: The blue line is the type-$\bar{\theta}$ native's productivity schedule $F(\bar{\theta}, e)$ and the red line the type-$\bar{\theta}$ immigrant's net productivity $F(\bar{\theta}, e) - d$. The purple curve is the natives' equilibrium wage schedule $W(e)$; each point on it is a job typically held by a worker with education e, and a type-$\bar{\theta}$ native holds $W(\bar{e})$. A type-$\bar{\theta}$ immigrant still chooses $\bar{e}$ to separate but is worth $W(\bar{e}) - d$ to the firm. Because the firm cannot pay them less in the job $W(\bar{e})$ (Assumption 2), it offers the job whose wage equals their net value, $W(\tilde{e}) = W(\bar{e}) - d$, which is typical for education $\tilde{e} < \bar{e}$.

The immigrant holds $\bar{e}$ only to be identified as $\bar{\theta}$; the job itself is one that natives with education $\tilde{e}$ hold. The friction is output the firm loses: its net output from the immigrant is $F(\bar{\theta}, \bar{e}) - d = W(\tilde{e})$, against $F(\bar{\theta}, \bar{e}) = W(\bar{e})$ from an otherwise identical native, which is why the immigrant is offered $W(\tilde{e})$. Appendix H interprets the lost output d in a welfare context.

[4] Internal pay structures, under which a posted position carries one wage whoever fills it, point the same way. Since 2016, employers of STEM-OPT workers must also attest that the worker's pay is commensurate with that of similarly situated US workers, report material changes to the worker's training plan, and accept DHS site visits that check the job against that plan.

2.2.1 From Theoretical Friction to Observable Choices

The under-placement contract is identified not by the worker's education level but by the distance between their education $\bar{e}$ and the typical education of their job $\tilde{e}$: an immigrant who holds $\bar{e}$ and works at $W(\tilde{e})$ is over-educated relative to a native with the same credential.

The under-placement contract also predicts how the gap varies with education. Consider a lower type who chooses $\bar{e}' < \bar{e}$ and, as an immigrant, is placed in the job $\tilde{e}'$ with $W(\bar{e}') - W(\tilde{e}') = d$. The friction is the same amount at every education level. When the equilibrium wage schedule is convex, as in Figure 1, a given wage shortfall d spans a shorter stretch of the schedule at higher education, so $\bar{e} - \tilde{e} < \bar{e}' - \tilde{e}'$: the more educated the immigrant, the smaller the distance between their education and that of their job. The over-education gap between immigrants and natives should therefore be positive and shrink with the degree. Convexity is not automatic: along the separating schedule $W''(e) = C_{ee} + C_{\theta e}\,\widehat{\Theta}'(e)$, and a sufficient condition is $C_{ee}F_\theta \geq -C_{\theta e}\,(C_e - F_e)$, so the curvature of the cost of education must outweigh sorting. Appendix B states the conditions of Mailath (1987) under which the separating schedule exists. The friction need not be the same for every group of immigrants: a friction in evaluating foreign credentials raises d for workers with foreign degrees, and professional degrees, which must be relicensed in the United States before they can be used, raise it further for foreign-trained holders.

2.3 Comparative Statics: Lowering the Friction

A policy that lowers d for one group of immigrants, holding their types fixed, has two predictions in the model. The first concerns the contract available to a type-$\bar{\theta}$ immigrant who takes a US job. Because the under-placement contract satisfies $W(\tilde{e}) = W(\bar{e}) - d$, lowering d raises $\tilde{e}$ toward $\bar{e}$: the job they are placed in moves closer to their credential. For a given worker, the job moves closer to the credential. That can show up in over-education, in whether the worker works in their field, or in the pay and rank of the occupation; because over-education is measured in degree levels, a small fall in d can improve the job without moving the worker across a degree boundary.

When d falls because work authorization lasts longer, the jobs that open up are likely to be those that require on-the-job training. Following Becker (1962), firms and workers share the cost of firm-specific training so that neither party wants to end the match, and whether they share it depends on how easily they can agree on the worker's value once the training is done (Hashimoto, 1981); where matches last longer, earnings rise more steeply with tenure, consistent with more such training (Hashimoto and Raisian, 1985). An OPT period that expires or an H-1B lottery that is lost ends the match for reasons neither party controls, so jobs that require substantial training are those a firm is least willing to offer an immigrant. Longer work authorization lengthens the expected match and raises the firm's return to training, which moves the immigrant toward these jobs.

The second prediction concerns who takes a US job at all. The model as stated assumes that every immigrant does. Suppose instead that each graduate i also has a best foreign option worth v_i, and takes the US contract only if it is worth at least v_i. Let

$$V(d) = W(\tilde{e}) = W(\bar{e}) - d \tag{3}$$

denote the value of that contract. Because $\tilde{e}$ rises as d falls, $V(d)$ is decreasing in d, and the share of graduates who take a US job, $\Pr(v_i \leq V(d))$, rises when d falls. The graduates it draws in are those whose foreign option exceeded $V(d)$ before the change and falls short of it after: the graduates whose friction-distorted US offers were closest to their foreign options. If d varies across workers within the group, the graduates the change draws in need not hold the same jobs as those who would have stayed regardless, so the average job among stayers can move by more or less than the job of any worker who would have stayed. Section 4.3 separates the two.

3 Data and Measurement

The empirical analysis combines longitudinal career histories with external sources that assign common values to jobs and credentials. The core source is a dataset of worker-level records assembled by People Data Labs from publicly available LinkedIn profiles (People Data Labs, 2022). It reports workers' employment histories, educational records, employer names, job titles, and locations. I map reported job titles to six-digit Standard Occupational Classification (SOC) codes using a classifier trained on the LinkUp job-title taxonomy (LinkUp, 2025); the occupation-coding procedure is described below. Occupational median wages come from the Occupational Employment and Wage Statistics (OEWS) (Bureau of Labor Statistics, 2026), job characteristics from O*NET (National Center for O*NET Development, 2025), and H-1B sponsorship measures from Labor Condition Applications (U.S. Department of Labor, 2026). The American Community Survey (ACS) supplies an external benchmark for the coverage of the LinkedIn panel (Ruggles et al., 2025). Integrated Postsecondary Education Data System (IPEDS) admissions and completions records and QS rankings supply the institutional information used to construct education-selectivity measures (Quacquarelli Symonds, 2019; U.S. Department of Education, National Center for Education Statistics, 2019).

The underlying panel contains approximately 44 million individuals who either report US work experience or a post-secondary degree from a US institution and have at least one job that can be assigned a valid SOC. I retain workers with a post-secondary degree and sufficient educational history to classify the location of their first and last post-secondary degrees. The analysis distinguishes three mutually exclusive education pathways. Fully US-educated (FUS) workers completed all observed post-secondary education at US institutions. Came-to-Study (CtS) workers completed their first post-secondary degree outside the United States and their terminal degree at a US institution. Came-to-Work (CtW) workers completed all observed post-secondary education outside the United States. These labels describe observed education histories, not birthplace or citizenship. I use FUS workers as the empirical comparison group corresponding to the model's native benchmark; FUS therefore proxies for, but does not identify, nativity.

I group terminal post-secondary degrees into four levels, aligned with the Census classification: bachelor's, master's, doctorate, and professional. Professional degrees, such as the MD, JD, DDS, and PharmD, are degrees in fields that require a US or state license to practice. The doctorate group contains the PhD and other research doctorates, such as the EdD and the DMA, that do not lead to a license. The two groups involve similar amounts of schooling; what separates them is whether a third party must certify the holder before practice, the feature that adds to the friction for foreign-educated professionals in the model. Institutional names in the education histories permit a measure of the selectivity of the terminal-degree institution. I assign each institution to one of nine

selectivity tiers; unmatched institutions form a separate unranked category (tier 0). US institutions are tiered using an admissions-based IPEDS index incorporating undergraduate acceptance rates and test-score distributions, with tier 9 denoting the top 1% most selective institutions. For foreign institutions, which lack comparable admissions data, I assign the tier of the US institution with the closest average QS World University Ranking. The resulting measure enters only as flexible degree-by-tier controls. Appendix G describes its construction and coverage.

For labor market entry analyses, I collapse each worker's history to the first observed US job after the relevant degree sequence. The longitudinal analyses retain later US job-year observations and record departures from US employment when the profile reports subsequent employment abroad. Thus, an observed exit is not mechanically treated as missing follow-up, although profile updating after exit may be incomplete. The entry, dynamic, and STEM-OPT specifications apply further restrictions appropriate to their estimators; Table 2 reports the resulting coverage for the two principal estimation pipelines. Of 8.9 million first jobs of CtS and FUS graduate-degree holders, 4.2 million, or 47.7 percent, enter the entry analysis; of 29.4 million first jobs in the CtW–FUS pipeline, which also includes bachelor's degrees, 11.4 million, or 38.9 percent, do. The main losses come from the employer measures, chiefly firm utility, which exists only for employers in the connected set of worker flows between firms; the CtW pipeline also loses workers with missing arrival or pre-arrival information.

The final analytic samples impose common support for the covariates used in each specification, rather than a common restriction on workers' origins. Sample sizes therefore vary across outcomes as the availability of occupation, employer, and task measures differs. The principal first-job specifications retain large samples in both the Came-to-Study and Came-to-Work pipelines; Table 2 reports the sequential attrition associated with each covariate block. Estimates should consequently be interpreted for the covered population of credentialed LinkedIn users, rather than as population-representative estimates for all US labor-market participants.

Applying the model to data requires locating the equilibrium wage schedule, and individual wages do not suffice: realized pay reflects bargaining over match rents (Card et al., 2014; Mortensen and Pissarides, 1994), compensating differentials for non-pecuniary attributes (Rosen, 1986), and unobserved preferences (Lachowska et al., 2026), all of which generate dispersion at a given productivity-education pair.[5] Because each point on the schedule corresponds to an occupation rather than to a worker, I use occupation-level wages from the Occupational Employment and

[5]Recent structural and empirical work suggests the confound is quantitatively first-order rather than a second-order nuisance. Taber and Vejlin (2020) estimate a model that nests Roy selection, search frictions, and compensating differentials, and find that non-pecuniary compensating differentials account for a substantial share of the variance of wages alongside human capital. Consistent evidence comes from the search literature: Sullivan and To (2014) and Hall and Mueller (2018) recover dispersion in the non-wage value of jobs that is comparable to, or larger than, the dispersion in offered wages, while the discrete-choice experiment of Mas and Pallais (2017) shows workers forgoing sizable fractions of their wage for amenities such as scheduling flexibility.

Wage Statistics as the empirical wage schedule, which strips out the individual bargaining and preference heterogeneity that contaminates worker-level pay. Revealed-preference rankings of firms and of occupations recovered from worker mobility (Sorkin, 2018) absorb the non-pecuniary job value that occupation wages alone would not, and an occupational quantile index, which orders occupations by the education distribution of citizen workers, both measures mismatch and places each worker on the schedule. Occupational task content enters through the first three principal components of the O*NET skill and ability measures, described in Section 3.3.

3.1 The LinkedIn Panel

Because job titles are mapped to occupation codes by a fine-tuned large language model that reads titles in any language, the panel retains the career records of workers who leave the US labor market rather than dropping them. Attrition is thereby rendered observable and can be characterized directly instead of silently biasing the assimilation estimates (Figures 22 and 23). Nevertheless, due to the self-reported nature of LinkedIn profiles, workers who exited the US and chose not to reflect that on their profile are silently regarded as staying in the US. This process freezes their occupational outcomes at time of exit and is likely more prominent in more recent graduation cohorts.

3.1.1 Matching Job Titles to SOCs

I frame the matching process as a natural language classification task and fine-tune a Llama-3.1 (8B parameter) large language model to perform the categorization (*Llama 3.1 (8B)* 2024). The model is fine-tuned using a proprietary dataset provided by LinkUp, which contains thousands of distinct, real-world job posting titles for each of the 1,016 detailed ONET SOC codes.

Of 76,084,134 unique job titles, 74,280,970, or 97.6 percent, are matched to a valid SOC at the 8-digit level with an average confidence score of 0.83 out of 1. I then collapsed the SOC codes to the 6-digit level for better alignment with ACS data.

3.1.2 Non-random Selection into LinkedIn

The LinkedIn panel is a non-random subsample of the US workforce, exhibiting a well-documented selection bias toward highly educated, white-collar professionals. As shown in Table 3, the sample over-represents master's degree holders, who are 31.9 percent of FUS workers against 26.0 percent in the ACS, and is weighted toward recent graduation cohorts: 51.0 percent of FUS workers graduated between 2015 and 2021.

Admittedly, the non-random selection is a cause for concern. Appendix D provides a detailed tier-field coverage comparison between the LinkedIn data used in this paper and the IPEDS completions data. The main findings are that: (1) the representativeness of the data increases in institu-

tional selectivity and reputation, and (2) field representation skew towards corporate and technical fields.

3.2 Over-education and Horizontal Mismatch

Measurement of over-education generally follows one of three approaches: (a) job analysis, which relies on objective occupational requirements (e.g., O*NET classifications); (b) worker self-assessment, based on perceived educational needs; and (c) realized matches, which compare individual schooling to the distribution of education within an occupation (Allen and Velden, 2001; Chevalier, 2003; Chiswick and Miller, 2009; Hartog, 2000; McGuinness et al., 2018). Within the realized-match framework, over-education is often defined as schooling exceeding the occupational mean by one standard deviation (Verdugo and Verdugo, 1989) or as holding a degree above the modal level (Kiker et al., 1997). Fredriksson et al. (2018) refine the realized-match benchmark by comparing a worker's skills with those of tenured workers in the same occupation and firm, and find that poor match quality at entry lowers entry wages and predicts later job mobility.

Horizontal mismatch has received less attention than over-education (Somers et al., 2019); among US graduates, working in a job unrelated to one's major lowers earnings, and the penalty varies by field (Robst, 2007). I classify a job as horizontally matched when it is among the occupations that FUS graduates of the same degree and major most commonly hold. For each degree-by-major group, I order occupations by their share of the jobs held by FUS workers and take the smallest set that covers at least half of those jobs; an occupation tied at the margin is included. A share-based cutoff, unlike a fixed number of occupations, covers the same fraction of workers in every group, so the FUS mismatch rate does not mechanically depend on how widely a major's graduates spread across occupations. The shares count every job held over the career, not only first jobs, so that an occupation reached through promotion counts as a match. Figure 17 shows what the definition picks out. Within a degree-by-major group, the occupations outside the matched set pay 27 and 24 log points less at the median than those inside it for FUS bachelor's and master's graduates, and 15 log points less, imprecisely, for doctorate and professional-degree graduates, and they score 0.34 to 0.45 standard deviations lower on cognitive and interpersonal competence; they do not differ systematically in attention and coordination or in practical capabilities. Mismatched jobs are not uniformly worse: management, which accounts for 14 to 21 percent of mismatched jobs, pays about 30 log points more than the matched set, and excluding it leaves a wage gap of 13 to 20 log points. In this paper, I define and horizontal mismatch to be:

- **Over-education**:Worker i is over-educated in job j if their terminal degree is above the modal entry education requirement of occupation j in O*NET. All occupations are kept, and a worker whose degree is below the requirement is coded as not over-educated.

- **Horizontal mismatch**: Worker i is horizontally mismatched in job j if j is outside the smallest set of occupations that together hold half of the jobs of FUS workers with i's degree and major.

Because most occupations require at most a bachelor's degree for entry, over-education is common among FUS workers with graduate degrees: at the first job, 79.6 percent of FUS master's holders, 55.6 percent of FUS doctorate holders, and 22.9 percent of FUS professional-degree holders are over-educated. On the other hand, 37.4 percent of FUS master's holders are horizontally mismatched. The object of interest is therefore the gap between immigrants and FUS workers with the same degree, not the level. A postdoctoral appointment counts as a job, and its education requirement is a doctorate, as its title states, so a doctorate holder in a postdoctoral first job is never over-educated.

I do not separately classify under-education. Under-educated workers typically substitute experience, tenure, and on-the-job training for the credential, and they move between firms within an occupation rather than across occupations (Alba-Ramirez, 1993; Hartog, 2000; Leuven and Oosterbeek, 2011; Sicherman, 1991); Appendix E.8 reviews this evidence. Under this definition, the two types of mismatch can correlate but they are independent qualifiers of each education-occupation pair. In Section 5, I show that they are highly correlated yet they affect occupational sorting through different channels.

3.3 O*NET Job Characteristics

Consistent with the literature on worker skill development, mismatches initiated by various labor market frictions may instead be sustained by occupation-specific task requirements (Gathmann and Schönberg, 2010; Lise and Postel-Vinay, 2020). Guvenen et al. (2020) measure mismatch as the distance between a worker's bundle of skills and the skill requirements of the occupation, and find that it lowers wages and slows wage growth. Consequently, accounting for the specific task and skill demands of an occupation is essential when characterizing mismatch.

To characterize occupational skill requirements, I follow the broad literature utilizing O*NET Skills and Abilities (SA) data, simplified via Principal Component Analysis (PCA). Given the high dimensionality of these attributes, the extant literature typically simplifies the data through a two-step process: first, sorting SAs into pre-determined categories (e.g., cognitive, manual, or routine) and then performing PCA within each category to extract the first principal component (Autor et al., 2003; Cassidy, 2019; Guvenen et al., 2020; Imai et al., 2019; Lise and Postel-Vinay, 2020; Warman and Worswick, 2015).

In contrast, I adopt a more agnostic approach that utilizes the underlying data structure to determine the relevant skill clusters without ex-ante categorization. I perform a PCA on the level

scores of all 87 SAs simultaneously, where the first three principal components account for approximately 73% of the total variance. To label each component, I give a large language model, Microsoft Copilot, the O*NET descriptions of the 20 skills and abilities with the highest loadings on that component (Microsoft, 2025); the labels are descriptive and do not enter the estimation. Table A3 lists the highest-loading skills and abilities for the three components: (1) Cognitive and Interpersonal Competence (CIC); (2) Attention and Coordination (AC); and (3) Practical Job Capabilities (PJC).

3.4 Occupation and Firm Sorting

I construct three measures: a firm-level utility ranking, an occupation-level ranking estimated within degree-major strata, and a descriptive measure of how general a major or occupation is.

Firm PageRank utility I order employers by revealed preference following Sorkin (2018). A worker employed at firm k who receives an offer from firm j accepts if $V_j + \varepsilon_{ijk} > V_k + \varepsilon_{ikk}$, where V_j is a valuation of firm j common to every worker who might compare the two and ε_{ijk} is idiosyncratic taste. Under this choice rule, a firm that draws a disproportionate share of leavers from firms that are themselves valuable must be highly valued; the resulting valuation $\tilde{V}$ is the leading eigenvector of the firm-to-firm transition-share matrix, and the index used below, $V_j^e \approx \tilde{V}_j + \log(g_j)$, augments it with the log of firm size g_j, on the view that part of a firm's realized flow reflects scale of opportunity rather than worker preference. Appendix C states the construction, each assumption, and what its failure would do to the ranking. I estimate on five-year rolling windows and interpolate the rankings for 2020-2022. All regressions use the rankings in logs (Appendix E).

Within-education SOC PageRank utility The occupation ranking applies the same revealed-preference logic to a worker's choice of occupation rather than employer, but restricted to workers who share a highest degree and field of study, on the view that the relevant "market" over which occupational value is comparable differs by credential. Because a given stratum's worker flows can be thin, the estimated version adds a small damped uniform baseline (a damping factor of 0.85, per Brin and Page (1998)), so that every occupation retains some minimal comparability even absent a direct or well-populated chain of observed moves; Appendix C states the construction.

Major/SOC generality Independent of the rankings, the major-to-occupation network underlying the SOC stratification also reports a descriptive generality statistic, DUH, for every major and every occupation within a degree group (Chen, 2024). Applied to degree-majors, a high DUH_{major}

describes a versatile major whose graduates are spread broadly across occupations rather than funneled into one; applied to occupations, a high DUH_{SOC} describes an occupation with an open entry path, drawing from many different degree-majors rather than one or two feeder education credentials. Given that DUH is a non-standard statistic, I also use the Gini-Simpson index (Gini, 1912; Simpson, 1949) in a set of robustness checks.

4 Empirical Strategy

This section describes the estimating equations, sample selection, and identifying assumptions used in the three results sections that follow: mismatch and the friction at labor market entry (Section 5), the dynamics of the friction (Section 6), and the STEM-OPT extension (Section 7).

4.1 Identifying Occupational Mismatch at Labor Market Entry

The model's signature is in mismatch incidence rather than in wages. Under the under-placement contract, point (b) of Figure 4, the worker retains the credential $\bar{e}$ but accepts an occupation whose typical requirement is $\tilde{e} < \bar{e}$, so the signature is a probability of over-education higher than that of observationally similar natives at the same degree level, by less at higher degrees. Table 1 lists the predicted sign of the gap by degree and pathway, with $\bar{e}$ a doctorate and $\bar{e}'$ a master's degree.

I examine whether immigrant workers, at labor market entry, are more likely to experience occupational mismatch than natives by estimating the following linear probability model for person i at their first job j. For each mismatch outcome $m \in$ {over-education, horizontal mismatch}, I estimate separately by immigration pathway with progressive controls added:

$$
\begin{aligned}
Y_{ij}^{m} = \alpha^{m} + \beta^{m} P_i + \sum_{e \neq \text{MA}} [\eta_e^{m} \mathbb{1}\{E_i = e\} + \beta_e^{m} P_i \mathbb{1}\{E_i = e\}] + \mu_{\text{grad}(i)} + \varepsilon_i^{m} \qquad (4)\\
+ \sum_{e,\tau} \kappa_{e\tau}^{m} \mathbb{1}\{E_i = e\} \times \mathbb{1}\{T_i = \tau\} \qquad Col.(2)\\
+ \lambda_1^{m} \text{STEM}_i + \lambda_2^{m} DUH_{f(i)} + \delta_{e,f(i)}^{m} \qquad Col.(3)\\
+ \phi^{m} F_j + \rho^{m} CE_j + \psi^{m} P_i \times CE_j \qquad Col.(4)\\
+ \xi^{m} M_{ij}^{-m}, \qquad Col.(5)
\end{aligned}
$$

where P_i denotes the CtS (or CtW) pathway and FUS workers are the omitted comparison pathway. E_i is terminal-degree category, indexed by e, the empirical counterpart of the education choice e in the model, with master's degree as the reference, and $\mu_{\text{grad}(i)}$ denotes the graduation cohort fixed-effects. T_i is terminal-institution tier. Tier effects are allowed to differ by degree but are common across pathways, consistent with the model's assumption that institutional quality carries the same signal regardless of where a worker was educated. $f(i)$ is the coarse field of worker i's terminal major: STEM_i indicates a major on the DHS STEM list, and $DUH_{f(i)}$ is the generality of the major described in Appendix E. $\delta_{e,f(i)}^{m}$ are degree-by-major-field fixed effects, so that Column 3 compares an immigrant with FUS workers who hold the same degree, from an institution of the same tier, in the same field of study. I take Column 3 as the main estimate, because every control in it is fixed before the worker enters the labor market. The object of interest is the gap for degree

e, $\beta^m + \beta^m_e$; for master's holders it is β^m. F_j is a vector containing the logged firm Sorkin score, the firm's 3-, 6-, 12-, and 24-month churn rates, and an indicator for whether the firm filed any Labor Condition Application between 2011 and 2025. CE_j indicates that the employer is a higher-education institution, which is exempt from the annual H-1B cap. M^{-m}_{ij} is the complementary mismatch measure: horizontal mismatch when Y^m_{ij} is over-education, and over-education when it is horizontal mismatch; Column 5 is estimated only for these two outcomes. For CtW specifications, I additionally include non-US experience and its square, an indicator for a same-employer transfer, arrival-cohort fixed effects, and region fixed effects.

The controls in Columns 2 and 3 are fixed before the worker enters the labor market. I do not condition on the characteristics of the occupation the worker takes, because both mismatch measures are defined by that occupation. Major generality stays because it describes the major rather than the job.

Column 4 is the exception: the employer is chosen in the same search that produces the mismatch. I include F_j and CE_j to test one alternative explanation: immigrant workers accept mismatched jobs in exchange for better or visa-sponsoring employers, a compensating differential rather than a friction. The test is conservative: if the friction steers immigrants toward sponsoring employers, the immigrants who reach high-utility firms are positively selected, which biases the Column 4 gap toward zero. Cap-exempt employers may require separate treatment. Institutions of higher education, their affiliated nonprofits, and nonprofit and government research organizations are exempt from the annual H-1B cap: they can file petitions at any time of year, and their hires face no lottery.[6] Column 5 conditions on the complementary mismatch, which is a joint outcome of the same placement; I interpret it descriptively, as whether the two mismatch types move together.

The first job is the first employed, non-intern job whose spell starts in or after the year of the terminal degree, within five years of it; for CtW workers, whose degrees are foreign, the five years run from their first year in the US labor market. A job carried over from school, such as a graduate teaching or research assistantship, does not count, nor do volunteer work, or "PhD student" titles, and a worker who only continues a pre-degree job has no first job. I keep first jobs located in the United States, because both mismatch measures are scored against US occupational benchmarks, and graduates of the 2003–2021 cohorts. The US restriction removes graduates who begin their

[6] 8 U.S.C. §1184(g)(5)(A)–(B). My indicator CE_j covers higher-education institutions only, so a small number of cap-exempt employers may be counted as cap-subject. At these employers the lottery, one source of the friction, is absent. Doctorate holders start their careers there far more often than master's holders, because postdoctoral and faculty positions are concentrated at universities: higher-education institutions account for 37 to 40 percent of the first jobs of doctorate holders in each pathway, against 2 to 7 percent of those of master's holders. A postdoctoral appointment also requires a doctorate, so its holder is never over-educated. Pooling the two types of employer therefore mixes, for doctorate holders, jobs where the lottery applies with jobs where it does not. Column 4 includes $P_i \times CE_j$, so its pathway coefficients measure the gap at cap-subject employers, but because ψ^m is common to all degrees, it imposes the same cap-exempt shift on master's and doctorate holders. I therefore also re-estimate Column 3 separately at cap-subject and cap-exempt first employers, which lets every coefficient differ by employer type.

careers abroad, so the estimates describe immigrants who begin their careers in the US. All columns are estimated on one sample: workers with non-missing values of every covariate in the richest column.

I then compare the utility and pay of the occupations immigrants and FUS workers hold. Because the question is about pay given the type of job, the occupation's characteristics belong in this specification:

$$
\begin{aligned}
Y_{ij} = \alpha + \beta P_i + \sum_{e \neq \text{MA}} \left[\eta_e \mathbb{1}\{E_i = e\} + \beta_e P_i \mathbb{1}\{E_i = e\}\right] + \mu_{\text{grad}(i)} + \varepsilon_i \qquad (5)\\
+ \zeta_1 OE_{ij} + \zeta_2 HM_{ij} + \zeta_3 OE_{ij} \times HM_{ij} \qquad Col.(2)\\
+ \sum_{e,\tau} \kappa_{e\tau} \mathbb{1}\{E_i = e\} \times \mathbb{1}\{T_i = \tau\} \qquad Col.(3)\\
+ \lambda Z_j + \omega_{o(j)} \qquad Col.(4)\\
+ \phi \ln(\text{Firm Sorkin}_j) \qquad Col.(5)\\
+ \rho CE_j + \psi P_i CE_j, \qquad \text{Last col.}
\end{aligned}
$$

with $Y_{ij} = \ln(\text{SOC PageRank}_j)$ and $Y_{ij} = \ln(\text{OEWS median wage}_j)$. Z_j contains the three standardized O*NET principal components and the standardized importance of certification, and $\omega_{o(j)}$ are fixed effects for the occupation's on-the-job training category. For the median wage regressions, $ln(\text{SOC PageRank}_j)$ is added as a control variable in an additional column to account for occupational prestige sorting. Unlike Column 3 of Equation (4), these specifications include degree-by-tier effects but not degree-by-field fixed effects, and for CtW workers every column also absorbs arrival-cohort and origin-region fixed effects and controls for non-US experience and a same-employer transfer, so their first columns are not raw gaps.

4.1.1 Using Oaxaca-Blinder Decomposition to the Effects of d

The type-of-job interpretation of the equilibrium wage schedule allows for minimal assumptions on which occupations belong in the same labor market. As such, instead of testing the model with individual-level wages, it is sufficient to compare the wage distributions of the occupations workers choose.

For each pathway $p \in \{\text{CtS}, \text{CtW}\}$ and terminal degree e, I decompose the immigrant–FUS gap in the residualized log median OEWS wage of the occupation held:

$$
\bar{Y}_{pe} - \bar{Y}_{FUS,e} = \underbrace{\left(\bar{X}_{pe} - \bar{X}_{FUS,e}\right)\hat{\beta}_{FUS,e}}_{\text{Explained}} + \underbrace{\bar{X}_{pe}\left(\hat{\beta}_{pe} - \hat{\beta}_{FUS,e}\right)}_{\text{Unexplained}}. \qquad (6)
$$

The decomposition is estimated separately for each graduate degree level. The covariate vector X_{ipe} contains degree-by-terminal-institution-tier indicators, field indicators, STEM status, an indicator for graduation after 2008 and its interaction with STEM, terminal-major generality, experience, and experience squared. For CtW workers it additionally contains non-US experience and its square and an indicator for a same-employer transfer.

$$\begin{aligned} X_{ipe} = \sum_{\tau} \kappa_{e\tau} \mathbb{1}\{T_i = \tau\} + \gamma_1 \text{Field}_i + \gamma_2 \text{STEM}_i + \gamma_3 \text{Post2008}_i \\ + \gamma_4 (\text{STEM}_i \times \text{Post2008}_i) + \gamma_5 \text{DUHMajor}_i + \gamma_6 \text{Exp}_i + \gamma_7 \text{Exp}_i^2 \\ + \mathbb{1}\{p = \text{CtW}\} \left[\gamma_8 \text{NonUSExp}_i + \gamma_9 \text{NonUSExp}_i^2 + \gamma_{10} \text{Transfer}_i \right]. \end{aligned} \tag{7}$$

For CtS, Y residualizes log OEWS median wages by terminal-degree completion cohort; for CtW, it residualizes by arrival cohort and origin region. The specification deliberately excludes occupation, firm, mismatch, and task characteristics because these may themselves be outcomes of the hiring friction. I refer to the unexplained component as the adjusted placement gap, U_{pe}. Under the underplacement contract, $W(\tilde{e}) = W(\bar{e}) - d$, so a positive d predicts a negative U_{pe}; a positive U_{pe} is an adjusted premium, which the model does not predict.

4.1.2 The Hidden Drivers of the Median Wage Gaps

I use the decomposition formalized by Gelbach (2016), which quantifies the unambiguous contribution of distinct covariate blocks to the raw occupational sorting gap between immigrant workers and FUS workers, avoiding the path-dependency inherent in sequential addition of controls. Figures 5 and 6 present the decomposition results for Came to Study (CtS) and Came to Work (CtW) workers across educational credentials, respectively. The covariates are grouped into eight blocks: degree major, worker characteristics, work location, English-language origin, job characteristics, firm sorting, job sorting, and occupational mismatch; Tables A5 and A6 report the contribution of each covariate within them.

4.2 Examining the Dynamics of the Hiring Friction

The components of d differ in persistence: the part that reflects employers' uncertainty about foreign credentials should fade as they learn a worker's productivity, whereas sponsorship costs and taste-based discrimination should not, so the path of the adjusted gap over a worker's career bears on what d contains. To examine how the hiring friction evolves over time, this subsection focuses on estimating the dynamics of the friction in Equation (6): tracking the adjusted placement gap year by year across workers' first decade in the US labor market. Specifically, I estimate U_{pet}^{stay}

from entry into the US labor market through year 10 for long-term stayers, workers who remain in the US labor market through their first ten years, relative to their FUS counterparts. I report the paths from year one: in the year of entry, many workers have not yet started their first post-degree job, so the gap at $t = 0$ reflects when workers start rather than which occupations they hold.[7]

For each pathway p, terminal-degree group e, and year $t \in \{0, \dots, 10\}$ since entry into the US labor market, I re-estimate Equation (6) on the job observations held at t:

$$\bar{Y}_{pet} - \bar{Y}_{FUS,et} = \underbrace{\left(\bar{X}_{pet} - \bar{X}_{FUS,et}\right)\hat{\beta}_{FUS,et}}_{\text{Explained}} + \underbrace{\bar{X}_{pet}\left(\hat{\beta}_{pet} - \hat{\beta}_{FUS,et}\right)}_{U^{\text{stay}}_{pet}}, \tag{8}$$

with the same pre-market covariates as Equation (6). The sample is restricted in three ways. Workers must have graduated in 2003 or later, so that every year of their career is priced with realized rather than back-filled OEWS wages. They must be in the United States ten years after entry. At each t, only US job observations enter, so the entire trajectory compares US jobs. For CtS and FUS workers, entry is the year of the terminal degree, so t counts years since graduation; CtW workers hold only foreign degrees, so for them t counts years since their first year in the US labor market, and the FUS benchmark keeps its own years since graduation.

One issue complicates this exercise: the model assumes that all immigrants can and want to find jobs in the US, while, in reality, many seek opportunities elsewhere, leading to attrition bias. Figure 22 shows that a large share of CtS graduates at every degree level leave the United States at graduation and that, among those who stay, annual exit hazards are low and fall with tenure. CtW workers exit at higher rates (Figure 23).

Focusing on within-cohort dynamics directly addresses classic cohort-composition critiques of cross-sectional estimates (Borjas, 1985). Because the year-by-year decompositions condition on terminal-degree graduation cohorts, temporal variation in U^{stay}_{pet} reflects within-person trajectories rather than shifts in entry-cohort quality over time. The stayer restriction limits, but does not remove, the remaining threat of compositional change across horizons: every worker in the sample is in the United States at year 10, but a worker enters the estimate at year t only if a US job is observed in that year, so employment gaps and temporary spells abroad still change who is observed at each t. The cost of the restriction is that U^{stay}_{pet} describes workers who stay for ten years, not all entrants; if lower-earning immigrants are the ones who leave, as Lubotsky (2007) finds, the stayer path is that of a positively selected group.

Section 4.3 faces the same selection problem in a different form. There, the policy itself changes

[7] The decomposition includes an indicator for holding the first job. At $t = 0$ it accounts for -11.8, -12.0, and -7.8 log points of the unexplained gap for CtS master's, doctorate, and professional-degree holders and $+4.2$, $+10.4$, and -1.9 for their CtW counterparts; from $t = 1$ on, its contribution is within one log point of zero in every group. The tables report $t = 0$ for completeness.

who stays, so instead of fixing the population I decompose the estimate into an effect on workers who would have stayed regardless and a composition term.

4.3 Difference-in-Differences: STEM-OPT Extension

Isolating the causal effect of friction requires a shock that changes the cost of hiring a foreign worker without changing worker quality. The obstacle is that location decisions are endogenous and driven by traits unobservable to the econometrician. The LinkedIn profile data speaks to that obstacle in one specific way: it retains the international trajectories of foreign talent who exit the United States, though selection into observable departure depends on post-exit profile updates.

The April 2008 interim rule extended Optional Practical Training (OPT) work authorization by 17 months, from 12 to 29 months, for F-1 graduates of STEM degree programs. It lengthened the low-hiring-cost time US employers have before needing to sponsor an H-1B for a STEM graduate, while leaving the analogous friction facing non-STEM foreign graduates unchanged. Since H-1B petitions are filed once a year against the cap, the extension gives a STEM graduate at least one additional filing season before work authorization lapses, which lowers the risk of involuntary attrition that an employer prices into the initial hiring decision, including its decision to train the worker (Hashimoto, 1981).[8]

Comparing STEM to non-STEM graduates, before and after 2008, therefore identifies the consequences of a reduction in hiring friction via difference-in-differences, with non-STEM graduates serving as the counterfactual trend. The two predictions of Section 2.3 map onto two margins: the share of graduates whose first job is in the United States, which should rise, and the jobs held by those who stay, which should move closer to their credentials. Because the extension also changes who stays, the second margin combines a change in jobs with a change in the composition of stayers (Section 4.3.2). This design speaks to hiring friction as it operates on CtS STEM and non-STEM graduates on the OPT margin. I restrict the sample to CtS workers whose terminal degree is a master's or a doctorate; few professional degrees are STEM-designated, so professional degree holders

[8]The same rule also extended status for all F-1 graduates with a pending H-1B petition (the "cap-gap" provision); because it applies to STEM and non-STEM graduates alike, it is differenced out below. A 2016 rule further lengthened the STEM extension to 24 months, so graduates from 2016 onward receive a larger dose of the same treatment.

are excluded. For worker i with terminal degree e, I estimate separately by degree:

$$\begin{aligned}\text{US_Job}_i = \alpha_e + \beta_e \,\text{STEM}_i + \gamma_e \,\text{Post}_{c(i)} + \delta_e \left(\text{STEM}_i \times \text{Post}_{c(i)}\right) \\ + \sum_{\tau} \kappa_{e\tau} \mathbb{1}\{T_i = \tau\} + \lambda_e \,\text{DUHMajor}_i + \varepsilon_i,\end{aligned} \tag{9}$$

$$\begin{aligned}Y_{it}^k = \alpha_{et}^k + \beta_{et}^k \,\text{STEM}_i + \gamma_{et}^k \,\text{Post}_{c(i)} + \delta_{et}^k \left(\text{STEM}_i \times \text{Post}_{c(i)}\right) \\ + \sum_{\tau} \kappa_{et\tau}^k \mathbb{1}\{T_i = \tau\} + \lambda_{et}^k \,\text{DUHMajor}_i + u_{it}^k, \qquad \text{for } S_{it} = 1.\end{aligned} \tag{10}$$

Here STEM_i indicates that the major of i's terminal degree is on the 2008 STEM Designated Degree Program List,[9] $c(i)$ is the year i completed the terminal degree, $\text{Post}_c = \mathbb{1}\{c \geq 2008\}$, and t is years since the terminal degree.

On the extensive margin, when the hiring friction results in equilibrium contracts worse than a CtS worker's best foreign option, the reduction of said friction will increase the US placement rate. I classify each graduate's first job as located abroad or in the United States, so that US_Job_i in Equation (9) indicates a first post-degree job inside the United States, and δ_e is the difference-in-differences in the share landing a first job in the US.

On the intensive margin, Equation (10) is estimated on graduates whose first job is in the United States, and separately at each t, on the subsample still US-resident at t ($S_{it} = 1$). The outcomes Y_{it}^k are the residualized log OEWS median wage of the occupation held, over-education, horizontal mismatch, log SOC PageRank, and log firm PageRank. I report Equation (10) only at values of t at which every STEM×Post group retains at least 100 workers and at least 20% of its $t = 0$ base, so that no estimate is calculated solely off a thin tail of long-tenured stayers.

Two further outcomes describe what happens after entry, among graduates whose first job is in the United States. Retention, Stay_{it}, indicates that i has not left the United States by year t. Employer attachment, NoSwitch_{it}, indicates that i has not yet moved from the first employer to another employer by year t while in the United States; I read employers from every job on the profile.

[9] I assign STEM status from the terminal degree's own major. A graduate whose earlier degree is in a STEM field but whose master's or doctorate is not, such as an engineer who completes a master's in management, is coded non-STEM, because under the 2008 rule the extension applied to the degree on which OPT was granted. From 2016, a graduate can also qualify on a STEM degree earned at a US institution in the previous ten years. In this case, some post-2016 graduates coded non-STEM may become eligible, which biases the estimates for those cohorts toward zero.

For $m \in \{\text{Stay}, \text{NoSwitch}\}$ I estimate, separately by degree and at each t,

$$m_{it} = \alpha_{et}^{m} + \beta_{et}^{m}\,\text{STEM}_i + \gamma_{et}^{m}\,\text{Post}_{c(i)} + \delta_{et}^{m}\left(\text{STEM}_i \times \text{Post}_{c(i)}\right) + \sum_{\tau} \kappa_{et\tau}^{m}\mathbb{1}\{T_i = \tau\} + \lambda_{et}^{m}\,\text{DUHMajor}_i + v_{it}^{m}, \tag{11}$$

on graduates whose status at t is known: those who have left (or switched) by t and those still observed at t.

A graduate whose profile ends before t without an exit is dropped from Equation (11) at t. For employer attachment this makes little difference, but for retention the dropped graduates are disproportionately stayers in recent cohorts, whose profiles are observed less often in the last years of the data. I therefore report retention with the Kaplan–Meier estimator: within each STEM$\times$Post group, the share still in the United States at t is the product of one-year survival rates $\prod_{j \le t}\left(1 - \text{exits}_j/\text{at risk}_j\right)$, where a graduate whose profile ends counts as at risk until the last year observed. The retention difference-in-differences is the same two-by-two contrast of these shares. A positive $\delta_{et}^{\text{NoSwitch}}$ means that the extension reduced employer switching.

The comparison cohorts graduated between 2003 and 2006. I drop the 2007 cohort because eligibility for the extension turned on the date a worker's initial OPT period ended, not on graduation year, so 2007 graduates still on OPT in April 2008 were partly treated. I report three post-period windows: all cohorts graduating in 2008 or later, a local window of 2008-2011 graduates, and a window of 2012-2015 graduates that places the post-period after the financial crisis.

Standard errors. STEM status is assigned by major: every graduate of a given major is either eligible or not, according to the 2008 list. I therefore cluster standard errors by terminal-degree major, of which there are about 550 in the master's samples and 415 in the doctorate samples, in the event study, the stayer outcomes, and employer switching (Abadie et al., 2023). For retention, I report Greenwood standard errors for the Kaplan–Meier shares, which treat graduates as independent.

4.3.1 Identifying Assumptions

To interpret δ_e as the causal effect of reduced hiring friction requires three assumptions.

Assumption 3 (Parallel trends)**.** *Absent the extension, the share of STEM graduates taking a first job in the US would have evolved in parallel with the share of non-STEM graduates, conditional on tier and major generality.*

Assumption 3 is not testable in the post-period, but it has a testable implication in the pre-period, which I examine with the event study in Equation (12) below.

Assumption 4 (No differential concurrent shocks)**.** *Nothing else that changed around 2008 moved the US prospects of STEM graduates relative to non-STEM graduates.*

Two possible violators of Assumption 4 are the 2008-2009 recession and the H-1B cap itself. If STEM employment was more resilient to the recession, δ_e would overstate the policy effect; the 2012-2015 window, which leaves out the recession cohorts entirely, is designed to address this. The recession can also bias the intensive-margin estimates. Graduating into a downturn raises the probability of over-education (Summerfield and Theodossiou, 2017), pushes new hires into less complex jobs (Baley et al., 2022), and lowers earnings for a decade or more (Kahn, 2010; Oreopoulos et al., 2012; Schwandt and Wachter, 2019). If this entry penalty was smaller for STEM graduates, the 2008–2011 cohorts would show a STEM advantage in the jobs they take with no change in d, so δ^k_{et} would overstate the policy effect; the 2012–2015 window addresses this channel as well. The H-1B cap was not reached in the first week of filing for fiscal years 2010 through 2013, so for the cohorts graduating in those years an additional filing season was worth less than when the cap binds. This works against finding an effect in the local 2008–2011 window.

Assumption 5 (Stable treatment assignment)**.** *The composition of STEM and non-STEM graduates does not respond to the policy.*

For the 2008-2011 cohorts this is plausible, since most of them had chosen a program before the April 2008 rule. It is less plausible for the 2012-2015 cohorts, who enrolled after the extension was known: if foreign students who intended to stay in the United States shifted into STEM programs, the STEM share in US would rise even with no change in hiring friction. This response is documented: Amuedo-Dorantes et al. (2019) find that foreign students became 18 percent more likely to major in STEM after 2008, in part through non-STEM bachelor's holders entering STEM master's programs, and Khoo (2025) finds higher STEM enrollment and completion among international students at Ohio public universities. The 2012-2015 window therefore addresses Assumption 4 at the cost of weakening Assumption 5, and the two windows are best interpreted together rather than as one robustness check.

Event study. To examine the timing of the extensive-margin effect and pre-trends, I replace the single post indicator with a full set of graduation-cohort fixed effects and interact STEM with each cohort:

$$\begin{aligned}\text{US_Job}_i =& \mu_{c(i)} + \beta_e \,\text{STEM}_i + \sum_{c \neq 2006} \phi_{ec} \,\text{STEM}_i \times \mathbb{1}\{c(i) = c\} \\ &+ \sum_{\tau} \kappa_{e\tau} \mathbb{1}\{T_i = \tau\} + \lambda_e \,\text{DUHMajor}_i + \varepsilon_i, \end{aligned} \tag{12}$$

with the 2006 cohort as the reference and 2007 excluded. The coefficients ϕ_{ec} for $c \leq 2005$ test Assumption 3: under parallel trends they should be close to zero, with a break at 2008, and I test them jointly with a Wald test. The coefficients for $c \geq 2008$ trace how the effect evolves across cohorts, including any change for cohorts from 2016 onward, who received the 24-month extension.

4.3.2 Selection on the Intensive Margin

Equation (10) conditions on $S_{it} = 1$, a US first job and US residence at t, both of which can respond to the treatment: the extension raised the share of STEM graduates who start in the United States (Figure 12) and, by less, how long master's graduates stay (Section 7.2).

Assumption 6 (Monotone selection)**.** *The extension never causes a worker who would otherwise have stayed to leave:* $S_{it}(1) \geq S_{it}(0)$.

Under Assumption 6, treated stayers are either always-stayers, with $S_{it}(0) = S_{it}(1) = 1$, or marginal stayers, with $S_{it}(0) = 0$ and $S_{it}(1) = 1$; let $\omega_t = \Pr(S_{it}(0) = 1 \,|\, S_{it}(1) = 1,\, D_i = 1)$ denote the always-stayer share.

Assumption 7 (Parallel trends among would-be stayers)**.** *Absent the extension, the mean outcome among STEM graduates who would have stayed would have changed across the 2008 cutoff by the same amount as the mean outcome among non-STEM graduates who would have stayed:* $E[Y^k_{it}(0) \mid S_{it}(0) = 1, \text{STEM}, \text{Post}] - E[Y^k_{it}(0) \mid S_{it}(0) = 1, \text{STEM}, \text{Pre}]$ *equals the corresponding change for non-STEM graduates.*

Assumption 7 is stated on stayers because the estimator uses only stayers, but it restricts two things at once: how untreated outcomes evolve, and how the composition of would-be stayers evolves. Appendix F.3 shows what this requires and how it can be checked.

Under Assumption 7, the difference-in-differences estimator identifies

$$\delta^k_{et} = E\big[Y^k_{it}(1) \,\big|\, S_{it}(1) = 1, D_i = 1\big] - E\big[Y^k_{it}(0) \,\big|\, S_{it}(0) = 1, D_i = 1\big]\,. \tag{13}$$

Writing the first term as an ω_t-weighted average over always-stayers and marginal stayers, and noting that under Assumption 6, $S_{it}(0) = 1$ selects exactly the always-stayers, gives

$$\begin{aligned}\delta^k_{et} = {} & \underbrace{E\big[Y^k_{it}(1) - Y^k_{it}(0) \,\big|\, \text{always-stayer},\, D_i = 1\big]}_{\text{effect on always-stayers}} \\ & + \underbrace{(1-\omega_t)\Big(E\big[Y^k_{it}(1) \,\big|\, \text{marginal},\, D_i = 1\big] - E\big[Y^k_{it}(1) \,\big|\, \text{always-stayer},\, D_i = 1\big]\Big)}_{\text{composition}}\,. \end{aligned} \tag{14}$$

Appendix F derives Equation (14) step by step. The composition term is zero only if marginal stayers hold jobs of the same quality as always-stayers, which the model implies when d is the same for every worker in the group. If d varies across workers, the sign of the term depends on how the friction and the foreign option are jointly distributed. If the foreign option is independent of d and has a log-concave distribution, marginal stayers faced on average a larger friction than always-stayers and hold lower-quality jobs; the composition term is then negative, and a positive $\widehat{\delta}^{k}_{et}$ understates the effect on always-stayers. Without such a restriction the sign is not determined. I therefore interpret Equation (10) as the effect of the extension on the job quality of the population that remains in the United States, not as the effect on a fixed set of workers.

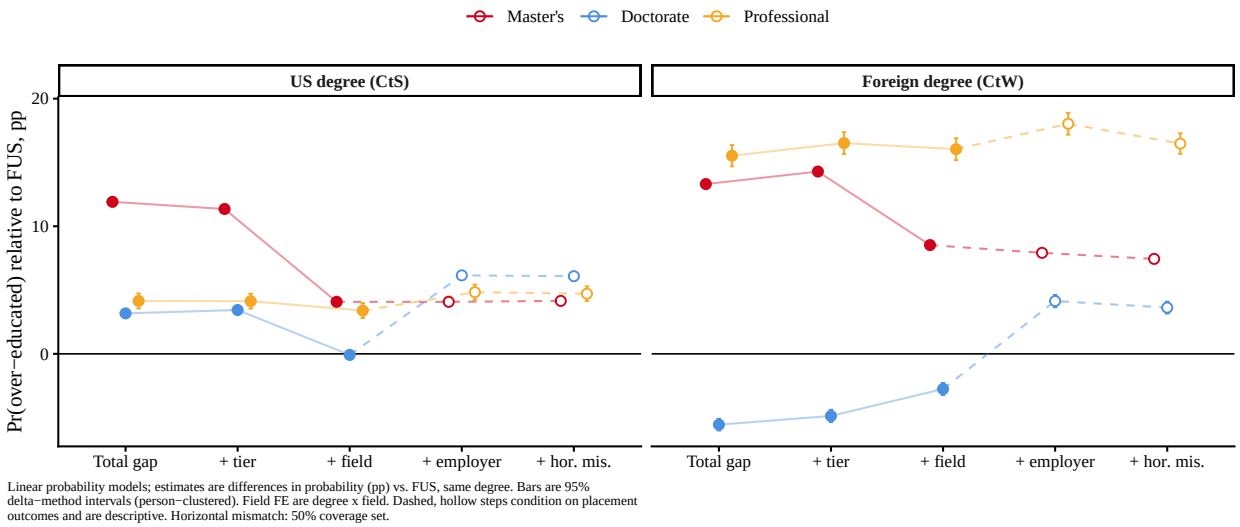
5 Occupational Mismatch at Labor Market Entry

This section reports what happens at the first US job: where immigrants sit relative to the contracts in Figure 4, and how the rank and pay of the occupations they hold differ from those of FUS workers.

5.1 Mismatched Contracts

Figures 2 and 3 summarize the over-education and horizontal-mismatch results, and Tables 6 through 9 report the estimates. Each point in the figure is the gap between an immigrant pathway and FUS workers with the same terminal degree, $\beta^m + \beta^m_e$, as blocks of controls are added.

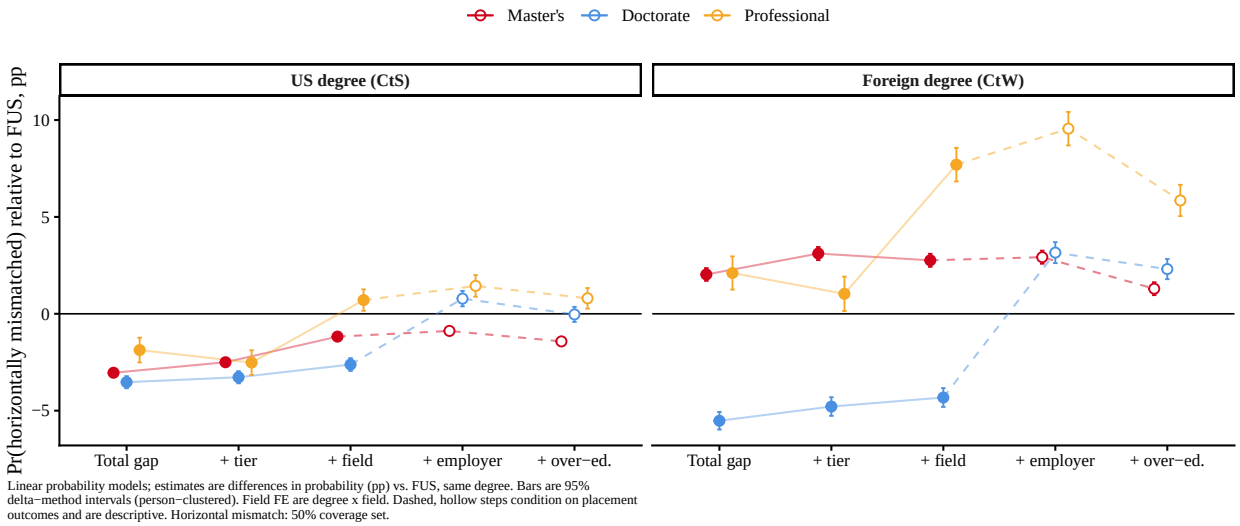
Figure 2: Over-Education Gap at the First US Job, Relative to FUS Workers with the Same Degree



Notes: Estimates of $\beta^{OE} + \beta^{OE}_e$ from Equation (4), in percentage points, with 95% confidence intervals. Moving right, the specification adds degree-by-tier effects, degree-by-field fixed effects with STEM and major generality, the employer block, and horizontal mismatch. Dashed, hollow steps condition on placement outcomes and are descriptive; from "+ employer" on, they are gaps at cap-subject employers. Doctorate pools the PhD and non-licensure doctorates; professional denotes licensure doctorates.

CtS Workers Without controls, CtS master's holders are 11.9 percentage points more likely than FUS master's holders to be over-educated at their first US job (Table 6). Institution tier barely changes the gap, but most of it reflects the fields CtS graduates study: within degree and field, it falls to 4.1 points. For doctorate holders, the Column 3 gap is −0.1 points and not statistically significant; for professional-degree holders it is 3.4 points. CtS master's and doctorate holders are slightly less likely than FUS workers to work outside their field (Table 7), by 1.2 and 2.6 points, while CtS professional-degree holders are 0.7 points more likely to do so. Optional Practical Training and the H-1B both require a job related to the degree, so part of this advantage in horizontal match may reflect visa rules rather than the friction.

Figure 3: Horizontal-Mismatch Gap at the First US Job, Relative to FUS Workers with the Same Degree



Notes: Same as in Figure 2, for horizontal mismatch. The last step adds over-education.

CtW Workers CtW master's holders are 8.5 percentage points more likely than FUS master's holders to be over-educated in Column 3, about twice the CtS gap in Table 8. On the other hand, CtW doctorate holders are 2.8 points less likely. CtW professional-degree holders are 16.0 points more likely, the largest gap of any group. Horizontal mismatch follows the same order (Table 9): 2.8 points for master's holders, −4.3 for doctorate holders, and 7.7 for professional-degree holders.

Employer Sorting If immigrants accepted mismatched jobs in exchange for better or visa-sponsoring employers, the gaps would shrink once the employer is held fixed. They do not: with firm utility, churn, any H-1B filing, and cap-exempt status added in Column 4, the master's over-education gaps are 4.1 points for CtS and 7.9 for CtW. Estimated separately at cap-subject and cap-exempt first employers (Table 10), the gap falls from the master's to the doctorate level at both types of employer, so the decline with the degree is not produced by academic jobs alone. Because the first employer is chosen in the same search as the job, the split locates the gap; it is not an effect of H-1B cap exemption.

Interpreting the Estimates through the Framework The positive master's gaps in both pathways place master's holders at the under-placement contract, point (b) of Figure 4, and the gap falls with the degree, as the convexity of the wage schedule implies. Pooled across employers, the doctorate gaps are −0.1 points for CtS and −2.8 for CtW, but the pool mixes cap-exempt employers, mostly universities, where the H-1B lottery does not apply. At cap-subject first employers (Table 10), CtS doctorate holders are 2.0 percentage points more likely than FUS doctorate holders

to be over-educated, about half the CtS master's gap of 3.8: positive and smaller, as the model predicts. CtW doctorate holders are the exception; their gap is -2.1 points even at cap-subject employers, a departure from Table 1. One possible explanation for this observation is positive selection at entry into the US labor market. CtW professional-degree holders have the largest gap of any group, 16.0 points, consistent with the relicensing that foreign-trained holders must complete, while the gap of CtS professional-degree holders, 3.4 points, is smaller than the master's gap. Two further features bear on what d contains. For master's and professional-degree holders, the gap is larger with a foreign degree than with a US degree, as a friction in evaluating foreign credentials would predict. And Column 5 shows that conditioning on horizontal mismatch leaves the master's over-education gaps almost unchanged, at 4.1 points for CtS and 7.4 for CtW, so over-education is not a by-product of working outside one's field.

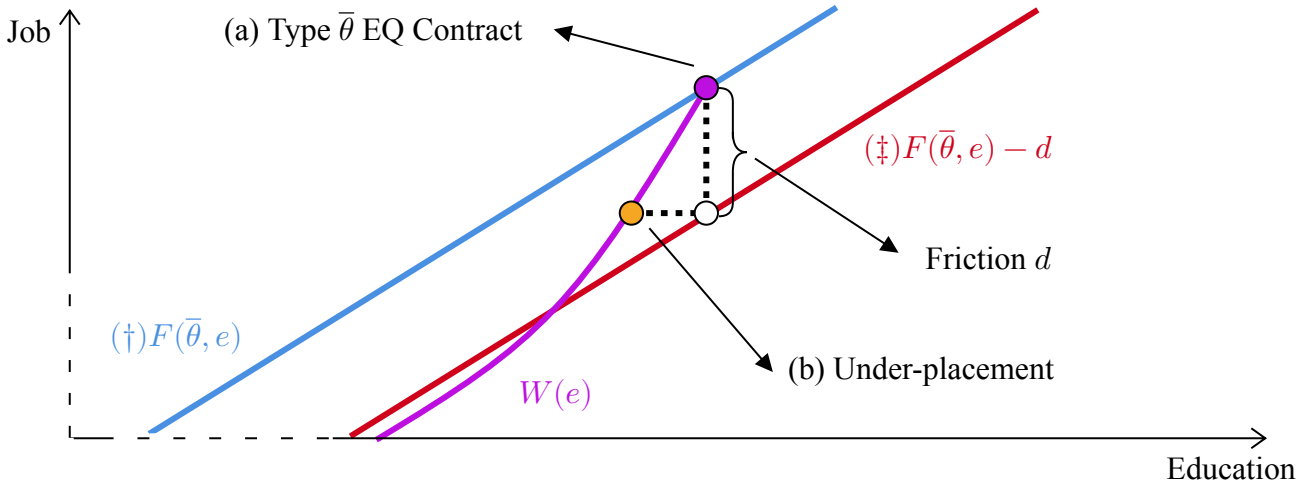
Selection at Two Gates Two CtW estimates sit awkwardly with the model: the negative over-education gap of doctorate holders and the very large positive gap of professional-degree holders. Both are likely to reflect who reaches the sample, and on what terms, rather than the friction facing a worker drawn from the same distribution as natives, which is what the framework assumes. The first gate is entry into the US labor market, and the two pathways clear it differently. A CtS graduate can start work on Optional Practical Training without an employer petition or a lottery, so the gate admits many graduates who seek a US job. On the other hand, a CtW worker may arrive on an employment-based visa, through an H-1B petition, a cap-exempt employer, an intra-company transfer, a visa for individuals of extraordinary ability, or on a family-based or other visa that does not depend on an employer's assessment of the worker. The mix of channels is likely to differ by degree.

For doctorate holders trained abroad, employment-based channels are likely to dominate, and the employers willing to sponsor them are recruiting researchers with established records. The CtW doctorate holders whom the model predicts would be over-educated are then the least likely to obtain a visa and never appear in US data; those who do are positively selected on productivity that the controls do not capture. Their negative gap describes who passes the gate rather than the friction faced by those who pass it: for them, the friction acts on whether they can work in the United States at all rather than on which job they hold. Two patterns rule out the simpler explanation that CtW doctorate holders work at different employers. Cap-exempt employers account for a similar share of doctorate first jobs in every pathway, 37 to 40 percent, and CtW doctorate holders are less over-educated than CtS doctorate holders at both types of employer: by 4.1 points at cap-subject employers and 4.6 at cap-exempt ones (Table 10). The selection account has a cost of its own, however. Among ten-year stayers, CtW doctorate holders trail CtS doctorate holders by 16.0 log points in occupational median wages one year after entry (Section 6), which positive selection

alone would not produce. Both facts would follow if the gate also steers CtW doctorate holders toward academic and research positions, which require a doctorate and are therefore rarely coded as over-educated, but which sit in occupations with low median pay. CtW doctorate holders are more concentrated in such positions: postdoctoral appointments account for 40 percent of their first jobs, against 27 percent for CtS and 21 percent for FUS doctorate holders.

CtW professional-degree holders face a second gate after they arrive. A doctorate earned abroad can be used in a US job immediately; a foreign medical, law, or dental degree cannot. Its holder must first pass US licensing examinations, and a physician must also complete a US residency. A holder who does not relicense works in an occupation that does not require the license, or requires a less costly one, and is over-educated by construction. Family-based channels, which do not screen on whether the credential can be used, are also more likely to admit them than an employer recruiting for a licensed position. Part of the CtW professional gap of 16.0 points may therefore reflect licensing rules rather than employers' assessments (Mattoo et al., 2008). Because relicensing is costly, and those who obtain a local license are a selected group (Kugler and Sauer, 2005), this component need not fade with US tenure: holders may stay in occupations outside the one they trained for. The panel does not record visa type or licensure, so I cannot observe either gate directly.

Figure 4: Separating Equilibrium with Immigrants



Notes: $(\dagger)$ marks the type-$\bar{\theta}$ native worker's net productivity schedule $F(\bar{\theta}, e)$. $(\ddagger)$ marks the type-$\bar{\theta}$ immigrant worker's net productivity schedule $F(\bar{\theta}, e) - d$. (a) marks the equilibrium contract for type-$\bar{\theta}$ native workers $W(\bar{e})$. (b) marks the under-placement contract for type-$\bar{\theta}$ immigrant workers, who hold $\bar{e}$ but work a job typical for $\tilde{e} < \bar{e}$ at $W(\tilde{e}) = W(\bar{e}) - d$; the vertical distance is the friction d.

5.2 Occupational Sorting

Tables 11 through 14 estimate Equation (5) for the rank (log PageRank) and the pay (log OEWS median wage) of the first US occupation. In the full specification, the occupations of CtS master's

holders rank 21.9 log points higher and pay 1.5 log points more than those of FUS master's holders; for CtW master's holders the figures are 14.0 and 4.1, and CtW professional-degree holders hold occupations ranked 55.2 log points below those of CtW master's holders. These regressions and the decomposition in Section 5.3 do not agree in sign for CtW master's holders: in the first column of Table 14, their occupations pay 5.7 log points more than those of FUS master's holders, whereas the raw gap in the decomposition is −7.8 log points. The two measure different outcomes. The regressions use the log median wage of the occupation itself and absorb graduation-cohort, arrival-cohort, and origin-region fixed effects, and their first column also holds non-US experience and same-employer transfers fixed. The decomposition uses the log median wage net of metropolitan-area, occupational-training, arrival-cohort, and region effects estimated on every job-year of the panel, so its raw gap already removes part of what distinguishes the occupations held. The samples differ as well: the regressions pool bachelor's and graduate degrees, whereas the decomposition compares master's holders alone on the sample in which every covariate block is observed. When Column 4 adds the occupation's training category and task content, the regression gap falls by half, from 11.0 to 5.5 log points, which suggests that the residualization accounts for part of the difference; I do not attribute it to any one choice. A pre-market decomposition of the first-job wage gap (Tables 15 to 17) gives similar raw gaps to the decomposition below, and Section 6 uses its unexplained components as the entry benchmark for the ten-year paths.

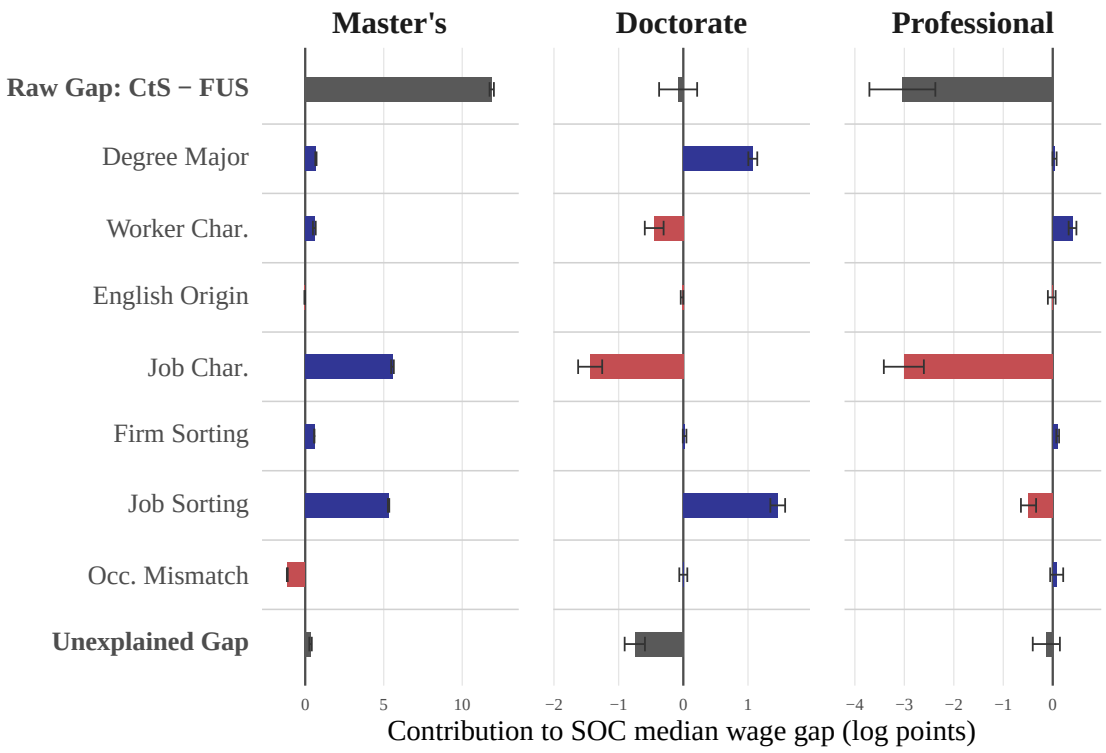
5.3 The Hidden Drivers Beyond Occupational Mismatch

Figures 5 and 6 decompose the gap in the log OEWS median wage of that occupation, by pathway and degree, into the contributions of eight covariate blocks (Gelbach, 2016); the last bar in each panel is the gap left once every block is held fixed. Because the blocks include the occupation's tasks and rank and the employer, which are chosen in the same search as the job, the contributions describe how the gap is composed, not why workers entered those occupations, and they do not measure d.

The occupations of CtS master's holders pay 11.9 log points more than those of FUS master's holders. Job characteristics account for 5.6 log points and job sorting, which contains the occupation's revealed-preference rank and its reliance on H-1B labor, for 5.3; the mismatch block accounts for −1.1 points, and 0.3 points remain. Tables A5 and A6 report the contribution of each covariate. For CtS doctorate holders the raw gap is −0.1 log points and not statistically different from zero, because job sorting contributes 1.5 log points and job characteristics −1.4. For CtS professional-degree holders the gap is −3.0 points, all of which the tasks of the occupations they hold account for.

The CtW gaps are negative at every degree: −7.8 log points for master's holders, −23.2 for

Figure 5: Decomposing First Job-Year Sorting Gap Drivers For CtS Workers

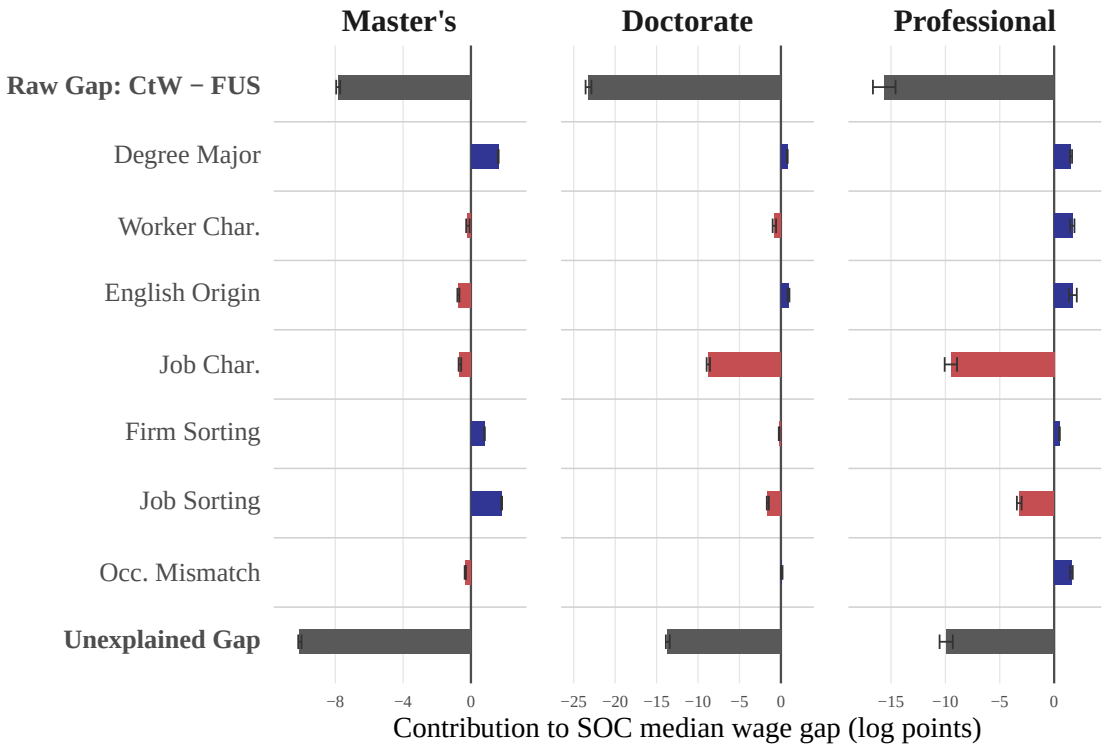


Notes: Gelbach (2016) decomposition of the gap between CtS and FUS workers with the same terminal degree in the log OEWS median wage of the occupation held at the first US job. Bars are block contributions in log points with 95 percent intervals; the last bar is the gap with every block held fixed. Doctorate pools the PhD and non-licensure doctorates.

doctorate holders, and −15.6 for professional-degree holders. For master's holders the covariates offset part of the gap, so it widens to −10.1 log points once every block is held fixed. For doctorate and professional-degree holders, job characteristics account for −8.8 and −9.5 log points, and −13.7 and −9.9 points remain. CtW master's holders are thus over-educated and in lower-paying occupations, whereas CtS master's holders are over-educated and in higher-paying ones.

The covariates account for most of the CtS gaps but little of the CtW gaps. Once the tasks, rank, and employer of the occupation are held fixed, little of the CtS gap remains at any degree: CtS workers earn what FUS workers in comparable occupations earn, and their premium reflects where they sort. CtW workers, by contrast, hold lower-paying occupations than FUS workers whose occupations have the same observed tasks, rank, and employer. CtS workers hold US degrees, which employers can evaluate as they evaluate those of FUS workers; CtW workers' foreign degrees and experience enter the decomposition only as characteristics, not at the value employers place on them, and foreign credentials are valued below their US equivalents (Friedberg, 2000; Oreopoulos, 2011). The pattern is consistent with a friction in evaluating foreign credentials, but the unexplained CtW gap can also reflect unobserved differences in skill.

Figure 6: Decomposing First Job-Year Sorting Gap Drivers For CtW Workers

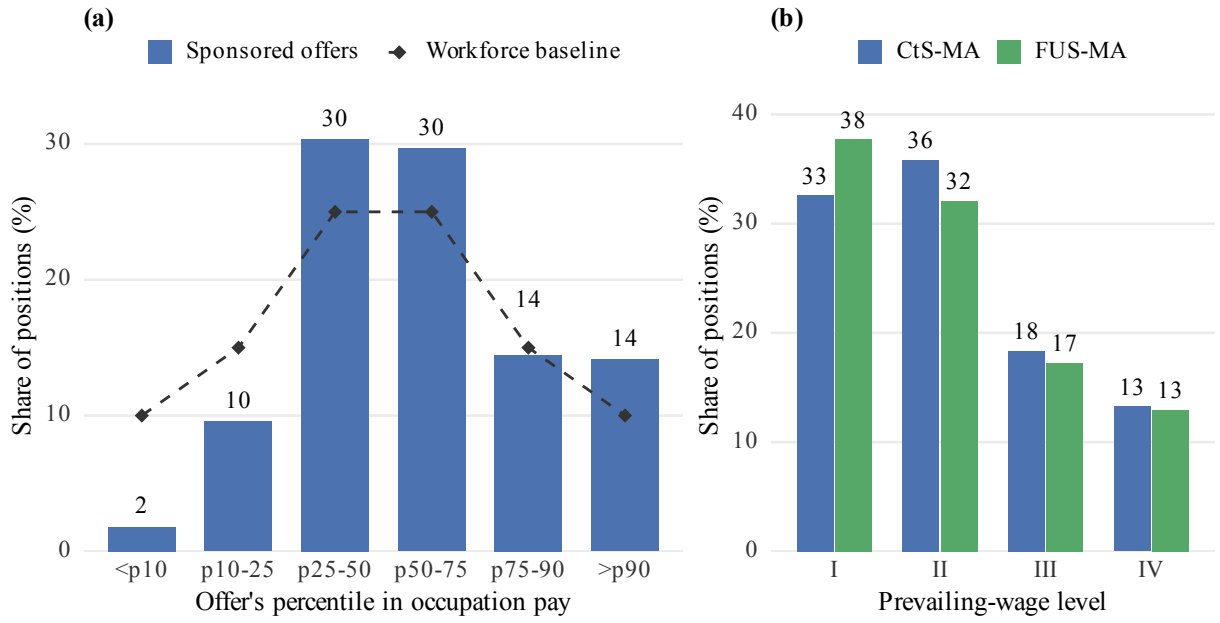


Notes: As in Figure 5, for CtW workers relative to FUS workers with the same terminal degree. Worker characteristics additionally include non-US experience and a same-employer transfer indicator.

5.4 Is the Friction a Discounted Wage? Evidence from the H-1B Wage Floor

This subsection reports the evidence behind Assumption 2. The friction d enters the model as a reduction in the net value of a foreign worker to the firm, and one interpretation of that reduction is that the worker is simply paid less for the same job. A natural reaction to the sorting result is that Came-to-Study (CtS) Master's workers occupy high-wage occupations only because H-1B sponsorship lets employers hire them below market price, displacing FUS workers within those occupations rather than competing on equal terms. The Labor Condition Application (LCA) disclosures speak to this directly, because they record the wage each certified H-1B position actually offers. If sponsored workers were hired as cheap substitutes, their offers would concentrate at the bottom of each occupation's wage distribution. They do not. Panel (a) of Figure 7 places every certified offer in the sponsoring occupation's national OES wage distribution—weighting occupations by the intensity with which CtS Master's workers sort into them—and compares the result against the distribution one would observe if sponsored workers were paid like the occupation's own workforce. Sponsored offers are scarce at the floor: only 2% fall below the tenth percentile against a ten-percent benchmark, and 11% below the twenty-fifth against twenty-five, with the mass instead displaced upward into the p25–median band and above.

Figure 7: **H-1B sponsored offers are not concentrated at the occupation wage floor.**

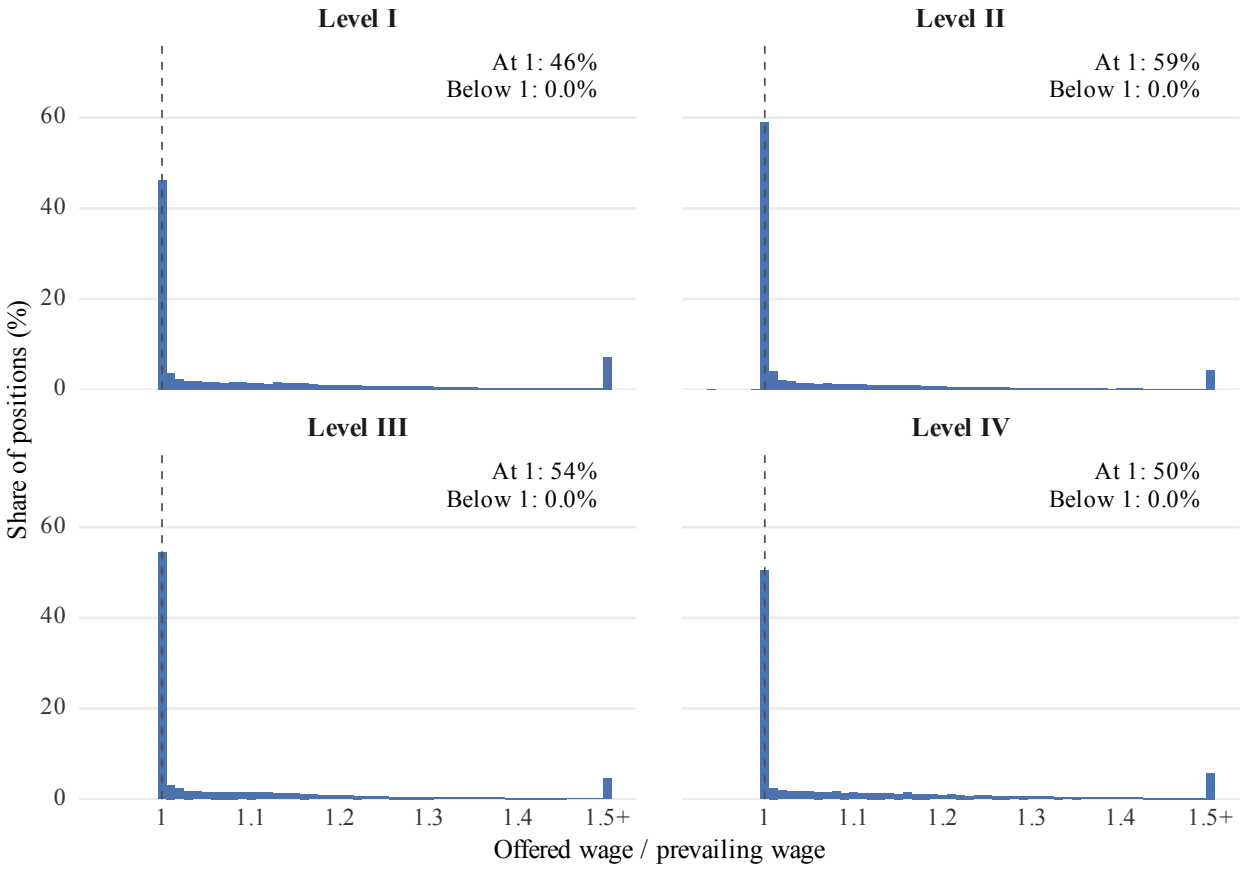


The prevailing-wage level tells the same story from the employer's side. Each LCA position is certified at one of four statutory wage levels, Level I being the floor—pinned near the OES seventeenth percentile—and the natural instrument for suppressing pay. Panel (b) shows that the

occupations CtS Master's workers enter are, if anything, slightly less Level-I-intensive than those entered by comparable FUS Master's workers: weighting each group's occupations by where its entrants sort, 33 percent of CtS filings are at Level I against 38 percent of FUS filings, and roughly two-thirds of CtS filings sit at Level II or higher. Foreign-educated workers are therefore not being channeled into the discounted tier of the occupations they share with natives—and where a position is filed at Level I, panel (a) shows its offered wage still clears the occupation's own floor, since the statutory Level I rate in these high-wage occupations sits well above their tenth percentile. Two qualifications bound the claim. The benchmark in panel (a) is the occupation's incumbent wage distribution rather than matched native new hires, and entrants of any origin sit below the occupation median early in their careers; the floor deficit and the level mix are thus consistent with, but do not on their own establish, within-occupation wage parity. A decisive test would compare sponsored offers to the realized wages of same-occupation, same-experience native entrants—a comparison the present panel, which imputes occupational medians rather than individual native wages, cannot support.

Panels (a) and (b) locate sponsored offers within the occupation; they do not show whether the attestation itself sets pay. Figure 8 does, by comparing each certified offer with the prevailing wage the employer attested to. If the floor were slack, offers would spread smoothly above it; if it binds, they pile up at it. They pile up. Across all certified H-1B positions, 46 to 59 percent at each wage level are offered within half a percent of the prevailing wage, and at most 0.5 percent below it. The mass at the floor does not come from a few large filings: counting each filing once rather than each position, the shares are 31 to 46 percent. For roughly half of these positions, the attested prevailing wage, not the employer's discretion, sets the offer, which is what Assumption 2 requires: an employer that would pay a sponsored worker less cannot do so in the offer it files. Because the prevailing wage rises with the level, from between the tenth and twenty-fifth OES percentiles for about half of Level I positions to above the seventy-fifth for most Level IV positions, bunching at a level's floor is not bunching at the bottom of the occupation. The attestation does not police the choice of level, the margin panel (b) addresses. The figure covers every certified position, not only those of CtS workers, because LCA records cannot be linked to individuals in the panel.

Figure 8: Certified H-1B Offers Relative to the Prevailing Wage, by Wage Level



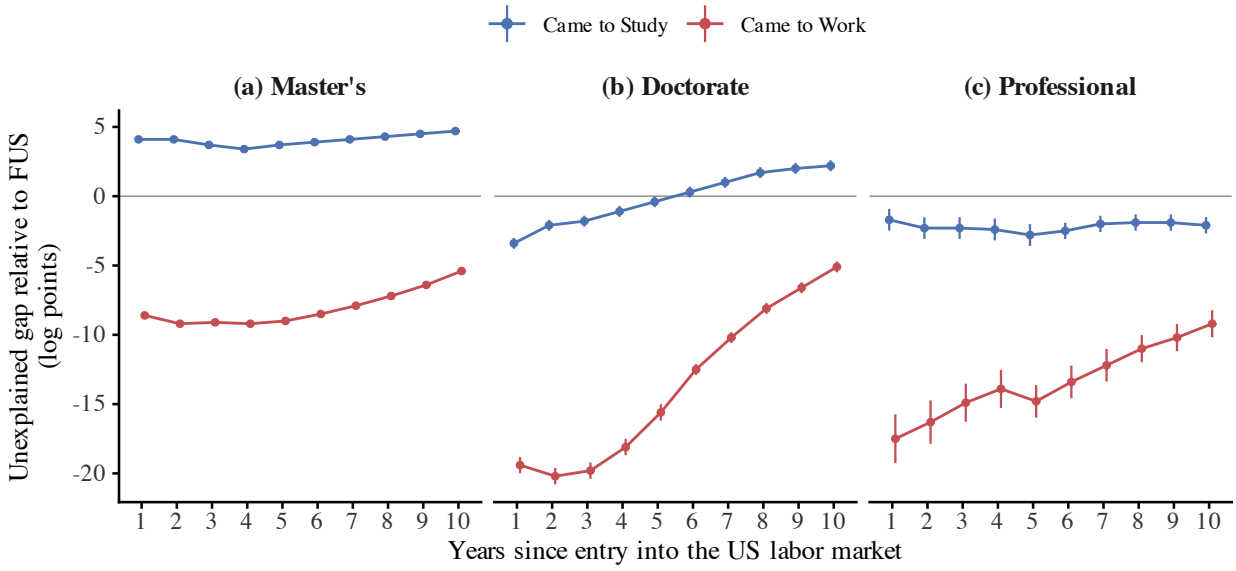
Notes: Distribution of the ratio of the offered wage to the prevailing wage, both annualized, for certified full-time H-1B positions in fiscal years 2015–2025, by prevailing-wage level. Each filing counts by the number of positions it requests; occupations are not reweighted. Bins have width 0.01; the bar at 1 covers ratios from 0.995 to 1.005, and ratios above 1.5 are pooled. Ratios below 0.5 or above 5, 0.1 percent of positions, are dropped. Fiscal years 2011–2014 are excluded because the wage level is not recorded.

6 The Dynamics of the Hiring Friction

The comparison between the two immigrant pathways is the test this section rests on. Workers in both pathways need sponsorship, but only CtW workers hold terminal credentials that US employers cannot readily evaluate. The CtS path therefore measures the part of d that the two pathways share, and the distance between the CtW and CtS paths measures what CtW workers face in addition: employers' uncertainty about a foreign credential, together with any US-specific skills that a US degree confers, and, for professional-degree holders, the need to re-license (Section 5). If that additional component fades as employers learn a worker's productivity or as professionals obtain US licenses, the CtW path should converge to the CtS path, not to zero; if it persists, the two paths should run parallel.

Figure 9 shows convergence of the unexplained component for doctorate and professional-degree holders and little for master's holders. In year one, CtW workers trail CtS workers with the same degree by 12.8, 15.5, and 17.2 log points for master's, doctorate, and professional-degree holders. By year ten the distance is 10.4, 7.0, and 7.9 log points. For doctorate holders, it widens to 17.5 log points in year two, holds through year three, and then closes steadily; for professional-degree holders, it closes from the start. For master's holders, it stays between 12.7 and 13.5 log points through year six and narrows only in the last four years. The CtS paths, by contrast, stay within five log points of zero throughout, so whatever part of the friction persists for workers with US degrees is small.

Figure 9: Unexplained Occupational-Wage Gap of Ten-Year Stayers, CtS and CtW



Notes: Unexplained components of the decompositions in Figure 10, with 95 percent confidence intervals. Each pathway's component is priced at the coefficients of its own FUS sample, and the CtW decompositions include additional covariates, so the vertical distance between the two paths is not itself an estimate. Years count from the terminal degree for CtS workers and from the first US labor-market year for CtW workers.

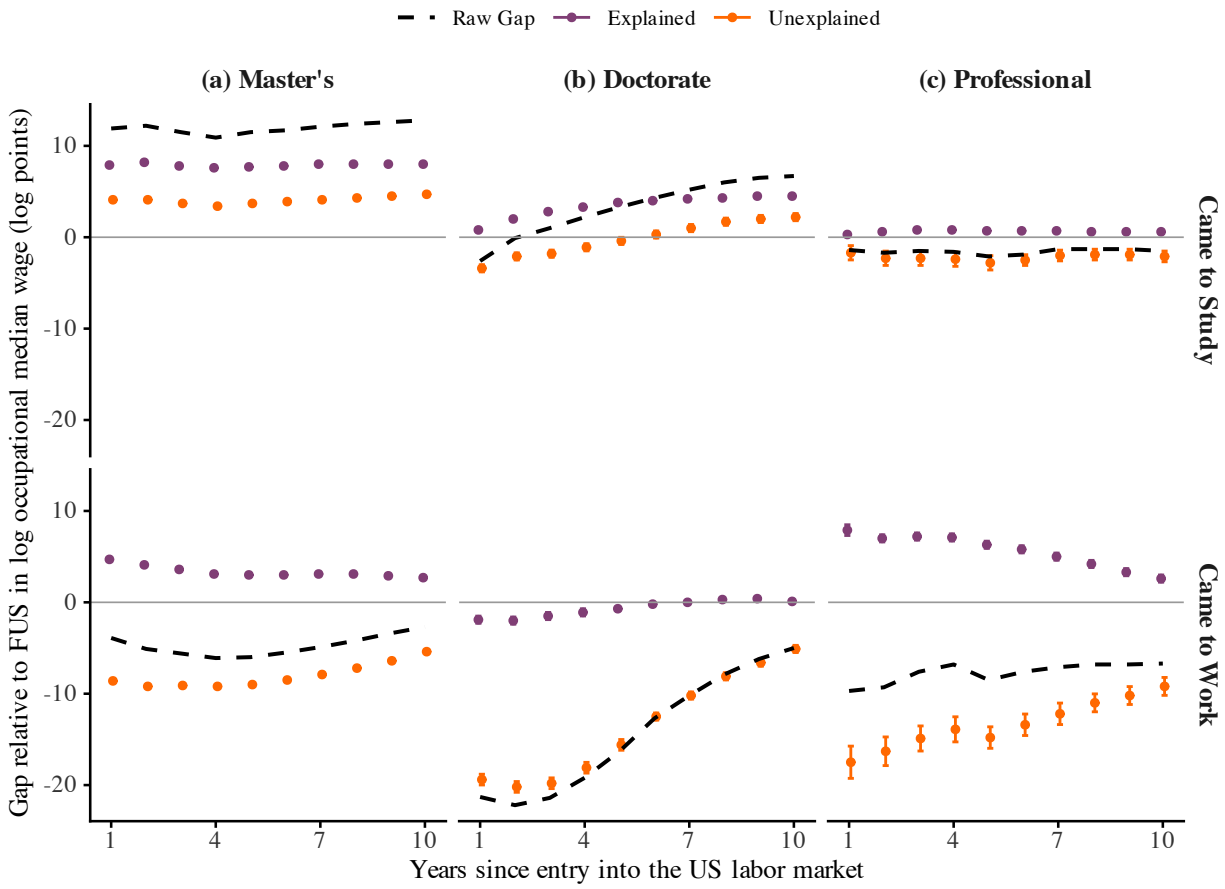
Figure 10 shows the full decomposition for each group. Among CtS stayers (top row), there is no common pattern of attenuation. For master's holders, the raw advantage in occupational median wages is 11.9 log points a year after graduation and 12.8 at year ten; the explained component is about 9 points throughout, and the unexplained premium rises slightly, from 3.3 to 4.0. For doctorate holders, the unexplained gap moves from a 3.9-point deficit at year one to a 1.9-point premium at year ten, crossing zero between years six and seven. Professional-degree holders keep a small deficit, 2.2 log points at year ten.

Among CtW stayers (bottom row), the unexplained deficit of master's holders is 9.5 log points at year one, stays near 10 through year five, and narrows to 6.4 by year ten. The doctorate deficit, 19.4 log points at year one and 20.2 at year two, then closes steadily to 5.1 at year ten, while the explained component rises from about -2 to about zero. For professional-degree holders, the unexplained deficit falls from 19.0 to 10.1 log points, but the explained component falls from $+9.6$ to $+3.9$ over the same years, so the raw gap narrows only from 9.4 to 6.3 log points. Their convergence in Figure 9 therefore overstates how much their occupational position improves relative to FUS workers.

Selection into Staying Every estimate in this section conditions on remaining in the United States through year ten, so the paths describe a selected group, and their starting points show how selected it is. Among CtS workers, the unexplained gap of ten-year stayers in year one is close to that of all entrants at their first US job: 3.3, -3.9, and -1.8 log points for master's, doctorate, and professional-degree holders, against 3.2, -2.3, and -3.1 (Tables A7 to A9 and 15). Among CtW workers it is smaller in magnitude: -9.5, -19.4, and -19.0 log points, against -11.5, -20.8, and -25.2 (Tables A10 to A12 and 16). CtW workers who remain for ten years thus began in somewhat better-paying occupations, relative to comparable FUS workers, than CtW entrants as a whole, by 1.4 to 6.2 log points, as expected if lower-earning immigrants are the ones who leave (Lubotsky, 2007). Their paths understate the gaps of all CtW entrants, most of all for professional-degree holders.

The comparison is approximate for two reasons. First, each decomposition prices covariates at the coefficients of its own FUS sample: FUS ten-year stayers in year t for the paths, and all FUS entrants at their first US job for the entry gaps. FUS stayers are themselves selected, so a difference between the two unexplained gaps combines selection among immigrants with a change in the benchmark. Second, the entry sample takes each worker's first US job within five years of entry rather than the job held in a fixed year. Neither decomposition holds the job or the employer fixed, so neither reproduces the final gaps in Figures 5 and 6.

Figure 10: Occupational-Wage Gap of Ten-Year Stayers, by Pathway and Degree



Notes: Oaxaca–Blinder decomposition, FUS coefficients as reference, of the gap relative to FUS workers in the log OEWS median wage of the occupation held in each year, among workers who remain in the United States through year ten; 2003–2021 graduation cohorts. Each pathway, degree, and year is estimated separately (Tables A7 to A12). Years count from the terminal degree for CtS workers and from the first US labor-market year for CtW workers; year zero is reported in the tables but not shown (Section 4.2). The CtW decompositions also hold fixed non-US labor-market experience, its square, and a same-employer transfer indicator. Bars are 95 percent confidence intervals. Doctorate pools the PhD and non-licensure doctorates; Professional denotes licensure doctorates.

Interpreting the Estimates through the Framework The doctorate and professional-degree paths converge at similar rates, but the framework says more about the former. For doctorate holders, whose credential can be used on arrival, convergence is what a component of d that fades with employer learning would produce: the uncertainty attached to a foreign credential is resolved as employers observe the worker. For professional-degree holders, the framework is less informative: convergence could reflect employer learning, relicensing, or movement between occupations that do not require a US license. The fall in the explained component, from $+9.6$ to $+3.9$ log points, indicates that the occupations they hold change along the way.[10] The stable distance for master's holders through year six is what a persistent component would produce, and it is still 10.4 log points at year ten. Four qualifications limit this interpretation. First, each pathway's unexplained gap is priced at the coefficients of its own FUS sample, and the CtW decompositions hold fixed additional covariates, so the distance between the paths is not a single estimate and has no standard error. Second, the clocks differ: t counts years since the degree for CtS workers and years since the first US labor-market year for CtW workers, who may arrive with years of experience abroad. The comparison aligns time in the US labor market, which is the relevant clock for employer learning, but not career stage. Third, convergence need not reflect learning. Sponsorship ends at permanent residence, which the two pathways reach at different points in their careers, and US-specific skills accumulate with tenure. Because the outcome is the occupation's median wage, it also moves only when a worker changes occupation, and a worker whose occupation pays less than that of comparable natives has an incentive to change it whatever the cause of the gap; a worker whose occupation pays more has none, which is consistent with the persistent CtS master's premium. Fourth, stayers are positively selected, more so in the CtW pathway, so the distances understate those among all entrants. Employer learning, job search, and worker-side assimilation all move workers into higher-paying occupations, and these estimates do not separate them.

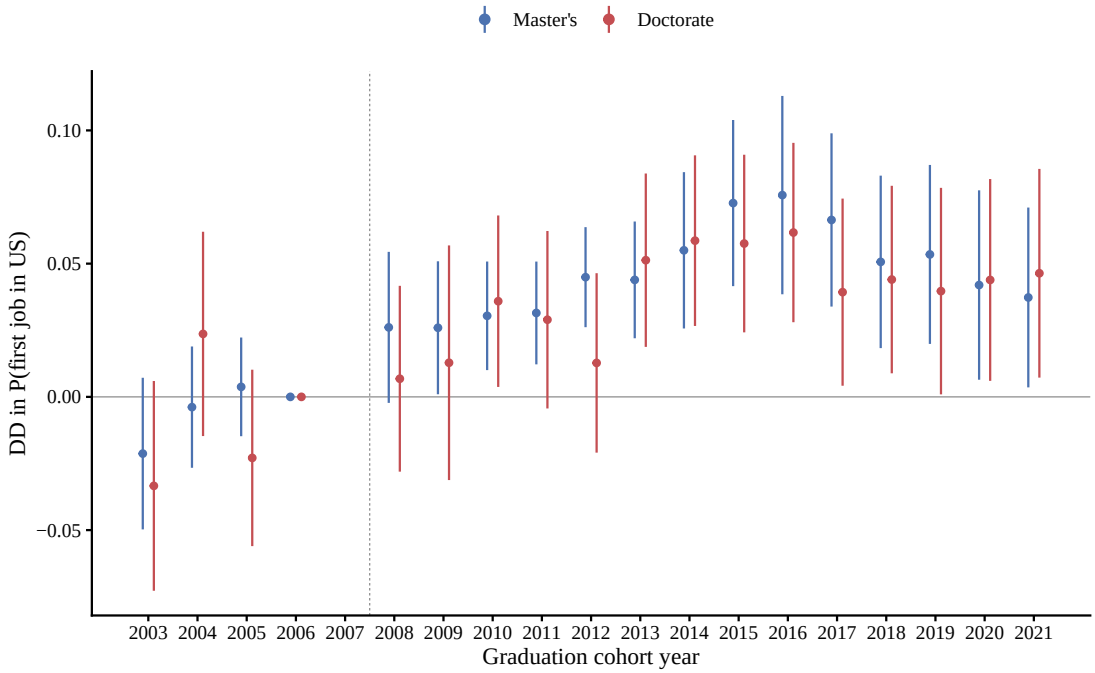
[10] Re-licensing is unlikely to be the main channel. I define licensed occupations as those that professional degrees train for and that require a US license to practice: physicians and surgeons, dentists, pharmacists, optometrists, podiatrists, chiropractors, physical therapists, audiologists, veterinarians, lawyers, and judges, identified by their SOC codes. Occupations that require a less costly license, such as nursing, are not included. Among ten-year stayers, the share of CtW professional-degree holders working in a licensed occupation is 52 percent in year one and 45 percent in year ten, against 60 and 58 percent for FUS professional-degree holders.

7 The Hiring Friction at the Margins

The STEM-OPT extension provides a test of whether lowering the cost of hiring an international graduate changes both the decision to begin work in the United States and the quality of jobs held there. The analysis concerns CtS master's and doctorate graduates, because few professional degrees were eligible for the extension. I first examine when the placement effect appears, then the jobs held by graduates who stay, and finally whether the effects are concentrated at the employers subject to the H-1B lottery.

7.1 US Placement and Its Timing

Figure 11: Cohort Event Study: STEM–Non-STEM Difference in the Probability of a First Job in the United States



Note: Coefficients ϕ_{ec} from Equation (12), CtS master's and doctorate graduates, estimated separately by degree and offset horizontally. Doctorate pools the PhD and non-licensure doctorates. The outcome is an indicator for a first job located in the United States, so a positive coefficient is a relative increase in US placement. STEM indicates that the graduate's own terminal-degree major is on the 2008 STEM Designated Degree Program List. The 2006 cohort is the reference and 2007 is omitted. Bars are 95 percent confidence intervals clustered by major.

Figure 11 plots the cohort coefficients ϕ_{ec} from Equation (12). The outcome is an indicator for a first job in the United States, as in Equation (9), so a positive coefficient means that STEM graduates of that cohort became more likely than non-STEM graduates to start in the United States, relative to the 2006 cohort. For master's graduates, the 2003–2005 coefficients are -2.1, -0.4,

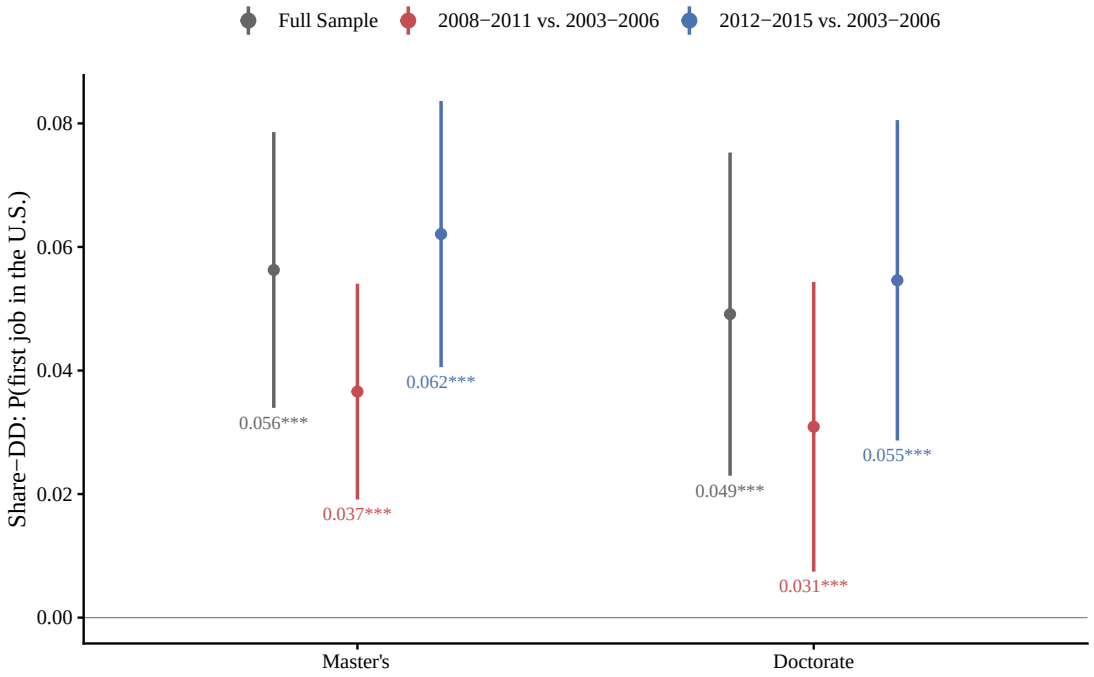
and $+0.4$ percentage points. They rise toward the 2006 reference, which, if it reflected a trend, would overstate the post-2008 estimates, but they are imprecise and jointly indistinguishable from zero. For doctorate graduates, the coefficients are -3.3, $+2.4$, and -2.3 percentage points. They alternate in sign rather than trending into 2008, but they are larger and less precise than the master's coefficients, because the doctorate sample is about a fifth the size of the master's sample, and I treat the doctorate estimates as suggestive.

The master's effect appears with the first treated cohort and grows. For 2008–2011 graduates, the increase in US placement ranges from 2.6 to 3.2 percentage points; the 2008 and 2009 graduates, who had enrolled before the rule was announced, are both at 2.6 points, and every coefficient but the 2008 one is statistically significant. The increase is 4.4 to 5.5 percentage points for 2012–2014 graduates, rises to 6.6 to 7.6 for 2015–2017 graduates, and is 5.1 and 5.4 for 2018 and 2019 and 4.2 and 3.7 for 2020 and 2021, all statistically significant. Two features of the policy environment line up with this path. First, the H-1B cap was not reached in the first week of filing for fiscal years 2010 through 2013 (Section 4.3), so an additional filing season was worth less to graduates of those years than later years. Second, the 2016 rule lengthened the extension from 17 to 24 months for graduates who applied after May 2016, a group that includes much of the 2015 cohort.

Both rules lower the same friction, the risk that a hire is lost before winning the lottery, so the rise from 2015 does not by itself contradict the mechanism. It does mean that the 2008 rule alone raised master's placement by about 3 percentage points, and that the cohorts that received the longer extension raise the pooled estimate below. These later cohorts also enrolled after the extension was known, and foreign students shifted into STEM majors after 2008 (Amuedo-Dorantes et al., 2019), so the post-2015 coefficients combine a larger dose of the treatment with any selection into STEM (Assumption 5). For doctorate graduates, the coefficients are 0.7, 1.3, 2.9, and 1.3 percentage points for 2008, 2009, 2011, and 2012 and not statistically significant, whereas the coefficient of 3.6 for 2010 and those of 3.9 to 6.2 for every cohort from 2013 on are all statistically significant.

Figure 12 shows that relative to non-STEM graduates, STEM graduates became more likely to take their first job in the US after the extension. Pooling the 2008–2021 cohorts, the difference-in-differences in US placement is 5.6 percentage points for master's graduates, about 6.5 percent of the pre-period rate of 86.3 percent at which STEM master's graduates started in the United States, and 4.9 percentage points for doctorate graduates. The 2008–2011 estimates are 3.7 and 3.1 percentage points, and for 2012–2015 graduates they are 6.2 and 5.5; all six estimates are statistically significant with standard errors clustered by major. Persistence beyond the recession cohorts weighs against an explanation confined to the immediate downturn. Taken together, the event study and the local windows show an increase in the relative US placement of master's graduates of about 3 to 3.5 percentage points under the 2008 rule and a larger one once the 2016 rule applies, and a doctorate increase of similar size that appears from the 2010 cohort on.

Figure 12: Increase in CtS US Placement Rate after STEM OPT Extension



Came to Study. Extensive−margin share−DD on P(first job in the U.S.): [STEM post−pre] − [Non−STEM post−pre], covariate−adjusted (school tier and major generality). POSITIVE = STEM grads became relatively MORE likely to land a first job in the U.S. post−extension. Labels = point estimate (*p<.05 **p<.01 ***p<.001). Pointranges 95%% CI.

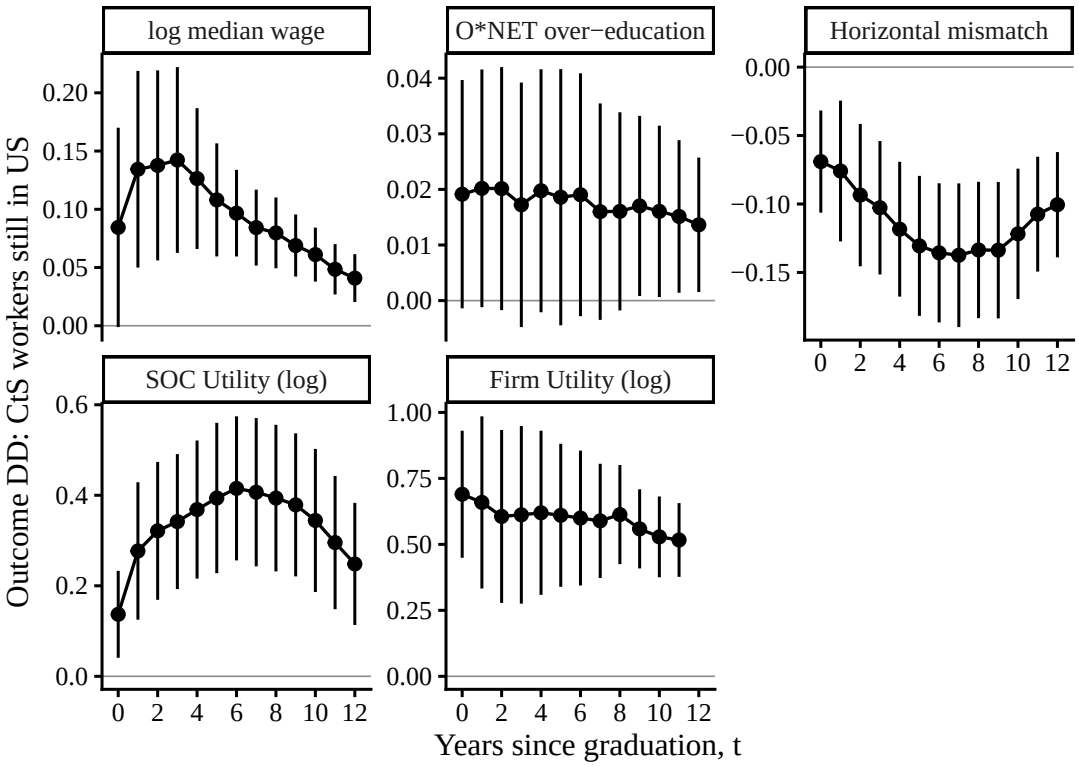
7.2 Job Quality among Stayers

On the intensive margin, any observed improvement in realized job quality could reflect a compositional shift in the marginal stayer rather than a change in contract terms available to a given worker; Equation (14) writes the estimate as the sum of the two. I estimate Equation (10) for the residualized log occupational median wage, over-education, horizontal mismatch, log SOC PageRank, and log firm PageRank among graduates still working in the United States at each year t since graduation (Figures 13 and 14).

Table 19 reports the difference-in-differences five years after graduation, and Figures 13 and 14 trace them year by year. I begin with the 2008–2011 window, which excludes the cohorts that received the longer extension and enrolled after it was known. Among master's graduates of those cohorts still working in the United States five years after graduation, STEM graduates are 7.6 percentage points less likely than non-STEM graduates to work outside their field. They hold occupations ranked 0.19 log points higher by revealed preference and paying 0.08 log points more at the median, and work at employers valued 0.21 log points higher. Over-education does not change: the estimate is 0.5 percentage points. Over the career, horizontal mismatch is lower from $t = 0$ on, by 4.9 to 8.3 percentage points from $t = 2$, and the occupational median wage is unchanged at $t = 0$ and 0.07 to 0.11 log points higher at $t = 1$ to $t = 4$.

Pooling the 2008–2021 cohorts roughly doubles the estimates. At $t = 5$, horizontal mismatch is 13.1 percentage points lower, occupational PageRank 0.39 and firm PageRank 0.61 log points higher, and the occupational median wage 0.11 log points higher; at $t = 11$ the mismatch and PageRank estimates are still -10.7 percentage points and 0.30 log points. Over-education again does not fall: it is 1.7 to 2.0 points higher in the first five years, not statistically significant, and 1.4 to 1.7 points higher, and statistically significant, at $t = 9$ to $t = 12$. The post-recession 2012–2015 window shows the same pattern: horizontal mismatch falls by 6.9 to 13.9 percentage points, PageRank rises by 0.26 to 0.37 log points from $t = 1$ on, firm PageRank is 0.64 log points higher at $t = 5$, and the occupational median wage 0.10 log points higher, while over-education is 1.1 to 1.9 points higher and statistically significant only at $t = 11$ and $t = 12$. Lower horizontal mismatch with no fall in over-education suggests that the master's gain is a move up within an education level, into the graduate's own field and into higher-ranked and higher-paying occupations, which over-education cannot see.

Figure 13: Trends of Occupational Outcome of Stayers, Master's Obtained 2008–2021

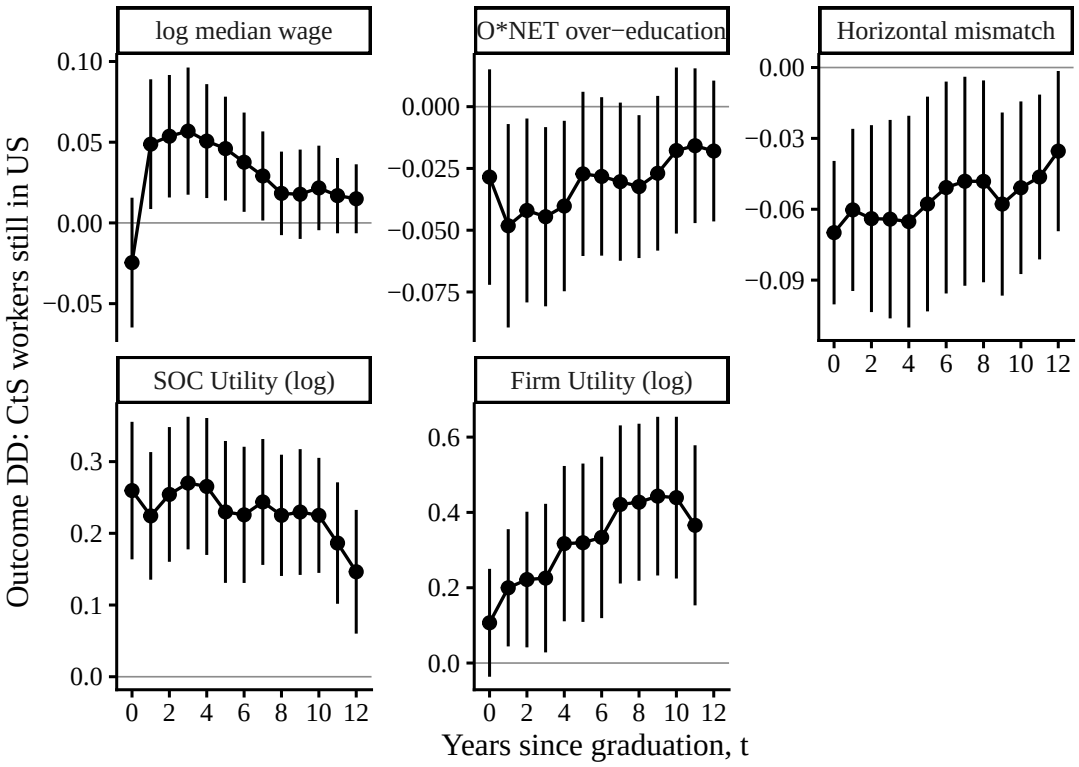


Came to Study, Master's. Common−horizon DD of the CtS workers still in US OUTCOME gap: [STEM post−pre] − [Non−STEM post−pre].Pointranges 95% CI.

For doctorate graduates the gains are similar, and over-education also falls. Pooling the 2008–2021 cohorts, horizontal mismatch is 3.5 to 7.0 percentage points lower at every horizon, occupational PageRank is 0.15 to 0.27 log points higher at every horizon, and the occupational median

wage is 0.04 to 0.06 log points higher from $t = 1$ to $t = 6$, though not statistically significant from $t = 8$. Firm PageRank rises later, to 0.43 and 0.44 log points at $t = 8$ and $t = 10$. Over-education is 4.0 to 4.8 percentage points lower from $t = 1$ to $t = 4$, not statistically significant at $t = 5$, and 3.2 points lower at $t = 8$. In the 2008–2011 window the doctorate estimates are smaller, and the mismatch estimates are not statistically significant after $t = 5$; firm PageRank, for example, is 0.31 and 0.32 log points higher at $t = 9$ and $t = 10$, against 0.54 and 0.51 in the 2012–2015 window. The evidence for improved placement is therefore similar for the two degrees in the full post-period and less precise for doctorate graduates within the local windows.

Figure 14: Trends of Occupational Outcome of Stayers, Doctorate Obtained 2008–2021



Came to Study, Doctorate. Common−horizon DD of the CtS workers still in US OUTCOME gap: [STEM post−pre] − [Non−STEM post−pre].Pointranges 95% CI.

These job-quality estimates condition on obtaining a US first job and remaining in the United States at each horizon. The extension changes that population. As Equation (14) shows, the estimates combine changes in jobs available to graduates who would have stayed anyway with differences in the composition of stayers. The composition term has two sources, more STEM graduates starting in the United States and more of them staying, and both are small relative to the population of stayers. Relative to non-STEM graduates, the share of STEM master's graduates still in the United States rose by 1.0 percentage point by $t = 5$ and 1.9 by $t = 10$ (standard errors 0.2 and 0.4); for 2008–2011 graduates the estimates are -0.1 and 0.6 (0.3 and 0.4). For doctorate grad-

uates they are 0.3 and 0.9 (0.5 and 0.7). Table F2 reports the estimates at other horizons and for the 2012–2015 window, together with a linear probability version of Equation (11) with standard errors clustered by major, which gives similar magnitudes. Counting both the additional entrants and the additional retention, marginal stayers are at most about one in eight of STEM master's stayers at year five. For composition alone to produce the 13.1-point fall in horizontal mismatch, they would need a mismatch rate about 100 percentage points below that of other stayers, which is impossible for a binary outcome. The sign also runs the other way: graduates who later leave the United States were more often mismatched in their US jobs than those who stay, 41.5 against 37.9 percent at $t = 0$ among post-2008 non-STEM master's graduates, so drawing in marginal stayers should raise measured mismatch, not lower it. The same argument does not rule out composition for over-education, whose estimate is small, or for the continuous outcomes, which have no natural bound. The extension therefore changed where graduates start more than how long they stay.

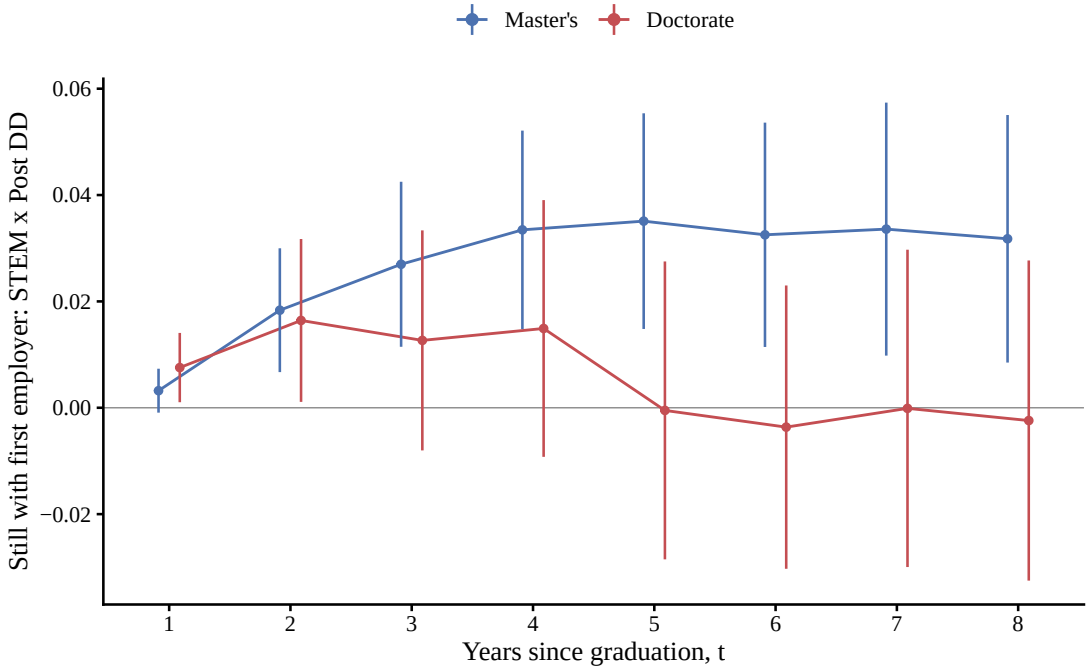
The extensive-margin results are consistent with the framework's prediction that a lower hiring friction makes US offers acceptable to more graduates (Section 2.3); the stayer results are consistent with improved placement for both degrees, and selection on staying is too small to generate the mismatch results, but they do not identify a job-quality gain for a fixed group of workers.

Retention measures whether graduates leave the United States; I next ask whether they leave their first employer. Figure 15 plots $\delta_{et}^{\text{NoSwitch}}$ from Equation (11). Graduates of later cohorts change employers sooner in every group: five years after graduation, 58 percent of post-2008 non-STEM master's graduates are still with their first employer, against 71 percent of pre-2008 graduates. The decline is smaller for STEM master's graduates, 9.4 percentage points against 12.7. As a result, STEM master's graduates of post-2008 cohorts are relatively more likely to still be with their first employer: by 0.3 percentage points at $t = 1$, 1.8 at $t = 2$, 2.7 to 3.5 from $t = 3$ to $t = 5$, and 3.2 to 3.4 through $t = 8$, with standard errors of 0.8 to 1.2 from $t = 3$ on. For doctorate graduates the estimates are 0.8 and 1.6 percentage points at $t = 1$ and $t = 2$, and within 1.5 points of zero from $t = 3$ on, with standard errors of 1.1 to 1.5.

Fewer switches have three interpretations, which differ in timing. Before 2008, a STEM graduate whose first H-1B petition failed lost work authorization twelve months after graduation and had to leave the employer, so part of the pre-period switching was forced; the extension removed some of these moves, which predicts an effect that appears at $t = 1$ and $t = 2$. Alternatively, a graduate who is waiting on the lottery, or whose employer has filed a petition, has reasons to stay with that employer until status is secured, and the extension keeps graduates in this position for up to three filing seasons, which predicts an effect that builds over the first three years. Finally, a longer expected match raises the employer's return to training the worker (Section 2.3), and when the two share the cost of firm-specific training, neither wants to end the match (Becker, 1962; Hashimoto, 1981); this predicts an effect that builds with tenure and persists after status is secured.

The master's path, small at $t = 1$, rising through $t = 5$, and still about 3.2 points at $t = 8$, fits the second and third interpretations; its persistence after the added filing seasons have passed is easier to reconcile with the third, although sponsorship for permanent residence can also keep a worker with the sponsoring employer. The doctorate path, small, early, and gone by $t = 5$, is closer to the first, although about a third of doctorate first jobs are at cap-exempt employers, where the first two mechanisms do not operate. Timing alone does not separate the three. Under any of them the extension acted on the first employment relationship, not on the decision to remain in the United States, which changed much less.

Figure 15: Employer Switching after a US First Job: STEM–Non-STEM Difference-in-Differences



Note: Estimates of $\delta_{et}^{\text{NoSwitch}}$ from Equation (11), CtS master's and doctorate graduates whose first job is in the United States, estimated separately by degree and offset horizontally. The outcome indicates that the graduate has not yet moved from the first employer to another employer by year t, so a positive estimate means less switching. Bars are 95 percent confidence intervals clustered by major. Doctorate is all doctoral degrees excluding professional degrees.

7.3 Cap-Exempt First Employers

This STEM extension relaxes the friction that operates through the H-1B lottery. A graduate whose first employer is cap-exempt faces no lottery, so if the extension works through increasing the number of lottery draws, it should change little for these graduates. I classify first employers with the indicator CE_j, which covers higher-education institutions only, and use it in two ways: as an outcome, the type of employer at which a graduate starts, and as a split of the stayer estimates.

Cap-exempt employers account for 3.6 percent of the US first jobs of pre-period STEM master's graduates and 34.5 percent of those of doctorate graduates. Panel A of Table 18 shows that the extension did not change these shares: the difference-in-differences in the cap-exempt share is -0.2 percentage points for master's graduates and 0.6 for doctorate graduates (standard errors 0.4 and 1.8). Because the shares are computed among graduates whose first job is in the United States, and the extension changed who takes one, these are not effects on the probability of starting at a cap-exempt employer among all graduates.

Panel B splits the stayer estimates by the type of first employer. Most of the gains are concentrated among graduates who started at cap-subject employers, as the lottery channel predicts. Five years after graduation, horizontal mismatch is 13.2 percentage points lower for master's graduates at cap-subject employers and 6.9 points lower at cap-exempt ones (standard errors 2.6 and 4.8); for doctorate graduates the estimates are -8.7 and -1.7 (3.0 and 1.9). Firm PageRank is 0.63 and 0.47 log points higher at cap-subject employers and unchanged at cap-exempt ones, and master's occupational PageRank rises by 0.41 log points at cap-subject employers against 0.08 at cap-exempt ones. Two outcomes do not follow the prediction: the occupational median wage of master's graduates rises at both types of employer, by 0.11 and 0.07 log points, and the doctorate PageRank gain is similar at both, 0.22 and 0.17. The comparison has three limits: the master's cap-exempt sample is small, about 400 workers in its smallest group; firm PageRank compares poorly across the split when most cap-exempt employers are universities; and the split conditions on where graduates start, although Panel A shows no shift.

Taken together, the extension changed where graduates start and the jobs they hold more than how long they stay. It raised US placement of STEM master's graduates and moved those who stayed into their field and into higher-ranked occupations and employers, while retention rose little and over-education did not fall. The match gains are concentrated where the lottery binds, which ties them to the hiring friction the extension relaxed rather than to a general improvement for STEM graduates.

8 Conclusion

I use a signaling model with an employer-side hiring friction to organize evidence on immigrant occupational placement. The model predicts that immigrant workers keep their degrees and accept jobs requiring less education, with a gap that shrinks with education. At the first US job, CtS and CtW master's holders are 4.1 and 8.5 percentage points more likely to be over-educated than FUS workers with the same degree, tier, and field. The gap is smaller for doctorate holders: CtS doctorate holders show no gap pooled across employers and a 2.0-point gap at cap-subject employers, while CtW doctorate holders are less often over-educated than FUS doctorate holders, even at cap-subject employers. The excess is largest for CtW professional-degree holders, 16.0 points, consistent with licensing rules that bar the use of foreign professional credentials until they are relicensed. Over-education need not imply a lower-paying occupation: the occupations of CtS master's holders pay 11.9 log points more than those of FUS master's holders, while in the decomposition those of CtW master's holders pay 7.8 log points less. These gaps describe sorting into occupations; they are not structural estimates of the hiring friction d. Differences in occupational tasks account for substantial portions of several raw gaps, but the decomposition does not establish why workers entered those jobs.

Over the next decade, the part of the gap tied to foreign credentials fades for some degrees and not for others. Immigrants in both pathways need sponsorship, but only CtW workers hold credentials that US employers cannot readily evaluate, so the distance between the CtW and CtS paths isolates what a foreign credential adds. Among workers observed in US jobs through year ten, that distance falls from 15.5 to 7.0 log points for doctorate holders and from 17.2 to 7.9 for professional-degree holders, but only from 12.8 to 10.4 for master's holders, while the CtS paths stay within five log points of zero. The two pathways' gaps are priced at different FUS benchmarks, so these distances are descriptive and carry no standard errors. These paths trace occupational placement among long-term stayers and cannot separately identify employer learning, occupational mobility, or changes in the underlying hiring friction.

The 2008 STEM-OPT extension provides evidence that work authorization affects both the margin of US entry and the jobs immigrants hold. For the 2008–2011 cohorts, the share of STEM CtS master's graduates taking a first job in the United States rose by 3.7 percentage points relative to non-STEM graduates, and by 5.6 points over all post-2008 cohorts; the doctorate estimates are similar but rest on less stable pre-period coefficients. Five years after graduation, 2008–2011 STEM master's graduates still in the United States were 7.6 percentage points less likely to work outside their field and held occupations paying 8 log points more, while over-education did not change; over all post-2008 cohorts, doctorate stayers also became less often over-educated. Retention changed little: by year ten it rose 1.9 percentage points for master's graduates over all post-2008 cohorts and

0.6 for the 2008–2011 cohorts, and an imprecise 0.9 for doctorate graduates, too little to produce these gains through a change in who stays. The stayer gains are concentrated among graduates whose first employer was subject to the H-1B cap, where the lottery the extension relaxed applies.

The table below sets these results against the model's predictions to ask what d contains. The answer matters for welfare: if d reflects lower productivity, under-placement is efficient; if it is a cost created by visa rules, under-placement is misallocation. An information friction, in which employers evaluate foreign credentials less accurately, predicts a gap that is larger for foreign degrees and fades with US experience. The first prediction holds for master's and professional-degree holders, whose gaps are larger with foreign degrees; the second, which implies that the CtW path converges to the CtS path, holds for doctorate and professional-degree holders, whose distance halves over ten years, but barely for master's holders, whose distance narrows by a fifth. Convergence is also consistent with sponsorship ending at permanent residence; for professional-degree holders it is unlikely to reflect re-licensing, because their share in licensed occupations does not rise. A hiring cost, such as sponsorship, predicts a gap that persists and that a longer work permit reduces. Both channels are present. The master's gap persists, and the 2008 rule, which lengthened work authorization without changing the training graduates had received, raised US placement and moved stayers into their field and into higher-ranked, higher-paying occupations. At least part of d is therefore a cost that policy can remove, not a productivity gap. The extension speaks only to the part of d it removed, and this interpretation rests on the design's identifying assumptions.

What the Evidence Says about the Friction d

Channel	Predicts	Finding	Evidence
Information (fades with experience)	Gap larger for foreign degrees	Yes for master's and professional degrees: CtW $>$ CtS at entry	Descriptive
	CtW gap converges to CtS gap with US experience	Doctorate and professional degrees: yes, the distance halves over ten years. Master's: little, it narrows by a fifth	Descriptive
Hiring cost (persists: sponsorship, licensing)	A longer work permit raises US placement	Yes: $+3.7$ pp for 2008–2011 master's graduates	Difference-in-differences
	…and improves stayers' jobs	Yes, except over-education: horizontal mismatch falls, occupational PageRank, median wage, and firm PageRank rise, mostly at cap-subject employers; too large to come from extra staying	Difference-in-differences

Notes: Summary of Sections 5.1, 6 and 7.

The occupational median wage decomposition does not identify d without assumptions about selection and unobserved productivity. It adjusts for characteristics measured before labor-market

matching and deliberately excludes the occupation and employer obtained. Its unexplained component can reflect differences other than the hiring friction, and its sign is not uniformly negative. The particularly large mismatch and adjusted deficit among CtW professional-degree holders are consistent with credential-recognition or licensing barriers, but the analysis does not distinguish those mechanisms. The pathways also differ in who can work in the United States at all. CtS graduates enter through Optional Practical Training with little screening; CtW workers enter through a mix of employer-sponsored and family-based visas that likely differs by degree, and foreign-trained professionals must also relicense before practicing. The negative over-education gap of CtW doctorate holders is consistent with selection at entry and with their concentration in postdoctoral positions rather than with the absence of a friction, and the large gap of CtW professional-degree holders may reflect licensing rules as well as employers. The panel observes neither visa type nor licensure, so both accounts remain to be tested directly. H-1B certified offers are not concentrated at the bottom of their occupations' wage distributions; that comparison does not establish pay parity with otherwise comparable FUS entrants. The STEM-OPT design provides a policy contrast for CtS STEM graduates, not an estimate of the level of d for all immigrant workers.

The wage and job-quality outcomes measure occupational or employer characteristics rather than workers' realized pay. They therefore establish differences in placement, not within-job wage effects. Departures from the United States are observed only when workers update their profiles, so the measured extensive margin can contain error. Finally, STEM status is coded from the 2008 designated-degree list.

The main unfinished task is to quantify the welfare cost framed in Appendix H. That requires an estimate of the displaced share $s(\theta)$ and a defensible $d = 0$ counterfactual. Further work should model selection into US employment and continued profile observation. Allowing d to vary with worker type or schooling would test the model's constant-friction restriction. IPEDS-based post-stratification weights would assess sensitivity to selection into the LinkedIn panel. CtW workers offer a further test: they never held student visas, so the extension did not cover them, but they draw from the same H-1B lottery, in which a 29-month OPT gives a STEM CtS graduate up to three draws instead of one. If employers treat the two groups as substitutes, cheaper CtS hiring should lower the placement and job quality of STEM CtW arrivals after 2008, which a triple difference against CtS graduates can check. These exercises would clarify the scope of the findings before using the model for quantitative policy counterfactuals.

Immigrant occupational mismatch differs sharply by credential pathway and degree, and its associated occupational median wage gaps do not follow a single career trajectory. The STEM-OPT evidence indicates that work authorization can change US entry and is consistent with improved placement among observed master's and doctorate stayers. How much of these differences represents a hiring friction, and the welfare gain from removing it, remain to be measured.

References

Abadie, Alberto et al. (2023). "When Should You Adjust Standard Errors for Clustering?" In: *Quarterly Journal of Economics* 138.1, pp. 1–35.

Alba-Ramirez, Alfonso (1993). "Mismatch in the Spanish labor market: overeducation?" In: *Journal of Human Resources*, pp. 259–278.

Aleksynska, Mariya and Ahmed Tritah (2013). "Occupation–education mismatch of immigrant workers in Europe: Context and policies". In: *Economics of Education Review* 36, pp. 229–244.

Allen, Jim and Rolf van der Velden (2001). "Educational Mismatches versus Skill Mismatches: Effects on Wages, Job Satisfaction, and On-the-Job Search". In: *Oxford Economic Papers* 53.3, pp. 434–452. DOI: 10.1093/oep/53.3.434.

Altonji, Joseph G. and Charles R. Pierret (2001). "Employer Learning and Statistical Discrimination". In: *The Quarterly Journal of Economics* 116.1, pp. 313–350. DOI: 10.1162/003355301556329.

Amuedo-Dorantes, Catalina, Delia Furtado, and Huanan Xu (2019). "OPT Policy Changes and Foreign Born STEM Talent in the U.S." In: *Labour Economics* 61, p. 101752.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo (2010). "Beyond Signaling and Human Capital: Education and the Revelation of Ability". In: *American Economic Journal: Applied Economics* 2.4, pp. 76–104.

Autor, David H, Frank Levy, and Richard J Murnane (2003). "Skill demand, inequality, and computerization: Connecting the dots". In: *Technology, growth, and the labor market*. Springer, pp. 107–129.

Baert, Stijn, Bart Cockx, and Dieter Verhaest (2013). "Overeducation at the Start of the Career: Stepping Stone or Trap?" In: *Labour Economics* 25, pp. 123–140. DOI: 10.1016/j.labeco.2013.04.013.

Baley, Isaac, Ana Figueiredo, and Robert Ulbricht (2022). "Mismatch Cycles". In: *Journal of Political Economy* 130.11, pp. 2943–2984.

Batalova, Jeanne and Michael Fix (2021). "Leaving Money on the Table: The Persistence of Brain Waste among College-Educated Immigrants." In: *Migration Policy Institute*.

Becker, Gary S. (1957). *The Economics of Discrimination*. Chicago: University of Chicago Press.

— (1962). "Investment in Human Capital: A Theoretical Analysis". In: *Journal of Political Economy* 70.5, Part 2, pp. 9–49.

Beenstock, Michael, Barry R Chiswick, and Ari Paltiel (2010). "Testing the immigrant assimilation hypothesis with longitudinal data". In: *Review of Economics of the Household* 8.1, pp. 7–27.

Beine, Michel, Giovanni Peri, and Morgan Raux (2023). "International College Students' Impact on the US Skilled Labor Supply". In: *Journal of Public Economics* 223, p. 104917. DOI: 10.1016/j.jpubeco.2023.104917.

Berry, Alexander, Elizabeth M. Maloney, and David Neumark (Mar. 2024). "The Missing Link? Using LinkedIn Data to Measure Race, Ethnic, and Gender Differences in Employment Outcomes at Individual Companies". In: *NBER Working Paper* 32294. DOI: 10.3386/w32294. URL: https://www.nber.org/papers/w32294.

Bohn, Sarah E (2010). "The quantity and quality of new immigrants to the US". In: *Review of Economics of the Household* 8.1, pp. 29–51.

Borjas, George J (1985). "Assimilation, changes in cohort quality, and the earnings of immigrants". In: *Journal of labor Economics* 3.4, pp. 463–489.

Bratsberg, Bernt and James F. Ragan Jr. (2002). "The Impact of Host-Country Schooling on Earnings: A Study of Male Immigrants in the United States". In: *Journal of Human Resources* 37.1, pp. 63–105.

Brin, Sergey and Lawrence Page (1998). "The anatomy of a large-scale hypertextual web search engine". In: *Computer networks and ISDN systems* 30.1-7, pp. 107–117.

Bureau of Labor Statistics (2026). *Occupational Employment and Wage Statistics*. Tech. rep. Accessed: May 15, 2026. U.S. Department of Labor. URL: https://www.bls.gov/oes/.

Card, David, Francesco Devicienti, and Agata Maida (2014). "Rent-sharing, Holdup, and Wages: Evidence from Matched Panel Data". In: *Review of Economic Studies* 81.1, pp. 84–111. DOI: 10.1093/restud/rdt030.

Cassidy, Hugh (2019). "Occupational attainment of natives and immigrants: A cross-cohort analysis". In: *Journal of Human Capital* 13.3, pp. 375–409.

Cassidy, Hugh and Amanda Gaulke (2024). "The increasing penalty to occupation-education mismatch". In: *Economic Inquiry* 62.2, pp. 607–632.

Chen, Willy (2024). "Descriptive Units of Heterogeneity: An Axiomatic Approach to Measuring Heterogeneity". In: *Available at SSRN 4952556*.

Chevalier, Arnaud (2003). "Measuring Over-Education". In: *Economica* 70.279, pp. 509–531. DOI: 10.1111/1468-0335.t01-1-00296.

Chevalier, Arnaud and Joanne Lindley (2009). "Overeducation and the Skills of UK Graduates". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 172.2, pp. 307–337.

Chiswick, Barry R and Paul W Miller (2008). "Why is the payoff to schooling smaller for immigrants?" In: *Labour Economics* 15.6, pp. 1317–1340.

— (2009). "The international transferability of immigrants' human capital". In: *Economics of Education Review* 28.2, pp. 162–169.

Chiswick, Barry R and Paul W Miller (2010). "Occupational language requirements and the value of English in the US labor market". In: *Journal of Population Economics* 23.1, pp. 353–372.

— (2013). "The impact of surplus skills on earnings: Extending the over-education model to language proficiency". In: *Economics of Education Review* 36, pp. 263–275.

Cho, In-Koo and David M Kreps (1987). "Signaling games and stable equilibria". In: *The Quarterly Journal of Economics* 102.2, pp. 179–221.

Collins, William J and Ariell Zimran (2023). "Working their way up? US immigrants' changing labor market assimilation in the age of mass migration". In: *American Economic Journal: Applied Economics* 15.3, pp. 238–269.

Coraggio, Luca et al. (2026). "Mismatch in the 21st Century: An Overview". In: *Labour Economics* 102, p. 102955. DOI: 10.1016/j.labeco.2026.102955.

Demirci, Murat (2019). "Transition of International Science, Technology, Engineering, and Mathematics Students to the U.S. Labor Market: The Role of Visa Policy". In: *Economic Inquiry* 57.3, pp. 1367–1391. DOI: 10.1111/ecin.12795.

Doran, Kirk, Alexander Gelber, and Adam Isen (2022). "The Effects of High-Skilled Immigration Policy on Firms: Evidence from Visa Lotteries". In: *Journal of Political Economy* 130.10, pp. 2501–2533. DOI: 10.1086/720467.

Duncan, Greg J. and Saul D. Hoffman (1981). "The Incidence and Wage Effects of Overeducation". In: *Economics of Education Review* 1.1, pp. 75–86. DOI: 10.1016/0272-7757(81)90028-5.

Fredriksson, Peter, Lena Hensvik, and Oskar Nordström Skans (2018). "Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility". In: *American Economic Review* 108.11, pp. 3303–3338.

Friedberg, Rachel M (2000). "You can't take it with you? Immigrant assimilation and the portability of human capital". In: *Journal of labor economics* 18.2, pp. 221–251.

Garibaldi, Pietro, Pedro Gomes, and Thepthida Sopraseuth (2025a). "Economic Consequences of Vertical Mismatch". In: *Quantitative Economics* 16.2, pp. 535–564.

— (2025b). "Output Costs of Education and Skill Mismatch in OECD Countries". In: *Economics Letters* 250, p. 112278.

Gathmann, Christina and Uta Schönberg (2010). "How general is human capital? A task-based approach". In: *Journal of Labor Economics* 28.1, pp. 1–49.

Gelbach, Jonah B. (2016). "When Do Covariates Matter? And Which Ones, and How Much?" In: *Journal of Labor Economics* 34.2, pp. 509–543. DOI: 10.1086/683668.

Gini, Corrado (1912). "Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche". Italian. In: *Studi Economico-Giuridici della R. Università de Cagliari* 3, pp. 3–159.

Guo, Xue, Jing Gong, and Min-Seok Pang (2024). "Creation or Destruction? STEM OPT Extension and Employment of Information Technology Professionals". In: *MIS Quarterly* 48.2, pp. 715–730.

Guvenen, Fatih et al. (2020). "Multidimensional Skill Mismatch". In: *American Economic Journal: Macroeconomics* 12.1, pp. 210–244.

Hall, Robert E. and Andreas I. Mueller (2018). "Wage Dispersion and Search Behavior: The Importance of Nonwage Job Values". In: *Journal of Political Economy* 126.4, pp. 1594–1637. DOI: 10.1086/697739.

Hartog, Joop (2000). "Over-education and earnings: where are we, where should we go?" In: *Economics of education review* 19.2, pp. 131–147.

Hashimoto, Masanori (1981). "Firm-Specific Human Capital as a Shared Investment". In: *American Economic Review* 71.3, pp. 475–482.

Hashimoto, Masanori and John Raisian (1985). "Employment Tenure and Earnings Profiles in Japan and the United States". In: *American Economic Review* 75.4, pp. 721–735.

Hsieh, Chang-Tai et al. (2019). "The Allocation of Talent and U.S. Economic Growth". In: *Econometrica* 87.5, pp. 1439–1474.

Hunt, Jennifer (2011). "Which Immigrants Are Most Innovative and Entrepreneurial? Distinctions by Entry Visa". In: *Journal of Labor Economics* 29.3, pp. 417–457.

Imai, Susumu, Derek Stacey, and Casey Warman (2019). "From engineer to taxi driver? Language proficiency and the occupational skills of immigrants". In: *Canadian Journal of Economics/Revue canadienne d'économique* 52.3, pp. 914–953.

Kahn, Lisa B. (2010). "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy". In: *Labour Economics* 17.2, pp. 303–316.

Kerr, William R. and William F. Lincoln (2010). "The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention". In: *Journal of Labor Economics* 28.3, pp. 473–508. DOI: 10.1086/651934.

Khanna, Gaurav and Nicolas Morales (Dec. 2024). *The IT Boom and Other Unintended Consequences of Chasing the American Dream*. Working Paper 460. Center for Global Development. URL: https://www.cgdev.org/publication/it-boom-and-other-unintended-consequences-chasing-american-dream.

Khoo, Pauline (2025). "If You Extend It, They Will Come: The Effects of the STEM OPT Extension". In: *Applied Economics*. Online first. DOI: 10.1080/00036846.2025.2551288.

Kiker, Billy Frazier, Maria C Santos, and M Mendes De Oliveira (1997). "Overeducation and undereducation: Evidence for Portugal". In: *Economics of Education Review* 16.2, pp. 111–125.

Kugler, Adriana D. and Robert M. Sauer (2005). "Doctors without Borders? Relicensing Requirements and Negative Selection in the Market for Physicians". In: *Journal of Labor Economics* 23.3, pp. 437–465.

Kwon, Seoyoung, Jongkwan Lee, and Joan Monràs (2026). *The Effect of Retaining High-Skilled International Graduates: Evidence from the STEM OPT Extension*. CEPR Discussion Paper 21224. Centre for Economic Policy Research.

Lachowska, Marta et al. (2026). "Work hours mismatch". In: *Econometrica* 94.3, pp. 991–1025.

Lang, Kevin and Michael Manove (2011). "Education and Labor Market Discrimination". In: *American Economic Review* 101.4, pp. 1467–1496.

Lee, David S. (2009). "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects". In: *Review of Economic Studies* 76.3, pp. 1071–1102.

Leuven, Edwin and Hessel Oosterbeek (2011). "Overeducation and mismatch in the labor market". In: *Handbook of the Economics of Education* 4, pp. 283–326.

Li, Xiaoguang and Yao Lu (2023). "Education–occupation mismatch and nativity inequality among highly educated US workers". In: *Demography* 60.1, pp. 201–226.

LinkUp (2025). *Job Records*. Dewey Data. DOI: 10.82551/V091-HN53. URL: https://doi.org/10.82551/V091-HN53.

Lise, Jeremy and Fabien Postel-Vinay (Aug. 2020). "Multidimensional Skills, Sorting, and Human Capital Accumulation". In: *American Economic Review* 110.8, pp. 2328–76. DOI: 10.1257/aer.20162002. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20162002.

Llama 3.1 (8B) (2024). Model Card. Version Release Date: July 23, 2024. Licensed under the Llama 3.1 Community License Agreement. URL: https://huggingface.co/meta-llama/Llama-3.1-8B.

Lu, Yao and Feng Hou (2020). "Immigration system, labor market structures, and overeducation of high-skilled immigrants in the United States and Canada". In: *International migration review* 54.4, pp. 1072–1103.

Lu, Yao and Xiaoguang Li (2021). "Vertical education-occupation mismatch and wage inequality by race/ethnicity and nativity among highly educated US workers". In: *Social Forces* 100.2, pp. 706–737.

Lubotsky, Darren (2007). "Chutes or ladders? A longitudinal analysis of immigrant earnings". In: *Journal of Political Economy* 115.5, pp. 820–867.

Mailath, George J (1987). "Incentive compatibility in signaling games with a continuum of types". In: *Econometrica: Journal of the Econometric Society*, pp. 1349–1365.

Mas, Alexandre and Amanda Pallais (2017). "Valuing Alternative Work Arrangements". In: *American Economic Review* 107.12, pp. 3722–3759. DOI: 10.1257/aer.20161500.

Mas-Colell, Andreu, Michael D Whinston, and Jerry R Green (1995). *Microeconomic theory*. Vol. 1. Oxford University Press New York.

Mattoo, Aaditya, Ileana Cristina Neagu, and Çağlar Özden (2008). "Brain waste? Educated immigrants in the US labor market". In: *Journal of development economics* 87.2, pp. 255–269.

McGuinness, Seamus, Konstantinos Pouliakas, and Paul Redmond (2018). "Skills Mismatch: Concepts, Measurement and Policy Approaches". In: *Journal of Economic Surveys* 32.4, pp. 985–1015. DOI: 10.1111/joes.12254.

Microsoft (2025). *Microsoft Copilot*. https://copilot.microsoft.com. Used to construct names for the three principal components of skills and abilities from O*NET.

Mortensen, Dale T and Christopher A Pissarides (1994). "Job creation and job destruction in the theory of unemployment". In: *The review of economic studies* 61.3, pp. 397–415.

National Center for O*NET Development (2025). *O*NET Database (Version 29.2)*. O*NET 29.2 released February 2025. Sponsored by the U.S. Department of Labor, Employment & Training Administration. URL: https://www.onetcenter.org/db_releases.html.

Nieto, Sandra, Alessia Matano, and Raul Ramos (2013). *Skill Mismatches in the EU: Immigrants vs. Natives*. Working Paper 18. Research Institute of Applied Economics (IREA), University of Barcelona.

Ordine, Patrizia and Giuseppe Rose (2009). "Overeducation and Instructional Quality: A Theoretical Model and Some Facts". In: *Journal of Human Capital* 3.1, pp. 73–105.

Oreopoulos, Philip (2011). "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes". In: *American Economic Journal: Economic Policy* 3.4, pp. 148–171. DOI: 10.1257/pol.3.4.148.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz (2012). "The Short- and Long-Term Career Effects of Graduating in a Recession". In: *American Economic Journal: Applied Economics* 4.1, pp. 1–29.

People Data Labs (2022). *Person*. Dewey Data. DOI: 10.82551/ZSJV-Z052. URL: https://doi.org/10.82551/ZSJV-Z052.

Pivovarova, Margarita and Jeanne M. Powers (2022). "Do Immigrants Experience Labor Market Mismatch? New Evidence from the US PIAAC". In: *Large-scale Assessments in Education* 10, p. 9.

Quacquarelli Symonds (2019). *QS World University Rankings (2008–2019)*. Tech. rep. Historical tracking panel covering cycles 2008 through 2019. URL: https://www.topuniversities.com/world-university-rankings.

Rendon, Silvio and Alfredo Cuecuecha (2010). "International Job Search: Mexicans in and out of the US". In: *Review of Economics of the Household* 8.1, pp. 53–82.

Robst, John (2007). “Education and job match: The relatedness of college major and work”. In: *Economics of Education review* 26.4, pp. 397–407.

Rosen, Sherwin (1986). “The Theory of Equalizing Differences”. In: *Handbook of Labor Economics*. Ed. by Orley C. Ashenfelter and Richard Layard. Vol. 1. Amsterdam: North-Holland. Chap. 12, pp. 641–692.

Ruggles, Steven et al. (2025). *IPUMS USA: Version 16.0*. Minneapolis, MN: IPUMS. DOI: 10.18128/D010.V16.0. URL: https://doi.org/10.18128/D010.V16.0.

Schwandt, Hannes and Till von Wachter (2019). “Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets”. In: *Journal of Labor Economics* 37.S1, S161–S198.

Sicherman, Nachum (1991). “” Overeducation” in the labor market”. In: *Journal of labor Economics* 9.2, pp. 101–122.

Simpson, Edward H. (1949). “Measurement of Diversity”. In: *Nature* 163, p. 688. DOI: 10.1038/163688a0.

Somers, Melline A. et al. (2019). “Horizontal Mismatch between Employment and Field of Education: Evidence from a Systematic Literature Review”. In: *Journal of Economic Surveys* 33.2, pp. 567–603.

Sorkin, Isaac (2018). “Ranking firms using revealed preference”. In: *The quarterly journal of economics* 133.3, pp. 1331–1393.

Spence, Michael (1978). “Job market signaling”. In: *Uncertainty in economics*. Elsevier, pp. 281–306.

Sullivan, Paul and Ted To (2014). “Search and Nonwage Job Characteristics”. In: *Journal of Human Resources* 49.2, pp. 472–507. DOI: 10.3368/jhr.49.2.472.

Summerfield, Fraser and Ioannis Theodossiou (2017). “The Effects of Macroeconomic Conditions at Graduation on Overeducation”. In: *Economic Inquiry* 55.3, pp. 1370–1387.

Taber, Christopher and Rune Vejlin (2020). “Estimation of a Roy/Search/Compensating Differential Model of the Labor Market”. In: *Econometrica* 88.3, pp. 1031–1069. DOI: 10.3982/ECTA14441.

Team, MLX Explore (2026). *mlx-lm*. Version 0.31.0. Python package for running, fine-tuning, and quantizing large language models on Apple Silicon using the MLX framework. URL: https://pypi.org/project/mlx-lm/.

U.S. Department of Education, National Center for Education Statistics (2019). *Integrated Postsecondary Education Data System (IPEDS): Completions Component*. Tech. rep. URL: https://nces.ed.gov/ipeds/datacenter/DataFiles.aspx.

U.S. Department of Labor (2026). *Labor Condition Application (LCA) Disclosure Data 2011-2025*. Employment and Training Administration (ETA), Office of Foreign Labor Certification. Ac-

cessed: 2025-12-15. URL: https://www.dol.gov/agencies/eta/foreign-labor/performance.

Van Dijcke, David, Florian Gunsilius, and Austin Wright (May 2024). *Return to Office and the Tenure Distribution*. Working Paper. University of Michigan.

Verdugo, Richard R and Naomi Turner Verdugo (1989). "The impact of surplus schooling on earnings: Some additional findings". In: *Journal of human resources*, pp. 629–643.

Visintin, Stefano, Kea Tijdens, and Maarten van Klaveren (2015). "Skill Mismatch among Migrant Workers: Evidence from a Large Multi-Country Dataset". In: *IZA Journal of Migration* 4, pp. 1–34.

Waldman, Michael (1984). "Job Assignments, Signalling, and Efficiency". In: *RAND Journal of Economics* 15.2, pp. 255–267.

Warman, Casey and Christopher Worswick (2015). "Technological change, occupational tasks and declining immigrant outcomes: Implications for earnings and income inequality in Canada". In: *Canadian Journal of Economics/Revue canadienne d'économique* 48.2, pp. 736–772.

Weber, Rosa, Mathieu Ferry, and Mathieu Ichou (2024). "Which degree for which occupation? Vertical and horizontal mismatch among immigrants, their children, and grandchildren in France". In: *Demography* 61.6, pp. 1923–1948.

9 Tables

Table 1: Predicted Equilibrium Contract and Over-Education Gap by Degree and Degree Origin

	Worker Type (Terminal Degree Origin)		
Pathway (Education location)	Native	Immigrant (US)	Immigrant (Foreign)
Master's ($\bar{e}'$) [Rate of Over-education relative to Natives]	$W(\bar{e}')$	$W(\tilde{e}')$ [++]	$W(\tilde{e}')$ [+ + +]
Doctorate ($\bar{e}$) [Rate of Over-education relative to Natives]	$W(\bar{e})$	$W(\tilde{e})$ [+]	$W(\tilde{e})$ [++]
Professional ($\bar{e}$) [Rate of Over-education relative to Natives]	$W(\bar{e})$	$W(\tilde{e})$ [+]	$W(\tilde{e})$, larger d [+ + +]

Notes: Every immigrant holds the under-placement contract $W(\tilde{e}) = W(\bar{e}) - d$ at their own degree. Signs give the predicted over-education gap relative to natives with the same degree; more signs denote a larger gap. The gap is smaller at higher degrees because $W(e)$ is convex, and larger with a foreign degree if the friction includes difficulty evaluating foreign credentials. Doctorate comprises the PhD and other research doctorates that do not lead to a professional license. Professional comprises degrees that require a US or state license to practice, such as the MD and the JD; a license earned abroad does not transfer, which adds to d for foreign-educated holders.

Table 2: Sample Coverage of the First-Job Analysis

Pipeline	Initial first jobs	Analytic sample	Retained
CtS–FUS, graduate degrees	8,873,285	4,431,443	49.9%
CtW–FUS, bachelor's and above	29,417,691	11,823,563	40.2%

Notes: USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. One observation per person. The analytic sample is the common-support sample used in every column of Tables 6 through 9. Common support requires the employer block (firm utility, firm churn, H-1B sponsorship, and cap-exempt status) and both mismatch measures; the CtW pipeline also requires arrival cohort, origin region, and non-US experience. The main losses come from the employer measures, chiefly firm utility, which exists only for employers in the connected set of worker flows between firms. Estimates apply to the covered population of credentialed LinkedIn users and are not population-representative.

Table 3: Degree and Graduation Year Distribution

	LinkedIn									ACS
Pathway	Fully US-educated			Came-to-Study			Came-to-Work			
Observations	17.0M			1.5M			9.4M			3.8 M
	Row %	N	Col.%	Row %	N	Col.%	Row %	N	Col.%	
Terminal Degree										
Bachelor's	66.2%	10.4M	60.8%	1.5%	0.2M	15.7%	32.3%	5.1M	53.7%	64.2%
Master's	52.5%	5.4M	31.9%	9.8%	1.0M	66.0%	37.7%	3.9M	41.4%	26.0%
Doctorate	46.4%	0.5M	3.2%	19.8%	0.2M	15.1%	33.8%	0.4M	4.2%	4.0%
Professional	85.5%	0.7M	4.2%	6.0%	0.05M	3.3%	8.5%	0.1M	0.8%	5.8%
Graduation/Entering Cohorts										
'03–'04	62.5%	0.8M	4.6%	4.7%	0.1M	3.9%	32.7%	0.4M	4.4%	–
'05–'09	61.9%	2.8M	16.4%	4.9%	0.2M	14.6%	33.2%	1.5M	16.0%	–
'10–'14	61.7%	4.8M	28.0%	5.2%	0.4M	26.2%	33.1%	2.6M	27.2%	–
'15–'19	60.0%	6.0M	35.5%	5.9%	0.6M	39.1%	34.0%	3.4M	36.4%	–
'20–'21	60.0%	2.6M	15.5%	5.6%	0.2M	16.1%	34.4%	1.5M	16.1%	–

Notes: USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Entries are row percentages, person counts, and column percentages. Counts use the terminal (highest-ranked) degree per person. Doctorate comprises the PhD and doctorates that do not lead to a professional license (EdD, DBA, DMin, DMA, DSc, DrPH); Professional comprises licensure doctorates (MD, JD, DDS, PharmD, and similar). The ACS column reports doctorate degrees and professional degrees beyond a bachelor's degree, which follow the same division. The LinkedIn universe is restricted to terminal-degree graduation cohorts 2003–2021. The ACS column is an independent extract and is never pooled with the LinkedIn columns; it is retained only for the terminal-degree comparison because the supplied ACS extract has no comparable graduation-cohort measure. Percentages may not sum to 100 because of rounding.

Table 4: Post-Graduation Observation Years by Cohort and Status

	LinkedIn								
	Fully US-educated 17.0M			Came-to-Study 1.5M			Came-to-Work 9.4M		
Graduation/Entering Cohorts	(Observed Years) Total	After Grad/Entry	in US After Grad/Entry						
'03–'04	26.2	23.5	11.3	26.6	23.5	7.7	25.0	23.5	2.5
'05–'09	23.1	19.8	9.9	23.6	19.8	6.7	21.8	19.8	2.0
'10–'14	19.2	14.8	7.8	19.6	14.8	5.1	17.7	14.8	1.2
'15–'19	15.5	10.0	5.3	15.7	10.0	3.5	13.6	9.9	0.5
'20–'21	12.4	6.5	3.3	13.2	6.5	2.1	10.9	6.5	0.2

Notes: USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Entries are mean distinct calendar years in which a person appears in the panel: total observed years, observed years at or after terminal-degree graduation, and the latter restricted to `live_in_us_raw = 1`. The universe and the inclusive 2003–2021 graduation-cohort restriction are identical to the first table. The ACS has no comparable longitudinal observation and is therefore omitted.

Table 5: Post-Secondary Institution Selectivity Bins

Tier	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Selectivity Ranking	Unranked	Open or not reported	Bottom 25%	75%–50%	50%–33%	33%–20%	20%–10%	10%–5%	5%–1%	Top 1%

Notes: A worker's tier is that of the institution that granted their terminal degree. Tier 0 (unranked): the terminal institution has no tier. Tier 1: the institution is matched but reports no admissions selectivity or admits all applicants.

Table 6: Probability of First Job Over-Education by Postgraduate Degree Holders, CtS vs. FUS

	(1)	(2)	(3)	(4)	(5)
			Pr(Over-Education)		
CtS	0.119***	0.114***	0.041***	0.041***	0.041***
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Doctorate	-0.249***	–	–	–	–
	(0.001)				
Professional	-0.572***	–	–	–	–
	(0.001)				
CtS × Doctorate	-0.087***	-0.079***	-0.042***	0.021***	0.019***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
CtS × Professional	-0.078***	-0.072***	-0.007*	0.008*	0.006
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Cap-Exempt				-0.231***	-0.227***
				(0.001)	(0.001)
CtS × Cap-Exempt				-0.173***	-0.165***
				(0.002)	(0.002)
Hor. Mis.					0.084***
					(0.000)
Baseline Pr. (FUS, Master's)	0.796	0.796	0.796	0.796	0.796
N	4,229,371	4,229,371	4,229,367	4,229,367	4,229,367
Cohort FE	✓	✓	✓	✓	✓
Degree × School Tier		✓	✓	✓	✓
Degree × field FE, STEM, generality			✓	✓	✓
Employer block				✓	✓

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Column 3 is the main specification. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). The employer block contains log firm utility, 3-, 6-, 12-, and 24-month churn, an indicator for any H-1B filing, cap-exempt status, and its interaction with the pathway, so pathway coefficients in Columns 4 and 5 are gaps at cap-subject employers. Degree main effects are absorbed by the degree-by-field fixed effects from Column 3 on. Baseline: share of FUS master's holders with the outcome.

Table 7: Probability of First Job Horizontal Mismatch by Postgraduate Degree Holders, CtS vs. FUS

	(1)	(2)	(3)	(4)	(5)
	Pr(Horizontal Mismatch)				
CtS	-0.030***	-0.025***	-0.012***	-0.009***	-0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Doctorate	-0.054***	–	–	–	–
	(0.001)				
Professional	-0.003**	–	–	–	–
	(0.001)				
CtS × Doctorate	-0.005**	-0.008***	-0.014***	0.017***	0.014***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
CtS × Professional	0.012***	-0.000	0.019***	0.023***	0.022***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Cap-Exempt				-0.051***	-0.020***
				(0.001)	(0.001)
CtS × Cap-Exempt				-0.092***	-0.069***
				(0.002)	(0.002)
Over Edu.					0.133***
					(0.001)
Baseline Pr. (FUS, Master's)	0.374	0.374	0.374	0.374	0.374
N	4,229,371	4,229,371	4,229,367	4,229,367	4,229,367
Cohort FE	✓	✓	✓	✓	✓
Degree × School Tier		✓	✓	✓	✓
Degree × field FE, STEM, generality			✓	✓	✓
Employer block				✓	✓

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Column 3 is the main specification. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). The employer block contains log firm utility, 3-, 6-, 12-, and 24-month churn, an indicator for any H-1B filing, cap-exempt status, and its interaction with the pathway, so pathway coefficients in Columns 4 and 5 are gaps at cap-subject employers. Degree main effects are absorbed by the degree-by-field fixed effects from Column 3 on. Baseline: share of FUS master's holders with the outcome.

Table 8: Probability of First Job Over-Education by Postgraduate Degree Holders, CtW vs. FUS

	(1)	(2)	(3)	(4)	(5)
	Pr(Over-Education)				
CtW	0.133***	0.143***	0.085***	0.079***	0.074***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Bachelor's	-0.438***	–	–	–	–
	(0.000)				
Doctorate	-0.249***	–	–	–	–
	(0.001)				
Professional	-0.572***	–	–	–	–
	(0.001)				
CtW × Bachelor's	-0.214***	-0.209***	-0.115***	-0.112***	-0.107***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
CtW × Doctorate	-0.189***	-0.192***	-0.113***	-0.038***	-0.038***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
CtW × Professional	0.022***	0.022***	0.075***	0.101***	0.090***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Cap-Exempt				-0.164***	-0.163***
				(0.001)	(0.001)
CtW × Cap-Exempt				-0.182***	-0.150***
				(0.002)	(0.002)
Hor. Mis.					0.163***
					(0.000)
Baseline Pr. (FUS, Master's)	0.796	0.796	0.796	0.796	0.796
N	11,430,012	11,430,012	11,430,007	11,430,007	11,430,007
Cohort FE	✓	✓	✓	✓	✓
Degree × School Tier		✓	✓	✓	✓
Degree × field FE, STEM, generality			✓	✓	✓
Employer block				✓	✓

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Column 3 is the main specification. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). The employer block contains log firm utility, 3-, 6-, 12-, and 24-month churn, an indicator for any H-1B filing, cap-exempt status, and its interaction with the pathway, so pathway coefficients in Columns 4 and 5 are gaps at cap-subject employers. Degree main effects are absorbed by the degree-by-field fixed effects from Column 3 on. Baseline: share of FUS master's holders with the outcome. Quadratic non-U.S. experience and within-employer U.S.-bound transfer are controlled for in all columns.

Table 9: Probability of First Job Horizontal Mismatch by Postgraduate Degree Holders, CtW vs. FUS

	(1)	(2)	(3)	(4)	(5)
	Pr(Horizontal Mismatch)				
CtW	0.020***	0.031***	0.028***	0.029***	0.013***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Bachelor's	0.044***	–	–	–	–
	(0.000)				
Doctorate	-0.054***	–	–	–	–
	(0.001)				
Professional	-0.004***	–	–	–	–
	(0.001)				
CtW × Bachelor's	-0.040***	-0.051***	-0.030***	-0.032***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
CtW × Doctorate	-0.075***	-0.079***	-0.071***	0.002	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
CtW × Professional	0.001	-0.021***	0.049***	0.066***	0.046***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Cap-Exempt				-0.006***	0.027***
				(0.001)	(0.001)
CtW × Cap-Exempt				-0.197***	-0.159***
				(0.003)	(0.002)
Over Edu.					0.206***
					(0.000)
Baseline Pr. (FUS, Master's)	0.374	0.374	0.374	0.374	0.374
N	11,430,012	11,430,012	11,430,007	11,430,007	11,430,007
Cohort FE	✓	✓	✓	✓	✓
Degree × School Tier		✓	✓	✓	✓
Degree × field FE, STEM, generality			✓	✓	✓
Employer block				✓	✓

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Column 3 is the main specification. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). The employer block contains log firm utility, 3-, 6-, 12-, and 24-month churn, an indicator for any H-1B filing, cap-exempt status, and its interaction with the pathway, so pathway coefficients in Columns 4 and 5 are gaps at cap-subject employers. Degree main effects are absorbed by the degree-by-field fixed effects from Column 3 on. Baseline: share of FUS master's holders with the outcome. Quadratic non-U.S. experience and within-employer U.S.-bound transfer are controlled for in all columns.

Table 10: Over-Education Gap Relative to FUS Workers with the Same Degree, by Type of First Employer

		Cap-exempt	Cap-subject	Difference
Master's	CtS	5.5 [4.8, 6.3]	3.8 [3.7, 3.9]	1.7 [1.0, 2.5]
	CtW	−1.5 [−3.1, 0.1]	8.2 [7.9, 8.5]	−9.8 [−11.4, −8.1]
Doctorate	CtS	−2.9 [−3.4, −2.4]	2.0 [1.6, 2.3]	−4.9 [−5.5, −4.3]
	CtW	−7.5 [−8.7, −6.3]	−2.1 [−2.6, −1.6]	−5.4 [−6.7, −4.1]
Professional	CtS	6.7 [4.5, 9.0]	3.2 [2.6, 3.8]	3.5 [1.2, 5.9]
	CtW	1.7 [−1.0, 4.4]	17.4 [16.5, 18.3]	−15.7 [−18.6, −12.9]

Notes: USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Column 3 of Tables 6 and 8, estimated separately on workers whose first employer is cap-exempt and cap-subject. Cap-exempt employers account for 7.1, 4.1, and 1.5 percent of the first jobs of FUS, CtS, and CtW master's holders and 40.3, 36.9, and 38.5 percent of those of doctorate holders. Estimates are in percentage points, with 95 percent confidence intervals from person-clustered standard errors in brackets. The difference assumes the two samples share no clusters. Cap-exempt employers are higher-education institutions. The first employer is chosen in the same search as the job, so the split locates the gap; it is not an effect of cap exemption.

Table 11: First Job Occupational Utility Sorting of CtS Workers (vs. FUS)

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(SOC PageRank Score)					
CtS	0.403***	0.320***	0.292***	0.230***	0.223***	0.219***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Doctorate	0.673***	0.618***	0.598***	0.571***	0.557***	0.555***
	(0.003)	(0.002)	(0.004)	(0.004)	(0.004)	(0.004)
Professional	1.167***	1.290***	1.335***	1.119***	1.126***	1.127***
	(0.003)	(0.002)	(0.006)	(0.005)	(0.005)	(0.005)
CtS × Doctorate	-0.219***	-0.211***	-0.191***	-0.163***	-0.167***	-0.195***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
CtS × Professional	-0.437***	-0.402***	-0.364***	-0.291***	-0.284***	-0.286***
	(0.012)	(0.009)	(0.009)	(0.007)	(0.007)	(0.007)
Cap-Exempt						0.014***
						(0.002)
CtS × Cap-Exempt						0.087***
						(0.005)
Over Edu.		0.182***	0.175***	0.070***	0.064***	0.068***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Hor. Mis.		-2.031***	-2.030***	-2.024***	-2.029***	-2.028***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
OE × HM		0.150***	0.157***	0.481***	0.486***	0.486***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
N	4,266,561	4,266,561	4,266,561	4,266,561	4,266,561	4,266,561
Cohort FE	✓	✓	✓	✓	✓	✓
Occ. Mis.		✓	✓	✓	✓	✓
Degree × School Tier			✓	✓	✓	✓
O*NET Job Char.				✓	✓	✓
Firm Utility.					✓	✓

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). Unlike Column 3 of Tables 6–9, these specifications control for degree-by-tier effects but not for field of study, so the pathway gaps are not directly comparable with those tables. Occupational mismatch controls are indicator variables for over-education, horizontal mismatch, and both over-education and horizontal mismatch.Firm utility controls include firm pagerank utility a la Sorkin (2018) and firm 3-, 6-, 12-, and 24-month churn rates. Cap-exempt refers to H-1B cap-exempt status of the employer, proxied by institution of higher education in this table.O*NET job characteristics include the on-the-job training fixed effects, standardized importance of certification, and the three principal components labeled Cognitive and Interpersonal Capabilities, Attention and Coordination, and Practical Job Capabilities. Quadratic non-U.S. experience and within-employer U.S.-bound transfer are controlled for in all columns.

Table 12: First Job Occupational Utility Sorting of CtW Workers (vs. FUS)

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(SOC PageRank Score)					
CtW	0.165***	0.262***	0.249***	0.167***	0.149***	0.140***
	(0.005)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Bachelor's	-0.238***	-0.363***	-0.449***	-0.062***	-0.058***	-0.067***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Doctorate	0.671***	0.464***	0.416***	0.540***	0.526***	0.571***
	(0.003)	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)
Professional	1.172***	0.900***	0.922***	1.075***	1.084***	1.079***
	(0.003)	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)
CtW × Bachelor's	0.029***	-0.140***	-0.139***	-0.083***	-0.082***	-0.078***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
CtW × Doctorate	0.000	-0.221***	-0.209***	-0.106***	-0.100***	-0.128***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
CtW × Professional	-0.750***	-0.734***	-0.670***	-0.563***	-0.558***	-0.552***
	(0.014)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)
Cap-Exempt						-0.148***
						(0.002)
CtW × Cap-Exempt						0.092***
						(0.005)
Over Edu.		-0.459***	-0.462***	-0.106***	-0.110***	-0.117***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hor. Mis.		-1.774***	-1.774***	-1.721***	-1.724***	-1.724***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
OE × HM		-0.004**	0.000	0.317***	0.321***	0.319***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Non-US Exp.	-0.050***	-0.032***	-0.030***	-0.021***	-0.018***	-0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Non-US Exp.^{2}	0.003***	0.004***	0.004***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Same Employer (Transfer)	0.253***	0.214***	0.217***	0.132***	0.105***	0.104***
	(0.009)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
N	11,547,735	11,547,735	11,547,735	11,547,735	11,547,735	11,547,735
Graduation Cohort FE	✓	✓	✓	✓	✓	✓
Arrival Cohort, Origin Region FE	✓	✓	✓	✓	✓	✓
Occ. Mis.		✓	✓	✓	✓	✓
Degree × School Tier			✓	✓	✓	✓
O*NET Job Char.				✓	✓	✓
Firm Utility.					✓	✓

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). Unlike Column 3 of Tables 6–9, these specifications control for degree-by-tier effects but not for field of study, so the pathway gaps are not directly comparable with those tables. Occupational mismatch controls are indicator variables for over-education, horizontal mismatch, and both over-education and horizontal mismatch.Firm utility controls include firm pagerank utility a la Sorkin (2018) and firm 3-, 6-, 12-, and 24-month churn rates. Cap-exempt refers to H-1B cap-exempt status of the employer, proxied by institution of higher education in this table.O*NET job characteristics include the on-the-job training fixed effects, standardized importance of certification, and the three principal components labeled Cognitive and Interpersonal Capabilities, Attention and Coordination, and Practical Job Capabilities. Quadratic non-U.S. experience and within-employer U.S.-bound transfer are controlled for in all columns.

Table 13: First U.S. Job Wage Sorting of CtS Workers (vs. FUS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(OEWS Median Wage)						
CtS	0.114***	0.110***	0.097***	0.030***	0.028***	0.013***	0.015***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Doctorate	-0.026***	-0.022***	-0.050***	-0.032***	-0.037***	-0.077***	-0.069***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Professional	0.337***	0.322***	0.306***	0.251***	0.254***	0.175***	0.172***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
CtS × Doctorate	-0.119***	-0.115***	-0.103***	-0.062***	-0.063***	-0.053***	-0.038***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
CtS × Professional	-0.147***	-0.136***	-0.125***	-0.046***	-0.044***	-0.025***	-0.023***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Cap-Exempt							-0.039***
							(0.001)
CtS × Cap-Exempt							-0.046***
							(0.001)
Over Edu.		0.074***	0.069***	0.099***	0.098***	0.092***	0.087***
		(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Hor. Mis.		0.149***	0.148***	0.072***	0.071***	0.209***	0.209***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
OE × HM		-0.371***	-0.366***	-0.111***	-0.110***	-0.144***	-0.145***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln(SOC PageRank)						0.069***	0.069***
						(0.000)	(0.000)
N	3,977,618	3,977,618	3,977,618	3,977,618	3,977,618	3,977,618	3,977,618
Cohort FE	✓	✓	✓	✓	✓	✓	✓
Occ. Mis.		✓	✓	✓	✓	✓	✓
Degree × School Tier			✓	✓	✓	✓	✓
O*NET Job Char.				✓	✓	✓	✓
Firm Utility.					✓	✓	✓

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). Unlike Column 3 of Tables 6–9, these specifications control for degree-by-tier effects but not for field of study, so the pathway gaps are not directly comparable with those tables. Firm utility controls include firm pagerank utility a la Sorkin (2018) and firm 3-, 6-, 12-, and 24-month churn rates. Cap-exempt refers to H-1B cap-exempt status of the employer, proxied by institution of higher education in this table.O*NET job characteristics include the on-the-job training fixed effects, standardized importance of certification, and the three principal components labeled Cognitive and Interpersonal Capabilities, Attention and Coordination, and Practical Job Capabilities. Quadratic non-U.S. experience and within-employer U.S.-bound transfer are controlled for in all columns.

Table 14: First U.S. Job Wage Sorting of CtW Workers (vs. FUS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(OEWS Median Wage)						
CtW	0.057***	0.115***	0.110***	0.055***	0.049***	0.042***	0.041***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Bachelor's	-0.183***	-0.345***	-0.351***	-0.017***	-0.016***	-0.013***	-0.019***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Doctorate	-0.026***	-0.129***	-0.178***	-0.072***	-0.076***	-0.106***	-0.085***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Professional	0.337***	0.115***	0.095***	0.210***	0.213***	0.152***	0.150***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
CtW × Bachelor's	0.042***	-0.050***	-0.056***	-0.025***	-0.025***	-0.021***	-0.020***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
CtW × Doctorate	-0.152***	-0.237***	-0.237***	-0.113***	-0.110***	-0.106***	-0.065***
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
CtW × Professional	-0.105***	-0.080***	-0.111***	-0.088***	-0.086***	-0.055***	-0.042***
	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Cap-Exempt							-0.069***
							(0.000)
CtW × Cap-Exempt							-0.099***
							(0.001)
Over Edu.		-0.347***	-0.347***	-0.015***	-0.016***	-0.011***	-0.016***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hor. Mis.		-0.039***	-0.039***	-0.016***	-0.017***	0.075***	0.074***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
OE × HM		-0.157***	-0.155***	0.008***	0.009***	-0.008***	-0.009***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln(SOC PageRank)						0.054***	0.053***
						(0.000)	(0.000)
Non-US Exp.	-0.021***	-0.020***	-0.019***	-0.006***	-0.006***	-0.005***	-0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Non-US Exp.^{2}	0.007***	0.007***	0.007***	0.004***	0.004***	0.003***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Same Employer (Transfer)	0.135***	0.133***	0.133***	0.061***	0.053***	0.047***	0.045***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
N	10,787,186	10,787,186	10,787,186	10,787,186	10,787,186	10,787,186	10,787,186
Graduation Cohort FE	✓	✓	✓	✓	✓	✓	✓
Cohort and Region FE	✓	✓	✓	✓	✓	✓	✓
Occ. Mis.		✓	✓	✓	✓	✓	✓
Degree × School Tier			✓	✓	✓	✓	✓
O*NET Job Char.				✓	✓	✓	✓
Firm Utility.					✓	✓	✓

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). Unlike Column 3 of Tables 6–9, these specifications control for degree-by-tier effects but not for field of study, so the pathway gaps are not directly comparable with those tables. Firm utility controls include firm pagerank utility a la Sorkin (2018) and firm 3-, 6-, 12-, and 24-month churn rates. Cap-exempt refers to H-1B cap-exempt status of the employer, proxied by institution of higher education in this table.O*NET job characteristics include the on-the-job training fixed effects, standardized importance of certification, and the three principal components labeled Cognitive and Interpersonal Capabilities, Attention and Coordination, and Practical Job Capabilities. Quadratic non-U.S. experience and within-employer U.S.-bound transfer are controlled for in all columns.

Table 15: Decomposing CtS Worker SOC Median Wage Sorting, First U.S. Job

	(1) Master's	(2) Doctorate	(3) Professional
CtS	0.053***	-0.081***	0.233***
	(0.001)	(0.001)	(0.003)
FUS	-0.065***	-0.087***	0.261***
	(0.000)	(0.001)	(0.001)
Difference	0.118***	0.005***	-0.028***
	(0.001)	(0.001)	(0.003)
Explained	0.086***	0.028***	0.003**
	(0.000)	(0.001)	(0.001)
Unexplained	0.032***	-0.023***	-0.031***
	(0.001)	(0.001)	(0.003)
Explained, by covariate			
STEM	0.006***	-0.022***	0.000***
	(0.000)	(0.001)	(0.000)
Grad $\geq$ 2008	0.002***	0.001***	-0.000***
	(0.000)	(0.000)	(0.000)
STEM $\times$ Grad $\geq$ 2008	0.016***	0.016***	-0.002***
	(0.000)	(0.001)	(0.000)
Major Generality	-0.000***	-0.007***	0.003***
	(0.000)	(0.000)	(0.000)
Years since Degree	-0.005***	-0.002***	-0.008***
	(0.000)	(0.000)	(0.001)
Years since Degree^2	0.001***	-0.001***	0.006***
	(0.000)	(0.000)	(0.000)
Unexplained, by covariate			
STEM	-0.010***	0.002	0.001
	(0.002)	(0.008)	(0.004)
Grad $\geq$ 2008	0.003	0.012	0.016*
	(0.002)	(0.007)	(0.008)
STEM $\times$ Grad $\geq$ 2008	0.025***	-0.001	-0.004
	(0.002)	(0.007)	(0.004)
Major Generality	0.041***	-0.025**	-0.078***
	(0.003)	(0.008)	(0.015)
Years since Degree	-0.010***	-0.003	-0.015*
	(0.001)	(0.004)	(0.007)
Years since Degree^2	0.002**	0.007***	0.002
	(0.001)	(0.002)	(0.004)
N	3,590,577	440,037	505,626

Notes: Standard errors in parentheses. All columns use only the first job-year observation of a worker after their terminal degree. First jobs in the United States, 2003–2021 graduation cohorts. The outcome is the log OEWS median wage of the occupation, net of graduation-cohort effects estimated on the full panel, as in Figures 5 and 6 and the ten-year decompositions. Covariates are STEM, graduation in 2008 or later and its interaction with STEM, major generality, years since degree and its square, school tier, and major field; school-tier and major-field contributions are included in the totals but not shown. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Table 16: Decomposing CtW Worker SOC Median Wage Sorting, First U.S. Job

	(1) Master's	(2) Doctorate	(3) Professional
CtW	-0.023***	-0.198***	0.238***
	(0.001)	(0.002)	(0.005)
FUS	0.052***	0.030***	0.378***
	(0.000)	(0.001)	(0.001)
Difference	-0.075***	-0.228***	-0.139***
	(0.001)	(0.002)	(0.005)
Explained	0.040***	-0.020***	0.113***
	(0.000)	(0.001)	(0.002)
Unexplained	-0.115***	-0.208***	-0.252***
	(0.001)	(0.002)	(0.005)
Explained, by covariate			
STEM	0.002***	-0.015***	0.001***
	(0.000)	(0.001)	(0.000)
Grad $\geq$ 2008	-0.002***	-0.003***	-0.000
	(0.000)	(0.000)	(0.000)
STEM $\times$ Grad $\geq$ 2008	0.006***	0.009***	-0.005***
	(0.000)	(0.001)	(0.000)
Major Generality	0.000*	-0.006***	0.011***
	(0.000)	(0.000)	(0.000)
Years since Degree	0.000***	-0.000**	-0.038***
	(0.000)	(0.000)	(0.001)
Years since Degree^2	-0.000***	-0.000	0.028***
	(0.000)	(0.000)	(0.001)
Non-US Labor Mkt. Exp.	-0.034***	-0.011***	0.009***
	(0.000)	(0.001)	(0.002)
Non-US Labor Mkt. Exp.^{2}	0.016***	0.005***	-0.003*
	(0.000)	(0.001)	(0.001)
Same Employer (Transfer)	0.002***	-0.000*	-0.000
	(0.000)	(0.000)	(0.000)
Unexplained, by covariate			
STEM	-0.001	0.007	0.027***
	(0.001)	(0.007)	(0.006)
Grad $\geq$ 2008	-0.003	0.016*	0.088***
	(0.002)	(0.007)	(0.014)
STEM $\times$ Grad $\geq$ 2008	0.012***	-0.014*	-0.020***
	(0.001)	(0.006)	(0.006)
Major Generality	0.023***	-0.067***	-0.021
	(0.003)	(0.011)	(0.023)
Years since Degree	0.048***	0.129***	0.063*
	(0.003)	(0.008)	(0.030)
Years since Degree^2	-0.022***	-0.043***	-0.031
	(0.002)	(0.005)	(0.023)
Non-US Labor Mkt. Exp.	-0.024***	-0.104***	-0.113***
	(0.002)	(0.007)	(0.026)
Non-US Labor Mkt. Exp.^{2}	0.009***	0.046***	0.071**
	(0.002)	(0.005)	(0.022)
Same Employer (Transfer)	0.000***	0.001***	0.000
	(0.000)	(0.000)	(0.000)
N	3,752,029	409,924	495,517

Notes: Standard errors in parentheses. All columns use only the first job-year observation of a worker after their terminal degree. First jobs in the United States, 2003–2021 graduation cohorts. The outcome is the log OEWS median wage of the occupation, net of arrival-cohort and origin-region effects estimated on the full panel, as in Figures 5 and 6 and the ten-year decompositions. Covariates are STEM, graduation in 2008 or later and its interaction with STEM, major generality, years since degree and its square, non-US experience and its square, a same-employer transfer indicator, school tier, and major field; school-tier and major-field contributions are included in the totals but not shown. Doctorate is all doctoral degrees excluding professional degrees; Professional denotes licensure doctorates (MD, JD, DDS, PharmD, and similar). USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Table 17: Decomposing CtS Master's Degree Holder SOC Wage Distribution, First U.S. Job

	(1)	(2)	(3)	(4)	(5)
OEWS Quantile	10	25	50	75	90
CtS	0.081***	0.093***	0.099***	0.099***	0.100***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
FUS	-0.012***	-0.013***	-0.014***	-0.014***	-0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Difference	0.092***	0.106***	0.113***	0.113***	0.113***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Explained	0.069***	0.079***	0.083***	0.081***	0.078***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained	0.023***	0.027***	0.030***	0.032***	0.035***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Explained, by covariate					
STEM	0.012***	0.010***	0.007***	0.003***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Grad $\geq$ 2008	-0.000***	-0.000***	-0.000***	-0.000***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
STEM $\times$ Grad $\geq$ 2008	0.008***	0.012***	0.015***	0.016***	0.016***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Major Generality	0.003***	0.001***	-0.000***	-0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years since Degree	-0.003***	-0.004***	-0.005***	-0.006***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years since Degree^2	-0.000***	-0.000	0.000***	0.001***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained, by covariate					
STEM	-0.008***	-0.009***	-0.010***	-0.009***	-0.010***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Grad $\geq$ 2008	-0.006**	-0.001	0.002	0.005*	0.007**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
STEM $\times$ Grad $\geq$ 2008	0.025***	0.025***	0.025***	0.024***	0.028***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Major Generality	0.041***	0.041***	0.034***	0.019***	0.015***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Years since Degree	-0.004**	-0.006***	-0.009***	-0.012***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years since Degree^2	-0.000	0.000	0.002*	0.004***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	3,590,577	3,590,577	3,590,577	3,558,157	3,125,708

Notes: Standard errors in parentheses. All columns use only the first job-year observation of a worker after their terminal degree. First jobs in the United States, 2003–2021 graduation cohorts; each OEWS percentile is net of graduation-cohort effects. School-tier and major-field contributions are included in the explained and unexplained totals but not shown. USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Mediators included are education school tiers, major generality, STEM major interaction with post-2008, and gap year(s). Outcome is net of terminal degree graduation cohort fixed-effects.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 18: STEM-OPT and Cap-Exempt First Employers

Panel A. Type of first US employer

	Master's		Doctorate	
	Pre-period share	DD	Pre-period share	DD
Cap-exempt	0.036	−0.002	0.345	0.006
		(0.004)		(0.018)
Cap-subject	0.849	0.011	0.569	−0.000
		(0.010)		(0.017)
Unidentified	0.115	−0.009	0.087	−0.006
		(0.007)		(0.007)

Panel B. Stayer outcomes at $t = 5$, by type of first employer

	Master's		Doctorate	
	Cap-subject	Cap-exempt	Cap-subject	Cap-exempt
Log occupational median wage	0.114	0.069	0.031	0.035
	(0.025)	(0.029)	(0.019)	(0.014)
Over-education	0.024	−0.034	−0.021	−0.046
	(0.013)	(0.028)	(0.017)	(0.017)
Horizontal mismatch	−0.132	−0.069	−0.087	−0.017
	(0.026)	(0.048)	(0.030)	(0.019)
Log occupational PageRank	0.409	0.075	0.224	0.169
	(0.084)	(0.114)	(0.062)	(0.047)
Log firm PageRank	0.633	0.108	0.473	−0.026
	(0.136)	(0.184)	(0.157)	(0.081)
Smallest STEM × Post group	9,858	410	1,181	1,173

Notes: USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. CtS graduates whose first job is in the United States, 2003–2006 and 2008–2021 graduation cohorts. Doctorate is all doctoral degrees excluding professional degrees. Panel A: share of US first jobs at each type of employer among pre-2008 STEM graduates, and the STEM × Post difference-in-differences in that share. The three difference-in-differences sum to zero. Panel B: Equation (10) at $t = 5$, estimated separately by the type of first employer; graduates whose first employer is unidentified are excluded. Standard errors, clustered by major, in parentheses. The smallest group is the smallest STEM × Post group across the five outcomes.

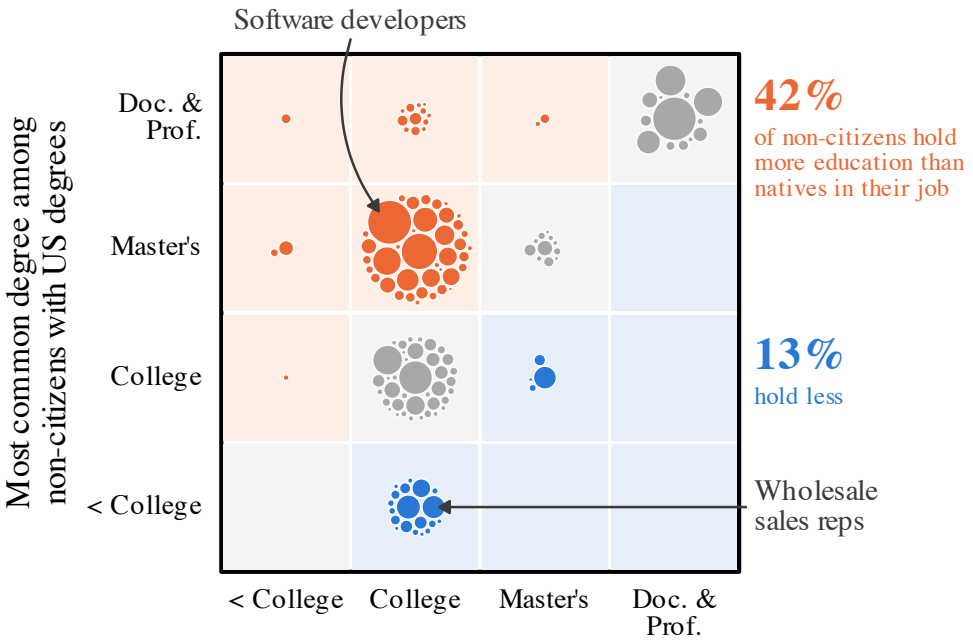
Table 19: STEM-OPT and the Jobs of Master's Stayers Five Years after Graduation

	Graduates of 2008–2011	Graduates of 2008–2021
Panel A. Stayer outcomes at $t = 5$		
Over-educated (pp)	0.5	1.9
	[−0.6, 1.6]	[−0.4, 4.2]
Horizontally mismatched (pp)	−7.6	−13.1
	[−10.6, −4.7]	[−18.2, −7.9]
Log occupational PageRank	0.19	0.39
	[0.10, 0.28]	[0.23, 0.56]
Log firm PageRank	0.21	0.61
	[0.10, 0.33]	[0.34, 0.88]
Log occupational median wage	0.08	0.11
	[0.05, 0.11]	[0.06, 0.16]
Panel B. Entry and retention		
US placement at graduation (pp)	3.7	5.6
	[1.9, 5.4]	[3.4, 7.9]
Still in the US at $t = 5$ (pp)	−0.1	1.0
	[−0.7, 0.5]	[0.5, 1.5]

Notes: USO = Only US post-secondary Education; CtS = foreign post-secondary education, then US terminal degree; CtW = only non-US post-secondary education. Difference-in-differences, STEM versus non-STEM CtS master's graduates, post- versus pre-2008 cohorts (2003–2006; 2007 excluded). Panel A: Equation (10) at $t = 5$ among graduates working in the United States. US placement: Equation (9). Still in the US: Kaplan–Meier retention among US starters (Table F2). 95 percent confidence intervals in brackets, clustered by terminal-degree major except for retention, which uses Greenwood standard errors.

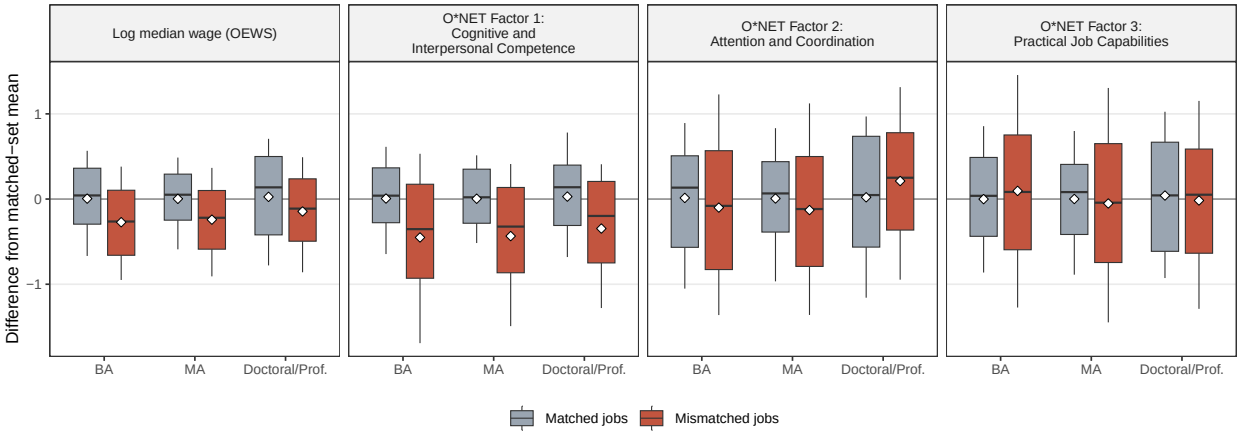
10 Figures

Figure 16: Occupational Mismatch of Non-Citizens with US Degrees



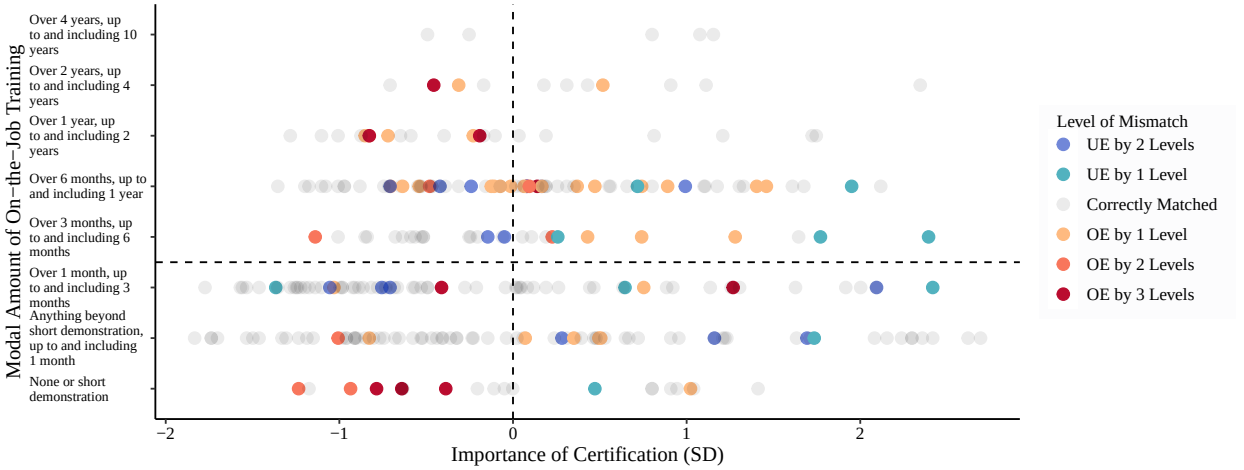
Notes: ACS 2010–2019. Wage and salary workers at work 50–52 weeks, ages 18–65, not enrolled in school. Natives are citizens by birth; non-citizens with US degrees are non-citizens who arrived before the typical completion age of their highest degree. Each circle is one of the 142 occupations with at least 30 respondents in each group and a most common degree of college or above in either group, sized by non-citizen employment. Orange circles: the non-citizens' most common degree is higher; blue: lower; gray: equal. Shares are weighted by non-citizen employment.

Figure 17: What Horizontally Mismatched Jobs Look Like



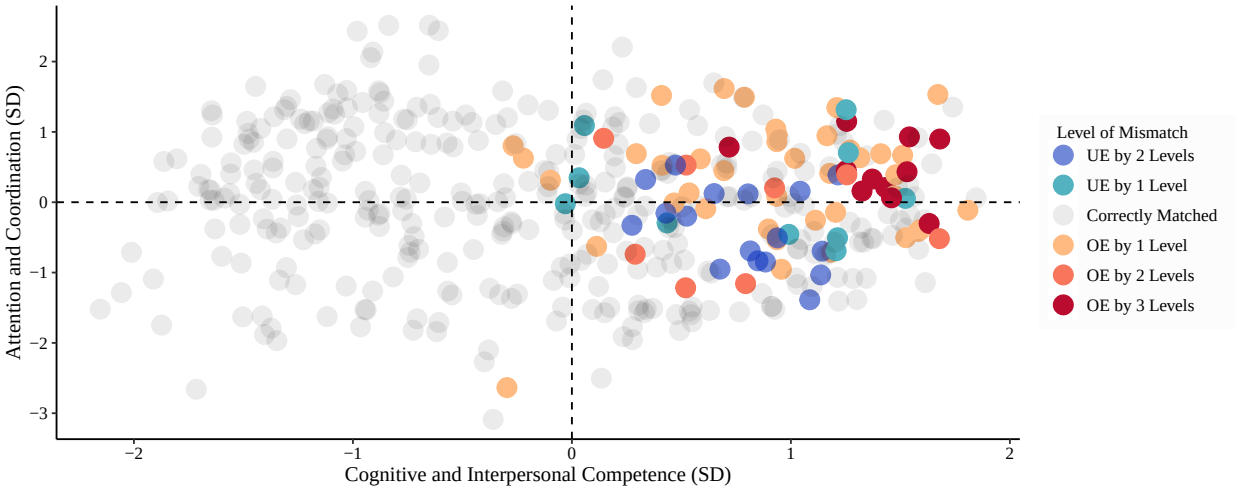
Notes: Occupations held by FUS workers, by degree-by-major group. Matched jobs are the occupations in the smallest set that holds half of the group's FUS jobs (Section 3.2); mismatched jobs are all others. Each outcome is the occupation's value minus the mean of the group's matched set, weighted by FUS holdings. Log median wage is the national OEWS median; the three O*NET factors (Section 3.3) and SOC generality are in standard-deviation units. Doctoral/Prof. pools doctorate and professional degrees. Boxes span the 25th to 75th percentiles, lines mark medians, whiskers the 10th and 90th percentiles, and diamonds means.

Figure 18: Educational Mismatch of Non-Citizens, Highest Degree Obtained in US



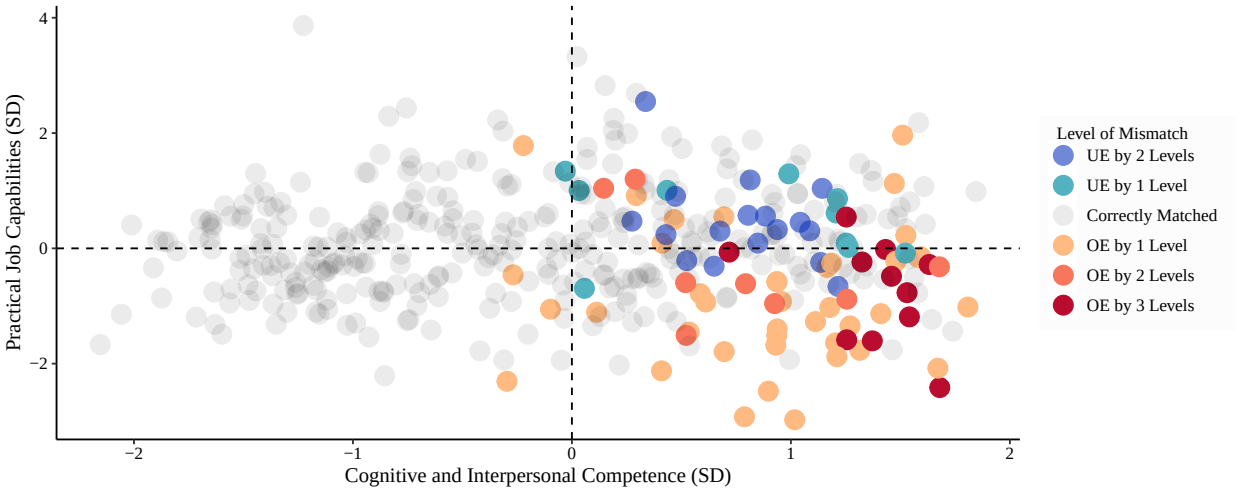
Notes: Level of mismatch is the difference between the modal degree of a non−citizen worker, whose highest degree is obtained inside the United States, and the modal degree of a citizen worker. The degree levels, from highest to lowest, are: Doctorate, Professional, Master's, Bachelor's, Associates, High School, and less than High School. Job characteristics are from the O*NET dataset. Mismatch realizations are from the American Community Survey 2010−2019.

Figure 19: Immigrant Workers Experience More Educational Mismatch at Jobs Requiring Above-Average Cognitive and Interpersonal Competence



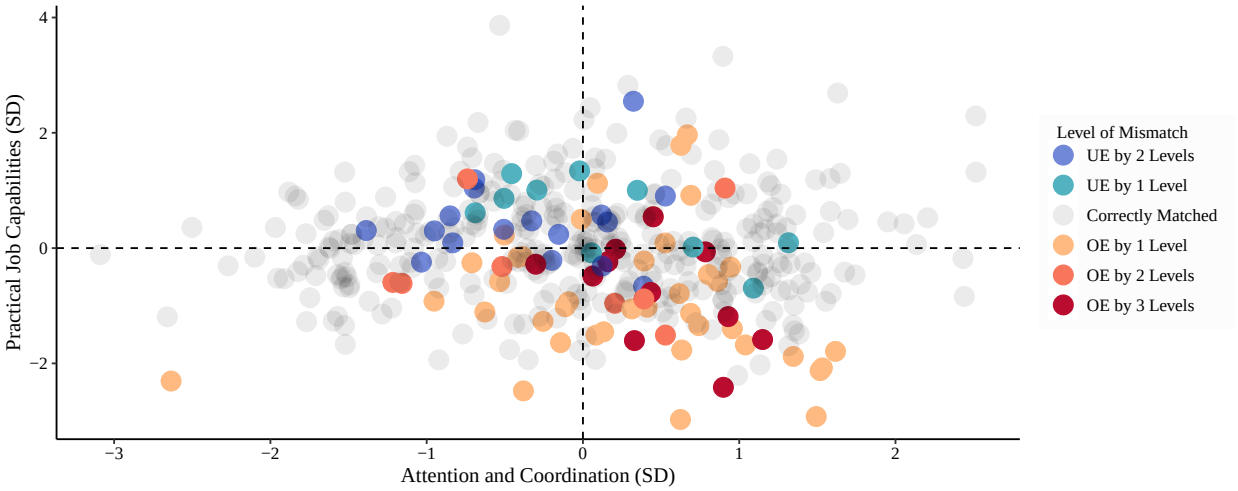
Notes: Level of mismatch is the difference between the modal degree of a non−citizen worker, whose highest degree is obtained inside the United States, and the modal degree of a citizen worker. The degree levels, from highest to lowest, are: Doctorate, Professional, Master's, Bachelor's, Associates, High School, and less than High School.

Figure 20: Negative Relationship between Over-Education and Job Requirements of Practical Capabilities



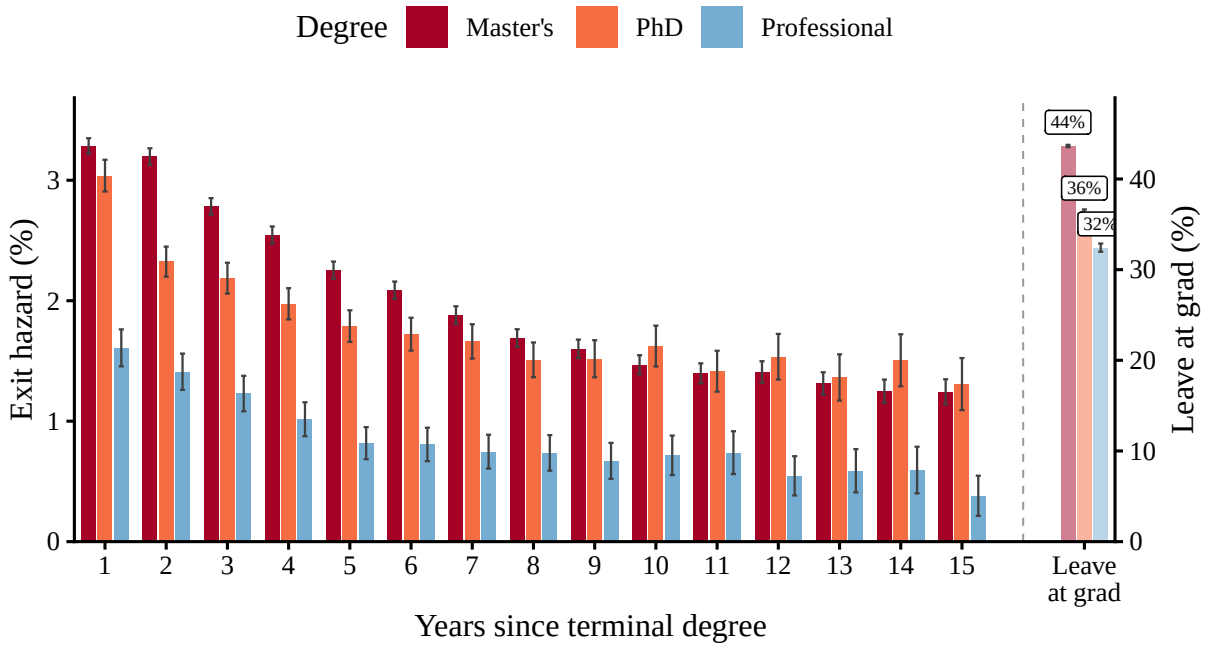
Notes: Level of mismatch is the difference between the modal degree of a non−citizen worker, whose highest degree is obtained inside the United States, and the modal degree of a citizen worker. The degree levels, from highest to lowest, are: Doctorate, Professional, Master's, Bachelor's, Associates, High School, and less than High School.

Figure 21: No Clear Relationship between Educational Mismatch and Job Requirements of Attention and Coordination



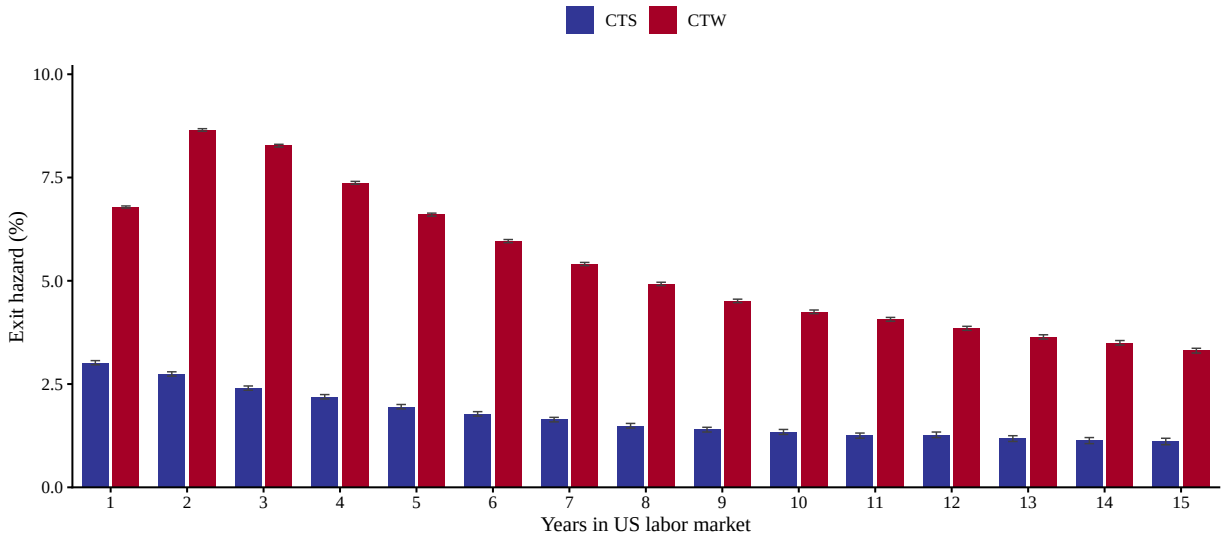
Notes: Level of mismatch is the difference between the modal degree of a non−citizen worker, whose highest degree is obtained inside the United States, and the modal degree of a citizen worker. The degree levels, from highest to lowest, are: Doctorate, Professional, Master's, Bachelor's, Associates, High School, and less than High School.

Figure 22: Low Exits After Initial US Placement, Came-to-Study, 1970-2025



Notes: Came to Study. Bars grouped by degree within each year. Years 1+: discrete−time exit hazard h(ce) = leavers / at−risk; at−risk conditions on being in US and observed (right−censored); left axis. Rightmost bars: leave−at−graduation, share not in US at graduation; right axis. Error bars 95% CI (binomial SE).

Figure 23: Exit Hazard Rates by Years in US Labor Market, 1970-2025



Notes: CTS (blue, grad deg 3/4/5) vs CTW (red, full BA+ deg 2/3/4/5). Discrete−time exit hazard h(t) = leavers / at−risk. Error bars 95% CI (binomial SE).

A Appendix Tables

Table A1: Occupational Quantile Example

Occupation	Doc./Pro.	MA	BA	Assoc.	HS	<HS	Occ. Quantile
Surgeon	99.5	0.5	0	0	0	0	100
Post-Secondary Teacher	37.8	33.2	19.4	2.4	6.8	0.4	97.1
Sales Managers	1.0	10.8	50.2	7.5	29.5	1.0	61.0
Respiratory Therapists	0.9	3.1	25.0	56.6	14.1	0.3	59.5
Fence Erectors	0	0.7	3.7	4.4	69.4	21.8	3.0
Earth Drillers, not Oil	0	0.2	4.0	6.0	76.3	13.5	2.4

Table A2: Age-Education Schedule

Highest Degree	Age at Completion
Kindergarten or below	6
First Grade	7
Second Grade	8
Third Grade	9
Fourth Grade	10
Fifth Grade	11
Sixth Grade	12
Seventh Grade	13
Eighth Grade	14
Ninth Grade	15
Tenth Grade	16
Eleventh Grade	17
Twelfth Grade or High School Diploma	18
GED or Some College	19
Associate's	20
Bachelor's Degree	22
Master's Degree	24
Professional Degree Beyond Bachelor's	28
Doctoral Degree	30

Table A3: List of Skills and Abilities used to Label Each Principal Component

Cognitive and Interpersonal Competence	**Attention and Coordination**	**Practical Job Capabilities**
Written Expression	*Visualization*	*Programming*
Speaking	*Operations Monitoring*	*Service Orientation*
Active Listening	*Quality Control Analysis*	*Social Perceptiveness*
Reading Comprehension	*Perceptual Speed*	*Time Sharing*
Writing	*Troubleshooting*	*Explosive Strength*
Active Learning	*Visual Color Discrimination*	*Stamina*
Written Comprehension	*Operation and Control*	*Coordination*
Oral Expression	*Equipment Selection*	*Gross Body Coordination*
Critical Thinking	*Depth Perception*	*Equipment Selection*
Judgment and Decision Making	*Hearing Sensitivity*	*Gross Body Equilibrium*
Oral Comprehension	*Auditory Attention*	*Negotiation*
Deductive Reasoning	*Far Vision*	*Repairing*
Speech Clarity	*Repairing*	*Static Strength*
Systems Evaluation	*Equipment Maintenance*	*Near Vision*
Complex Problem Solving	*Reaction Time*	*Technology Design*
Systems Analysis	*Selective Attention*	*Equipment Maintenance*
Fluency of Ideas	*Response Orientation*	*Installation*
Speech Recognition	*Flexibility of Closure*	*Trunk Strength*
Monitoring	*Technology Design*	*Persuasion*
Inductive Reasoning	*Rate Control*	*Speech Recognition*

Table A4: Example of Predicted SOC Codes using Preliminarily Fine-Tuned Llama-3.1

LinkedIn Title	**O*NET SOC**	**Corresponding O*NET Title**
director-field support	27-2012.05	Media Technical Directors/Managers
senior page designer and copy editor	27-3041.00	Editors
gfliwn askdm	00-0000.00	NA
taylor mattis fellow	51-9195.05	Potters, Manufacturing
solar analyst intern	17-2199.11	Solar Energy Systems Engineers
international visiting attorney	23-1011.00	Lawyers
director of lower school admission	11-9032.00	Education Administrators, Kindergarten through Secondary
asset & reliability engineer	17-2199.00	NA
sales / acct manager	41-1011.00	First-Line Supervisors of Retail Sales Workers
attest intern	13-2099.00	NA
immersion associate	25-1194.00	Career/Technical Education Teachers, Postsecondary
medicaid waiver case manager	11-9151.00	Social and Community Service Managers
網頁工程師	15-1299.08	Computer Systems Engineers/Architects
博士生	25-9044.00	Teaching Assistants, Postsecondary

Table A5: Gelbach Decomposition of the First-Job Occupational-Wage Gap, CtS vs. FUS

	Master's	Doctorate	Professional
Raw gap (no controls)	0.1188*** (0.0007)	-0.0008 (0.0015)	-0.0304*** (0.0034)
Major Generality	-0.0028*** (0.0000)	-0.0017*** (0.0001)	0.0037*** (0.0002)
STEM	-0.0044*** (0.0003)	-0.0162*** (0.0005)	-0.0003*** (0.0001)
Grad $\geq$ 2008	0.0024*** (0.0001)	0.0007*** (0.0001)	0.0004*** (0.0001)
STEM $\times$ Grad $\geq$ 2008	0.0082*** (0.0003)	0.0135*** (0.0005)	-0.0005*** (0.0001)
English-Speaking Origin	-0.0003*** (0.0001)	-0.0002 (0.0001)	-0.0002 (0.0004)
Partial English-Speaking Origin	-0.0000 (0.0000)	0.0000 (0.0000)	-0.0000 (0.0000)
Years since Degree	-0.0034*** (0.0001)	-0.0012*** (0.0001)	-0.0021*** (0.0002)
Years since Degree^2	0.0011*** (0.0001)	0.0004*** (0.0001)	0.0016*** (0.0002)
Skill Factor 1	0.0079*** (0.0003)	-0.0129*** (0.0005)	-0.0062*** (0.0012)
Skill Factor 2	0.0232*** (0.0002)	0.0228*** (0.0006)	-0.0104*** (0.0013)
Skill Factor 3	0.0319*** (0.0002)	0.0089*** (0.0004)	0.0046*** (0.0004)
Cert. Importance	-0.0067*** (0.0001)	-0.0155*** (0.0002)	-0.0092*** (0.0009)
SOC Generality	-0.0007*** (0.0000)	-0.0177*** (0.0003)	-0.0089*** (0.0004)
H-1B Demand Decile 1 (of 10)	0.0004*** (0.0000)	-0.0004*** (0.0000)	-0.0001 (0.0001)
H-1B Demand Decile 2 (of 10)	0.0001*** (0.0000)	-0.0005*** (0.0001)	-0.0001 (0.0001)
H-1B Demand Decile 3 (of 10)	0.0006*** (0.0000)	-0.0025*** (0.0002)	-0.0004*** (0.0001)
H-1B Demand Decile 4 (of 10)	0.0019*** (0.0000)	-0.0016*** (0.0002)	-0.0006*** (0.0001)
H-1B Demand Decile 5 (of 10)	0.0037*** (0.0000)	0.0011*** (0.0001)	-0.0001 (0.0001)
H-1B Demand Decile 6 (of 10)	0.0006*** (0.0000)	-0.0030*** (0.0001)	0.0004*** (0.0001)
H-1B Demand Decile 7 (of 10)	0.0011*** (0.0000)	-0.0001 (0.0001)	0.0004*** (0.0001)
H-1B Demand Decile 8 (of 10)	0.0015*** (0.0000)	-0.0025*** (0.0001)	-0.0015*** (0.0004)
H-1B Demand Decile 9 (of 10)	-0.0018*** (0.0000)	-0.0026*** (0.0001)	-0.0000 (0.0003)
H-1B Demand Decile 10 (of 10)	0.0208*** (0.0001)	0.0120*** (0.0002)	-0.0004* (0.0002)
Firm Util.	0.0033*** (0.0000)	-0.0011*** (0.0001)	0.0002 (0.0001)
Churn 3mo	0.0000 (0.0000)	0.0000** (0.0000)	0.0000 (0.0000)
Churn 6mo	0.0004*** (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
Churn 12mo	-0.0001*** (0.0000)	0.0003*** (0.0000)	0.0000 (0.0000)
Churn 24mo	0.0003*** (0.0000)	-0.0002*** (0.0000)	-0.0000 (0.0000)
Employer Ever Sponsors H-1B	0.0001 (0.0000)	0.0010*** (0.0001)	-0.0001* (0.0000)
SOC Util.	0.0242*** (0.0002)	0.0147*** (0.0004)	-0.0025*** (0.0005)
OE$\times$ ln(Edu. Signal)	0.0000 (0.0000)	0.0043*** (0.0002)	-0.0001 (0.0002)
HM$\times$ ln(Edu. Signal)	-0.0001*** (0.0000)	0.0001 (0.0001)	0.0017*** (0.0004)
Over Edu.	-0.0091*** (0.0001)	0.0007*** (0.0001)	0.0006* (0.0003)
Hor. Mismatch	-0.0022*** (0.0001)	-0.0051*** (0.0002)	-0.0014*** (0.0004)
Cap-Exempt Employer (PDL Educational)	0.0018*** (0.0000)	0.0002*** (0.0000)	0.0009*** (0.0001)
Total Explained ($b_0 - b_{\text{full}}$)	0.1154*** (0.0006)	0.0067*** (0.0013)	-0.0291*** (0.0031)
Final gap (full controls)	0.0034*** (0.0004)	-0.0075*** (0.0008)	-0.0013 (0.0014)
N (common sample)	3,115,337	397,523	412,613

Notes: Gelbach (2016) decomposition of the gap in the log OEWS median wage of the occupation held at the first US job between CtS and FUS workers with the same terminal degree. Rows are the contributions of each covariate, in log points, with standard errors in parentheses; the raw gap minus the total explained equals the final gap. School-tier and major-field contributions are included in the total but not shown. Figure 5 groups the contributions into blocks. Doctorate pools the PhD and non-licensure doctorates; Professional denotes licensure doctorates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A6: Gelbach Decomposition of the First-Job Occupational-Wage Gap, CtW vs. FUS

	Master's	Doctorate	Professional
Raw gap (no controls)	-0.0783*** (0.0006)	-0.2322*** (0.0018)	-0.1563*** (0.0053)
Major Generality	-0.0002*** (0.0000)	-0.0009*** (0.0001)	0.0119*** (0.0003)
STEM	-0.0020*** (0.0001)	-0.0116*** (0.0004)	-0.0010*** (0.0001)
Grad $\geq$ 2008	-0.0029*** (0.0001)	-0.0016*** (0.0001)	0.0011*** (0.0002)
STEM $\times$ Grad $\geq$ 2008	0.0029*** (0.0001)	0.0077*** (0.0003)	-0.0016*** (0.0002)
Same Employer (Transfer)	0.0008*** (0.0000)	0.0001*** (0.0000)	0.0000 (0.0000)
English-Speaking Origin	-0.0076*** (0.0003)	0.0091*** (0.0006)	0.0172*** (0.0018)
Partial English-Speaking Origin	0.0002*** (0.0000)	0.0000 (0.0000)	0.0001 (0.0001)
Non-US Labor Mkt. Exp.	-0.0102*** (0.0003)	-0.0047*** (0.0006)	-0.0061*** (0.0006)
Non-US Labor Mkt. Exp.^{2}	0.0074*** (0.0002)	0.0047*** (0.0004)	0.0067*** (0.0005)
Skill Factor 1	-0.0145*** (0.0003)	-0.0349*** (0.0006)	-0.0599*** (0.0016)
Skill Factor 2	-0.0059*** (0.0002)	-0.0103*** (0.0006)	0.0302*** (0.0016)
Skill Factor 3	0.0171*** (0.0001)	0.0059*** (0.0003)	0.0254*** (0.0006)
Cert. Importance	-0.0072*** (0.0001)	-0.0227*** (0.0003)	-0.0397*** (0.0014)
SOC Generality	0.0040*** (0.0000)	-0.0256*** (0.0004)	-0.0510*** (0.0010)
H-1B Demand Decile 1 (of 10)	0.0001*** (0.0000)	-0.0002*** (0.0000)	-0.0010*** (0.0001)
H-1B Demand Decile 2 (of 10)	-0.0001*** (0.0000)	-0.0012*** (0.0001)	-0.0014*** (0.0002)
H-1B Demand Decile 3 (of 10)	0.0002*** (0.0000)	-0.0045*** (0.0002)	-0.0026*** (0.0002)
H-1B Demand Decile 4 (of 10)	0.0009*** (0.0000)	-0.0060*** (0.0002)	-0.0036*** (0.0003)
H-1B Demand Decile 5 (of 10)	0.0017*** (0.0000)	0.0009*** (0.0000)	-0.0014*** (0.0002)
H-1B Demand Decile 6 (of 10)	0.0001*** (0.0000)	-0.0046*** (0.0001)	0.0003 (0.0001)
H-1B Demand Decile 7 (of 10)	0.0006*** (0.0000)	-0.0022*** (0.0002)	0.0012*** (0.0001)
H-1B Demand Decile 8 (of 10)	0.0008*** (0.0000)	-0.0022*** (0.0001)	0.0043*** (0.0004)
H-1B Demand Decile 9 (of 10)	-0.0021*** (0.0000)	-0.0049*** (0.0002)	0.0095*** (0.0005)
H-1B Demand Decile 10 (of 10)	0.0110*** (0.0001)	0.0030*** (0.0002)	-0.0072*** (0.0003)
Firm Util.	0.0040*** (0.0001)	-0.0019*** (0.0002)	0.0029*** (0.0001)
Churn 3mo	-0.0000 (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)
Churn 6mo	-0.0000 (0.0000)	-0.0001*** (0.0000)	0.0002*** (0.0000)
Churn 12mo	0.0002*** (0.0000)	0.0007*** (0.0001)	-0.0001** (0.0000)
Churn 24mo	0.0001*** (0.0000)	-0.0009*** (0.0001)	0.0000 (0.0000)
Employer Ever Sponsors H-1B	0.0002*** (0.0000)	0.0001*** (0.0000)	-0.0001 (0.0001)
SOC Util.	0.0049*** (0.0001)	0.0061*** (0.0004)	-0.0302*** (0.0007)
OE$\times$ ln(Edu. Signal)	-0.0007*** (0.0001)	0.0037*** (0.0003)	-0.0005*** (0.0001)
HM$\times$ ln(Edu. Signal)	0.0013*** (0.0001)	0.0004* (0.0002)	-0.0007*** (0.0002)
Over Edu.	-0.0109*** (0.0001)	-0.0009*** (0.0001)	0.0105*** (0.0005)
Hor. Mismatch	0.0070*** (0.0001)	-0.0022*** (0.0003)	0.0065*** (0.0004)
Cap-Exempt Employer (PDL Educational)	0.0035*** (0.0000)	0.0001*** (0.0000)	0.0021*** (0.0001)
Total Explained ($b_0 - b_{\text{full}}$)	0.0226*** (0.0007)	-0.0957*** (0.0017)	-0.0570*** (0.0049)
Final gap (full controls)	-0.1009*** (0.0005)	-0.1365*** (0.0012)	-0.0993*** (0.0031)
N (common sample)	3,269,102	369,951	404,794

Notes: Gelbach (2016) decomposition of the gap in the log OEWS median wage of the occupation held at the first US job between CtW and FUS workers with the same terminal degree. Rows are the contributions of each covariate, in log points, with standard errors in parentheses; the raw gap minus the total explained equals the final gap. School-tier and major-field contributions are included in the total but not shown. Figure 6 groups the contributions into blocks. Doctorate pools the PhD and non-licensure doctorates; Professional denotes licensure doctorates. Worker characteristics additionally include non-US experience and a same-employer transfer indicator. $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Table A7: Decomposing the Occupational-Wage Gap by Year since Degree, CtS Master's Holders, Ten-Year Stayers

	Years since terminal degree										
	0	1	2	3	4	5	6	7	8	9	10
CtS	-0.144***	-0.001	0.029***	0.038***	0.050***	0.070***	0.081***	0.092***	0.101***	0.107***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
FUS	-0.189***	-0.121***	-0.094***	-0.077***	-0.060***	-0.045***	-0.036***	-0.029***	-0.023***	-0.018***	-0.016***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Difference	0.046***	0.119***	0.122***	0.115***	0.109***	0.115***	0.117***	0.121***	0.124***	0.126***	0.128***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Explained	0.064***	0.086***	0.089***	0.085***	0.082***	0.084***	0.085***	0.087***	0.087***	0.088***	0.087***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained	-0.018***	0.033***	0.034***	0.030***	0.028***	0.031***	0.032***	0.034***	0.036***	0.038***	0.040***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Explained, by covariate											
First Job (this spell)	-0.002***	-0.002***	-0.000***	0.000***	0.000***	0.000**					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
STEM	0.013***	0.009***	0.009***	0.009***	0.008***	0.009***	0.010***	0.011***	0.011***	0.012***	0.011***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Grad $\geq$ 2008	0.001***	0.002***	0.002***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
STEM $\times$ Grad $\geq$ 2008	0.004***	0.011***	0.013***	0.011***	0.011***	0.010***	0.009***	0.007***	0.006***	0.005***	0.004***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Major Generality	-0.002***	-0.001***	-0.001***	-0.001***	-0.001***	-0.000***	-0.000	0.000***	0.000***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained, by covariate											
First Job (this spell)	-0.114***	0.004***	-0.000	-0.000***	-0.000**	-0.000***					
	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
STEM	-0.010**	-0.011***	-0.013***	-0.008***	-0.006*	-0.004	0.000	0.001	0.003	0.004	0.005*
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Grad $\geq$ 2008	-0.001	0.007*	0.004	0.001	0.001	0.001	0.000	-0.001	-0.003	-0.003	-0.002
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
STEM $\times$ Grad $\geq$ 2008	0.020***	0.026***	0.027***	0.021***	0.016***	0.014***	0.012***	0.011***	0.009***	0.009***	0.008***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Major Generality	0.020***	0.029***	0.021***	0.016***	0.018***	0.018***	0.022***	0.027***	0.023***	0.022***	0.021***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
N	1,899,315	2,050,563	2,139,411	2,163,412	2,182,373	2,245,379	2,282,945	2,318,256	2,352,860	2,390,462	2,490,133

Notes: Oaxaca–Blinder decomposition, FUS coefficients as reference, of the log OEWS median wage of the occupation held in each year since the terminal degree, net of graduation-cohort effects, among workers who remain in the United States through year ten; 2003–2021 graduation cohorts. Each column is estimated separately. Year zero is reported for completeness; the text and Figures 10 and 9 start at year one (Section 4.2). Covariates are STEM, graduation in 2008 or later and its interaction with STEM, major generality, an indicator for the first job spell, school tier, and major field; tier and field contributions are included in the totals but not shown. Standard errors in parentheses. Doctorate pools the PhD and non-licensure doctorates; Professional denotes licensure doctorates. $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Table A8: Decomposing the Occupational-Wage Gap by Year since Degree, CtS Doctorate Holders, Ten-Year Stayers

	Years since terminal degree										
	0	1	2	3	4	5	6	7	8	9	10
CtS	-0.390***	-0.169***	-0.100***	-0.055***	-0.015***	0.018***	0.039***	0.057***	0.069***	0.076***	0.078***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
FUS	-0.261***	-0.143***	-0.098***	-0.065***	-0.037***	-0.015***	-0.004***	0.005***	0.009***	0.012***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Difference	-0.129***	-0.026***	-0.001	0.010***	0.022***	0.033***	0.043***	0.052***	0.060***	0.065***	0.067***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Explained	-0.045***	0.013***	0.025***	0.033***	0.038***	0.042***	0.044***	0.046***	0.047***	0.048***	0.049***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unexplained	-0.083***	-0.039***	-0.027***	-0.022***	-0.016***	-0.009***	-0.001	0.007***	0.013***	0.016***	0.019***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Explained, by covariate											
First Job (this spell)	-0.001***	-0.000	-0.000***	-0.000***	-0.000	-0.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
STEM	-0.015***	-0.014***	-0.012***	-0.006***	-0.001	0.004***	0.008***	0.012***	0.014***	0.018***	0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Grad $\geq$ 2008	0.000***	0.001***	0.001***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
STEM $\times$ Grad $\geq$ 2008	-0.009***	0.003*	0.006***	0.006***	0.007***	0.007***	0.007***	0.006***	0.004***	0.002**	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Major Generality	-0.011***	-0.008***	-0.007***	-0.006***	-0.006***	-0.005***	-0.004***	-0.004***	-0.004***	-0.004***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained, by covariate											
First Job (this spell)	-0.119***	0.001	0.000	0.000	0.000	-0.000					
	(0.015)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)					
STEM	-0.001	-0.004	-0.006	-0.013	-0.007	-0.010	-0.011	-0.008	-0.007	-0.008	-0.009
	(0.011)	(0.010)	(0.009)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Grad $\geq$ 2008	-0.035***	-0.001	0.002	-0.004	-0.004	-0.005	-0.008	-0.008	-0.010	-0.013*	-0.015**
	(0.010)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
STEM $\times$ Grad $\geq$ 2008	-0.006	-0.008	-0.002	0.002	-0.002	0.002	0.008	0.006	0.008	0.010*	0.012*
	(0.009)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Major Generality	-0.024	-0.008	0.002	0.012	0.001	0.009	0.009	0.011	0.017	0.017	0.016
	(0.013)	(0.012)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)
N	234,328	248,052	257,085	260,850	263,284	269,288	273,186	277,223	281,464	285,515	297,095

Notes: Oaxaca–Blinder decomposition, FUS coefficients as reference, of the log OEWS median wage of the occupation held in each year since the terminal degree, net of graduation-cohort effects, among workers who remain in the United States through year ten; 2003–2021 graduation cohorts. Each column is estimated separately. Year zero is reported for completeness; the text and Figures 10 and 9 start at year one (Section 4.2). Covariates are STEM, graduation in 2008 or later and its interaction with STEM, major generality, an indicator for the first job spell, school tier, and major field; tier and field contributions are included in the totals but not shown. Standard errors in parentheses. Doctorate pools the PhD and non-licensure doctorates; Professional denotes licensure doctorates. $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Table A9: Decomposing the Occupational-Wage Gap by Year since Degree, CtS Professional Degree Holders, Ten-Year Stayers

	Years since terminal degree										
	0	1	2	3	4	5	6	7	8	9	10
CtS	0.138***	0.222***	0.243***	0.252***	0.254***	0.245***	0.246***	0.248***	0.248***	0.248***	0.246***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
FUS	0.141***	0.236***	0.260***	0.267***	0.270***	0.266***	0.265***	0.261***	0.261***	0.261***	0.261***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Difference	-0.003	-0.014**	-0.017***	-0.015***	-0.016***	-0.021***	-0.019***	-0.013***	-0.013***	-0.013***	-0.015***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Explained	0.002	0.004**	0.007***	0.009***	0.009***	0.008***	0.008***	0.008***	0.007***	0.007***	0.007***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unexplained	-0.005	-0.018***	-0.024***	-0.024***	-0.025***	-0.029***	-0.026***	-0.021***	-0.020***	-0.019***	-0.022***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Explained, by covariate											
First Job (this spell)	0.001	-0.000**	-0.000***	0.000***	0.000	-0.000					
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
STEM	0.000*	0.000	0.000	-0.000	-0.000**	-0.000***	-0.000***	-0.000***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Grad $\geq$ 2008	-0.001	-0.001***	-0.001***	-0.000***	-0.000	0.000**	0.000**	0.000***	0.000***	0.000***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
STEM $\times$ Grad $\geq$ 2008	-0.001	-0.000	-0.000	-0.000	-0.001**	-0.000**	-0.000**	-0.000*	-0.000***	-0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Major Generality	0.003***	0.002***	0.003***	0.003***	0.003***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained, by covariate											
First Job (this spell)	-0.082***	-0.000	-0.002*	-0.001	-0.001	-0.001					
	(0.017)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)					
STEM	-0.001	-0.001	0.004	0.006	0.005	0.003	0.000	-0.001	-0.002	-0.001	0.000
	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Grad $\geq$ 2008	0.026*	0.022*	0.013	0.022**	0.019*	0.013	0.011	0.013*	0.016*	0.014*	0.014*
	(0.012)	(0.010)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
STEM $\times$ Grad $\geq$ 2008	-0.010	-0.007	-0.006	-0.005	-0.003	0.000	0.002	0.004	0.004	0.003	0.002
	(0.006)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Major Generality	-0.039	-0.033	-0.022	-0.037	-0.017	-0.031	-0.013	0.009	-0.008	-0.015	-0.022
	(0.025)	(0.022)	(0.021)	(0.020)	(0.020)	(0.019)	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)
N	234,526	273,801	294,486	307,050	314,564	321,862	327,617	332,837	337,497	342,626	355,495

Notes: Oaxaca–Blinder decomposition, FUS coefficients as reference, of the log OEWS median wage of the occupation held in each year since the terminal degree, net of graduation-cohort effects, among workers who remain in the United States through year ten; 2003–2021 graduation cohorts. Each column is estimated separately. Year zero is reported for completeness; the text and Figures 10 and 9 start at year one (Section 4.2). Covariates are STEM, graduation in 2008 or later and its interaction with STEM, major generality, an indicator for the first job spell, school tier, and major field; tier and field contributions are included in the totals but not shown. Standard errors in parentheses. Doctorate pools the PhD and non-licensure doctorates; Professional denotes licensure doctorates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A10: Decomposing the Occupational-Wage Gap by Year since Degree, CtW Master's Holders, Ten-Year Stayers

	Years since first US labor-market year										
	0	1	2	3	4	5	6	7	8	9	10
CtW	-0.065***	-0.036***	-0.020***	-0.008***	0.004***	0.021***	0.035***	0.050***	0.064***	0.076***	0.085***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
FUS	-0.074***	-0.006***	0.021***	0.037***	0.054***	0.069***	0.077***	0.085***	0.091***	0.095***	0.097***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Difference	0.009***	-0.030***	-0.041***	-0.046***	-0.050***	-0.048***	-0.042***	-0.035***	-0.027***	-0.020***	-0.013***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Explained	0.068***	0.065***	0.060***	0.055***	0.051***	0.051***	0.053***	0.054***	0.055***	0.054***	0.051***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unexplained	-0.060***	-0.095***	-0.101***	-0.101***	-0.101***	-0.100***	-0.095***	-0.089***	-0.082***	-0.074***	-0.064***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Explained, by covariate											
First Job (this spell)	0.002***	-0.001***	0.000***	0.000***	0.001***	0.000***					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
STEM	0.006***	0.005***	0.004***	0.005***	0.004***	0.005***	0.006***	0.006***	0.007***	0.007***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Grad $\geq$ 2008	-0.001***	-0.001***	-0.002***	-0.002***	-0.003***	-0.004***	-0.004***	-0.004***	-0.004***	-0.003***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
STEM $\times$ Grad $\geq$ 2008	0.002***	0.005***	0.005***	0.005***	0.005***	0.004***	0.003***	0.003***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Major Generality	-0.001***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Non-US Labor Mkt. Exp.	-0.000***	-0.003***	-0.009***	-0.014***	-0.017***	-0.018***	-0.018***	-0.018***	-0.017***	-0.019***	-0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Non-US Labor Mkt. Exp.^{2}	-0.000	0.000***	0.001***	0.002***	0.003***	0.003***	0.003***	0.002***	-0.000	-0.001	-0.010***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Same Employer (Transfer)	-0.000	0.001***	0.001***	0.002***	0.002***	0.002***	0.003***	0.004***	0.005***	0.005***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained, by covariate											
First Job (this spell)	0.044**	-0.005***	-0.003***	-0.002***	-0.003***	-0.002***					
	(0.015)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)					
STEM	-0.008**	-0.006**	-0.008***	-0.005**	-0.005**	-0.007***	-0.007***	-0.008***	-0.007***	-0.006***	-0.005***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Grad $\geq$ 2008	0.022***	0.010***	0.009***	0.006**	0.005**	0.002	-0.002	-0.006***	-0.008***	-0.007***	-0.004**
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
STEM $\times$ Grad $\geq$ 2008	0.022***	0.018***	0.018***	0.014***	0.012***	0.013***	0.013***	0.013***	0.013***	0.012***	0.011***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Major Generality	0.009	0.012*	0.007	0.001	-0.000	0.006	0.015***	0.018***	0.020***	0.022***	0.024***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Non-US Labor Mkt. Exp.	0.001*	0.000	0.002	0.005***	0.008***	0.008***	0.008***	0.012***	0.013***	0.020***	0.026***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Non-US Labor Mkt. Exp.^{2}	-0.001	0.000	0.001*	0.001**	0.003***	0.006***	0.008***	0.011***	0.014***	0.014***	0.015***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Same Employer (Transfer)	-0.000	0.000	0.000*	0.000	0.000	0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	1,847,886	2,017,467	2,127,672	2,180,375	2,226,860	2,313,211	2,375,420	2,441,256	2,511,625	2,590,801	2,740,678

Notes: Oaxaca–Blinder decomposition, FUS coefficients as reference, of the log OEWS median wage of the occupation held in each year since the worker's first year in the US labor market (for FUS workers, since the terminal degree), net of graduation-cohort effects, among workers who remain in the United States through year ten; 2003–2021 graduation cohorts. Each column is estimated separately. Year zero is reported for completeness; the text and Figures 10 and 9 start at year one (Section 4.2). Covariates are STEM, graduation in 2008 or later and its interaction with STEM, major generality, an indicator for the first job spell, school tier, and major field, and for CtW workers non-US labor-market experience, its square, and a same-employer transfer indicator; tier and field contributions are included in the totals but not shown. Standard errors in parentheses. Doctorate pools the PhD and non-licensure doctorates; Professional denotes licensure doctorates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A11: Decomposing the Occupational-Wage Gap by Year since Degree, CtW Doctorate Holders, Ten-Year Stayers

	Years since first US labor-market year										
	0	1	2	3	4	5	6	7	8	9	10
CtW	-0.264***	-0.241***	-0.204***	-0.163***	-0.112***	-0.061***	-0.014***	0.021***	0.048***	0.067***	0.078***
	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
FUS	-0.146***	-0.028***	0.016***	0.049***	0.077***	0.099***	0.110***	0.119***	0.123***	0.125***	0.124***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Difference	-0.119***	-0.213***	-0.221***	-0.212***	-0.189***	-0.160***	-0.124***	-0.098***	-0.075***	-0.058***	-0.046***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Explained	-0.020***	-0.020***	-0.019***	-0.013***	-0.008***	-0.004**	0.001	0.004**	0.007***	0.008***	0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unexplained	-0.099***	-0.194***	-0.202***	-0.199***	-0.181***	-0.156***	-0.125***	-0.103***	-0.081***	-0.066***	-0.051***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Explained, by covariate											
First Job (this spell)	-0.001***	-0.002***	0.000***	0.000**	-0.000	0.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
STEM	-0.008***	-0.010***	-0.009***	-0.005***	-0.000	0.003***	0.006***	0.009***	0.011***	0.014***	0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Grad $\geq$ 2008	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
STEM $\times$ Grad $\geq$ 2008	-0.004***	0.001*	0.003***	0.003***	0.004***	0.004***	0.003***	0.003***	0.002***	0.001*	0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Major Generality	-0.007***	-0.006***	-0.005***	-0.005***	-0.005***	-0.004***	-0.004***	-0.003***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Non-US Labor Mkt. Exp.	-0.000	0.006***	0.000	0.001	0.001	0.001	0.003*	0.006***	0.007***	0.006**	0.012***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Non-US Labor Mkt. Exp.^{2}	0.000	-0.000***	0.000	0.000	-0.000	-0.001	-0.003**	-0.005***	-0.007***	-0.008***	-0.019***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Same Employer (Transfer)	-0.000	-0.000*	-0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained, by covariate											
First Job (this spell)	0.107	0.002	0.002**	0.002**	0.001**	-0.000					
	(0.056)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)					
STEM	-0.041***	-0.027**	-0.030**	-0.027**	-0.030***	-0.029***	-0.029***	-0.028***	-0.027***	-0.029***	-0.023***
	(0.012)	(0.010)	(0.009)	(0.009)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)
Grad $\geq$ 2008	0.033**	0.017	0.011	0.014	0.012	0.005	0.002	-0.001	-0.005	-0.008	-0.003
	(0.012)	(0.010)	(0.009)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)
STEM $\times$ Grad $\geq$ 2008	0.036***	0.011	0.005	-0.006	-0.006	-0.001	0.005	0.008	0.012**	0.016***	0.013***
	(0.009)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
Major Generality	-0.121***	-0.075***	-0.038*	-0.018	-0.006	0.017	0.022	0.019	0.011	0.019	0.016
	(0.022)	(0.018)	(0.017)	(0.016)	(0.015)	(0.014)	(0.013)	(0.012)	(0.011)	(0.011)	(0.010)
Non-US Labor Mkt. Exp.	0.001	-0.032***	-0.024***	-0.018***	-0.012**	-0.010*	-0.011*	-0.004	0.006	0.019***	0.027***
	(0.001)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Non-US Labor Mkt. Exp.^{2}	-0.000	0.004***	0.005***	0.008***	0.011***	0.013***	0.016***	0.015***	0.012***	0.006	0.006
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Same Employer (Transfer)	-0.000	0.001***	0.001***	0.002***	0.002***	0.002***	0.002***	0.002***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	200,757	221,459	234,289	241,376	246,817	255,096	261,521	268,080	275,179	282,460	297,518

Notes: Oaxaca–Blinder decomposition, FUS coefficients as reference, of the log OEWS median wage of the occupation held in each year since the worker's first year in the US labor market (for FUS workers, since the terminal degree), net of graduation-cohort effects, among workers who remain in the United States through year ten; 2003–2021 graduation cohorts. Each column is estimated separately. Year zero is reported for completeness; the text and Figures 10 and 9 start at year one (Section 4.2). Covariates are STEM, graduation in 2008 or later and its interaction with STEM, major generality, an indicator for the first job spell, school tier, and major field, and for CtW workers non-US labor-market experience, its square, and a same-employer transfer indicator; tier and field contributions are included in the totals but not shown. Standard errors in parentheses. Doctorate pools the PhD and non-licensure doctorates; Professional denotes licensure doctorates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A12: Decomposing the Occupational-Wage Gap by Year since Degree, CtW Professional Degree Holders, Ten-Year Stayers

	Years since first US labor-market year										
	0	1	2	3	4	5	6	7	8	9	10
CtW	0.247***	0.257***	0.282***	0.306***	0.316***	0.296***	0.303***	0.305***	0.308***	0.308***	0.310***
	(0.010)	(0.008)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
FUS	0.257***	0.350***	0.374***	0.381***	0.383***	0.378***	0.377***	0.373***	0.373***	0.373***	0.373***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Difference	-0.010	-0.094***	-0.092***	-0.074***	-0.067***	-0.082***	-0.074***	-0.069***	-0.065***	-0.065***	-0.063***
	(0.010)	(0.009)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Explained	0.140***	0.096***	0.086***	0.088***	0.087***	0.079***	0.072***	0.064***	0.055***	0.047***	0.039***
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Unexplained	-0.149***	-0.190***	-0.178***	-0.163***	-0.153***	-0.161***	-0.146***	-0.133***	-0.120***	-0.112***	-0.101***
	(0.010)	(0.009)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Explained, by covariate											
First Job (this spell)	0.013***	0.003***	0.003***	0.005***	0.002***	0.001**					
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
STEM	0.001*	0.000*	0.000	0.000	-0.000**	-0.001***	-0.001***	-0.001***	-0.001***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Grad $\geq$ 2008	-0.001	-0.000**	-0.000*	0.000	0.000	-0.000*	-0.001*	-0.001*	-0.001*	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
STEM $\times$ Grad $\geq$ 2008	-0.002	-0.001	-0.001*	-0.002***	-0.002***	-0.001***	-0.001***	-0.001***	-0.001***	-0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Major Generality	0.010***	0.008***	0.011***	0.012***	0.010***	0.008***	0.007***	0.007***	0.006***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Non-US Labor Mkt. Exp.	-0.000	-0.010***	-0.017***	-0.018***	-0.009***	-0.002	0.006***	0.011***	0.018***	0.022***	0.028***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Non-US Labor Mkt. Exp.^{2}	-0.000	0.001	0.001***	0.000	-0.003***	-0.008***	-0.014***	-0.021***	-0.030***	-0.038***	-0.047***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Same Employer (Transfer)	-0.000	-0.000	-0.000	-0.000	-0.000	0.000	0.000	0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained, by covariate											
First Job (this spell)	-0.019	-0.000	0.009**	0.004	0.005**	0.004**					
	(0.078)	(0.006)	(0.003)	(0.003)	(0.002)	(0.002)					
STEM	0.027*	0.035***	0.027**	0.018*	0.017**	0.013*	0.013**	0.015**	0.018***	0.011**	0.013***
	(0.013)	(0.010)	(0.008)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
Grad $\geq$ 2008	0.145***	0.136***	0.130***	0.085***	0.045***	0.029*	0.019	0.024*	0.031***	0.014	0.016*
	(0.024)	(0.020)	(0.017)	(0.015)	(0.013)	(0.011)	(0.011)	(0.010)	(0.009)	(0.008)	(0.008)
STEM $\times$ Grad $\geq$ 2008	-0.022*	-0.021*	-0.012	0.001	0.002	0.002	0.000	-0.003	-0.005	-0.001	-0.002
	(0.011)	(0.009)	(0.007)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Major Generality	-0.046	-0.043	0.063	0.139***	0.167***	0.128***	0.124***	0.148***	0.141***	0.141***	0.160***
	(0.043)	(0.039)	(0.037)	(0.035)	(0.033)	(0.032)	(0.030)	(0.028)	(0.027)	(0.025)	(0.024)
Non-US Labor Mkt. Exp.	0.006	-0.010	-0.034***	-0.053***	-0.061***	-0.060***	-0.065***	-0.069***	-0.062***	-0.061***	-0.059***
	(0.005)	(0.008)	(0.008)	(0.009)	(0.010)	(0.011)	(0.011)	(0.012)	(0.013)	(0.014)	(0.014)
Non-US Labor Mkt. Exp.^{2}	-0.002	0.002	0.004*	0.007**	0.008*	0.010	0.021***	0.027***	0.031***	0.032***	0.037***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.004)	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	(0.010)
Same Employer (Transfer)	0.000	0.000	-0.000	-0.000	-0.001	-0.001	-0.001	-0.000	0.000	0.001	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
N	227,793	266,285	286,928	299,691	307,426	315,004	320,950	326,365	331,464	337,072	350,304

Notes: Oaxaca–Blinder decomposition, FUS coefficients as reference, of the log OEWS median wage of the occupation held in each year since the worker's first year in the US labor market (for FUS workers, since the terminal degree), net of graduation-cohort effects, among workers who remain in the United States through year ten; 2003–2021 graduation cohorts. Each column is estimated separately. Year zero is reported for completeness; the text and Figures 10 and 9 start at year one (Section 4.2). Covariates are STEM, graduation in 2008 or later and its interaction with STEM, major generality, an indicator for the first job spell, school tier, and major field, and for CtW workers non-US labor-market experience, its square, and a same-employer transfer indicator; tier and field contributions are included in the totals but not shown. Standard errors in parentheses. Doctorate pools the PhD and non-licensure doctorates; Professional denotes licensure doctorates. $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

B Existence of a Natives' Separating Equilibrium (Mailath, 1987)

The general model with minimal assumptions on functional form follows the model outlined in (Mailath, 1987). The original paper made no assumptions on whether a high θ denotes higher productivity. For the purpose of this paper, I make this assumption and specify the conditions from the paper as such.

Let $U\left(\theta, \hat{\theta}, e\right)$ be the utility function of a type θ worker with education level e when the firms believe that they are type $\hat{\theta}$. Conditions 1 through 5 below are regularity conditions:

Condition 1 (Smoothness)**.** $U\left(\theta, \hat{\theta}, e\right)$ *is twice differentiable on* $\Theta^2 \times \mathcal{E}$.

Condition 2 (Belief Monotonicity)**.** $\dfrac{\partial}{\partial \hat{\theta}} U\left(\theta, \hat{\theta}, e\right) > 0.$

In other words, fixing the education level and actual productivity type, being recognized as having higher productivity is strictly better. The condition incentivizes high productivity workers to signal their productivity type.

Condition 3 (Type monotonicity)**.** $\dfrac{\partial^2}{\partial \theta \partial e} U\left(\theta, \hat{\theta}, e\right) > 0$

Given that $U\left(\theta, \hat{\theta}, e\right) = w\left(\hat{\theta}\right) - C\left(\theta, e\right)$, this condition ensures that the marginal cost of education decreases in productivity.

Condition 4 ("Strict" Quasi-Concavity[11])**.** *The equation* $\frac{\partial}{\partial e} U\left(\theta, \hat{\theta}, e\right) = 0$ *has a unique solution in* e*, denoted* $\phi(\theta)$ *such that* $\phi(\theta) \in \underset{e}{argmax}\ U\left(\theta, \hat{\theta}, e\right)$*, and* $\dfrac{\partial^2}{\partial e^2} U\left(\theta, \hat{\theta}, e\right)\Big|_{e=\phi(\theta)} < 0$

This condition ensures that there is a unique optimal education level for each productivity type, fixing $\hat{\theta}$.

Condition 5 (Boundedness)**.** $\exists k > 0$ *such that*

$$\forall (\theta, e) \in \Theta \times \mathcal{E}, \frac{\partial^2}{\partial e^2} U(\theta, \hat{\theta}, e) \geq 0 \ \Rightarrow\ \left|\frac{\partial}{\partial e} U(\theta, \hat{\theta}, e)\right| > k.$$

Conditions 4 and 5 together implies that, for arbitrary $\sigma(\theta)$, the set of possible signals that are reasonable, given the firms' belief update, is bounded (Mailath, 1987). In other words, condition 4 ensures the existence of an optimal signaling education level while condition 5 ensures that signals must not be free.

These next two conditions help characterize the separating equilibrium in the game:

Condition 6 (Initial Value [IV])**.** $\sigma\left(\tilde{\theta}\right) = \phi\left(\tilde{\theta}\right)$*, where*

$$\tilde{\theta} = \begin{cases} \sup \Theta & \text{if } \frac{\partial}{\partial \hat{\theta}} U(\theta, \hat{\theta}, e) < 0 \\ \inf \Theta & \text{if } \frac{\partial}{\partial \hat{\theta}} U(\theta, \hat{\theta}, e) > 0 \end{cases} \tag{B.1}$$

[11] Mailath (1987) notes that this condition is stronger than the typical strict quasi-concavity but weaker than concavity.

In other words, IV is satisfied if the marginal utility with respect to the worker's true type being negative (positive) implies that the lowest (highest) productivity type chooses the the optimal education level in equilibrium. This condition helps with the solvability of Equation (B.3) below.

Condition 7 (Single Crossing [SC])**.**

$$\frac{d}{d\theta}\frac{\frac{\partial}{\partial e}U(\theta,\hat{\theta},e)}{\frac{\partial}{\partial\hat{\theta}}U(\theta,\hat{\theta},e)} \gtrless 0 \tag{B.2}$$

SC ensures that the indifference curves of any two informed workers of different types crosses at most once.

Mailath (1987) theorem 1: If conditions 1 through 5 are satisfied and $\sigma(\theta)$ is one-to-one and satisfies incentive compatibility, then $\sigma(\theta)$ has at most one discontinuity on Θ and where it is continuous, it is differentiable and satisfies:

$$\frac{d\sigma}{d\theta} = -\frac{\frac{\partial}{\partial\hat{\theta}}U(\theta,\hat{\theta},e)}{\frac{\partial}{\partial e}U(\theta,\hat{\theta},e)}. \tag{B.3}$$

Mailath (1987) theorem 2: If conditions 1 through 5 are satisfied, and $\sigma(\theta)$ is one-to-one and satisfies incentive compatibility, then

SC $\Rightarrow$ σ is strictly monotonic on Θ and σ satisfies Eq. (3)

IV $\Rightarrow$ σ is strictly monotonic on Θ, satisfies Eq. (3), and $sign\left(\frac{d\sigma}{d\theta}\right) = sign\left[\frac{\partial^2}{\partial\theta\partial e}U\left(\theta,\hat{\theta},e\right)\right]$

Mailath (1987) corollary: Suppose conditions 1 through 5 are satisfied and IV holds, then the equilibrium σ is the optimal σ, which uniquely solves Equation (B.3).

Definition 2 (Strict Incentive Compatibility [SIC])**.** *The strategy $e(\theta)$ satisfies strict incentive compatibility if*

$$\{\sigma(\theta)\} = \underset{e}{argmax}\ U(\theta,\hat{\Theta}(e),e),\ \forall\theta\in\Theta \tag{B.4}$$

Mailath (1987) theorem 3: Suppose that conditions 1 through 5 are satisfied. A strictly increasing function σ, continuous at $\underline{\theta}$, satisfies strict incentive compatibility if and only if

(i) σ solves Equation (B.3).

(ii) When $\frac{\partial}{\partial\hat{\theta}}U(\theta,\hat{\theta},e) > (<)0$, $\frac{\frac{\partial}{\partial e}U(\theta,\hat{\theta},e)}{\frac{\partial}{\partial\hat{\theta}}U(\theta,\hat{\theta},e)}$ is a strictly increasing (decreasing) function of θ for $\left(\hat{\theta},e\right)$ in the graph of σ.

B.1 Graphical derivation of the continuous type model

Step 1: Suppose $\Theta = \{\bar{\theta}, \underline{\theta}\}$.

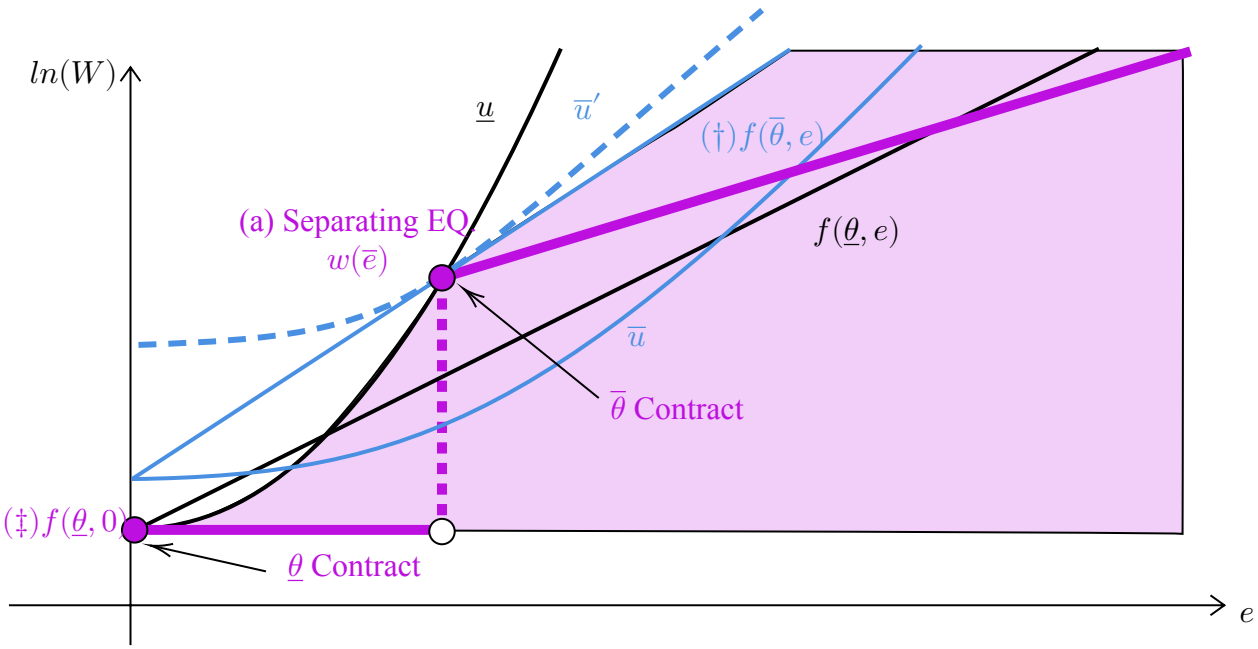
In Figure B1 below, the curves $\underline{u}$ and $\bar{u}$ indicates the indifference curve of a worker of type $\underline{\theta}$ and $\bar{\theta}$ with no education, respectively. The linear production functions of workers of respective types are indicated on the graph. Where $\underline{u}$ intersects $f(\bar{\theta}, e)$ creates a separating contract where the type $\underline{\theta}$ worker is indifferent between getting the education level $\bar{e}$ required for that contract and accepting a contract that requires no education.

At that same point, the $\bar{\theta}$ worker will obtain the education level $\bar{e}$ for the contract, as indicated by the $\bar{\theta}$ worker's higher level indifference curve $\bar{u}'$. A full equilibrium wage schedule is thus represented by the purple discontinuous function on the graph:

$$w(e) = \begin{cases} f(\underline{\theta}, 0) & e < \bar{e} \\ f(\bar{\theta}, \bar{e}) & e = \bar{e} \\ f(\theta, e) & e > \bar{e},\ \theta \in \Theta \end{cases}$$

Note that the equilibrium wage schedule can be any wage schedule that is in the shaded area as long as it passes through $f(\underline{\theta}, 0)$ and $f(\bar{\theta}, \bar{e})$

Figure B1: Separating Equilibrium with $\Theta = \{\bar{\theta}, \underline{\theta}\}$



Notes: (†) marks the type-$\bar{\theta}$ worker's productivity schedule $f(\bar{\theta}, e)$. (‡) marks the type-$\underline{\theta}$ worker's productivity schedule $f(\underline{\theta}, e)$. (a) marks the separating equilibrium contract for type-$\bar{\theta}$ workers, $w(\bar{e})$. $\bar{u}'$ represents type-$\bar{\theta}$ workers' indifference curve at the separating contract. $\bar{u}$ represents type-$\bar{\theta}$ workers' indifference curve at the $\underline{\theta}$ contract.

Step 2: Add an intermediate productivity: $\Theta = \{\bar{\theta}, \theta_M, \underline{\theta}\}$ where $\bar{\theta} > \theta_M > \underline{\theta}$.

The result here is achieved in the same way as step 1. In Figure B2 below, the curves $\underline{u}$, $\underline{u}_M$, and $\bar{u}$ indicates the indifference curve of a worker of type $\underline{\theta}$, θ_M, and $\bar{\theta}$ with no education, respectively. The linear production functions of workers of respective types are indicated on the graph. Where $\underline{u}$ intersects $f(\theta_M, e)$ creates a separating contract $w(e_M)$where the type $\underline{\theta}$ worker is indifferent between getting the education level e_M required for that contract and accepting a contract that requires no education.

At the contract $w(e_M)$, the θ_M worker will obtain the education level e_M for the contract, as indicated by the θ_M worker's higher level indifference curve $\bar{u}_M$.

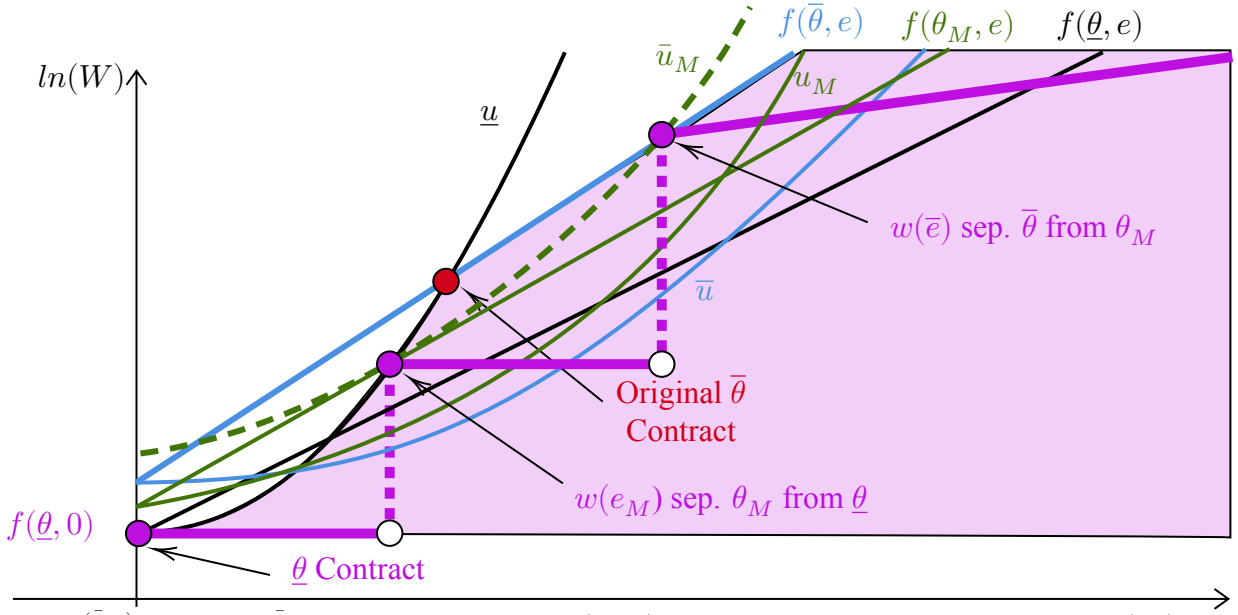
Similarly, the indifference curve $\bar{u}_M$ intersects $f\left(\bar{\theta}, e\right)$ at the contract $w(\bar{e})$. Since both $w(\bar{e})$ and $w(e_M)$ are on $\bar{u}_M$, the θ_M worker is indifferent between the two education levels while the $\bar{\theta}$ worker will choose $\bar{e}$ as that contract achieves a higher utility for them (indifference curve not shown).

A full equilibrium wage schedule is thus represented by the purple discontinuous function on the graph:

$$w(e) = \begin{cases} f\left(\underline{\theta}, 0\right) & e < e_M \\ f\left(\theta_M, e_M\right) & e_M \leq e \leq \bar{e} \\ f\left(\bar{\theta}, \bar{e}\right) & e = \bar{e} \\ f\left(\theta, e\right) & e > \bar{e},\ \theta \in \Theta \end{cases}$$

Note that the equilibrium wage schedule can be any wage schedule that is in the shaded area as long as it passes through $f\left(\underline{\theta}, 0\right)$, $f\left(\theta_M, e_M\right)$ and $f\left(\bar{\theta}, \bar{e}\right)$.

Figure B2: Separating Equilibrium with $\Theta = \{\bar{\theta}, \theta_M, \underline{\theta}\}$ where $\bar{\theta} > \theta_M > \underline{\theta}$



Notes: $f(\bar{\theta}, e)$ marks the type-$\bar{\theta}$ worker's productivity schedule. $f(\theta_M, e)$ marks the type-θ_M worker's productivity schedule. $f(\underline{\theta}, e)$ marks the type-$\underline{\theta}$ worker's productivity schedule. $w(\bar{e})$ marks the separating equilibrium contract for type-$\bar{\theta}$ workers, $w(\bar{e})$; $w(e_M)$ marks the separating equilibrium contract for type-θ_M workers, $w(e_M)$. $\bar{u}_M$ represents type-θ_M workers' indifference curve at the separating contracts where they are indifferent between obtaining e_M and $\bar{e}$. $\bar{u}$ represents type-$\bar{\theta}$ workers' indifference curve at the $\underline{\theta}$ contract.

Step 3: A Continuum of Types: $\Theta = \left[\underline{\theta}, \bar{\theta}\right]$.

Notice that because each point on the separating wage schedule is determined by the ordered productivity functions, the straight line between $w(e_M)$ and $w(\bar{e})$ necessarily lies below the $w(\bar{e})$ in step 1, as indicated on the Figure B2. This implies that, as more and more types are added to Θ, the equilibrium wage schedule will conform to a convex function as shown in Figure B3 below.

The equilibrium wage schedule is thus:

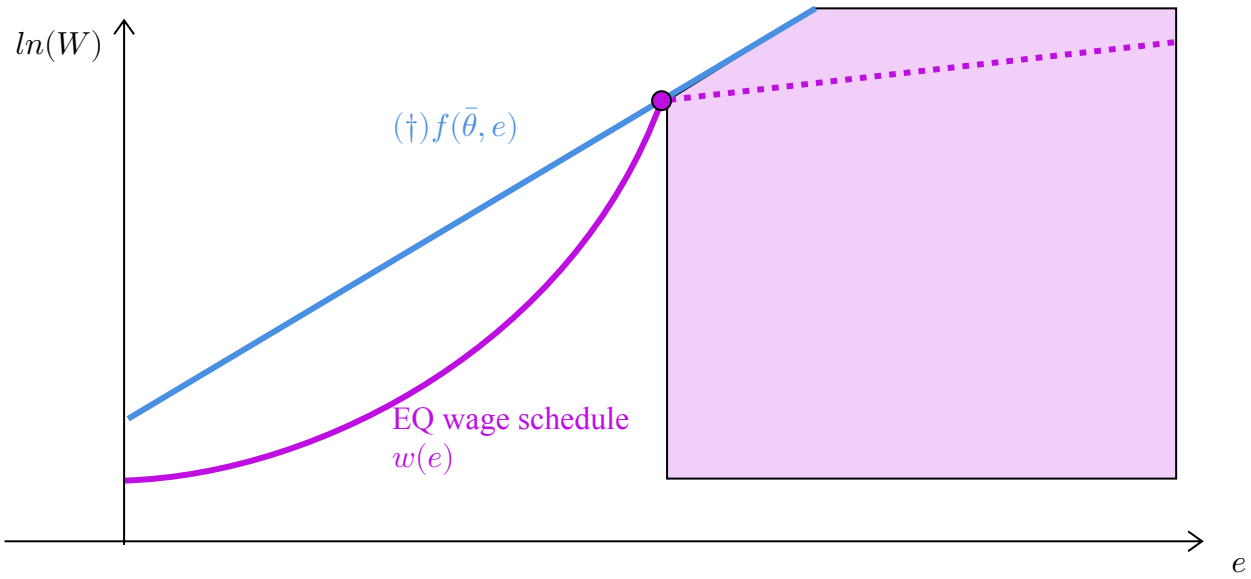
$$w(e) = \begin{cases} f(\theta, \sigma(\theta)) & , e \in \left[\sigma\left(\underline{\theta}\right), \sigma\left(\bar{\theta}\right)\right] \\ f(\theta, e) & , otherwise. \end{cases}$$

Contracts where $e > \sigma\left(\bar{\theta}\right)$—represented by the dotted wage schedule and shaded region—exhibit over-signaling, as the marginal wage return to education strictly exceeds its marginal cost. The shaded region spans the complete set of equilibrium wages these contracts can sustain, while the dotted schedule serves as an illustrative benchmark.

Assumption 8 (Occupation-Level Posted Wages)**.** *The contract $w(e)$ attaches to the occupation rather than to the worker: a firm filling a position whose equilibrium contract is $w(e)$ pays $w(e)$ to whomever it hires, and cannot adjust that wage downward for an immigrant.*

Assumption 2 is what makes the immigrant extension bite, and it is worth stating why. Firms directly observe immigrant status, including foreign education and sponsorship needs. Absent a

Figure B3: Separating Equilibrium with $\Theta = [\underline{\theta}, \bar{\theta}]$



Notes: $(\dagger)$ marks the type-$\bar{\theta}$ worker's productivity schedule $f(\bar{\theta}, e)$. The intersection of $(\dagger)$ and the equilibrium wage schedule is the separating equilibrium contract for type-$\bar{\theta}$ workers.

restriction on wage setting, Bertrand competition among firms that observe it would simply offer an immigrant the contract $f(\theta, e) - d$ at every e: a smooth downward shift of the schedule that leaves no education level unavailable and generates no discrete mismatch. Under Assumption 2 the firm cannot make that offer. Employing a type-$\bar{\theta}$ immigrant in an occupation whose contract is $W(e)$ yields $f(\bar{\theta}, e) - d - w(e)$, and where this is negative the firm excludes the immigrant from the occupation rather than underpaying her. The H-1B program makes the assumption concrete rather than merely convenient: a sponsoring employer must attest to paying at least the Department of Labor prevailing wage for the occupation, so nativity-specific downward wage adjustment within an occupation is legally foreclosed and exclusion into a lower-education occupation becomes the binding margin. Posted occupational wage scales and equal-pay norms deliver the same restriction more broadly.

C Theoretical Foundation of the PageRank Score

Firm PageRank utility follows Sorkin (2018), who utilizes systematic mobility between firms to create an order over employers under the revealed-preference framework.

Denote the set of workers $\mathcal{I}$ and the set of firms $\mathcal{J} = \{1, \dots, J\}$. Consider a worker $i \in \mathcal{I}$, currently employed at firm k, who receives an offer from firm j; the worker accepts if $V_j + \varepsilon_{ijk} > V_k + \varepsilon_{ikk}$, where V_j and V_k are common utility attached to firms j and k—the same for every worker who might compare them—and ε_{ijk} is worker i's idiosyncratic taste for firm j given current employer k, unrelated to V_j on average across workers. Let M_{jk} be the number of workers observed transitioning from firm k to firm j over the sample period, and let $S_k = \sum_{j \in \mathcal{J}} M_{jk}$ be firm k's total observed outflow. The observed share of k's leavers who land at j is then $A_{jk} = M_{jk}/S_k$, and under the choice rule above this share is informative about V_j relative to all other firms that k's workers are observed moving to: firms that draw a disproportionate share of leavers, from firms that are themselves valuable, must be highly valued. This is a self-referential definition—a firm's value is revealed by the value of the firms whose workers it attracts—and it is exactly encoded in the fixed point $\tilde{V}_j = \sum_k A_{jk}\tilde{V}_k$, where $\tilde{V} = (\tilde{V}_1, \dots, \tilde{V}_J)$ is the leading eigenvector of the $J \times J$ matrix A. The size-adjusted index used in the analysis, $V_j^e \approx \tilde{V}_j + \log(g_j)$, augments the pure choice-revealed value $\tilde{V}_j$ with the log of firm j's size g_j, on the view that part of a firm's realized flow reflects scale of opportunity rather than worker preference.

This logic requires three assumptions, each with a direct economic interpretation:

- **A common value index**: Any worker moving from k to j is assumed to agree, up to the idiosyncratic term ε_{ijk}, that j is worth V_j—i.e., firms are vertically ranked (some firms are simply better places to work), rather than workers sorting on match-specific comparative advantage. If firm value were instead worker-specific, the share A_{jk} would reveal relative match quality for the mover, not a firm attribute, and no single ranking $\tilde{V}$ would be recoverable.

- **Selection on value, not on outside constraints**: A transition captured in M_{jk} is assumed to reflect the worker choosing the best available offer, not being forced to move (e.g., by a layoff) or by a search friction unrelated to V_j. To the extent involuntary separations are common and correlated with firm quality, A_{jk} conflates "workers prefer j" with "workers were pushed out of k," attenuating the revealed-preference interpretation of $\tilde{V}$.

- **A connected and stable labor market**: Every firm's value V_j must be comparable to every other's through some chain of observed moves captured in M (workers routed through the non-employment node count as one such link), and the ranking $\tilde{V}$ is assumed stable over the window used to estimate it. To account for moving market conditions, I estimate the ranking using 5-year rolling windows and interpolate the rankings for the period 2020-2022.

C.1 Within-education occupation ranking mechanics.

Within a given degree-major pair containing n distinct occupations, the estimated ranking adds a small baseline term representing an occupation's share of workers independent of where they came

from,

$$r = d\,A\,r + \frac{1-d}{n}\mathbf{1}, \tag{C.1}$$

where $A = [A_{ba}]$ is the stratum's occupation-to-occupation share matrix, $\mathbf{1}$ is the n-vector of ones, and the damping factor $d\,(= 0.85$ per Brin and Page (1998)) is the weight placed on the observed-transition term relative to the uniform baseline term. This ensures every occupation retains some minimal comparability even absent a direct or well-populated chain of observed moves connecting it to the rest of the stratum's network.

Feasible-set truncation and the estimation sample Revealed-preference rankings presume that observed moves reveal preference over a commonly valued bundle. For immigrant workers that presumption is not innocuous: their feasible set is truncated, so a move into a lower-quantile occupation may reflect the constraint rather than the ranking, compressing the ranking across the interval of interest and attenuating d toward zero. I therefore re-estimate the occupation ranking on native workers alone, whom the model places on the undistorted schedule and whose mobility reveals unconstrained preference. The two rankings agree closely—person-year-weighted Spearman of 0.979 across the 37 postgraduate strata, with no loss of network connectivity — so the full-sample ranking is retained. Firm rankings use the full sample: exclusion operates at the occupation contract rather than the firm, and immigrant moves supply the panel's only coverage of foreign employers.

D LinkedIn Coverage with IPEDS Benchmark

This appendix evaluates the representativeness of the People Data Labs (PDL) LinkedIn education records by benchmarking it against the Integrated Postsecondary Education Data System (IPEDS) Completions census, maintained by the National Center for Education Statistics. Because IPEDS tracks the universe of degrees conferred by Title IV-eligible institutions within the United States, this exercise establishes the extent to which the LinkedIn panel reflect the true population of U.S. graduates. International institutions fall outside the scope of this benchmark as they lack a corresponding IPEDS institutional identifier (UNITID).

The rest of this appendix proceeds as follows: Appendix D.1 outlines the alignment and filtering protocols used to construct the comparative sample. Appendix D.2 assesses absolute coverage across institutional tiers. Appendix D.3 examines the fidelity of the within-school field mix. Appendix D.4 documents systematic field-level representation biases; and finally, I present data-quality diagnostics and robustness checks.

D.1 Data Alignment and Sample Construction

I compare the population census (IPEDS) and the sample panel (LinkedIn/PDL) at the cell level, defined as the institution$\times$degree level$\times$graduation year$\times$field of study level.

- **Population Data:** The IPEDS Completions baseline is restricted to bachelor's degrees (award level 5) and master's degrees (award level 7), capturing first majors conferred between 2008 and 2019.
- **Sample Data:** The LinkedIn/PDL panel is structured at the individual degree level. Each record is resolved to an IPEDS UNITID, and free-text major fields are mapped onto standardized two-digit Classification of Instructional Programs (CIP2) codes via a curated crosswalk. Fields of study are evaluated as CIP2 share vectors within each institution-year cell.

To ensure statistical reliability, two primary filters are applied to the data:

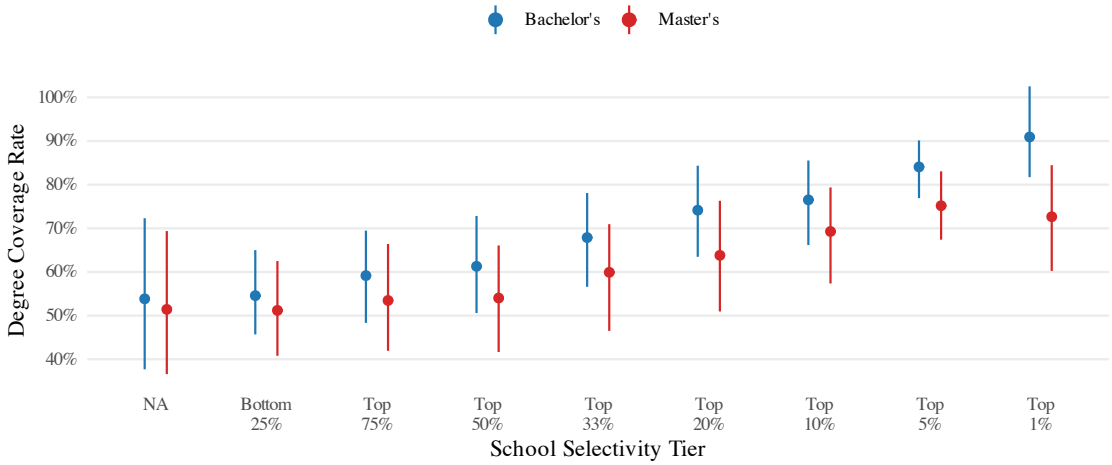
1. **Overlap requirement:** An institution-year cell is retained only if it contains at least 50 graduates in the IPEDS census and 20 matched profiles in the LinkedIn panel. Out of 54,277 potential school-year cells, 29,536 satisfy this criterion. While this percentage can seem low, the majority of drops can be attributed to misclassified schools in the PDL data.
2. **Crosswalk Collision Filter:** Institutional cells subject to crosswalk attribution errors—where a single institutional identifier in LinkedIn erroneously aggregates a multi-campus system or an online-only university chain—are identified by an impossible coverage rate exceeding 100%. These collided cells are excluded from the primary analytic samples to preserve the integrity of campus-level tiering.

D.2 Coverage and Institutional Selectivity

Coverage is defined as the ratio of captured LinkedIn profiles to the total IPEDS graduate count. Aggregated across all fields and graduation years, the panel captures **56.3%** of U.S. bachelor's graduates and **50.6%** of master's graduates within the collision-free, volume-weighted sample.

However, coverage is not uniform across higher education institutions. Figure D1 demonstrates that the median school-level coverage rate rises monotonically with institutional admissions selectivity. Median coverage ranges from approximately 57% for institutions in the lowest selectivity tier (open or unreported admissions) to 83% for institutions in the most selective tier, indicating that graduates of elite institutions are significantly more likely to maintain professional online profiles.

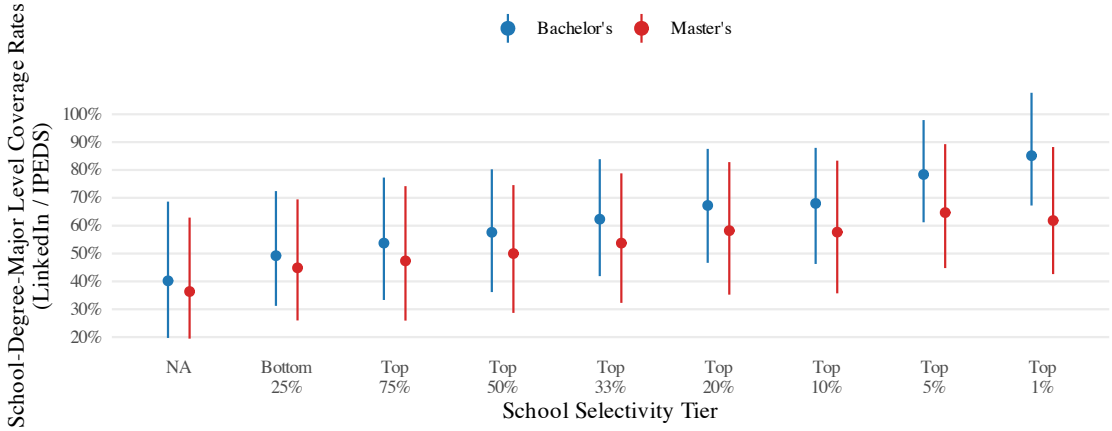
Figure D1: **Institutional Coverage by Admissions Selectivity Tier**



Median (point) with IQR (range). Tier 1 = open / not reported.

This institutional gradient persists at the finer granular level at which the dataset is utilized (institution × degree × major), as illustrated in Figure D2. Volume-weighted cell-level coverage climbs from 26% in the lowest selectivity tier to 88% in the most selective tier for bachelor's degrees, and from 28% to 77% for master's degrees.

Figure D2: **Granular Cell-Level Coverage by Selectivity Tier**

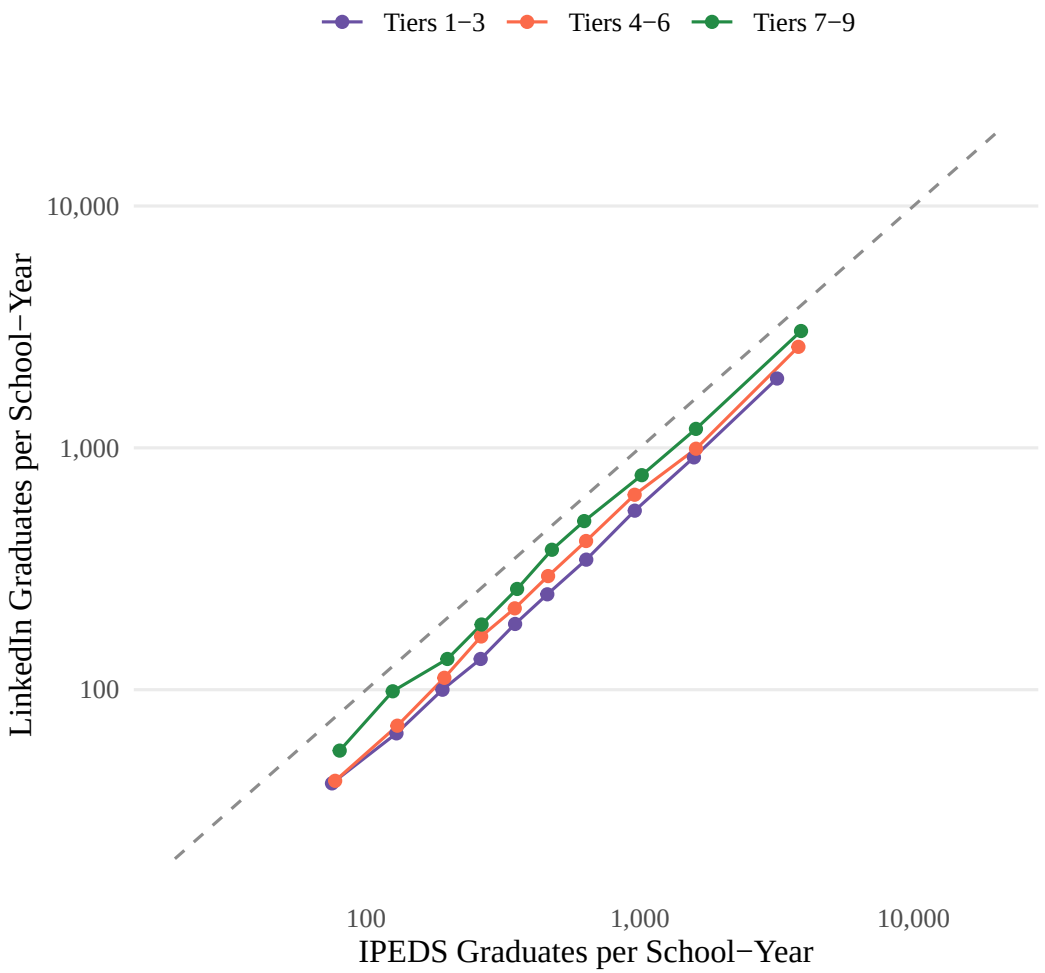


Median (point) + IQR (range) of cell coverage rates among school x major cells with at least one LinkedIn grad (IPEDS >= 25, CIP 15/46 excluded). Zero-coverage cells are excluded and are common at low tiers (~half of bachelor's cells at tier 1 (NA)), so these are coverage rates CONDITIONAL on any LinkedIn presence.

Crucially, within any given selectivity tier, coverage exhibits size-neutrality. Figure D3 plots log LinkedIn graduate counts against log IPEDS graduate counts. Within each selectivity band,

the relationship runs strictly parallel to the 45-degree identity line. A formal elasticity estimate indicates that LinkedIn volume scales one-for-one with cohort size (pooled elasticity $\approx$ 1.02, $R^2 \approx 0.90$). Consequently, the proportional under-reporting shortfall is invariant to the size of an academic program, and the vertical dispersion between bands reflects the pure institutional selectivity gradient.

Figure D3: Size-Neutrality of Coverage Within Selectivity Bands



Notes: Binned medians of graduate counts within IPEDS−size deciles by tier band

D.3 Within-School Composition: The Field Mix

While absolute coverage captures level differences, establishing the representativeness of the panel also requires verifying that the internal *distribution* of academic fields within schools is accurately reproduced. Isummarize the fidelity of this within-school major composition using the Dissimilar-

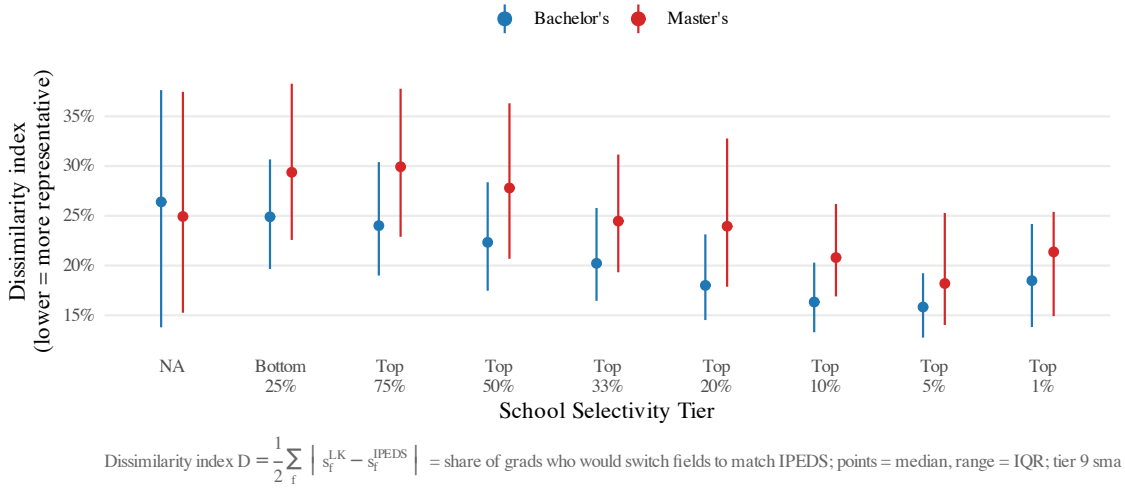
ity Index (D), which is mathematically equivalent to the total variation distance:

$$D = \frac{1}{2} \sum_{f} \left| s_f^{LK} - s_f^{IPEDS} \right| \tag{D.1}$$

where s_f^{LK} and s_f^{IPEDS} represent the shares of field f within an institution-year in LinkedIn and IPEDS, respectively. The index represents the proportion of graduates in the LinkedIn sample who would need to reallocate to alternative fields for the sample distribution to perfectly replicate the population baseline. As such, a higher dissimilarity index value corresponds to an institution where certain fields are reported disproportionally higher than others.

As shown in Figure D4, internal compositional divergence falls as institutional selectivity increases. The median dissimilarity index decreases from approximately 26% in the lowest tier to 16% in the highest tier. Master's programs exhibit slightly higher compositional divergence than bachelor's programs (overall medians of 27% versus 22%), which is consistent with smaller cell counts and greater heterogeneity in specialized professional program titles.

Figure D4: Within-School Field Mix Dissimilarity Across Selectivity Tiers



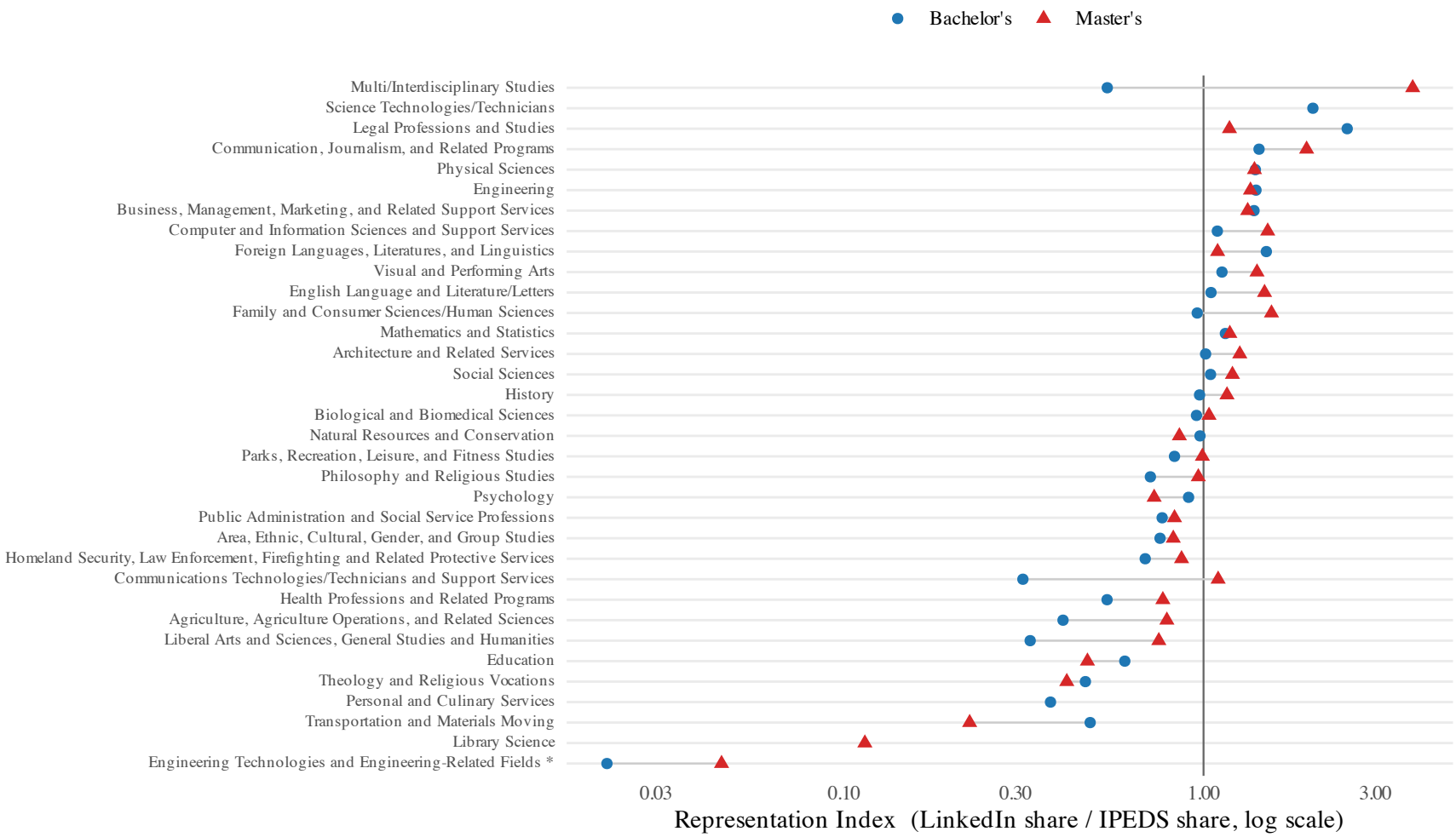
D.4 Field Bias and Classification Artifacts

A main concern with using LinkedIn data is that compositional errors are systematic and driven by field-level preferences for using LinkedIn. Figure D5 tracks the representation index, defined as the ratio of a field's share in the LinkedIn panel to its share in the IPEDS census. Values above unity signify that a field is over-represented in the sample relative to the population field mix.

The panel consistently over-represents professional, commercial, and quantitative fields, such as Business, Engineering, Communications, and Legal Studies. Conversely, it under-represents care-, teaching-, and public-oriented disciplines, including Health Professions, Education, and Liberal Arts/General Studies. This pattern holds true across both bachelor's and master's degree levels. Importantly, the nature of mapping free-text majors to CIP2 codes result in mis-categorization

in certain cases such as inputting ”Engineering Technologies and Engineering-Related Fields” as ”Engineering” and hence getting mapped to Engineering (CIP2-14) instead of the proper CIP2-15). However, the large distortions in representation seen in Figure D5 does not translate to the overall sample as major-relevant variables are calculated using the set of canonical majors in Appendix I.

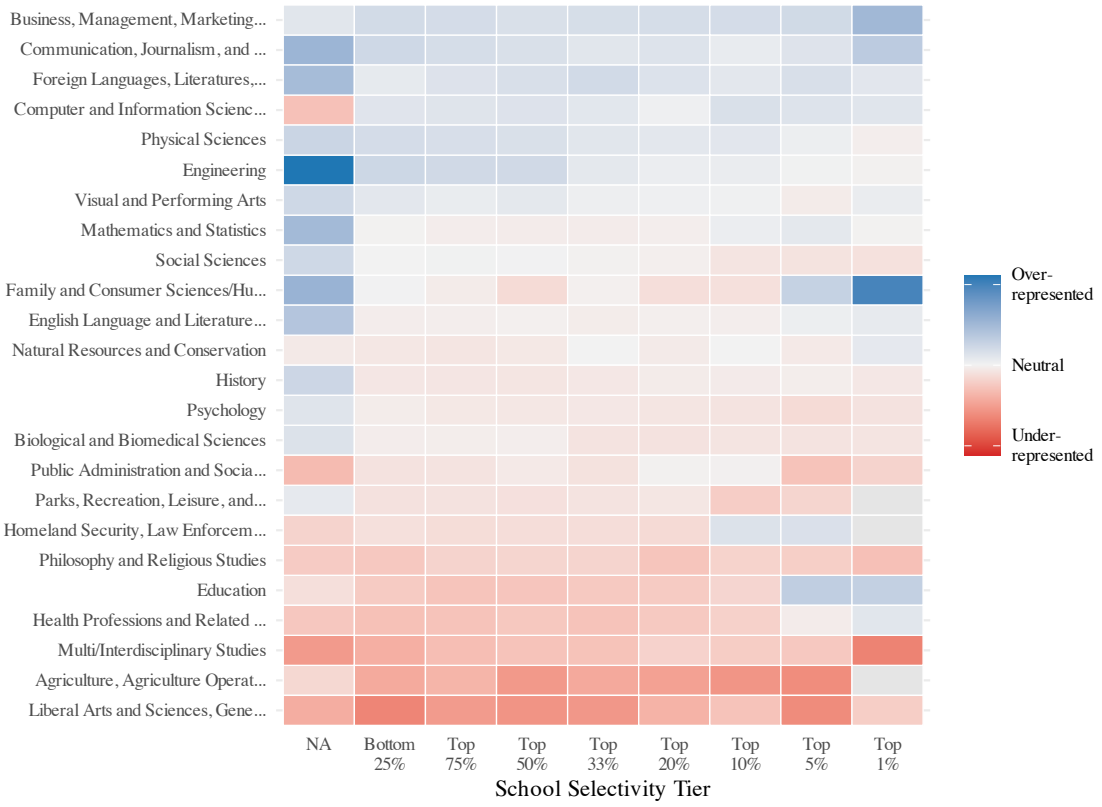
Figure D5: Field-Level Representation Index Relative to IPEDS Census Baseline



*Pooled over years, collision-free, IPEDS >= 5,000 grads. >1 over-, <1 under-represented. Circles = bachelor's, triangles = master's. * = major->CIP self-report seam (CIP 15 eng-tech, 46 construction): extreme values are matching artifacts, not sampling, showing free-text major->CIP matching is hard; Engineering (CIP 14) is correspondingly inflated.*

This field bias is uniform across institutional tiers rather than driven by specific subsets of schools. As shown in Figure D6, the institutional selectivity gradient and the field-level composition bias operate as roughly additive components: Health and Education remain under-represented across virtually all tiers, while Business and Communications are consistently over-represented.

Figure D6: Joint Distribution of Representation Across Fields and Tiers



Representation index = within-tier LinkedIn field share / IPEDS field share. Blue over-, red under-represented. Eng-tech (CIP 15) excluded; Engineering (CIP 14) remains inflated as eng-tech grads self-label as 'Engineering'.

The cell-grain coverage heatmap in Figure D7 synthesizes these dual margins. While coverage generally increases from left to right with selectivity across all fields, care and teaching fields remain compressed across nearly all tiers. Notably, several cells in the highest selectivity tiers exhibit coverage rates exceeding 1.0, most visibly in Business (CIP 52), which reaches 1.27 and 2.71 in the top two tiers.

This phenomenon does not reflect genuine over-capture of business degrees. Rather, elite institutions confer very few undergraduate business degrees, resulting in an exceptionally small IPEDS denominator. Concurrently, LinkedIn users who graduate from elite institutions with degrees in Economics or other quantitative social sciences and subsequently enter corporate or financial sectors frequently self-label their major as "Business." For example, Harvard reportedly conferred 0 business BAs from 2009 to 2018, but there are 3,641 LinkedIn users reporting obtaining a business bachelor's in the same period. This inflation in CIP 52 is mirrored by a corresponding deflation in Social Sciences (CIP 45), which remains below unity at those same elite tiers.

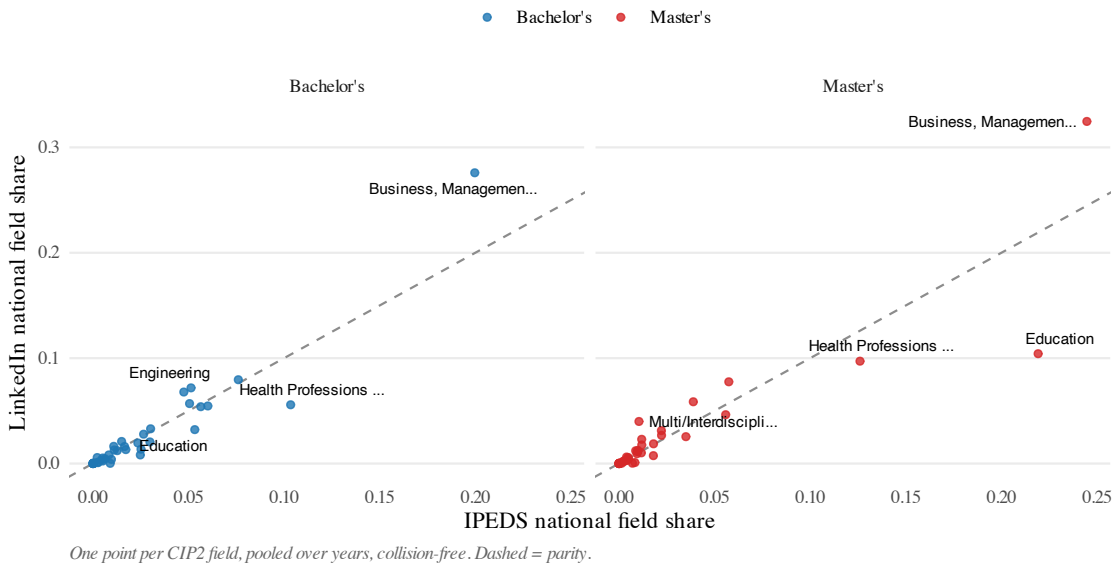
Figure D7: **Volume-Weighted Cell Coverage Heatmap**

Figures reported are volume-weighted coverage rates; Grey = fewer than 50 IPEDS grads (too thin). Engineeting tech (CIP 15) and construction (CIP 46) are excluded due to string matching misrepresentation.

D.5 Aggregate Robustness and Data Quality Diagnostics

National Portfolio Tracking: As a broad aggregate validation, Figure D8 plots the national CIP2 field shares in LinkedIn against those in IPEDS, pooling all available years. The overall composition tracks the population distribution closely, yielding Pearson correlation coefficients of 0.948 for bachelor's degrees and 0.897 for master's degrees.

Figure D8: **National Field Shares: LinkedIn versus IPEDS Baseline**



One point per CIP2 field, pooled over years, collision-free. Dashed = parity.

Institutional Crosswalk Collisions: As noted in Appendix D.1, coverage rates above 100% at the institutional level indicate mapping errors where LinkedIn's institutional identifiers merge distinct campuses or online chains into a single corporate hub. These account for 14.1% of the raw LinkedIn degree records and are concentrated heavily among lower-tier institutions (representing 57% of the total volume in the lowest selectivity tier). Rather than aggregating these records and losing institutional granularity, these cells are dropped from the primary sample.

Classification Seams: Certain severe field deviations stem from mapping artifacts rather than true sampling bias. For instance, "Engineering Technologies" (CIP 15) approaches zero in LinkedIn because graduates likely systematically label their majors as "Engineering" (CIP 14), causing a corresponding artificial inflation in the latter. Similarly, the text string "construction" initially inflated "Construction Trades" (CIP 46) before being algorithmically corrected to Construction Management under Business (CIP 52).

Sensitivity Checks: Three additional tests verify that the core results are robust:

1. *Match Quality:* Every in-scope degree is resolved via direct, un-ambiguous institutional identifiers rather than fuzzy name matching, ensuring that coverage patterns are not artifacts of alignment errors.

2. *Date Imputation:* Restricting the sample to the 2.8% of records with precise, month-resolved graduation dates leaves the national field-share correlations virtually unchanged (the bachelor's correlation shifts minimally from 0.948 to 0.967).

3. *Profile-Completion Lag:* Over time, total LinkedIn degree counts remain below IPEDS and exhibit an inverted-U trajectory that peaks around 2015-2016. This pattern is the signature of a multi-year lag in profile creation and completion by recent graduates, rather than a shifting structural sampling bias. Manual spot-checks against public institutional records confirm the validity of the underlying counts.

D.6 Methodological Implications

Because absolute profile coverage scales with institutional selectivity and the sample composition tilts toward corporate and quantitative fields, the raw LinkedIn panel over-represents graduates from elite schools and professional programs relative to the true U.S. graduate population. For research designs that are sensitive to these margins, these systematic discrepancies can be corrected. Researchers can construct post-stratification weights that align LinkedIn cells (at the institution $\times$ CIP2 $\times$ year level) to the true IPEDS population margins, effectively eliminating first-order selection bias.

Table D1: **Cell-Level Coverage and Capture Rates by Selectivity Tier (Bachelor's Degrees)**

Selectivity Tier	Number of Cells	Volume-Weighted Coverage	% IPEDS Captured
1 (Open / Unreported)	39,843	0.265	26.5
2 (Bottom 25%)	59,259	0.479	47.9
3 (Top 75%)	74,124	0.557	55.7
4 (Top 50%)	50,202	0.577	57.7
5 (Top 33%)	40,316	0.622	62.2
6 (Top 20%)	27,166	0.635	63.5
7 (Top 10%)	14,705	0.658	65.8
8 (Top 5%)	12,665	0.798	79.8
9 (Top 1%)	3,087	0.882	88.2
All Tiers	**321,367**	**0.563**	**56.3**

Notes: Volume-weighted coverage measures the percentage of the total IPEDS graduation cohort successfully matched to corresponding profiles in the professional panel.

E Data Construction

E.1 Why a Longitudinal Panel

Employer-learning is a within-worker process, and the workhorse sources for immigrant labor-market research—repeated cross-sections such as the decennial Census and the ACS—cannot follow the same individual across waves. Estimated experience-earnings and assimilation profiles therefore identify the true within-person path only if cohort composition is stable across waves, and that assumption fails when out-migration is selective. Lubotsky (2007) matches survey respondents to longitudinal Social Security earnings records and shows that lower-earning immigrants disproportionately leave the United States, so the surviving cross-section mechanically appears to assimilate; the longitudinal profile is markedly flatter than its cross-sectional counterpart, which overstates immigrant earnings growth by roughly half. Panel data does not fully resolve this, since longitudinal samples are themselves subject to non-random attrition (Beenstock et al., 2010), and linked historical records reveal analogous selection among earlier immigrant waves (Collins and Zimran, 2023). I therefore construct a novel longitudinal dataset.

E.2 Classifying Job Titles

To operationalize the LinkedIn panel, it is necessary to map self-reported, unstructured job titles into the standardized O*NET Standard Occupational Classification (SOC) codes. Because user-generated profiles frequently contain idiosyncratic, highly specific, or non-standard occupational phrasing, traditional fuzzy-matching algorithms or dictionary-based approaches yield low match rates.

During training, the fine-tuned model is prompted to evaluate the self-reported job title and output only the corresponding 8-digit SOC code. To prevent forced matches on irrelevant data, the training data explicitly instructs the model to return a null identifier (00-0000.00) for gibberish or un-mappable titles (e.g., "gfliwn askdm"). An analysis of the preliminary predictions demonstrates the model's robust semantic understanding. It successfully maps highly specific titles (e.g., "asset & reliability engineer" to Industrial Engineers) and academic appointments (e.g., "international visiting attorney" to Lawyers). Crucially for a study examining immigrant labor market trajectories, the model inherently processes and classifies non-English job titles without requiring ex-ante translation; for example, successfully mapping "網頁工程師" (Website Engineer) directly to Computer Systems Engineers/Architects and "博士生" (PhD Student) to Post-secondary Teaching Assistants. More examples are reported in Table A4.

E.3 Online-Profile Data in Labor Research

The tradeoff between platform selection and panel coverage is well understood in the emerging literature that repurposes online-profile and resume data for labor-market research. Berry et al. (2024) show that LinkedIn-derived samples misrepresent the demographic composition of employment in known and measurable ways, while demonstrating that the data can nonetheless be benchmarked against administrative population totals, and resume data at sufficient scale now supports credible causal designs (Van Dijcke et al., 2024). The analysis here follows that logic, trading the

return-migration bias that contaminates survey data for a platform selection bias whose structure is observable against the IPEDS benchmark and can be addressed with post-stratification weights.

E.4 The Conventional Mismatch Taxonomy

The main taxonomy of occupational mismatch a la Hartog (2000) can be visualized in the following way. Let the xy-plane represent the educational modality of an occupation and relabel it the fa-plane. The f-axis represents the field of specialization and the a-axis represents education attainment. For worker $i \in \boldsymbol{\mathcal{I}}$ with education $e_i = (f_i, a_i)$. The typical education requirement for job j is defined by the modal education e_M^j of workers in job j, i.e., $e_M^j = \underset{e}{argmax} \sum_{i \in \{i \in \boldsymbol{\mathcal{I}} | i \text{ works at job } j\}} \mathbb{1}\{e = e_i\}$. The conventional way to define vertical and horizontal mismatch is: (1) Worker i is vertically mismatched in job j if $a_i \neq \bar{a}^j$; and (2) Worker i is horizontally mismatched in job j if $f_i \neq \bar{f}^j$.

However, this taxonomy introduces two imprecision: (1) It implicitly assumes weak correlation between these two mismatch types, which can cause a bias in estimation using equilibrium observational data; (2) it assumes that education attainment is fixed upon job placement. To account for (1), it is prudent to define the two mismatch types separately and evaluate the correlation between them. To account for (2), I use the O*NET occupation-level modal required education for entry to define over-education.

A reasonable way to define horizontal mismatch, adopted by Li and Lu (2023), is to look at mismatches in the underlying tasks requirement of a previous job. However, it is difficult to cleanly separate transient jobs from skill-forming jobs in observational data. Moreover, Lise and Postel-Vinay (2020) show that workers skill-sorting towards occupations that are desirable is a non-negligible mechanism in determining their career paths. This contaminates the identification of horizontal mismatch via task comparisons in observational data.

E.5 Horizontal-Mismatch Cutoff

The matched set for each degree-by-major group is defined by the share of FUS jobs it covers rather than by a fixed number of occupations. An earlier version of the measure counted a job as matched if it was among the 20 occupations most commonly held by FUS graduates with the same degree and major. A fixed count covers a different share of workers in every group: graduates of a major that leads to few occupations are almost always matched, and graduates of a major that spreads across many are often mismatched, so the FUS mismatch rate under a count rule mainly reflects how dispersed a major's graduates are. Dispersion also falls with the degree. The median matched set under the share rule contains 34 occupations for bachelor's groups, 22 for master's groups, and 11 for doctoral groups, so a 20-occupation cutoff was strict for bachelor's and lenient for doctoral graduates, and a lower doctoral mismatch rate under it would have been mechanical. The share rule fixes the matched set at half of each group's FUS jobs, with occupations tied at the margin included, and counts jobs over the whole career, so that an occupation reached through promotion counts as a match (Section 3.2). Figure 17 shows which jobs it classifies as mismatched.

E.6 Certification and On-the-Job Training

I hypothesize that the presence of institutional entry barriers, specifically professional certifications and on-the-job training (OJT) requirements, alters the signaling value of formal education. If a professional certification is a prerequisite for employment, the certification itself serves as a more informative signal of worker ability than the underlying degree. Conversely, in occupations where OJT is essential, educational attainment can be more informative to the employer. This assumes a structural link between formal education and a worker's latent productivity type (θ), where higher attainment signals a lower cost of training or a higher rate of skill acquisition during the OJT phase. While Figure 18 explores the relationship between occupational mismatch and the requirements for certification and OJT, the empirical patterns yield few deterministic implications.

E.7 Principal Components of O*NET Skills and Abilities

O*NET provides both level and importance scores for 87 distinct SAs associated with each occupation; for example, the occupation "Chief Executives" requires an average level score of 4.88 (on a 0–7 scale) for oral comprehension.

To interpret these factors, I extract the detailed O*NET descriptions for the 20 SAs with the highest loadings on each component. I then utilize a Large Language Model (LLM), Microsoft Copilot, to synthesize these descriptions into encompassing labels, weighting the semantic content by the magnitude of the respective loadings (Microsoft, 2025).

These components account for 73% of the variation in SAs with the next principal component accounting for roughly 3%.

By comparing the modal degree of citizens vs. non-citizens in an occupation, we can explore the relationship between these characteristics and the probability of mismatches. Figure 20 illustrates the relationship between these skill dimensions and occupational mismatch. Both over-education (OE) and under-education (UE) occur almost exclusively in occupations requiring above-average CIC. Furthermore, the direction of mismatch is highly correlated with PJC requirements: occupations with high PJC demands are significantly more likely to be characterized by under-education. In contrast, Figure 19 shows that the second principal component, AC, exhibits no discernible relationship with occupational mismatch.

E.8 Why Under-Education Is Not Classified

The analysis keeps every occupation, and a worker whose degree falls below the occupation's entry requirement is coded as not over-educated (Section 3.2). Under-education is not treated as a separate outcome because the literature suggests that it arises from a different process than over-education. In stable under-education matches, workers substitute unobserved existing human capital, such as innate ability, extensive tenure, or OJT, for formal academic credentials (). This asymmetry aligns with the classical human capital theory: over-educated workers accept lower-level positions to acquire general experience, whereas under-educated workers accumulate occupation-specific human capital that tightly binds them to their current roles (). Mobility patterns corroborate this distinction; Alba-Ramirez (1993) demonstrates that under-educated workers are far more likely to change firms within the same occupation rather than switch occupational categories en-

tirely. Consequently, under-education reflects a robust match where a worker's specific skill bundle satisfies occupational demands, largely eliminating the incentive for occupational switching.

Empirical training patterns provide further evidence of this structural divergence. As Hartog (2000) notes, under-educated workers are significantly more likely to receive job-specific training, whereas over-educated workers are less likely to require or receive it. Because this training constitutes a direct investment in occupation-specific human capital, the duration of an under-educated match positively correlates with specific training accumulation. This reinforces the premise that under-educated workers sort into higher-level roles requiring targeted learning, and that the resulting skill accumulation further entrenches them within their occupational track. Collectively, this evidence implies that under-educated workers operate under a distinct search and matching process governed by occupational lock and the relative permanence of specific human capital investments.

E.9 Ranking Details

Recovering $\tilde{V}$ requires that observed moves reflect choice rather than displacement, and that the mobility network is connected and stable over the window used to estimate it.

All regressions use the rankings as $\ln(\text{PageRank score})$. Log-transformation removes the mechanical dependence of raw scores on stratum size, converting a multiplicative scaling into an additive shift absorbed by stratum fixed effects; it compresses the heavy right skew that concentrated worker flows induce, preventing hub occupations from dominating; and, unlike ordinal ranks, it preserves the ratio comparisons that a recursively multiplicative ranking makes meaningful.[12]

Within a given degree-major pair, a worker observed leaving occupation a and entering occupation b is treated as having revealed that b was the most attractive reachable alternative at the time.

The generality statistic DUH is not a revealed-preference object and does not feed into the PageRank estimation above. This statistic is between 0 and 1, and it reflects the distribution of realized SOC/degree-major within a degree-major/SOC.

[12] Ordinal ranks discard magnitude and impose equal spacing between adjacent positions, neither of which the network construction supports. A level specification, by contrast, has no interpretable units. In the log-median-wage regressions the log specification yields a standard semi-elasticity: the percentage change in SOC median wage associated with a one percent increase in an occupation's or firm's utility.

F Selection on the Intensive Margin of the STEM-OPT Design

Table F1 collects the notation of Section 4.3, which this appendix uses throughout. Symbols used only in this appendix are defined where they first appear.

Table F1: Notation for the STEM-OPT Design

Symbol	Meaning
i	Worker
e	Terminal degree (master's or doctorate, where doctorate pools the PhD and non-licensure doctorates), the empirical counterpart of the education choice e in the model
$c(i)$	Year worker i completed the terminal degree (graduation cohort)
Post_c	$\mathbb{1}\{c \geq 2008\}$
STEM_i	Terminal-degree major is STEM
D_i	Treatment: $\text{STEM}_i \times \text{Post}_{c(i)}$
T_i, τ	Terminal-institution tier; τ indexes tiers
DUHMajor_i	Major generality
t	Years since the terminal degree
k	Job-quality outcome
Abroad_i	First post-degree job is outside the United States
S_{it}	Holds a US first job and is still US-resident at t; $S_{it}(1)$ and $S_{it}(0)$ with and without the extension
Y_{it}^k	Outcome k at t; $Y_{it}^k(1)$ and $Y_{it}^k(0)$ with and without the extension
ω_t	Share of treated stayers at t who would have stayed without the extension
δ_e	Extensive-margin difference-in-differences
δ_{et}^k	Intensive-margin difference-in-differences among stayers at t
ϕ_{ec}	Event-study coefficient for graduation cohort c

This appendix derives Equation (14) and states what Assumption 7 requires. The difference-in-differences in Equation (10) is estimated on workers who hold a US first job and are still US-resident t years after graduation. The extension changes who is in that sample, so the four group means average over different populations, and the estimate cannot be interpreted as an average effect on a fixed set of workers without further steps. Two distinct selection problems arise. The first is that who stays may change across the 2008 cutoff for reasons unrelated to the policy; this enters the identifying assumption. The second is that the policy itself changes who stays; this enters the interpretation of the estimand. I treat them in that order.

Throughout, I drop the outcome, degree, and time indices k, e, and t, and I ignore covariates. The derivation is therefore exact for the difference-in-differences of group means. With tier and major generality as controls, the same argument applies among workers with the same characteristics, and the regression estimate combines these comparisons with weights determined by OLS, so the decomposition holds for it approximately.

F.1 Setup

Index the four groups by (s, a), where $s = 1$ for STEM and $a = 1$ for graduation in or after 2008, and let $D_i = \mathbb{1}\{s = 1,\, a = 1\}$. Each worker has two potential outcomes, $Y_i(1)$ and $Y_i(0)$, with

and without the extension, and two potential stayer indicators, $S_i(1)$ and $S_i(0)$, equal to one if the worker would hold a US first job and still be US-resident at t under each regime. The data reveal $(Y_i, S_i) = (Y_i(1), S_i(1))$ in group $(1,1)$ and $(Y_i, S_i) = (Y_i(0), S_i(0))$ in the other three groups.

Each outcome is a benchmark attached to an occupation or an employer: the OEWS median wage, SOC PageRank, firm value, over-education, and horizontal mismatch. $Y_i(0)$ and $Y_i(1)$ are therefore defined for jobs held abroad as well as in the United States. The estimator never uses these outcomes: it averages only over stayers, and Assumption 7 is a statement about stayer means. Their role is interpretive. They make the mean outcome over all graduates in a group, stayers and leavers together, a well-defined object, which lets me write each stayer mean as a group mean plus a selection term in Appendix F.3, and they make that selection term observable in the untreated groups. For an outcome defined only for US jobs, the split would not exist, and Assumption 7 would have to be read directly as a restriction on stayer means, with the same content.

F.2 What the estimator recovers

Let $m_{sa} = E[Y_i \mid S_i = 1,\ (s,a)]$ denote the observed mean among stayers in group (s,a). The estimator is

$$\delta = (m_{11} - m_{10}) - (m_{01} - m_{00}). \tag{F.1}$$

Because no worker outside group $(1,1)$ is treated,

$$m_{sa} = E[Y_i(0) \mid S_i(0) = 1,\ (s,a)] \qquad \text{for } (s,a) \neq (1,1), \tag{F.2}$$

$$m_{11} = E[Y_i(1) \mid S_i(1) = 1,\ (1,1)]. \tag{F.3}$$

Define the unobserved counterfactual

$$m_{11}^0 \equiv E[Y_i(0) \mid S_i(0) = 1,\ (1,1)], \tag{F.4}$$

the mean outcome that STEM graduates in the post-period would have had without the extension, among those who would have stayed without it. Adding and subtracting m_{11}^0 gives the identity

$$\delta = \underbrace{(m_{11} - m_{11}^0)}_{\text{(A)}} + \underbrace{\left[(m_{11}^0 - m_{10}) - (m_{01} - m_{00})\right]}_{\text{(B)}}. \tag{F.5}$$

I define the counterfactual by conditioning on $S_i(0)$ rather than $S_i(1)$ for two reasons. First, by Equation (F.2), every term in (B) is then a mean of untreated outcomes among workers who would stay without the extension, so (B) is a difference-in-differences of like objects. Second, conditioning on $S_i(1) = 1$ would require the untreated outcome of workers who stay only because of the extension.

Assumption 7 states that (B) is zero. (A) is the object of interest and is taken up in Appendix F.4.

F.3 What (B) $= 0$ requires

The standard parallel-trends assumption restricts means of untreated outcomes over whole groups. Every term in (B) is instead a mean over a selected subset of each group: the workers who would stay without the extension. This subsection separates the two.

Step 1: a stayer mean is a group mean plus a selection term. Fix a group (s, a) and suppress it in the notation. Write $\rho = \Pr(S_i(0) = 1)$ for the untreated stay rate and

$$\nu^S = E[Y_i(0) \mid S_i(0) = 1], \qquad \nu^L = E[Y_i(0) \mid S_i(0) = 0], \qquad \nu = E[Y_i(0)] \tag{F.6}$$

for the mean untreated outcome among would-be stayers, among would-be leavers, and over the whole group. By the law of total expectation,

$$\nu = \rho\, \nu^S + (1 - \rho)\, \nu^L. \tag{F.7}$$

Subtracting this from ν^S,

$$\nu^S - \nu = \nu^S - \rho\, \nu^S - (1 - \rho)\, \nu^L = (1 - \rho)\, (\nu^S - \nu^L). \tag{F.8}$$

Define the selection term

$$\sigma \equiv \nu^S - \nu = (1 - \rho)\, \Delta, \qquad \Delta \equiv \nu^S - \nu^L. \tag{F.9}$$

The stayer mean is then

$$E[Y_i(0) \mid S_i(0) = 1] = \nu + \sigma. \tag{F.10}$$

σ is how far would-be stayers sit above the average graduate in their group. It is the product of two things: Δ, how different stayers are from leavers, and $1 - \rho$, how many leavers there are. If nobody leaves ($\rho = 1$) or stayers and leavers are alike ($\Delta = 0$), the stayer mean equals the group mean.[13]

Step 2: (B) splits into outcome trends and selection trends. Applying Equation (F.10) to each of the four terms in (B), with ν_{sa} and σ_{sa} indexing groups,

$$\text{(B)} = \underbrace{[(\nu_{11} - \nu_{10}) - (\nu_{01} - \nu_{00})]}_{\text{(B1): parallel trends in outcomes}} + \underbrace{[(\sigma_{11} - \sigma_{10}) - (\sigma_{01} - \sigma_{00})]}_{\text{(B2): parallel trends in selection}}. \tag{F.11}$$

(B1) is the standard parallel-trends condition, stated over whole groups. (B2) is new. Assumption 7 is the statement that (B1) + (B2) $= 0$. The natural sufficient condition is that each is zero separately.

[13] Equivalently, $\sigma = \text{Cov}(Y_i(0), S_i(0))/\rho$, using $E[Y \mid S = 1] = E[YS]/\Pr(S = 1)$ and $E[YS] = E[Y]\Pr(S = 1) + \text{Cov}(Y, S)$, together with $\text{Cov}(Y, S) = \rho(1 - \rho)\Delta$ for binary S. The Δ form is easier to interpret.

Offsetting violations would also satisfy the assumption, but I do not rely on them.

Step 3: necessary conditions for (B2) $= 0$. (B2) is zero when the change in σ across the 2008 cutoff is the same in STEM as in non-STEM:

$$(1-\rho_{11})\Delta_{11} - (1-\rho_{10})\Delta_{10} = (1-\rho_{01})\Delta_{01} - (1-\rho_{00})\Delta_{00}. \tag{F.12}$$

Three observations clarify what this does and does not require.

First, it does not require that stayers be the same kind of worker before and after 2008. σ may change over time within each field, provided it changes by the same amount in both. Only the difference across fields has to cancel, exactly as for outcomes in (B1).

Second, σ has two channels: how many workers would stay, ρ, and how selected the stayers are, Δ. Within field s, the change across the cutoff separates exactly into the two:

$$\sigma_{s1} - \sigma_{s0} = -\Delta_{s0}\,(\rho_{s1} - \rho_{s0}) + (1-\rho_{s1})\,(\Delta_{s1} - \Delta_{s0}). \tag{F.13}$$

(B2) can fail through either channel. It fails through Δ if, for example, the 2008–2009 recession pushed less productive non-STEM graduates out of the United States while leaving the composition of STEM stayers unchanged. Non-STEM stayer quality would then rise through selection alone, making (B2) negative even when (B1) holds.

Third, parallel trends in stay rates is not enough. Suppose untreated stay rates change by the same amount r in both fields, $\rho_{11} - \rho_{10} = \rho_{01} - \rho_{00} = r$, and the stayer–leaver gap is stable over time within each field, $\Delta_{s1} = \Delta_{s0} = \Delta_s$. By Equation (F.13), (B2) $= -(\Delta_1 - \Delta_0)\,r$, which is zero only if staying is equally selective in the two fields, $\Delta_1 = \Delta_0$. If staying in the United States is more selective among STEM graduates than among non-STEM graduates, a common shift in stay rates alone violates (B2).

Step 4: sufficient conditions. Each of the following makes (B2) zero:

(i) Staying is unrelated to untreated outcomes in every group: $\Delta_{sa} = 0$. This is the case in which conditioning on stayers is innocuous; I do not expect it to hold.

(ii) Selection into staying is stable over time within each field: $\rho_{s1} = \rho_{s0}$ and $\Delta_{s1} = \Delta_{s0}$ for $s \in \{0, 1\}$.

(iii) Staying is equally selective in both fields and stable over time ($\Delta_{sa} = \Delta$), and untreated stay rates follow parallel trends.

Condition (iii) is the closest to what the design already assumes elsewhere, and it shows what Assumption 7 adds beyond that assumption.

Relation to the extensive-margin analysis. The extensive-margin difference-in-differences is identified under parallel trends in untreated stay rates, $(\rho_{11} - \rho_{10}) = (\rho_{01} - \rho_{00})$. Two implications follow.

First, the extensive-margin estimate does not show that (B) is nonzero. That estimate measures workers with $S_i(1) = 1$ and $S_i(0) = 0$ in group $(1, 1)$, stayers induced by the extension. Only

$S_i(0)$ appears in (B). Treatment-induced staying enters (A), not (B). Reading the extensive-margin estimate as causal already assumes that untreated stay rates trend in parallel, so the same data cannot also be read as evidence of differential untreated selection.

Second, the extensive-margin assumption does not imply (B2). By Step 3, parallel stay rates deliver (B2) $= 0$ only when staying is equally selective in the two fields and stable over time. Assumption 7 restricts the composition of stayers, which the extensive-margin assumption leaves unrestricted.

Step 5: what can be checked. Because outcomes are defined for jobs held abroad, ρ_{sa}, Δ_{sa}, and therefore σ_{sa} are observed in the three untreated groups, and across graduation cohorts within the pre-period. Stability of Δ within the pre-period, and similar Δ across fields, are informative about condition (iii). An event study of stayer means over pre-period cohorts tests (B1) and (B2) jointly. The comparison with US natives, who face the same field-specific shocks to staying but no OPT, is informative about (B2) to the extent that such shocks affect CtS and native selection similarly. None of these checks the counterfactual for group $(1, 1)$ directly, which is unobserved by construction. Two caveats apply to the observed Δ: leavers' outcomes rely on profiles that may be updated incompletely after departure, and workers without an observed job at t contribute to ρ but not to Δ.

F.4 From δ to an effect on a fixed group

Under Assumption 7, Equation (F.5) reduces to

$$\delta = E[Y_i(1) \mid S_i(1) = 1,\ D_i = 1] - E[Y_i(0) \mid S_i(0) = 1,\ D_i = 1]. \tag{F.14}$$

The two terms average over different workers: those who stay with the extension and those who would stay without it. To relate them, classify treated workers by $(S_i(0), S_i(1))$:

Type	$S_i(0)$	$S_i(1)$	In the estimation sample?
Always-stayer	1	1	Yes
Marginal stayer	0	1	Yes, because of the extension
Pushed out	1	0	No
Never-stayer	0	0	No

Assumption 6 rules out the pushed-out type. The workers with $S_i(1) = 1$ are then the always-stayers and the marginal stayers, and the workers with $S_i(0) = 1$ are exactly the always-stayers. Let

$$\omega \equiv \Pr(S_i(0) = 1 \mid S_i(1) = 1,\ D_i = 1) \tag{F.15}$$

denote the share of treated stayers who would have stayed without the extension, and write:

$$\bar{y}_A^1 = E[Y_i(1) \mid \text{always-stayer},\ D_i = 1], \tag{F.16}$$

$$\bar{y}_A^0 = E[Y_i(0) \mid \text{always-stayer},\ D_i = 1], \tag{F.17}$$

$$\bar{y}_M^1 = E[Y_i(1) \mid \text{marginal},\ D_i = 1]. \tag{F.18}$$

By the law of total expectation over the two types in the sample,

$$E[Y_i(1) \mid S_i(1) = 1,\ D_i = 1] = \omega\, \bar{y}_A^1 + (1-\omega)\, \bar{y}_M^1. \tag{F.19}$$

Under Assumption 6, the second term of Equation (F.14) is $\bar{y}_A^0$. Substituting and adding and subtracting $\bar{y}_A^1$,

$$\delta = \omega\, \bar{y}_A^1 + (1-\omega)\, \bar{y}_M^1 - \bar{y}_A^0 = (\bar{y}_A^1 - \bar{y}_A^0) + (1-\omega)\,(\bar{y}_M^1 - \bar{y}_A^1)\,, \tag{F.20}$$

which is Equation (14):

$$\delta = \underbrace{E[Y_i(1) - Y_i(0) \mid \text{always-stayer},\ D_i = 1]}_{\text{effect on always-stayers}} \tag{F.21}$$

$$+\ \underbrace{(1-\omega)\Big(E[Y_i(1) \mid \text{marginal},\ D_i = 1] - E[Y_i(1) \mid \text{always-stayer},\ D_i = 1]\Big)}_{\text{composition}}. \tag{F.22}$$

The first term is an average effect on a fixed group. The second compares the two types at the outcomes they hold under the extension, the US jobs they hold, so it never uses the US outcome that marginal stayers would have had without the extension. It is zero only if marginal stayers hold jobs of the same quality as always-stayers. Without Assumption 6, the workers with $S_i(0) = 1$ would include the pushed-out type, whose treated outcome is never observed, and Equation (F.22) would acquire a third term of unknown sign.

A numerical illustration. Suppose 80 percent of STEM post-period stayers would have stayed anyway ($\omega = 0.8$), always-stayers score $\bar{y}_A^0 = 9$ without the extension and $\bar{y}_A^1 = 10$ with it, and marginal stayers score $\bar{y}_M^1 = 6$. The effect on always-stayers is 1, but $\delta = 0.8(10) + 0.2(6) - 9 = 0.2$, and Equation (F.22) gives $1 + 0.2(6 - 10) = 0.2$. When marginal stayers hold lower-quality jobs than always-stayers, a positive δ understates the effect on always-stayers.

F.5 Identifying ω and bounding the effect on always-stayers

Because $\{S_i(0) = 1\} \subseteq \{S_i(1) = 1\}$ under Assumption 6,

$$\omega = \frac{\Pr(S_i(0) = 1 \mid D_i = 1)}{\Pr(S_i(1) = 1 \mid D_i = 1)}. \tag{F.23}$$

The denominator is the observed stayer share in group $(1, 1)$. The numerator is identified by a difference-in-differences on stayer shares under parallel trends in untreated stay rates, the same assumption that identifies the extensive-margin estimate. Given ω, trimming the upper and lower $1-\omega$ tails of the treated stayers' outcome distribution bounds $\bar{y}_A^1$ in the spirit of Lee (2009). Since $\bar{y}_A^0 = m_{11}^0$ is identified under Assumption 7, this bounds the effect on always-stayers. I leave these bounds to future drafts.

F.6 Retention estimates

Table F2 reports the retention difference-in-differences that enter the bound on composition in Section 7.2, for each degree and post-period window. Panel A is the Kaplan–Meier estimate described in Section 4.3; its Greenwood standard errors treat graduates as independent. Panel B is the linear probability version of Equation (11), which drops graduates whose profile ends before t without an exit, with standard errors clustered by major. Where the two can be compared, the clustered standard errors are up to about one and a half times as large.

Table F2: STEM-OPT Extension and Retention in the United States

	Master's			Doctorate		
Post cohorts	2008–2021	2008–2011	2012–2015	2008–2021	2008–2011	2012–2015
Panel A. Kaplan–Meier, Greenwood standard errors						
$t = 1$	0.15	0.07	0.14	0.01	-0.05	-0.03
	(0.05)	(0.06)	(0.05)	(0.09)	(0.12)	(0.10)
$t = 2$	1.09	0.29	1.33	0.30	-0.28	0.30
	(0.11)	(0.15)	(0.14)	(0.21)	(0.28)	(0.25)
$t = 3$	1.23	0.06	1.45	0.09	-1.06	0.12
	(0.16)	(0.21)	(0.19)	(0.33)	(0.43)	(0.39)
$t = 5$	0.99	-0.09	1.28	0.32	-0.67	0.43
	(0.24)	(0.30)	(0.28)	(0.47)	(0.61)	(0.54)
$t = 8$	2.08	0.95	2.28	0.65	-0.61	0.55
	(0.33)	(0.40)	(0.37)	(0.65)	(0.80)	(0.72)
$t = 10$	1.86	0.61	1.96	0.94	-0.54	0.75
	(0.37)	(0.44)	(0.41)	(0.72)	(0.88)	(0.79)
Panel B. Linear probability, standard errors clustered by major						
$t = 1$	0.17	0.09	0.16	0.01	-0.05	-0.03
	(0.04)	(0.06)	(0.05)	(0.08)	(0.11)	(0.10)
$t = 2$	1.16	0.35	1.42	0.31	-0.29	0.31
	(0.18)	(0.19)	(0.23)	(0.24)	(0.29)	(0.28)
$t = 3$	1.36	0.17	1.64	0.12	-1.09	0.14
	(0.27)	(0.27)	(0.34)	(0.34)	(0.36)	(0.40)
$t = 5$	1.44	0.13	1.62	0.25	-0.67	0.48
	(0.37)	(0.37)	(0.40)	(0.51)	(0.59)	(0.56)
$t = 8$	2.43	1.30	2.79	0.30	-0.52	0.60
	(0.39)	(0.46)	(0.42)	(0.61)	(0.67)	(0.70)
$t = 10$	1.75	1.01	2.38	-0.17	-0.53	0.22
	(0.46)	(0.49)	(0.53)	(0.76)	(0.82)	(0.89)
Graduates	479,668	115,319	163,398	123,678	32,460	44,113
Majors	552	522	526	409	362	390

Notes: Difference-in-differences, STEM versus non-STEM CtS graduates whose first job is in the United States, post- versus pre-2008 graduation cohorts (2003–2006; 2007 excluded), in the share that has not left the United States by t, in percentage points. Doctorate is all doctoral degrees excluding professional degrees. Panel A: Kaplan–Meier estimates; a graduate whose profile ends counts as at risk until the last year observed. Panel B: linear probability version of Equation (11), which drops graduates whose profile ends before t without an exit. Graduates: graduates in the four groups at $t = 0$. Majors: terminal-degree majors at $t = 1$. At $t = 10$, the 2012–2015 window contains only the 2012–2014 cohorts.

G School Selectivity Tier Construction

The analysis assigns each degree-granting institution to one of nine selectivity tiers (1 = least selective, 9 = most selective) drawn from two sources joined by institution type. U.S. institutions are tiered using the admissions-based index constructed from undergraduate admission selectivity data maintained by the Integrated Postsecondary Education Data System (IPEDS). Non-U.S. institutions are tiered by a nearest-U.S.-comparator procedure that maps onto the domestic selectivity tiers using the QS World University Rankings (). In the panel, the two sources are merged via a conditional prioritization on institutional identifier, with a source indicator column recording which authority assigned each observation's tier.

U.S. Tier: Admissions Index The domestic tier is drawn from a composite admissions index constructed from institutional undergraduate admission characteristics—acceptance rate and SAT/ACT score distributions—for all U.S. degree-granting institutions. Schools are sorted on the composite and divided into nine tiers, with tier 9 corresponding to the most selective institutions. The index is keyed by a standard institutional identifier and joined to the education records through a school-name crosswalk maintained separately.

The QS rankings are deliberately used only to assign tiers for non-U.S. schools. Empirically, all 204 ranked U.S. institutions in QS ranking already carry an admissions-based tier, so global rankings would add no domestic coverage. More importantly, global rankings exhibit a systematic research-prestige bias that demotes highly selective liberal arts colleges relative to their true admissions standing (e.g., Swarthmore College, Williams College, and Lehigh University each fall two to four tiers lower on global rankings than on the admissions index).

Ranking Crosswalk and Bridge Sample Extending the tier to international schools first requires identifying which global ranking entries correspond to which domestic institutions, so that the U.S. bridge sample can be constructed. This crosswalk is built using fuzzy name matching against the full institutional characteristics directory (≈6,500 institutions). A match is accepted automatically only under high confidence: exact name agreement, a high similarity ratio with a margin over the runner-up, or agreement between the fuzzy match and an independent institutional data chain.

Ambiguous cases are exported for manual review; 41 institutions required hand curation (verified by state and city), including several flagship universities whose common names the matcher could not surface unambiguously (e.g., University of Virginia, University of Washington—Seattle, and Ohio State University). University system aggregate entries with no single identifier are excluded. Of 228 U.S. entries in the ranking file, 227 are resolved to a distinct identifier, yielding 204 bridge institutions that carry both a global average rank and a domestic admissions tier.

International Tier: Nearest-U.S.-Comparator For each non-U.S. institution in the crosswalk, the selectivity tier is assigned by identifying the U.S. bridge school whose global average rank is closest in absolute distance and inheriting that school's admissions tier:

$$\hat{\tau}_i = \tau_{j^*(i)}, \quad j^*(i) = \arg\min_{j \in \mathcal{B}} \left| R_i - R_j \right|, \tag{G.1}$$

where R_i is the global average rank of international school i, $\mathcal{B}$ is the set of 204 U.S. bridge schools, and τ_j is the domestic admissions tier of bridge school j. The matching is implemented using a nearest-neighbor matching strategy on the sorted average rank.

The procedure yields tiers for 3,964 non-U.S. institutions across 116 countries. Sanity checks on named institutions are consistent with expectations: Oxford, Cambridge, ETH Zürich, and the National University of Singapore map to tier 9; the University of Toronto, Tsinghua University, University of Tokyo, and the University of Melbourne to tier 8; IIT Delhi to tier 7.

Panel Integration and Country Assignment The panel joins domestic and international tiers by institutional identifier via a prioritization of the domestic tier over the international tier. Because the international file contains only non-U.S. institutions, no school can be claimed by both sources.

Whether a worker's degree is domestic or foreign—a key stratum in the analysis—is determined by a school-level country field. The resolver applies a four-step precedence rule:

1. If the institution has a domestic institutional identifier, classify as U.S.
2. If it appears in the global ranking crosswalk with a non-U.S. country, classify as non-U.S.
3. Otherwise, take a modal vote over all raw school location records for that school identifier (≈590 million observations), accepting the modal value if it clears a minimum purity threshold.
4. If the modal country lies outside North America, classify as non-U.S.

Schools that remain unresolved after all four steps are excluded from stratum-specific analyses. Validation against the two authoritative sources yields a 100 percent agreement rate for the domestic subsample and 95.7 percent for the global ranking non-U.S. subsample.

Limitations Several caveats apply to the international tier. First, global rankings score only approximately 1,500 institutions worldwide; unranked foreign schools receive no tier assignment. Second, the nearest-U.S.-comparator procedure assumes that global rank is comparable across countries as a signal of institutional reputation—an assumption that holds approximately for research-intensive universities but may be less accurate in national systems where ranking coverage is sparse or systematically biased. Third, the name match is the sole authority for non-U.S. schools, so match errors propagate directly into tier assignments. Fourth, the modal-country vote is approximately 96 percent reliable against the authoritative sources; the residual uncertainty is concentrated among small or obscure institutions with few records.

G.1 From Admissions Scores to the Continuous Education Signal.

This subsection details the reputation score underlying the tiered signal used in Section 3.1. To capture school reputation, I assign each U.S. degree-granting institution a composite score based on the following specification:

$$\text{score} = 0.65 \times \text{Avg. Std. Midpoint SAT/ACT Scores} + 0.35 \times \text{Std. Rejection Rate}. \quad \text{(G.2)}$$

This score serves as a composite measure of institutional selectivity, proxying the cost of attaining an undergraduate degree at a given institution. Due to program and degree availability, these scores are interpreted relative to their respective degree levels. Nevertheless, estimating signal strength using just undergraduate admission statistics directly requires a strong *Independence of Irrelevant Alternatives* (IIA) assumption across degrees. Because the scores are derived from undergraduate admission data, institutions providing other degree levels are considered subsets of the universe used to establish the score's induced order. Under IIA, this ordering can be applied within each degree level to compute a standardized reputation score for worker i's education attainment a_i, denoted as $\tilde{s}_{i,a} = \frac{score_i - \bar{score}_a}{sd(score_a)}$. The latent education signal is then constructed as follows in Equation (G.3):

$$\hat{s}_i = \hat{\beta}_0 + \sum_{a=\text{Master}}^{\text{PhD}} \hat{\delta}_a \cdot \mathbb{1}\{Deg_i = a\} + \sum_{a=\text{Bachelor}}^{\text{PhD}} \hat{\beta}_a \cdot \mathbb{1}\{Deg_i = a\} \cdot \tilde{s}_{i,a}. \tag{G.3}$$

Recognizing that the IIA assumption may be difficult to satisfy when comparing institutions in specialized fields, I implement a tiered classification of institutions instead of using the raw scores. This discretization ensures that the induced order is robust to minor fluctuations in selectivity and mitigates the potential for IIA violations when the choice set includes institutions with highly similar signaling values. While the constructed $\hat{s}$ remains a coarse measure, it more flexibly accounts for the heterogeneous cost structure $c(e)$ by allowing different degree levels to yield equivalent education signals via the coefficients β_a, thereby avoiding the biases inherent in a strictly monotonic ordering of degree-institution pairs. "'

H Welfare Implications and Limitations

The friction d distorts the allocation of $\bar{\theta}$ immigrants away from the contract they would obtain absent the friction, making explicit the welfare loss to the market. The logic is that of Hsieh et al. (2019), who attribute between 20 and 40 percent of US growth in aggregate market output per person from 1960 to 2010 to the decline in frictions that women and Black men faced in the labor market and in accumulating human capital. Without d, a type-$\bar{\theta}$ immigrant faces the natives' equilibrium schedule and, like a native of the same type, chooses $\bar{e}$ and is employed at $W(\bar{e})$. I use this contract as the benchmark rather than the full-information choice $e^*(\bar{\theta}) \in \underset{e}{argmax}\ F(\bar{\theta}, e) - C(\bar{\theta}, e)$: natives already over-invest in education to separate, and that distortion is not caused by d. Introducing d removes the jobs on the equilibrium wage schedule between $W(\tilde{e})$ and $W(\bar{e})$ from the choice set of a type-$\bar{\theta}$ immigrant, who keeps the credential $\bar{e}$ and is placed at $\tilde{e}$ (Figure 1). Because their education is unchanged, the distortion falls entirely on the job. The per-worker welfare loss, the induced reduction in output, is

$$L(\bar{\theta}) = F\left(\bar{\theta}, \bar{e}\right) - \left[F\left(\bar{\theta}, \bar{e}\right) - d\right] = d = W(\bar{e}) - W(\tilde{e}). \tag{H.1}$$

The firm loses output d on the immigrant, and the immigrant bears it as a wage sacrifice $W(\bar{e}) - W(\tilde{e})$ in an occupation for which less education is typical. Empirically, the loss is concentrated where the over-education gap is largest: among master's holders in both pathways and foreign-trained professional-degree holders (Section 5). Because the wage schedule is convex, the same d moves a doctorate holder a shorter distance down the schedule: the output lost per worker is the same at every degree, but the mismatch it produces is smaller at the top.

Crucially, the magnitude, and even the interpretation, of $L(\bar{\theta})$ depends on what d represents. If d is a real resource cost of sponsorship, such as legal and processing costs, then not hiring the immigrant at $W(\bar{e})$ frees up resources the firm would have spent, and the resulting allocation is efficient given that cost; L then measures what the sponsorship system costs, which policy can reduce. If d is a transfer, such as filing fees paid to the government, it acts as a tax and L is deadweight loss. If d is taste-based discrimination, the firm saves nothing by displacing the worker and the match surplus is simply destroyed: pure allocative deadweight with no offsetting benefit. If d instead reflects statistical discrimination, the distortion is transient: as employer-learning erodes d, the worker's allocation converges toward $\bar{e}$, so the present value of L is smaller and front-loaded. The dynamics of $\hat{d}$ documented in Section 6 therefore bear on the size of the welfare loss, not merely on its source. A persistent $\hat{d}$ implies a durable allocative cost, whereas an attenuating one implies a loss that employer learning partly undoes.

The STEM-OPT extension offers the closest evidence on how the loss responds when d falls. The extension lowered the cost of hiring a STEM graduate without changing who the graduate is, and Section 7 estimates two responses that map onto Equation (H.1). The extensive-margin estimate δ_e measures how many graduates, who would otherwise have taken a first job abroad instead, take one in the US. The intensive-margin estimates δ^k_{et} measure how the jobs held by those who remain change, which bears on the first term of $L(\bar{\theta})$. Because Equation (14) interprets δ^k_{et} as an effect on the population that remains rather than on a fixed set of workers, these estimates provide evidence on the direction in which the allocation moves when d falls, not a measure of

$L(\bar{\theta})$ for any one worker.

The aggregate welfare cost scales the per-worker loss by the mass of displaced high-type immigrants,

$$\mathcal{L} = \int_{\Theta} s(\theta)\, L(\theta)\, dG(\theta), \tag{H.2}$$

where $s(\theta)$ is the share of type θ displaced off the natives' equilibrium schedule and G is the type distribution. Evaluating $\mathcal{L}$ requires the displaced share $s(\theta)$ and a $d = 0$ counterfactual, which in turn rest on the estimates of d in Section 5.2 and the stayer restriction discussed in Section 4.2. Equation (H.2) is a partial-equilibrium object: it counts the allocative loss borne by displaced immigrants, holding fixed the native equilibrium schedule, firm entry, and output prices. A full accounting would net it against the general-equilibrium consequences of high-skilled inflows, which Khanna and Morales (2024) find are substantial and mixed in sign: the same inflows that raise firm profits and lower output prices depress the wages and employment of competing natives. Removing d would move high-type immigrants up into occupations where they compete more directly with natives, so an efficiency gain in the sense of Equation (H.1) need not be a welfare gain once those margins are counted. The comparison with US natives in Section 7 gives one direct check on this margin for a single policy: it measures whether the extension moved the job outcomes of STEM natives relative to non-STEM natives. Calibrated models of vertical mismatch put its output cost at 0.26 percent of US output (Garibaldi et al., 2025a) and at 3 to 4 percent of output on average across 17 OECD economies (Garibaldi et al., 2025b); both are calibrated to the workforce as a whole. I therefore treat Equation (H.2) as the framework that the estimated friction populates, and reserve its full quantification, together with the associated immigration-policy counterfactuals, for future work.

Limitations. Several features of the data limit what the estimates can show. First, every outcome is an occupation- or employer-level benchmark (the OEWS median wage, SOC PageRank, firm value, and the two mismatch measures) rather than the worker's realized wage, so the estimates describe where immigrants are placed, not what they are paid within a placement. Second, the dynamic estimates in Section 6 and the intensive-margin estimates in Section 7 condition on staying in the United States, so both describe the population that stays rather than all entrants. Third, departures are observed only if workers update their profiles after leaving, so exits are measured with error. Fourth, the STEM indicator marks majors that appear on the DHS STEM list in any year through 2024, so the few majors added after 2008 are coded as STEM for cohorts that were not yet eligible, which biases δ_e toward zero.

I Miscellaneous Tables