The Labor Hub

Job Market Paper Series No. 11

Advisor: Arnaud Maurel

VISA POLICY, GLOBAL EDUCATION AND LABOR MARKETS

Zhangchi Ma and Yangfan Wang

Visa Policy, Global Education and Labor Markets *

Zhangchi Ma[†] Yangfan Wang[‡]

Job Market Paper

November 9, 2026

Click here for the latest version

Abstract

Visa policy shapes the supply of foreign skilled labor through international students' study and work decisions. Exploiting the UK's 2012 visa reform, which abolished an unconditional two-year post-study work visa, we use LinkedIn profile data to quantify the effects of reduced post-graduation job access on international students' enrollment and employment. Focusing on Indian students, the largest group with LinkedIn profiles, we find that the reform deterred enrollment in UK master's programs, even among students who would have qualified for high-skill work. To connect these enrollment responses to labor supply, we develop a sequential model of education and post-graduation work. The model estimates indicate that increased job-matching frictions account for two-thirds of the decline in foreign labor supply in occupations requiring a bachelor's degree. By contrast, reduced student enrollment is the main driver of the corresponding decline in occupations requiring an advanced degree. These results highlight how a reform targeting low-skill employment can reduce high-skill labor supply through both enrollment and employment responses.

Keywords: Immigration Policy, Global Education, High-Skill Immigrants

JEL Codes: F22, J61, I23

*We are deeply grateful to Arnaud Maurel, Peter Arcidiacono, Michael Dinerstein, Matt Masten, Pengpeng Xiao, and Xavier D'Haultfoeuille for their invaluable guidance and mentorship. We also thank Tirza Angerhofer, Jason Baron, Luca Braghieri, Taha Choukhmane, Michael Clemens, Zihan Hu, Kyle Jurado, Faqiang Li, Chenran Liu, Steven Medema, Michael Pollmann, Adam Rosen, Philipp Sadowski, Haoqi Tong, Daniel Waldinger, Xi Weng, Wei You, Xiaohan Zhong, and many seminar participants at Duke University, Peking University, Renmin University of China, and Tsinghua University for useful comments and suggestions. We are grateful to Yizhi Su for her outstanding research assistance.

[†]Duke University. Email: zhangchi.ma@duke.edu

[‡]Duke University. Email: yangfan.wang@duke.edu

1 Introduction

Between 2004 and 2022, the number of international students in higher education from low- and middle-income countries increased from 1.5 million to nearly 4.6 million (UNESCO Institute for Statistics, 2025), and this growth is largely attributable to master's students. For example, in the United States, the largest host country for international students, the number of master's students from developing countries nearly tripled between 2004 and 2022.[1] For these students, a foreign master's degree typically serves as both a human capital investment and a strategic pathway to high-wage labor markets and permanent migration (Bound et al., 2021; Khanna, 2025). For host countries, international graduate education is a major export industry that cross-subsidizes domestic institutions (Shih, 2017; Bound et al., 2020; Zhu, 2024) and serves as a primary pipeline for high-skill labor (Bound et al., 2015; Ruiz and Budiman, 2018). Aiming to retain tuition revenue and global talent while restricting overall immigration, governments have implemented post-study work visa policies like the OPT and H-1B regimes in the United States. These policy instruments are currently the subject of active debate and revision across the world.

The United Kingdom's 2012 abolition of the Post-Study Work (PSW) visa represents a particularly consequential policy shock. Until 2012, international graduates of any UK higher-education institution were eligible for the Tier 1 PSW visa, which granted two years of unconditional residence with full labor-market access. However, the 2012 reform eliminated this route. Recent graduates had four months to secure a job offer from a government-certified sponsoring employer for a "graduate-level" position paying above a salary threshold; otherwise, they had to leave the country. While curbing visa abuse in low-skill sectors was a major objective of the reform, the government also explicitly committed to continuing to attract international students and retain the "best and brightest" in skilled employment.[2]

To study the impacts of the reform, we use LinkedIn profile data to construct a novel longitudinal dataset tracking the education and career paths of 177,683 Indian students who pursued master's degrees abroad between 2008 and 2013. India is one of the largest source countries for international master's students, and our sample covers approximately

[1] Authors' calculations from administrative F-1 student records in the Student and Exchange Visitor Information System (SEVIS), obtained under the Freedom of Information Act (2023-ICFO-43657). A student is counted in fiscal year t if her master's program spans any part of that year (October of year $t-1$ through September of year t) and she is not on post-completion OPT. We define developing countries as economies outside the current World Bank high-income group and apply this classification consistently to both years. Net education export revenue is the tuition and fees certified on each student's Form I-20, net of institutional funding, summed over enrolled students.

[2] See Home Office, *Final Impact Assessment: Reform of the Points Based Student Immigration System* (2011), pp. 1 and 6–11; and Theresa May, "Student Visas," House of Commons Debates, 22 March 2011, vol. 525.

65 percent of all Indian international master's students during this period.[3] Descriptively, UK master's enrollment among Indian students fell by 53% between 2010 and 2013, while their high-skill employment in the UK fell by 77%.

To identify the causal effect of the reform-induced reduction in employment access on students' master's enrollment, we exploit variation in exposure to the visa reform across undergraduate institutions, measured by each institution's pre-policy UK stay rate. Our difference-in-differences estimates reveal a substantial deterrent effect: the policy led to a 50% larger decline in UK master's enrollment among students from undergraduate institutions in the top decile of exposure relative to that among students from institutions in the bottom decile. Notably, our heterogeneity analysis suggests that the reform also discouraged prospective high-skill workers from enrolling in UK master's programs.

We then develop a two-stage discrete choice model to disentangle the mechanisms driving the declines in master's enrollment and labor supply. In the first stage, students choose their master's destinations around the world, taking into account the future value of job opportunities after graduation. Then, in the second stage, students choose whether to stay in the host country, return to their home country, move to a third country, or continue their education. In particular, we do not assume that all post-graduation options are available to every student, and we explicitly model job choice sets as stochastic across individuals. We establish a constructive identification strategy for unobserved stochastic choice sets by exploiting the joint response of master's enrollment and post-graduation work decisions to the policy.

Our model estimates imply that the 2012 abolition of the PSW visa reduced UK master's enrollment by 53%, corresponding to an annual tuition revenue loss of £74 million from Indian students alone, or 2.2% of total tuition fee revenue from international students.[4] The reform also shrank the labor supply of Indian master's graduates in two types of high-skill jobs—those requiring bachelor's-level skills and those requiring advanced-degree-level skills—by 80% and 61%, respectively. Using an approach similar to that of Clemens et al. (2025), we estimate that these losses could reduce UK GDP by a cumulative £2.5–5.0 billion over ten years, equivalent to 12–25% of annual UK exports to India. Consistent with our reduced-form evidence, these findings suggest that the reform weakened the UK's ability to attract and retain skilled international graduates, despite the government's stated goal of

[3]China is the largest source country for international students, but LinkedIn's coverage of the Chinese labor market was small during the period we focus on. Before launching its Chinese-language version in February 2014, LinkedIn had approximately 4 million members in China, compared with 24 million in India. It also discontinued its localized social networking service in China in 2021 and closed its local job-search app in 2023. Including Chinese students would therefore substantially understate the share returning to China.

[4]See Table 7 of the Department for Education's *UK Revenue from Education Related Exports and Transnational Education Activity 2010–2014* (2017).

retaining them.

Furthermore, our decomposition reveals that the channels driving the decline in labor supply differ substantially across the two high-skill job types. There are two channels through which the reform can operate. First, it *directly* reduces job-matching probabilities through both the supply and demand sides of the labor market: students' job-search window shrinks from 24 months to at most 4 months, while employers must bear the administrative costs of visa sponsorship and pay salaries above specified minimum thresholds. Second, it *indirectly* reduces skilled labor supply by lowering master's enrollment, thereby shrinking the pool of prospective foreign skilled workers. For bachelor-degree-level jobs, the direct reduction in job access accounts for 67% of the decline in labor supply. By contrast, for advanced-degree-level jobs, the direct effect accounts for only one-third of the decline, while the indirect effect through students' education choices accounts for the remaining two-thirds. Students who might qualify for these jobs face uncertainty about the type of job they will obtain after graduation. Reduced employment opportunities in bachelor-degree-level jobs therefore lower the expected value of a UK master's degree even for potential advanced-degree-level workers. As these students choose master's programs in other countries, the pool of candidates for advanced-degree-level jobs in the UK shrinks.

Finally, we simulate a counterfactual policy in which the UK government conditions PSW visa eligibility on undergraduate characteristics rather than abolishing the visa for all graduates. Conditioning PSW eligibility on undergraduate characteristics rather than the UK master's institution attended can mitigate concerns about unequal access to costly UK programs and widening resource disparities across UK universities. We find that, compared with abolishing PSW eligibility for all graduates, retaining eligibility for STEM students entails smaller percentage-point reductions in enrollment and high-skill labor supply per percentage-point reduction in low-skill employment. However, further screening based on undergraduate institution tier yields limited efficiency gains, except for retaining advanced-degree-level labor.

Our findings carry two main implications for the design of post-study immigration policy. First, eliminating visa pathways associated with lower-skill employment, particularly post-study work visas that are not tied to employer sponsorship, can substantially increase search and hiring frictions in the high-skill sector. Flexible work visas give qualified graduates time to search for high-skill jobs without the pressure of leaving the country immediately. They can also reduce the administrative frictions employers face when hiring foreign graduates, preventing losses in employer–employee match surplus. Immigration policy should therefore preserve sufficient time and flexibility for job search and hiring to support the retention of high-skill talent.

Second, our findings emphasize the sensitivity of international education choices to future job prospects upon graduation. By discouraging enrollment, restrictive post-study work policies further reduce the host country's ability to retain international talent. In addition, reduced employment opportunities in bachelor-degree-level jobs can substantially lower labor supply in advanced-degree-level jobs, generating cross-sector effects. Students face uncertainty about the types of jobs they will obtain after graduation, so declining opportunities in one type of job can make studying in the destination country less attractive. The resulting decline in enrollment further shrinks the pool of potential workers in other, possibly higher-value-added jobs.

This paper connects to four strands of literature. First, we contribute to the literature examining how immigration policy shapes international student enrollment. While it is well established that open immigration policies attract international students and restrictive reforms deter them (Shih, 2016; Amuedo-Dorantes et al., 2019; Demirci, 2019; Amuedo-Dorantes et al., 2023; Chen et al., 2023; Abarcar and Theoharides, 2024; Hu and Winters, 2025),[5] previous studies primarily focus on the downstream impacts on the host country, such as generating college revenue (Bound et al., 2021; Zhu, 2024), stimulating local employment (Jia et al., 2025), and fostering regional entrepreneurship and innovation (Hunt, 2011; Stuen et al., 2012; Crown et al., 2020; Beine et al., 2024; Ganguli and MacGarvie, 2025). We complement this literature by explicitly tracing the channel through which post-study immigration policy first alters access to the labor market and then shapes international education choices in a structural model. Previous reduced-form estimates of a given reform are difficult to extrapolate to policies that have never been implemented. By separating students' education preferences from the value of post-study labor-market access, our model instead recovers the elasticity of enrollment with respect to labor-market access itself. This micro-founded elasticity allows us to capture these interconnected dynamics and to simulate enrollment responses to counterfactual work visa policies.

Furthermore, our results speak to a related literature on how immigration policies alter the composition of international students by field of study and ability (Kato and Sparber, 2013; Chen et al., 2023; Amuedo-Dorantes et al., 2023; Flynn et al., 2024).[6] Documenting

[5] A related literature examines other determinants of international education choice, including destination wages, migrant networks, human-capital portability, education quality, tuition and financial support, and completion rates (Rosenzweig et al., 2006; Beine et al., 2014, 2020; Arenas, 2021; Batista et al., 2025). Our structural model incorporates most of these determinants. We discuss this further in Section 5.

[6] Amuedo-Dorantes et al. (2023) find that, following the OPT STEM extension, international matriculation in STEM bachelor's and master's programs increased by 18% and 30%, respectively, relative to non-STEM programs. They find that these effects are most pronounced at highly selective institutions. However, they also mention that these compositional changes may partially reflect demand-side responses (e.g., capacity expansion) rather than heterogeneity in students' preferences.

these heterogeneous responses matters because the success of immigration policies aimed at attracting and retaining talent depends largely on how they reshape the composition of the international student body. Our approach allows us to investigate the mechanisms driving heterogeneous impacts across student groups and to predict heterogeneous responses to other potential policy designs.

Perhaps closest to our paper is Clemens et al. (2025), who similarly examine how immigration policy deters international students and subsequently shrinks the skilled workforce in the context of the United States. Our approach differs from theirs in two main aspects. Whereas they use survey methods to elicit self-reported responses to hypothetical work visa restrictions (e.g., OPT rescission), we employ a quasi-experimental approach to evaluate a past immigration reform in the UK. A further distinction is that Clemens et al. (2025) impute the resulting labor supply shock by applying constant historical stay rates to a declining student pool. However, because stay rates themselves are endogenous to immigration policies, such an approach may yield biased estimates of the true labor market impact. Instead, we introduce a structural framework that jointly models education enrollment and post-graduation labor supply decisions, capturing how policy reforms affect both student inflows and their subsequent retention. This approach is in the spirit of Khanna and Morales (2025), who model forward-looking college major choices under uncertainty about post-graduation labor market access and immigration visa constraints.

Second, we contribute to the literature on immigration policy and the skilled labor supply. Past studies have extensively investigated how visa policies affect the stock of skilled workers and subsequently impact native-born employment and wages (Peri et al., 2015a,b; Kerr et al., 2015; Bound et al., 2015; Demirci, 2019, 2020), firm behavior and performance (Dimmock et al., 2022; Glennon, 2024), host country innovation (Kerr and Lincoln, 2010; Bernstein et al., 2022), and generate spillovers to third countries (Brinatti and Guo, 2025). Related work also uses LinkedIn data to examine global talent flows and high-skill immigrants' location choices (Amanzadeh et al., 2024), but focuses on migration decisions rather than preceding education choices. In this paper, we examine how restrictive immigration policies affect skilled labor supply through two distinct channels for migrants who initially enter on student visas: the education margin (the indirect, or deterrence, effect) and the post-graduation work-location margin (the direct effect). These study-to-work migrants are a major and rapidly growing source of skilled labor in the United States (Bound et al., 2015; Ruiz and Budiman, 2018; Beine et al., 2023), and they exhibit exceptional innovation performance upon entering the labor market (Hunt, 2011).[7] Distinguishing between these

[7] Bound et al. (2015) report that in 2010, 27% of all foreign-born workers and 35% of foreign-born IT workers in the United States initially arrived on a student visa.

two margins is essential for policy design, as a dominant direct effect calls for corrections during labor-market entry such as easing sponsorship and reducing search frictions, whereas a dominant deterrence effect calls for education-side instruments such as tuition subsidies. We quantify the contribution of each margin and find that their relative importance varies across job types.

In addition, our paper distinguishes students' transitions into different types of jobs, recognizing their distinct policy implications. This finer categorization allows us to uncover an important cross-sector spillover: reduced employment opportunities in bachelor-degree-level jobs can substantially lower labor supply in advanced-degree-level jobs. This finding highlights the importance of accounting for linkages across job types when evaluating the effects of immigration restrictions on skilled labor supply.

Third, we contribute to a literature on graduate education that remains comparatively limited, despite the rapid global expansion of advanced degrees (Altonji et al., 2016). Most existing work studies the returns to graduate degrees for domestic students (Altonji and Zhong, 2021; Altonji and Zhu, 2025), while the returns for international master's students, which are potentially much larger due to access to higher-wage labor markets, remain relatively understudied. One exception is Chen (2024), who documents substantial returns to US credentials in the Chinese labor market. While our primary focus is not on directly estimating these returns, our framework implicitly accounts for heterogeneous returns to a master's degree across different destination countries by linking expected wages to post-graduation work locations.[8]

Finally, from a methodological perspective, we contribute to the structural econometrics literature by providing a new identification strategy for models with heterogeneous but unobserved choice sets. Generally, observed choice probabilities alone cannot separately identify the arrival probability (e.g., job offer probability) from the preference parameters governing acceptance (e.g., worker preferences over jobs). The literature typically restores identification through "identification at infinity" (Heckman, 1990) using an excluded shifter that drives the arrival or acceptance probability to the limits (French and Taber, 2011; Arcidiacono et al., 2025), or by imposing some structure on choice decisions to achieve partial identification (Manzini and Mariotti, 2014; Barseghyan et al., 2021).[9] We propose a new point identification strategy in a dynamic context, with an excluded policy shock affecting

[8]Recent work has explored "push factors" of international educational mobility from the perspective of source countries, showing how China's economic development and higher education expansion have fueled the surge in graduate enrollment at US universities (Khanna et al., 2025; Jia et al., 2025).

[9]Similarly, in job search models, the offer arrival rate is confounded with the flow value of nonemployment on the unemployment side (Flinn and Heckman, 1982), and with switching costs on the on-the-job-search side; the literature separates them using moments from the observed wage and accepted-wage distributions or from firms' equilibrium wage-setting response (Bontemps et al., 2000).

only the probability that a choice becomes available. We show that, under this exclusion restriction, the ratio of changes in observed choice probabilities across different stages cancels out the change in the choice arrival probability. This "relative sensitivity" across stages isolates the role of preferences, allowing us to identify preference parameters separately from choice availability. The identification framework can be applied more broadly to other sequential-choice settings. For example, tighter antitrust enforcement (the excluded policy) may reduce startups' probabilities of exiting through acquisition (the downstream effect), thereby discouraging entrepreneurial entry (the upstream effect).

The rest of the paper is organized as follows. Section 2 describes the institutional background of UK higher education and the 2012 reform. Section 3 introduces the LinkedIn profile data and provides descriptive facts about Indian master's students. Section 4 reports the quasi-experimental evidence. Section 5 develops the model, and Section 6 describes identification and estimation. Section 7 reports the parameter estimates and model fit. Section 8 presents the decomposition and counterfactual results, and Section 9 concludes.

2 Institutional Context

This section details the institutional background of the international education market and work visas for high-skill sectors in the United Kingdom around 2010. In Section 2.1, we first describe the UK international education market and the position of Indian students within it. Then, we outline the pre-2012 post-study work visa regime in Section 2.2 and the 2012 policy reform that structurally altered the school-to-work transition in Section 2.3.

2.1 UK Higher Education System and Taught Master's Program

The United Kingdom was the second-largest host country for international tertiary students in 2009, accounting for 10% of international students worldwide, following the United States with 18%.[10] Taught master's programs were particularly prominent among non-EU students. In the 2010/11 academic year, among all first-year, full-time, non-EU students enrolled in UK higher education, taught master's programs accounted for 58.6%, followed by first-degree programs (bachelor degree, 33.9%), research degrees (research master & doctor, 6.2%), and other postgraduate qualifications (1.3%).[11] International students also constituted the

[10] See OECD (2011), *Education at a Glance 2011: OECD Indicators*, Chart C3.2. OECD data for 2023 show that the United Kingdom remained the second-largest host country in absolute terms, with nearly 749,000 international tertiary students, behind the United States with approximately 957,000; see OECD (2025), *Education at a Glance 2025: United Kingdom*.

[11] Shares are calculated from the Higher Education Statistics Agency (HESA) Student Record, Table 11b of the 2010/11 release, which reports first-year students at UK higher education institutions by level of study,

majority of full-time taught master's enrollment: non-UK-domiciled students accounted for 70% of enrollment, with non-EU-domiciled students alone accounting for 57%.[12]

A non-negligible share of higher education for international students was provided by private institutions prior to the 2012 reform. A government report focusing on non-EU students in the UK shows that in the 2009-2010 academic year, 72% of institution-issued visa documents (formally, Confirmation of Acceptance for Studies, or CAS) among bachelor-degree-level or above programs were for students studying at universities, 25% were for study at private institutions and 4% were for study at non-university public institutions.[13] Most privately funded higher education providers could not award UK degrees, so they offered non-degree diplomas, certificates, and professional qualifications.[14] However, they could partner with degree-awarding universities to provide higher-education degrees. These weakly regulated private institutions and the collaborative-provision pattern were vulnerable to low-quality education provision and serving as "diploma mills". Combined with the Post Study Work (PSW) visa described later, this route allowed some international students to get access to low-skill jobs in the UK

In academic year 2010/11, 39,090 Indian students were enrolled at UK higher education institutions across all levels of study, making India the second-largest source country behind China (67,325). Indian enrollment was overwhelmingly concentrated at the taught postgraduate level: an estimated 79 percent of first-year entrants enrolled in taught master's programs in 2010.[15] Indian students were also disproportionately represented at privately funded education providers. More details and evidence about the UK taught master's program are provided in Appendix A.1.

mode of study, and domicile. The definition of higher education here excludes sub-degree qualifications, including HNDs, HNCs, foundation degrees, and other certificates and diplomas of higher education, which are broadly comparable to sub-bachelor's or associate-level qualifications in the United States.

[12]Calculations based on the Higher Education Statistics Agency (HESA), *Students in Higher Education Institutions 2010/11*, Tables 5a and 5b. Full-time enrollment in the HESA category "higher degree (taught)" totaled 197,685 students, of whom 59,900 were UK-domiciled, 24,685 were domiciled in other EU countries, and 113,100 were domiciled outside the EU. By contrast, UK-domiciled students accounted for 125,640 of the 147,610 students enrolled part-time in higher degrees (taught), or 85%.

[13]Source: *Overseas students in the immigration system: Types of institution and levels of study*; UK Border Agency Occasional Paper 90

[14]At the end of 2009, only four private providers held degree-awarding powers: BPP Ltd, the College of Law, the University of Buckingham, and Ashridge Business School. See Crossick (2010), *The Growth of Private and For-Profit Higher Education Providers in the UK*, p. 19. BPP did not obtain university status until later.

[15]HESA Student Record, Table 8 of the 2010/11 release (non-UK domiciled students by country of domicile and level of study).

2.2 The Pre-2012 Work Visa Regime

Prior to the 2012 reform, the UK operated a relatively permissive post-study immigration regime that offered international higher education graduates of UK institutions multiple pathways into the labor market. Other pathways accounted for a negligible share (e.g. Tier 1 entrepreneurship visas), and the two primary routes were the Tier 1 Post-Study Work (PSW) visas and Tier 2 (General).

2.2.1 Tier 1 PSW Visa

The Tier 1 Post-Study Work (PSW) visa was the principal route through which non-EU international graduates transitioned from study into the UK labor market. In 2011, it accounted for 43,319 of the 46,875 (92%) grants of working visas to student-visa holders switching into work routes.[16] International students who had recently obtained a UK higher education degree, including a recognized bachelor's, master's, or doctoral degree, could apply for a Tier 1 Post-Study Work visa subject to additional eligibility requirements.[17] Successful applicants received a two-year visa with broadly unrestricted access to the labor market. They did not need a job offer to legally stay in the UK, and their employers were not required to sponsor their work visas. Therefore, PSW visa holders could stay unemployed and search for work, accept employment at any skill level, or become self-employed. Firms were also able to hire international graduates without administrative burdens related to sponsorship. The PSW visa was only temporary for two years and did not serve as a direct route to settlement. To legally work in the United Kingdom after it expired, graduates had to qualify other visa eligibility, such as Tier 1 (General) or Tier 2 (General).

However, the PSW route led to two connected policy concerns that contributed to the 2012 reform: the integrity of some qualifying educational provision and the prevalence of employment below the skill level expected for university graduates.[18] As discussed in the

[16] See Parliament Record, *Past and present post-study work routes*.

[17] For an initial application, an applicant generally had to have obtained the qualification at a qualifying UK institution, held student status or qualifying dependant status while studying, applied within 12 months of the award, and satisfied the maintenance requirement. The list of eligible qualifications also included a Postgraduate Certificate in Education (PGCE), a Professional Graduate Diploma in Education (PGDE) obtained in Scotland, and a Higher National Diploma (HND) awarded by a Scottish institution. See Migration Advisory Committee, *Analysis of the Points Based System: Tier 1* (December 2009), paras. 2.50–2.62 and Table 2.2, pp. 25–27.

[18] The Government's impact assessment identified both abuse of the student visa and the employment of some PSW migrants in less-skilled occupations as rationales for reform. See Home Office, *Reform of the Points Based Student Immigration System: Final Impact Assessment* (2011), pp. 1, 8, and 36. The House of Commons Home Affairs Committee likewise supported action against bogus colleges and considered restricting PSW eligibility by awarding institution, while cautioning that the available evidence on abuse remained limited. See House of Commons Home Affairs Committee, *Student Visas* (2011), paras. 58–62 and 77.

previous section, private providers could offer university-cooperated programs through franchise or validation arrangements. Students completing such programs could potentially obtain a university-awarded degree and thereby satisfy the eligibility for the PSW visa. As a result, an international student graduating with such a degree may enter the UK labor market without bachelor-level human capital and work in a low-skill sector. Additional evidence is presented in Appendix A.2.

2.2.2 Tier 2 Visa

The Tier 2 (General) visa coexisted with the Tier 1 PSW route and constituted UK's principal employer-sponsored immigration route for skilled non-EEA workers. It was broadly analogous to the H-1B visa in the US, although the two systems differed substantially in their eligibility criteria and allocation mechanisms. Prior to the reform, an eligible applicant for a Tier 2 (General) visa had to meet English-language proficiency and minimum salary requirements and secure a job requiring skills above the high-school level. Details are written in Appendix A.3.

Although Tier 2 was an available route for eligible international graduates before the reform, the vast majority of graduates entering the UK labor market between 2008 and 2011 used PSW rather than transitioning directly to Tier 2. As shown in Figure 1, among Tier 4 student visa holders who transited into a work visa in 2011, 92.4% entered Tier 1 PSW, whereas only 3.6% transitioned directly to Tier 2. The PSW visa was a substantially lower-cost instrument for employers to hire international graduates. It required neither employer sponsorship nor a minimum salary. For graduates, the PSW visa provided them with a two-year search window to find desired work. By contrast, Tier 2 visa restricted graduates to licensed employers and imposed salary and occupational skill constraints. The PSW visa was therefore a better hiring option for both employers and graduates.

2.2.3 Other Visa Routes

A small number of graduates entered the UK labor market through other Tier 1 sub-routes, most notably Tier 1 (General). Tier 1 (General) typically required a prior earnings record that most international students entering the labor market for the first time were not eligible for. Figure 1 shows that the Tier 1 PSW and the Tier 2 routes nearly covered all Tier 4 visa holders who transited into a work visa in 2011.

2.3 The 2012 Student Immigration Reform

In March 2011, the UK government announced a package of student-immigration reforms that would be implemented in stages over 2011 and 2012. The reforms sought to remove education providers acting as immigration loopholes rather than providing genuine education, protect the reputation of the UK education sector, and curb the abuse of student and post-study visas as pathways into low-skill employment.[19] At the same time, policymakers maintained a stated commitment to retaining highly skilled graduates. Official statements emphasized attracting the "best and the brightest" and allowing "only the best of them" to remain and enter the UK labor market after graduation.[20] The reform package can be generally classified into education and work visa aspects. The education-related measures primarily targeted low-quality private colleges and their international students, who are outside the main focus in this research. The education-side measures sharply reduced the volume of CAS issued by private institutions after 2012, and the decline was particularly salient for Indian students. The share of Indian CAS registrations sponsored by non-university institutions fell from 55.7% in 2011 to 9.6% in 2014, with the raw count falling from 22,266 to 1,233, a decline of 94 percent.[21] More details about specific education-side measures taken in the reform are written in Appendix A.4.

The most influential reform for genuine UK higher-education graduates was the abolition of the Tier 1 PSW visa. It was announced in March 2011 and fully implemented in April 2012. Notice that this timeline created a sharp exposure cut-off for international master's cohorts. Students entering in the fall of 2010 were largely unaffected, as they graduated in time to secure the PSW visa before its closure. Conversely, the next cohort enrolling in the fall of 2011 realized that PSW would be unavailable upon their graduation in 2012. That motivates us to treat 2011 as the first treated cohort in subsequent analysis. Figure A2 in the Appendix uses student visa transition data and shows that transitions from the study route into work within two years fall sharply across the reform. The situation of the 2010 cohort is similar to that of the 2009 cohort. However, for all international students, the number of transitions falls from approximately 29,100 transitions for the 2010 student cohort to 4,600 for the 2011 cohort and fewer than 2,500 for the 2012 cohort. For Indian students, the number falls from 11,400 for the 2010 cohort to 1,800 for the 2011 cohort and only 545 for

[19] See Home Office, *Final Impact Assessment: Reform of the Points Based Student Immigration System* (2011), pp. 1 and 6–11; and Theresa May, "Student Visas," House of Commons Debates, 22 March 2011, vol. 525.

[20] Theresa May, "Student Visas," House of Commons Debates, 22 March 2011, vol. 525. Both quoted phrases are excerpted from this statement.

[21] Author's calculations from the Home Office sponsorship dataset (CAS_D02).

the 2012 cohort.[22] These drops were primarily driven by the loss of the Tier 1 PSW route, as there was no significant drop in other types of study-to-work visa issuance.

After the closure of the PSW route, all fresh graduates could only rely on the Tier 2 visa to secure legitimate work rights in the UK labor market. Thus, they faced tight time constraints to obtain a skilled job offer from a licensed employer. Depending on the length of their academic program, graduates were granted a grace period of only two months (for programs lasting 6 to 12 months) or four months (for programs lasting 12 months or more) to secure a sponsored job offer before being required to leave the country.

Beyond the compressed window, Tier 2 itself became more restrictive, particularly due to the change in Tier 2 Visa eligibility. The occupational skill threshold rose from NQF Level 3 (high school level) to NQF Level 4 (associate degree level) after April 2011, and subsequently to NQF Level 6 after June 2012 (bachelor's degree level), meaning that only graduate-level (bachelor's-equivalent) occupations qualified.[23] The latter change excluded a wide range of associate-degree-level occupations, such as "office managers" and "IT operations technicians", that had previously been Tier-2-eligible. Yet, because Tier 2 visa still covered bachelor's-degree-level occupations (NQF Level 6) rather than restricting eligibility to only postgraduate-level jobs (NQF Levels 7 and 8), the increased skill threshold was unlikely to be a significant binding factor for graduates of taught master's programs. The Tier 2 Visa sponsorship also became more costly for employers. UK Employers needed to pay £1,200-2,200 for each worker they would like to sponsor, including the legal costs for lawyers' assistance. They also needed to bear the compliance costs during sponsorship. More details are illustrated in Appendix A.5.

Overall, the 2012 reform replaced the flexible two-year unconditional stay option with a constrained, heavily audited employer-sponsorship regime. As shown in Figure 1, the number of visas issued for the transition from a Tier 4 study visa to a work visa dropped sharply following the implementation of the policy bundle. While there was a marginal increase in Tier 2 General visa issuances, this substitution was entirely insufficient to compensate for the massive loss of the Tier 1 PSW visa, resulting in a severe aggregate decline in the retention of international graduates.

[22] This shows that some spring 2011 entrants completed their one-year programs in time and managed to claim the PSW visa before the formal closure in April 2012, while the majority entering in fall 2011 and graduating in late 2012, faced a regime in which the employer-sponsored Tier 2 visa was the only work visa route.

[23] See UK Border Agency, *Tier 2 of the Points Based System* (2011), and UK Border Agency, *Statement of Intent and Transitional Measures: Tier 2* (2012).

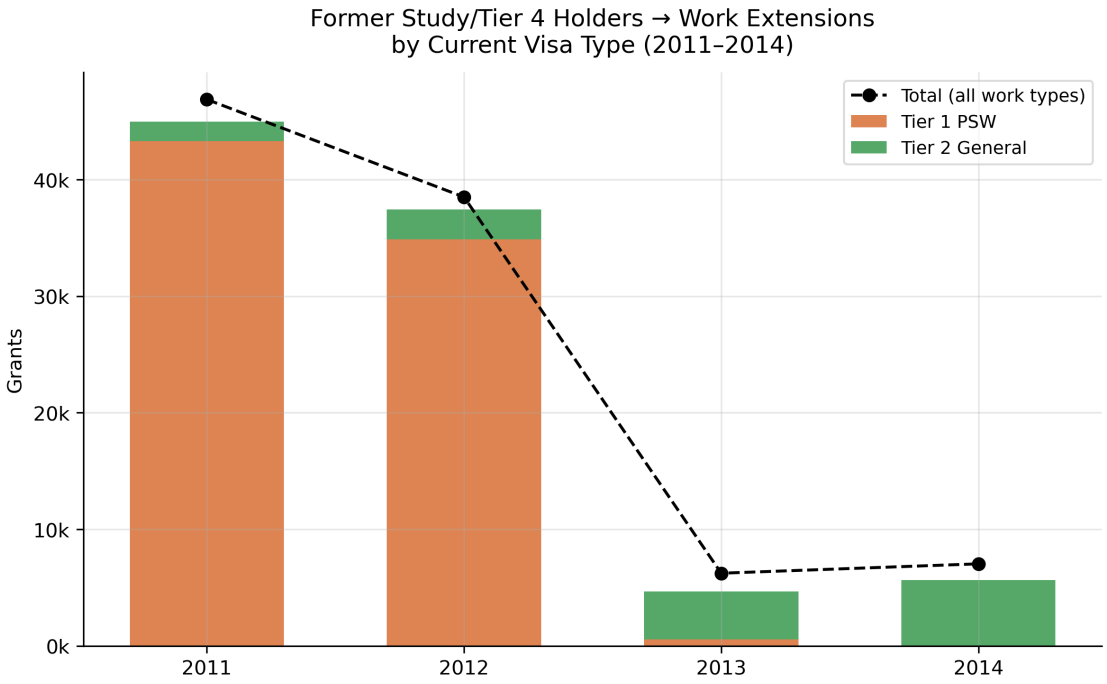


Figure 1: Study-to-Work Visa Transitions in the United Kingdom, 2011–2014

Notes: The figure reports annual study-to-work transition visas. Bars represent two of the most common visas– Tier 1 (Post-Study Work) and Tier 2 (General)—and the dashed line plots the total across *all* work categories. The residual consists of other routes such as Tier 1 (General). All figures exclude dependants and count main applicants only; they also exclude EEA and Swiss nationals, and exclude asylum cases and withdrawn applications while including reconsiderations and successful appeals. *Source:* Home Office, *Immigration Statistics, October to December 2015*.

2.4 Summary and Implications for Master's Students

The closure of low-quality private education providers and the removal of post-study work access to low-skill jobs substantially curtailed the private education markets in which the student visa system was most susceptible to misuse. However, these education-side reforms should have small effects on our population of interest: international students enrolled in taught master's programs at UK universities. The main policy shock relevant to this group was instead the abolition of the PSW visa.

If a non-trivial share of these taught master's graduates entered occupations requiring less than a bachelor's degree (below NQF Level 6), the abolition of the PSW visa meant they could no longer secure these jobs and remain in the UK labor market, forcing them to pursue alternative education or career pathways. Furthermore, even if the majority of graduates targeted occupations that remained Tier 2 eligible, the abolition could still reduce access to high-skill jobs due to two distinct matching frictions. On the employer side, hiring an international graduate now required firms to bear the administrative costs of sponsorship. On the student side, graduates faced a substantially shorter search window to secure an eligible offer from a licensed employer. Together, these frictions made matches between international graduates and UK high-skill jobs more difficult.

3 Data and Descriptive Statistics

3.1 LinkedIn Profile Data

Studying how visa policy jointly shapes education and employment choices requires longitudinal microdata that follow individuals across countries and over time. Conventional data sources, such as cross-sectional censuses and aggregate visa statistics, generally cannot link pre-migration education to post-migration labor-market outcomes. We address this limitation using LinkedIn profile data from Revelio Labs to construct education and employment trajectories for international master's students. LinkedIn is one of the most widely used job search platforms for high-skill workers, on which job seekers worldwide post their education and work histories. The data cover millions of worldwide workers and provide detailed, timestamped records of education and employment experiences. For each education spell, we observe institution names, degree types, fields of study, and dates of enrollment and graduation. For each employment spell, we observe job titles, employers, start and end dates, and geographic locations.

3.2 Sample Construction

We construct our primary analysis sample by identifying Indian students on LinkedIn who pursued master's degrees abroad. This sample selection is motivated by a few key factors. As discussed above, India is one of the largest source countries for international master's students and one of the most responsive to destination-country visa policy. Relative to the other major source country, China, Indian workers adopt LinkedIn at substantially higher rates, which yields better coverage of the population we study.[24] We also restrict attention to a single origin country so that undergraduate educational background and attainment are more comparable, whereas cross-country comparison would be more difficult. We acknowledge that Indian students might not represent the choice pattern for other international students. Nevertheless, Indian students constitute a large and policy-responsive population for studying the joint determination of international education and migration choices.

Because self-reported nationality is not available in the LinkedIn data, we proxy for Indian origin using the location of an individual's undergraduate institution. Specifically, we restrict the sample to users whose most recent bachelor's degree was granted by an Indian institution and who subsequently report enrollment in a foreign master's program. We further limit the sample to cohorts starting their first foreign master's degree between 2008 and 2013. We additionally exclude individuals pursuing a master's degree en route to a Ph.D. by dropping observations where the reported start year of a master's program coincides with the start year of a doctoral program. After these restrictions, we have a final sample of 177,683 unique individuals.

To evaluate the representativeness of our LinkedIn sample, we compare the destination shares and yearly trends from 2010 to 2018 of Indian master's students with administrative data from the United States, the United Kingdom, Australia, and Canada. Differences in program definitions and reporting conventions, such as enrollment stocks versus entrant flows, require harmonization assumptions, which we detail in Appendix Section B.1.1. Figure A1 shows that the LinkedIn data closely tracks the administrative destination shares over time, particularly the sharp decline in the UK share starting in 2011, reflecting the impact of the PSW abolition reform. More validation evidence and discussion are provided in the Appendix B.1.

Another potential concern lies in differential coverage across labor markets, rather than education markets. Specifically, although LinkedIn is a widely used job-seeking platform among high-skill workers in India, it is still less prevalent than in developed countries such

[24] In 2014, India had 24 million LinkedIn users while China only had 4 million. LinkedIn announced the closure of its localized social networking platform in October 2021 and discontinued its local job-search service in August 2023.

as the UK and the United States,[25] potentially resulting in an underestimate of the share of students returning to India after completing their master's degrees. To address this issue, we use the UK administrative dataset *Migration Journey*, which tracks the visa status of individuals who entered on Tier 4 student visas. These data allow us to estimate LinkedIn coverage among Indian students who stay in the UK for work post-graduation. Combining this estimate with the overall coverage rate for Indian taught master's students also allows us to infer coverage among those who leave the UK for work. We estimate that coverage is approximately 87% for Indian "stayers" and 62% for Indian "leavers". Details of this calculation are in Appendix Section B.2. Following Amanzadeh et al. (2024), we then assign individual weights to students equal to the inverse coverage rate for their combination of master's destination and work location in the model estimation. Details of the weighting procedure are provided in Appendix Section B.3.

3.3 Variable Construction

We construct the following variables from the LinkedIn dataset for our analysis.

Undergraduate Major: A central feature of global high-skill migration is the disproportionate representation of students in technical fields. We construct a binary indicator equal to one if the student's undergraduate field of study is classified within a Science, Technology, Engineering, or Mathematics (STEM) discipline. This classification is based on the standardized major text (e.g., "BEng", "Computer Science", "Electrical Engineering") reported for their last Indian bachelor's field and degree. We exclude students with undergraduate degrees in medicine, pharmacy, law, agriculture, or architecture, as career trajectories in these regulated professions differ systematically from those in standard STEM and non-STEM fields.

Undergraduate Institution Quality: To proxy for students' baseline ability, we categorize individuals based on the quality of their undergraduate institution in India. Our primary measure utilizes both the overall and discipline-specific rankings from the National Institutional Ranking Framework (NIRF). Introduced in 2015 by the Indian Ministry of Education, the NIRF provides a standardized, annual assessment of higher education institutions across the country (see Appendix C for detailed ranking criteria). Using these data, we categorize STEM students, accounting for around 80% of our sample, into three ability tiers based on the rank of their bachelor's institution. Specifically, Tier 1 comprises

[25] Using LinkedIn-based professional profiles from Revelio Labs, Amanzadeh et al. (2024) report coverage ratios of approximately 40% for India, compared with 100% for the United Kingdom and 86% for the United States (Table A1). These ratios compare professional profile counts with ILO estimates of the college-educated population.

Institutes of National Importance and the top-50 colleges in either the NIRF Overall or NIRF Engineering rankings;[26] Tier 2 comprises remaining colleges ranked by either of these systems; and Tier 3 comprises students whose undergraduate institution is unranked.

For non-STEM students, we do not further classify bachelor's institutions, for three reasons. First, non-STEM students are fewer than STEM students, so a finer partition would reduce the precision of the estimates. Second, no credible ranking of undergraduate education exists for non-STEM fields.[27] Third, some Indian students only report the degree-awarding university, whereas admission and teaching take place at the level of constituent or affiliated colleges whose selectivity varies widely within the same university.[28] The self-reported LinkedIn education profile does not identify which college. This ambiguity is more consequential for non-STEM fields. A large share of STEM graduates hold degrees from standalone technical institutions, like the Indian Institutes of Technology (IIT), which appear on the profile under their own name and are ranked individually in the engineering table of the NIRF. Thus, reporting affiliating universities is more common among Non-STEM individuals, which motivates us not to classify non-STEM students' bachelor's tier.

Undergraduate Attainment: Institutional tier captures selectivity at the point of undergraduate admission, but it does not capture students' academic performance once enrolled. We therefore construct a separate individual-level indicator of undergraduate academic attainment, "**Honour**", equal to one if a student's LinkedIn bachelor's education profile reports distinction on any of three self-reported dimensions. The first is an honours designation on the degree title. The second is an explicit distinction marker, including first class, first division, distinction, a university rank or institute medal, *summa cum laude*, valedictorian, or dean's list. The third is a numeric grade at or above the conventional distinction threshold on the reported scale.[29]

First Job Placement: To capture early-career labor market outcomes, we identify the first employment or the next education spell that commences after the completion of the foreign master's degree. We extract the location, job title, and company name of this

[26] Institutions of National Importance are public institutions of prominent Indian higher-education institutions designated by an act of the Indian Parliament. The category includes many of the country's most selective institutions, most notably the Indian Institutes of Technology (IITs).

[27] The NIRF does publish a management ranking, but it evaluates postgraduate management programs rather than undergraduate commerce education. Several of the most selective undergraduate business colleges (e.g., Narsee Monjee and Jai Hind College) do not appear in it.

[28] For example, a profile listing "University of Delhi" is consistent with St. Stephen's College or Shri Ram College of Commerce, among the most selective undergraduate institutions in the country, and equally with a lower-tier affiliated college.

[29] These thresholds are 8.5 on a ten-point scale, 3.7 on a four-point scale, 75 on a percentage or hundred-point scale, and 5.25 on a seven-point scale. Where a grade point average is reported without a stated denominator, we assign the scale by magnitude and leave ambiguous values unassigned; we exclude five-point scales, on which marks in several grading systems run in reverse order.

initial placement, allowing us to classify whether the student entered the host country's labor market, returned to India, or migrated to a third country. We further classify each placement along the following two dimensions.

1. Job Level. We classify each job into one of three tiers: "*advanced-degree-level*" that typically require graduate school education, "*bachelor-degree-level*" that typically require bachelor's education, and "*low-skill*" occupations. Throughout the paper, we define the "*high-skill*" occupation as the combination of "*advanced-degree-level*" and "*bachelor-degree-level*" occupations. The classification rests primarily on the O*NET Job Zone associated with the occupation code that Revelio Labs imputes from job title and employer. The classification details are described in Appendix D.

2. Wages. Wages are not directly observed from LinkedIn profiles. We therefore use salary estimates imputed by Revelio Labs based on job title, seniority, employer, location, tenure with the employer, and year. Revelio Labs constructs these estimates using machine learning models trained on salary information from Labor Condition Applications, online job postings, and employee reports on platforms such as Glassdoor and Levels.[30] This imputed wage data have also been exploited in Amanzadeh et al. (2024), who demonstrate that their model estimates using Revelio Labs salaries yield qualitatively identical conclusions to those using raw salary datasets not relying on imputation.

A non-trivial share of master's graduates immediately enter another educational program rather than the labor market. Amuedo-Dorantes et al. (2019) find that when relaxing visa restrictions makes it easier for STEM students to obtain job offers, PhD degrees become relatively less attractive. Thus, international students might be more likely to enter doctoral programs after the reform. We separately classify students who transited from the first master's program directly into a doctoral program or into a second master's program. These students do not appear as workers in our placement classification but are included as separate post-graduation outcomes in the structural model.

3.4 Summary Statistics

Table A1 reports summary statistics for our sample of 85,038 Indian international master's students enrolled between 2008 and 2010 before the reform, separately for three tiers of STEM students and non-STEM students. International Indian master's students were skewed toward STEM fields, accounting for 79% of students. Approximately 20% of students completed their bachelor's degree at a Tier 1 or Tier 2 institution in STEM fields.

[30] See Revelio Labs' salary methodology and the 2023 JUST Capital, Revelio Labs methodology report, pp. 2 – 5. The models predict salaries using job and employment characteristics; individual education, skills, gender, and ethnicity do not enter directly.

Women were a minority of the sample (27.5%) but constituted a substantially larger share in non-STEM fields (35.2%). More discussion on Table A1 is provided in Appendix E.

3.4.1 Descriptive Evidence about the Policy Shock

Before proceeding to the formal econometric and structural analysis, we present descriptive evidence highlighting the sharp aggregate and compositional shifts brought by the 2012 UK visa reform. The descriptive analysis shows three stylized facts regarding the policy's impact.

Fact 1: UK master's enrollment dropped by more than half.

Following the announcement of the reform in March 2011, UK master's programs experienced a sharp decline in their ability to attract students from India. As shown in Figure A3, master's enrollment of Indian students in the UK plummeted from nearly 9,800 students before 2011 to around 4,600 students in 2013. In contrast, enrollment trends in competing destinations exhibited slight growth during the same period, suggesting that the UK visa reform may have driven Indian students from the UK to other destinations.

Fact 2: STEM students' share declined after the policy.

Figure A4 plots the STEM share of Indian master's students at each destination and each year over the sample period. Before the policy announcement, the UK already had a lower share of STEM students than the US, Europe, and Canada for taught master's students from India. After the reform, the STEM share in the UK fell further, from 67% in the 2010 cohort to 59% in the 2013 cohort. We also observe no significant changes in other destination countries over this period, suggesting that the observed change reflects the policy rather than broader shifts in college education or study-abroad patterns in India. Figures A5 and A6 in the Appendix plot time trends in student composition by undergraduate institution quality and degree honors. While the figures show the increase in the share of STEM students from Tier 1 or Tier 2 undergraduate institutions, as well as the increase in the share of students with honors degrees, for the UK before and after the policy, it is difficult to draw a credible conclusion from the descriptive evidence. These shares also increased among master's students in other destination countries, and some preexisting trends appear between 2009 and 2010.

Fact 3: "bachelor-degree-level" and "low-skill" jobs drive the decline in UK stay rate.

Figure 2 shows the share of Indian taught master's students who remained in the UK for work. We further disaggregate stayers by occupational skill level: advanced-degree-level jobs, bachelor-degree-level jobs, and low-skill jobs (see Section 3.3). The share of UK master's graduates whose first job was in the UK (the "Work in Host" category) fell from 33% in

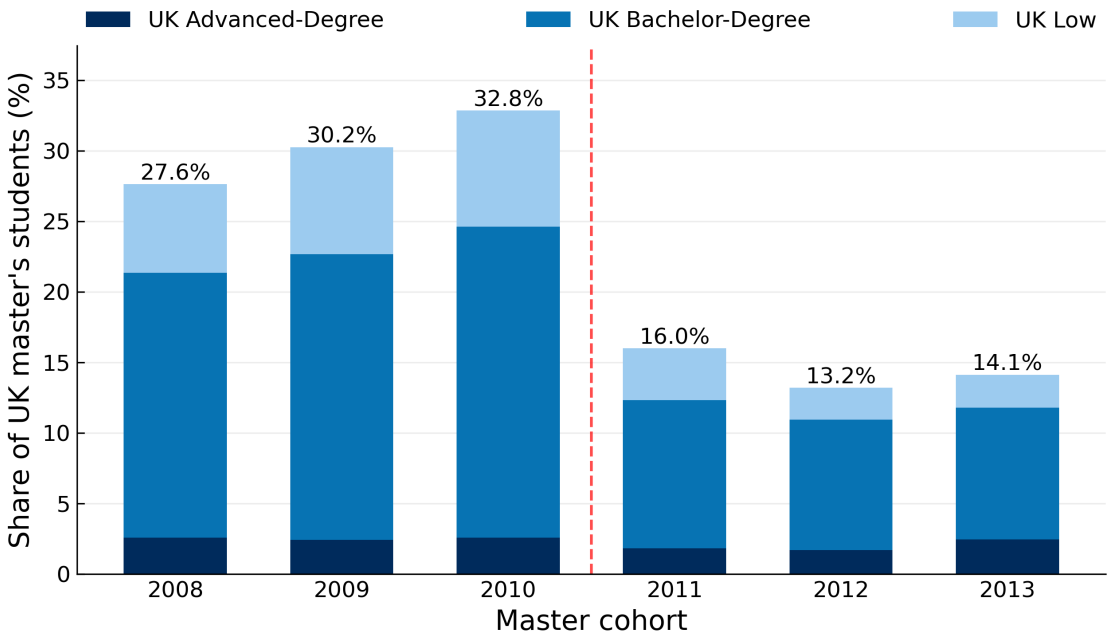


Figure 2: Stay Rate among UK Masters by Occupation Groups

Notes: The figure plots the share of Indian masters studying in the UK whose first job is in the UK, split by the skill tier of that job. Students who returned to India, took a job in a third country, enrolled in a doctoral or a second master's programme, or are not observed constitute the remainder and are not drawn. The dashed line marks the March 2011 announcement of the abolition of the UK Post-Study Work visa.

the 2010 cohort to 16% in the 2011 cohort and remained at a similarly low level through the 2013 cohort. Most of this decline is driven by the share entering bachelor-degree-level jobs, which fell from around 22% to 11%. The share entering low-skill jobs also decreased substantially, from 8% to 4%, while the share entering advanced-degree-level jobs declined only from 2.6% to 1.8% before returning to 2.5% by 2013. Figure A7 in the Appendix shows that the share returning to India or working in a third country rose after the reform. The total counts, which also incorporate the decline in taught master's enrollments, are shown in Figure A8 in the Appendix.

This figure has four implications for this research. First, most UK master's students in our LinkedIn sample who remained in the UK secured high-skill rather than low-skill jobs upon graduation, even when the PSW visa was still available before 2011. This suggests that our LinkedIn sample is more representative of master's students in universities than of those in low-quality franchised programs. Second, both before and after the reform, most master's graduates who remained in the UK worked in bachelor-degree-level rather than advanced-degree-level jobs, suggesting occupational downgrading for international master's graduates.[31] Third, the decline in low-skill employment reflects the direct effect of the

[31]This pattern has also been documented in broader policy reports and empirical literature. For instance,

abolition of the PSW visa, as students working in low-skill jobs were no longer eligible for Tier 2 visas after the reform. The existence of some low-skill employment after the reform may reflect alternative visa routes (e.g., family visas) or the misclassification of some high-skill occupations as low-skill due to self-reporting errors from LinkedIn users or classification errors from Revelio Labs.

Most importantly, the reform also reduced UK retention in the high-skill sector, beyond the low-skill category targeted by the policy. In other words, the policy generated spillovers to the high-skill sector, leading to the loss of some of the "best and brightest students" who would otherwise have been eligible for Tier 2 visas but ultimately did not remain in the UK. The heterogeneity in policy effects between advanced-degree-level and bachelor-degree-level jobs also sheds light on the policy's mechanisms. bachelor-degree-level jobs may face greater policy frictions because the Tier 2 salary requirement is more likely to bind for these positions. This pattern is the central descriptive fact motivating our structural analysis: the policy targeted low-skill employment but also substantially reduced high-skill retention. Whether this decline in high-skill retention reflects increased frictions at the labor-market margin (UK enrollees becoming less able to secure graduate-level employment) or changes in observed and unobserved student composition is the question we examine formally in Section 5.

4 Quasi-Experimental Evidence

Although the descriptive evidence shows a substantial decline in UK master's enrollment following the announcement of the 2012 UK immigration policy, this decline coincides with several potential confounding factors, most notably the 2008 global financial crisis and the associated worldwide fluctuations in macroeconomic conditions and exchange rates. To causally identify the effect of the 2012 UK immigration policy change on international student flows, we employ an event-study framework that exploits cross-sectional variation across undergraduate institutions in their exposure to the UK labor market. We measure exposure using the pre-policy share of UK master's graduates from each institution who remained in the UK for work after graduation. The identification strategy rests on the intuition that students from Indian undergraduate institutions with a higher historical propensity to remain in the UK for work are more exposed to visa restrictions on this pathway, and therefore more re-

UK Migration Advisory Committee (2018) analyzes UK HMRC tax records (LEO data) and reports that post-graduation earnings for non-EU Level 7 taught Master's graduates are virtually indistinguishable from those of non-EU undergraduates. Similarly, Demirci (2020) utilizes SEVIS and ACS data to show that international STEM Master's graduates on OPT accept significantly lower initial salaries than their native-born counterparts and are disproportionately concentrated in routine STEM and IT tasks.

sponsive to the policy change. Consequently, we expect institutions with greater exposure to experience a larger decline in UK master's enrollment following policy implementation. As is common in empirical designs without a pure, unexposed control group (e.g., Dustmann et al., 2022), our identification strategy requires a linear dose-response assumption in addition to the standard parallel-trends assumption. Specifically, we assume that the treatment effect on the outcome variable, the log of enrollment, is linear in the continuous measure of exposure, which in our setting is the historical stay rate. Details regarding econometric assumptions and identification are described in Appendix F.1.

Specifically, we specify our primary event-study model as follows:

$$E\big[Y_{ct}\big|D_c, \alpha_c, \delta_t, \kappa_{st}\big] = \exp\left(\sum_{\tau=2008, \tau\neq 2010}^{2013} \beta_\tau\big(D_c \times \mathbb{1}\{t=\tau\}\big) + \alpha_c + \delta_t + \kappa_{st}\right) \tag{1}$$

where:

- Y_{ct} is the number of students from institution c enrolling in a Master's program in the United Kingdom in year t.
- D_c represents treatment exposure, defined as, among UK master's students from undergraduate institution c who matriculated during the pre-reform period (2001–2007), the share who remained in the UK for work upon graduation.
- α_c and δ_t are institution and year fixed effects, respectively. These account for time-invariant institution characteristics and aggregate shocks to the demand for international education.
- κ_{st} denotes state-by-year fixed effects, which absorb time-varying state-level conditions such as economic growth. In an alternative specification, we also replace it with X_{st}, the natural logarithm of gross state domestic product (GSDP), as a robustness check.

The coefficient of interest, β_τ, captures the causal effect of UK job accessibility on students' decisions to enroll in a UK master's program. Specifically, it measures the differential change in enrollment for undergraduate institutions with varying levels of historical reliance on the UK post-study work pathway. We expect β_τ to be statistically indistinguishable from zero for $\tau < 2011$. For $\tau \geq 2011$, a negative and significant β_τ would indicate that students from institutions more reliant on the PSW pathway were more responsive to the policy change. Since the policy was announced in March 2011, we expect the coefficient β_{2011} to

capture a partial effect of the reform, as some students may have adjusted their educational choices in response to the announcement, while others may not yet have done so.

Our main empirical estimation uses Poisson pseudo-maximum likelihood (PPML), as the dependent variable Y_{ct} is a count of students and some institution-year cells have zero counts (i.e., no UK master's program enrollment records on LinkedIn for students from institution c and cohort t). We also consider alternative functional forms and specifications as robustness checks. First, we estimate a weighted OLS model using the logarithm of student counts as the dependent variable. We also turn to student-level regressions, where the dependent variable is an indicator for whether the student chose a UK master's program.

One potential source of bias is sampling error in the treatment exposure D_c. When the number of UK master's students from institution c in the pre-reform period is small, the exposure measure of the stay rate may be noisy. We address this concern in two ways. First, we restrict the sample to institutions with at least 30 pre-policy UK master's students. For institutions with fewer than 30 students, we pool them within an Indian state and bachelor's-institution tier (see Appendix F.2) to form a credible cluster. Second, we construct an alternative measure of the stay rate using Empirical Bayes (EB) shrinkage. This approach is in the same spirit as the methods used in the teacher value-added literature (e.g., Kane and Staiger, 2008; Chetty et al., 2014) and shrinks the institution-level stay rate toward the state-level stay rate, with the weight determined by the relative amount of sampling noise in the two measures. Details are provided in Appendix F.3.

4.1 Main Results

Figure 3 plots the event-study coefficients (β_τ) from equation (1), along with their associated 95% confidence intervals, with standard errors clustered by undergraduate institution. The coefficients are statistically indistinguishable from zero in 2008 and 2009, providing no evidence of an anticipation effect. There is also no evidence of a differential pre-trend prior to the policy change. Following the announcement of the PSW visa abolition in March 2011, we observe a statistically significant decline in UK master's enrollment among more highly exposed undergraduate institutions relative to less exposed institutions. The smaller point estimate in 2011 relative to 2012 and 2013 is consistent with the limited time available for the 2011 cohort to adjust its educational choices following the March announcement.

To summarize the event-study estimates, we combine the pre- and post-reform periods and estimate a standard difference-in-differences (DiD) specification:

$$E\left[Y_{ct}\middle|D_c, \alpha_c, \delta_t, \kappa_{st}\right] = \exp[\beta\big(D_c \times \text{Post}_t\big) + \alpha_c + \delta_t + \kappa_{st}] \tag{2}$$

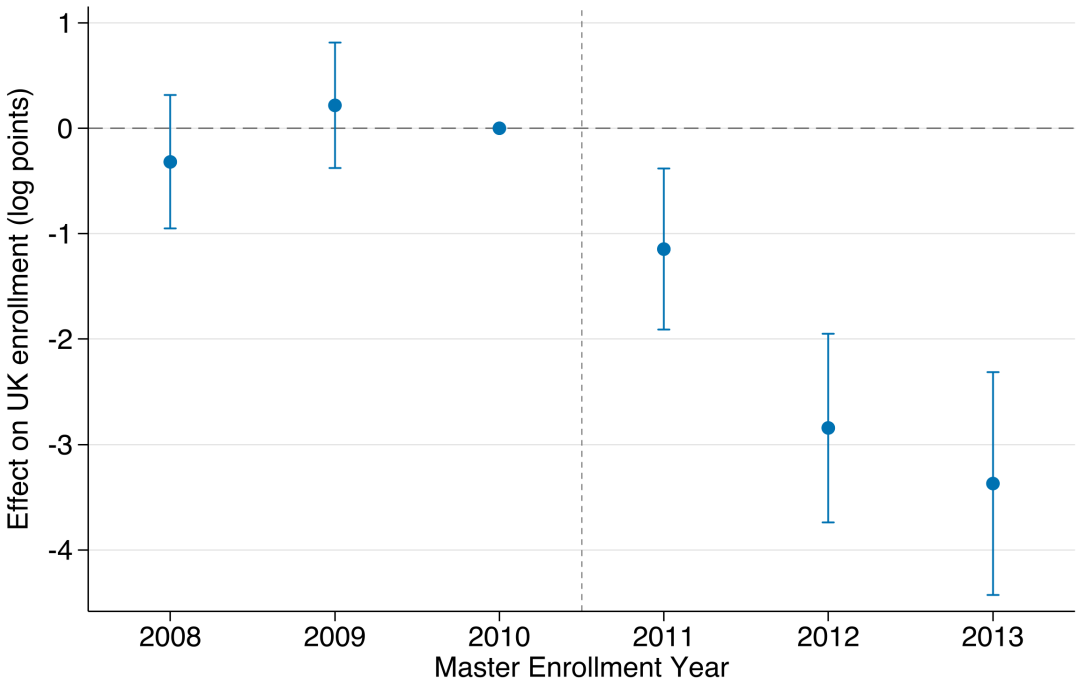


Figure 3: Event Study: Effect of Pre-Reform Exposure on UK Master's Enrollment Change

Notes: The figure plots the coefficients β_τ from a Poisson pseudo-maximum-likelihood regression in equation (1), with cluster and state-by-cohort fixed effects. The pre-reform UK stay rate is the share of the institution's UK master's students in the 2001–2007 cohorts whose first job after the degree was in the United Kingdom. Vertical bars are 95% confidence intervals based on standard errors clustered by undergraduate institutions. An undergraduate institution with at least 30 UK master's students in the 2001–2007 cohorts forms its own cluster, and smaller institutions are pooled together within Indian state and bachelor's-institution tier.

Table 1: Effect of Pre-Reform Exposure on UK Master's Enrollment Change

	(1)	(2)	(3)	(4)
Stay Rate × Post	-3.130***	-3.074***	-1.579***	-4.488***
	(0.490)	(0.486)	(0.483)	(0.739)
Estimator	PPML	Weighted log	Logit	PPML
Exposure	Raw	Raw	Raw	EB
Cluster FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
State × Year FE	Yes	Yes	Yes	Yes
Observations	690	686	137,554	690
Clusters	138	138	143	138

Notes: The dependent variable is the number of UK master's entrants from an undergraduate-institution cluster in a cohort (columns 1–2 and 4) or an indicator for choosing a UK master's degree at the student level (column 3). "Post" is an indicator for the 2012–2013 cohorts and the 2011 cohort, which was partially treated, is excluded; the pre-period is 2008–2010. Columns 1 and 4 report Poisson pseudo-maximum-likelihood estimates; column 2 reports OLS on the log count, weighted by the institution's 2001–2007 UK headcount; column 3 reports student-level logit coefficients (log-odds) with controls for gender, an honours degree, and bachelor's-institution tier. Column 4 replaces the raw stay rate with an empirical-Bayes estimate that shrinks each institution's rate toward the mean of the other institutions in its state, with weights given by the ratio of signal to total variance. In the columns with state-by-cohort fixed effects, the five clusters that are the only cluster in their state are absorbed by the fixed effects and excluded. Standard errors in parentheses are clustered by undergraduate institutions (cluster). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

where Post_t is an indicator variable equal to one for the years 2012 and 2013. The 2011 cohort is excluded since it was not fully treated. In all specifications, we cluster standard errors at the institution level to account for arbitrary serial correlation within institutions. The DiD estimates are reported in Table 1. The specification in column (1) is the same as that underlying Figure 3. In column (2), we use the logarithm of student counts as the dependent variable and estimate a weighted OLS regression, weighting by the number of UK master's students from each institution who stayed in the UK during the pre-policy period. In column (3), we turn to a student-level regression, where the dependent variable is an indicator for whether the student chooses a UK master's program. In column (4), we use the Empirical Bayes-adjusted stay rate as the measure of treatment exposure.

Across all specifications, the DiD estimates are negative and statistically significant at the 1 percent level. Our preferred specification in column (1) implies that a one-percentage-point higher pre-reform stay rate is associated with an approximately 3.1% larger post-reform decline in the number of students from that undergraduate institution enrolling in UK master's programs. The implied magnitude is economically substantial: comparing institutions in the same state, the estimate implies a 48% relative decline in UK master's enrollment for institutions in the top decile of exposure (a pre-reform stay rate of 39.4%)

compared with those in the bottom decile (18.6%). Thus, the reform disproportionately reduced enrollment from institutions whose graduates had previously been more likely to remain in the UK for work, consistent with post-study employment opportunities being an important determinant of international education choices.

4.2 Heterogeneous Responses by Job Tier

Our main results may mask important heterogeneity in policy responses across institutions, since exposure depends not only on how likely an institution's graduates were to remain in the UK before the reform, but also on what jobs they found upon graduation. Consider two institutions with identical pre-reform stay rates, where graduates of the first stayed predominantly for high-skill jobs and graduates of the second predominantly for low-skill jobs. Because the policy specifically targets low-skill employment, we would expect students from the latter institution to lose a larger share of the post-graduation opportunities and thus experience a larger decline in UK master's enrollment. The estimates in Table 1, however, capture only the average response across these different institutions that might be unequally exposed to the reform. Therefore, in Table 2, we decompose the pre-reform UK stay rate into three mutually exclusive components based on the occupation category of graduates' first jobs, advanced-degree-level Level, bachelor-degree-level Level, and low-skill, each expressed as a share of the institution's 2001–2007 UK master's cohorts, and examine the policy response associated with each component. If the policy successfully achieved its intended goal of curbing visa abuse in low-skill employment, we would expect the differences in enrollment response to be concentrated on the low-skill margin, weaker on the bachelor-degree-level margin, and close to zero on the advanced-degree-level margin, whose graduates are most aligned with "best and brightest" talent that the UK government sought to retain.

The results reflect that gradient, but not significantly. Columns (1) and (3) of Table 2 replicate the estimates in columns (1) and (4) of Table 1, respectively. Column (2) shows magnitudes of the enrollment response across all three types of stay rates: -2.21 on the advanced-degree-level share, -3.05 on the bachelor-degree-level share, and -4.40 on the low-skill share. Wald tests reject neither the equality of the bachelor-degree-level and low-skill coefficients ($p = 0.16$) nor that of the advanced-degree-level and bachelor-degree-level coefficients ($p = 0.56$). Column (4) uses the Empirical Bayes measure of the stay rate and shows larger responses across all three groups, and the corresponding coefficient differences remain statistically insignificant ($p = 0.30$ and $p = 0.86$). These results suggest that the policy did not only displace students on the low-skill employment pathway; rather, it also reduced enrollment among students with higher human capital who would go on to work in

high-skill occupations.

There are two potential explanations for this unintended consequence for the UK government. First, access to high-skill jobs might also have declined substantially following the closure of the PSW visa route, leading some of these students to leave the UK. Second, students faced uncertainty about which type of job offer they would ultimately receive. Thus, even if the policy did not directly reduce access to high-skill jobs, removing the low-skill alternative lowered the option value of a UK master's even for students who would ultimately have found high-skill work. As a result, they were deterred and became less likely to apply for UK master's programs. The two mechanisms operate at different stages and imply different policy counterfactuals. In the next section, we turn to our structural model to distinguish between and quantify these two channels.

Table 2: Effect of Exposure by Job Tier on UK Master's Enrollment Change

	(1)	(2)	(3)	(4)
Stay Rate × Post	-3.130***		-4.488***	
	(0.490)		(0.739)	
Adv. Job Stay Rate × Post		-2.211		-3.823
		(1.442)		(2.696)
Bach. Job Stay Rate × Post		-3.050***		-4.282***
		(0.508)		(0.791)
Low. Job Stay Rate × Post		-4.395***		-5.654***
		(0.982)		(1.383)
Exposure	Raw	Raw	EB	EB
p-value: Bach. = Low.		0.163		0.304
p-value: Adv. = Bach.		0.555		0.860
Observations	690	690	690	690
Clusters	138	138	138	138

Notes: The dependent variable is the number of UK master's entrants from an undergraduate-institution cluster in a cohort; all columns are Poisson pseudo-maximum-likelihood estimates with cluster and state-by-cohort fixed effects, on the 2008–2010 and 2012–2013 cohorts. Columns 2 and 4 decompose the pre-reform UK stay rate into three mutually exclusive components by the first job's occupation: advanced-degree-level, bachelor-degree-level, and low-skill, each expressed as a share of the institution's 2001–2007 UK master's cohorts. Columns 3 and 4 replace each raw share with its empirical-Bayes counterpart. The rows labeled p report Wald tests of equal coefficients on the indicated pair of components. Standard errors in parentheses are clustered by undergraduate institutions (clusters). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5 Model

The reduced-form evidence in the previous section suggests that Indian students do react to the change in employment perspectives by adjusting their masters' enrollment in the

UK. Motivated by this discovery, we build a two-stage dynamic discrete choice model in this section to account for students' joint decision of master's study location and post-graduation work destinations.

5.1 Two-stage Model Structure

In our model, each individual solves a sequential two-stage discrete choice problem for her master's and work destinations, as illustrated in Figure 4. In the first stage, the individual chooses a destination country j_1 from the master's choice set $\mathcal{J}_1$. After enrolling in a master's program in the first stage, the individual graduates in L_{j_1} years and enters the second stage, where L_{j_1} is the average master's program length in country j_1. In the second stage, the individual with a first-stage choice j_1 chooses from the second-stage choice set $\mathcal{J}_W(j_1)$, which consists of all the available work locations for this individual plus the choices of pursuing a second master's or a doctoral degree. The exact specifications of $\mathcal{J}_1$ and $\mathcal{J}_W$ will be described in more details later. Individuals solve the sequential two-stage problem with the awareness that their first-stage choice will also affect their second-stage choice set. Therefore, a student choosing where to enroll is also implicitly choosing a set of post-graduation opportunities.

At the high level, three features of the model carry the analysis. First, the two stages are linked by an option value. The expected value of the post-graduation choice set enters the first-stage utility through its inclusive value, so labor-market prospects shape enrollment even though they are realized only after graduation. Second, as shown later, labor-market choice is not a matter of preference alone. We represent it as a stochastic choice set: a host-country job offer is available to a given student only if it is realized, so the second-stage choice set itself is random. Finally, we assume the abolition of the post-study work visa operates on these realization probabilities rather than on the payoffs of the jobs. The policy lowers the probability that a host-country job is realized, which directly reduces the rate at which graduates remain in the host country. Furthermore, it also operates on the first stage by shifting the inclusive value, changing both master's enrollment and the composition of master's students.

5.2 Individual Characteristics

Each individual i is characterized by a vector of observed discrete state variables Ω_i and an unobserved latent ability type z_i. The observed states $\Omega_i = (M_i, G_i, O_i, t_i)$ consist of the individual's bachelor's institution-tier-major (hereafter referred to as tier-major) type (M_i), bachelor's attainment G_i, Indian state of undergraduate school (O_i), and the master's enrollment year (t_i). Appendix G.1 provides detailed definitions of these observed state

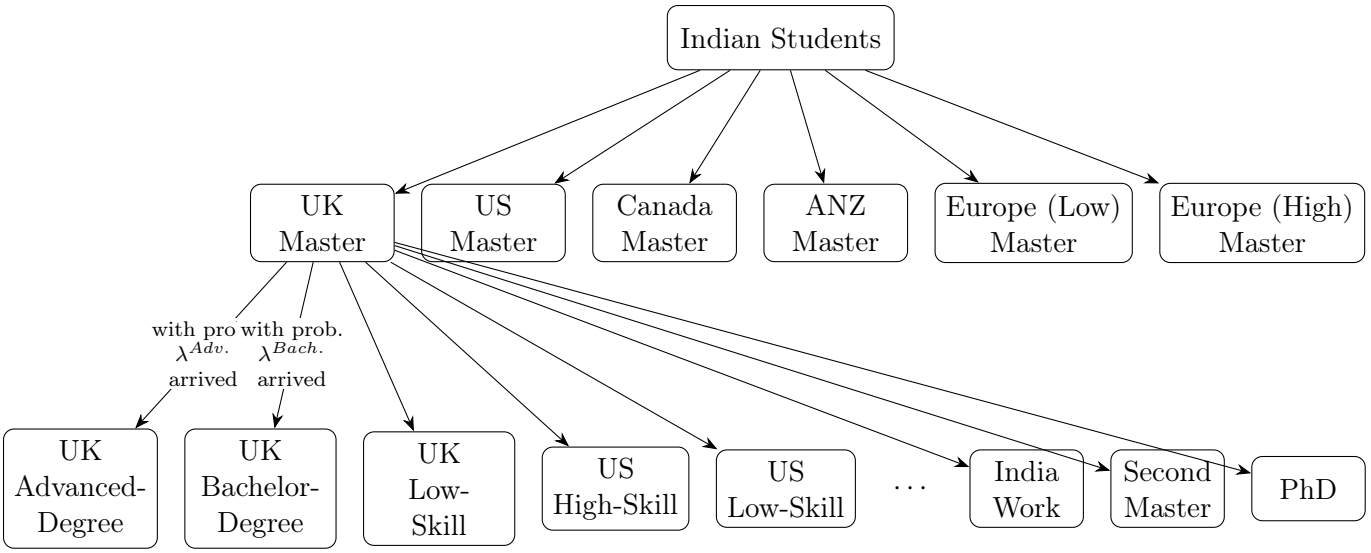


Figure 4: Two-stage sequential choice structure.

variables, along with how they enter the model. Compared to our reduced-form analysis in Section 4, we load the heterogeneity in the individual's bachelor institution into the institution-tier-major type M_i. We choose this method mainly for two reasons. First, we would like to preserve enough statistical power for our estimation, as further disaggregation by bachelor's institution could yield small cell sizes. Second, clustering at the bachelor's institution level would lead to a significantly larger parameter space, imposing a larger challenge to computational tractability.

The unobserved latent type $z_i \in \{H, L\}$ is a binary variable and corresponds to either high ability (H) or low ability (L), with probability $(\pi_z(m, g), 1 - \pi_z(m, g))$, depending on the tier-major type m and undergraduate academic attainment g. This unobserved type $z_i \in \{H, L\}$ is introduced to capture the correlation that students with higher ability may be more likely to receive a host-country job offer and to remain, may enjoy a larger foreign wage premium, and may hold different preferences over master's destinations.

Apart from the observed state variables Ω_i and the unobserved latent ability type z_i, there is one additional latent policy-adjustment type for the 2011 enrollment cohort. As discussed in Section 2.3, the 2011 cohort might or might not have adjusted to the reform given the timing of the announcement in March 2011. Thus, we model the 2011 cohort as a mixture of two types, with shares τ and $1 - \tau$ respectively. Share τ of the 2011 cohort made their decisions as if the old regime still persisted, while $1 - \tau$ of them adjusted to the new visa regime in their education decisions. The two latent types in the 2011 cohort only differ in their awareness of the new visa policy in the first-stage decision, and none of them could apply for the PSW visa when they graduated in 2012.

5.3 The Second-stage Problem

In this section we describe the individual's second-stage problem. In the second stage, the individual graduates from the master's program in country j_1 and chooses a post-graduation destination.

5.3.1 Second-stage Choice Sets

Based on the popular work destinations, we classify the world as a set $\mathcal{J}$ of the following regions:

$$\mathcal{J} = \{\text{UK, US, Canada, Australia and New Zealand, Europe, Others, India}\}$$

We also classify the skill requirements for jobs into three tiers, $\mathcal{Q} = \{A, B, L\}$, where A refers to Advanced-Degree Level Jobs, B refers to Bachelor-Degree Level Jobs, and L refers to Low-skill Jobs, as discussed in Section 3.3.

To focus on our question of interest and keep the model at the minimum necessary complexity, we make a few simplifying assumptions on the job types in each country. First, since the country of interest is the UK, we merge the Advanced-degree level job type A and Bachelor-degree level job type B into one category H in all non-UK countries. This simplification reduces the parameter dimension while keeping the low-skill and high-skill distinction. Additionally, we assume the same merge of A and B jobs in the UK if the individual finished master's education outside of the UK. These simplifying assumptions allow us to reduce the dimension of our problem while keeping our focus on the employment and enrollment prospects of the UK master's students.

The second-stage choice set $\mathcal{J}_W(j_1)$ consists of all the available job options mentioned above to the individual with first-stage choice j_1, plus the mutually exclusive options of pursuing a second master's or a doctoral degree. Each job option k is represented by a pair (j_2, q) of the work location $j_2 \in \mathcal{J}$ and the job skill requirement q. We take a parsimonious approach to the options of a second master's or a doctoral degree. We do not distinguish the location of further degrees, but allow their values to depend on the location j_1 of the first master's program.

The set of available second-stage job options depends on the individual's first-stage choice j_1. A student with non-UK master's education from the first stage ($j_1 \neq$ UK) faces a deterministic second-stage choice set. She can choose either a high-skill (H) or a low-skill (L) job in any country in $\mathcal{J}$, and she may also choose to pursue a second master's degree or a PhD degree. Therefore, her second-stage choice set $\mathcal{J}_W(j_1)$ is deterministic and specified

as follows.

$$\mathcal{J}_W(j_1 \neq \text{UK}) = \{(j_2, H), (j_2, L) : j_2 \in \mathcal{J}\} \cup \{\text{2nd Master}, \text{PhD}\}$$

A student with a UK master's degree ($j_1 = \text{UK}$) faces a different situation and chooses from a stochastic choice set. Her available choices in non-UK countries are still deterministic: she can choose either an H or an L job in any non-UK country. However, her choice set for UK jobs is random. The low-skill job in the UK is always available, but the high-skill jobs A and B only show up stochastically. Specifically, we assume A and B jobs arrive independently conditional on an individual's characteristics, and the arrival probabilities are $(\lambda^{Adv.}_{pre}, \lambda^{Bach.}_{pre})$ for the pre-reform cohorts ($t \leq 2010$) and $(\lambda^{Adv.}_{post}, \lambda^{Bach.}_{post})$ for the post-reform cohorts ($t \geq 2011$). Therefore, the set of available UK jobs for the UK masters is stochastic, and is specified as J^r_{UK} below.

$$\mathcal{J}^r_{UK} = \begin{cases} \{(\text{UK}, A), (\text{UK}, L)\}, \text{ with probability } \lambda^{Adv.}(1 - \lambda^{Bach.}) \\ \{(\text{UK}, B), (\text{UK}, L)\}, \text{ with probability } \lambda^{Bach.}(1 - \lambda^{Adv.}) \\ \{(\text{UK}, A), (\text{UK}, B), (\text{UK}, L)\}, \text{ with probability } \lambda^{Adv.}\lambda^{Bach.} \\ (\text{UK}, L), \text{ with probability } (1 - \lambda^{Adv.})(1 - \lambda^{Bach.}) \end{cases}$$

$$\mathcal{J}_W(j_1 = \text{UK}) = \underbrace{\{(j_2, H), (j_2, L) : j_2 \neq \text{UK}\} \cup \{\text{2nd Master}, \text{PhD}\}}_{\text{Deterministic: Non-UK Jobs and Further Education}} \cup \underbrace{\mathcal{J}^r_{UK}}_{\text{Stochastic}}$$

It is worth noting that we expect $\lambda_{post} < \lambda_{pre}$ due to the policy-induced increase in hiring frictions (See Section 2.4). Additionally, we assume that the job-finding probabilities vary by the tier-major type M_i, bachelor's attainment G_i, and the unobserved ability type z_i, with the following specification for each job type $q \in \{Adv., Bach.\}$:

$$\lambda^q_{pre}(m, g, z) = \Lambda(\varrho^q_m + \varrho^q_g + \varrho^q_Z \mathbf{1}\{Z = H\}), \quad \lambda^q_{post}(m, g, z) = \gamma_q \cdot \lambda^q_{pre}(m, g, z) \tag{3}$$

where ϱ^q_m, ϱ^q_g, ϱ^q_Z are parameters that govern the job-finding probability with respect to different individual characteristics. $\Lambda(\cdot)$ is the logistic cdf, and $\gamma_q \in (0, 1)$ is the post-reform proportional multiplier.[32]

[32]This specification imposes two restrictions which are essential for the identification. *1. Proportionality*: The reform multiplies each job-finding probability by a scalar γ_q that is common to every (m, g, z). This is the model's formalization of a "major-blind" policy: a shortened search window and a reduction in vacancies scale an arrival rate by the same factor for everybody. *2. Additive ability shift.* ϱ^q_Z does not vary with the observed type m: $\Lambda^{-1}\lambda^q_{pre}(m, g, H) - \Lambda^{-1}\lambda^q_{pre}(m, g, L) = \varrho^q_Z$ for all observed types m, g.

5.3.2 Wage Equations and Beliefs

One of the main incentives for Indian students to study and work abroad is to earn higher expected wages upon graduation, so an important component of the model is the specification of the post-graduation wage structure. Since individuals decide on their master's destination without knowing the labor market circumstances at the time they graduate, on which they need to form expectations, we model both the workers' realized wages and their beliefs. We make these expectations operational by assuming a parsimonious stochastic structure for wages. In our specification, the deterministic component is determined by country, occupation, ability, and master-employment match fixed effects, and the stochastic component captures the macroeconomic conditions and follows an AR(1) process with zero mean. Details about the wage equation specification and associated wage beliefs are written in Appendix G.2.

5.3.3 Second-stage Utility Functions

In this section we describe our specification of the second-stage utilities and derive the associated value functions. We assume the utility of any job consists of a deterministic component v_2 and an i.i.d. EVT1 shock ε. Specifically, for every feasible work alternative $k = (j_2, q) \in \mathcal{J}_W(j_1)$,

$$u_{2t}(k|j_1, m, g, r, z) = \underbrace{\beta^W \bar{w}_t^{(2)}(j_2, q; j_1, m, g, z) + \xi_t(k|j_1, m, g, r, z)}_{v_{2t}(k|j_1,m,g,r,z)} + \varepsilon_{2t} \tag{4}$$

where ε_{2t} is an i.i.d. EVT1 error term with unit variance[33], and it is independent across choices and time. $\bar{w}_t^{(2)}(j_2, q; j_1, m, z)$ is the expected log wage of the first job of type q in country j_2 after graduation, as described in Section 5.3.2. β^W is a multiplier that translates the log wage of this first job into expected *lifelong* earning utility. The subscript r denotes Indian regions constructed from their bachelor institution's states.[34] We take an agnostic approach towards the non-wage component of the second-stage utility and absorb it into a reduced-form parameter $\xi_t(k|j_1, m, g, r, z)$. The parametrization and interpretation of this term ξ_t are illustrated in Appendix G.3.

Apart from landing on a job, some individuals may choose to pursue further education in the same country. For the two continued-education choices in $\mathcal{J}_W$, we collapse their utilities into preference parameters $\xi^{PhD}(j_1, m)$ and $\xi^{2ndMA}(j_1)$ respectively. For simplicity,

[33] We normalize the variance of this EVT1 error to 1, so all parameters should be interpreted in units of its variance.

[34] Southern is taken as the reference region in the specification below.

we assume their utilities are time-invariant, and they absorb all systematic determinants of the continued-education choice conditional on the master's destination j_1 and observed individual characteristics (m, g).

$$\begin{aligned} u_{2t}(PhD|j_1, m, g) &= \xi^{PhD}(j_1, m) + \xi_g^{PhD} + \varepsilon_{2t}^{PhD} \\ u_{2t}(2ndMA|j_1) &= \xi^{2ndMA}(j_1) + \varepsilon_{2t}^{2ndMA} \end{aligned} \tag{5}$$

where ε_t^{PhD} and ε_t^{2ndMA} are idiosyncratic EVT1 shocks i.i.d. across time and individuals, and they are independent of the work-alternative shocks in (4).

5.3.4 Second-stage Conditional Choice Probabilities

The EVT1 error terms in the utility functions give us an analytical expression for the second-stage conditional choice probabilities. The specific functional forms of the conditional choice probabilities differ between the UK master's and non-UK master's graduates. For ease of notation, we suppress the individual characteristics Ω in the expression for the conditional choice probabilities. However, it should be noted that all the second-stage value functions v_{2t} and parameters, and thus the conditional choice probabilities, vary across cells of individual characteristics.

For non-UK master's graduates ($j_1 \neq$ UK), their choice set is deterministic, and the EVT1 errors yield the standard logit conditional choice probabilities.

$$P_{2t}(k|j_1) = \frac{\exp v_{2t}(k|j_1)}{\sum_{k' \in \mathcal{J}_2(j_1)} \exp v_{2t}(k'|j_1)}, \quad j_1 \neq \text{UK} \tag{6}$$

The observed second-stage conditional choice probability of the UK master's graduates is different from non-UK graduates. Using $\pi_J(\cdot; \lambda^{Adv.}, \lambda^{Bach.})$ to denote the distribution of the stochastic choice set $\mathcal{J}_W(j_1 = \text{UK})$, this observed conditional choice probability can be written as below.

$$P_{2t}(k|j_1 = \text{UK}) = \sum_{\mathcal{J}_W \in supp(\pi_J)} \pi_J(\mathcal{J}_W; \lambda_t^{Adv.}, \lambda_t^{Bach.}) P_{2t}(k|j_1 = \text{UK}, \mathcal{J}_2) \tag{7}$$

$$P_{2t}(k|j_1 = \text{UK}, \mathcal{J}_2) = \begin{cases} \frac{\exp v_{2t}(k|j_1)}{\sum_{k' \in \mathcal{J}_2} \exp v_{2t}(k'|j_1)}, & k \in \mathcal{J}_2 \\ 0, & k \notin \mathcal{J}_2 \end{cases} \tag{8}$$

Since the choice set of UK master's students is stochastic, the observed choice probability of UK master's students contains two layers of variations: the random realization of the choice set $\mathcal{J}_W(j_1 = \text{UK})$; and the choice probability conditional on its realization. Each

choice probability conditional on the choice set still has the standard logit structure, and the observed choice probability is the expectation taken over $\pi_J(\cdot; \lambda^{Adv.}, \lambda^{Bach.})$.

5.4 The First-stage Problem

In this section we describe the students' first-stage problem, where they choose a country for master's education.[35]

5.4.1 First-stage Choice Sets and Utility Functions

The first-stage choice set is deterministic. We choose five of the most popular destinations of master's study as the choice set $\mathcal{J}_1$:

$\mathcal{J}_1 = \{$UK, US, Canada, ANZ(Australia and New Zealand), Europe (Low), Europe (High)$\}$

The master's destination "Europe (Low)" comprises countries charging very low tuition during the period, such as France and Germany; "Europe (High)" comprises countries charging a non-negligible amount of tuition for master's programs, such as Ireland and Sweden.[36]

Similar to the second-stage specification, we assume the utility of a master's choice consists of a deterministic component v_{1t} and an idiosyncratic i.i.d. EVT1 error ε_{1t} with scale σ_1.

$$u_{1t}(j_1|m,g,r,z) = \underbrace{(\beta_1 + \beta_2 W_t(O)) P_{j_1,t} \chi_{j_1,t} L_{j_1}}_{\text{tuition cost}} + \underbrace{\varphi(j_1,m,g,r,z)}_{\text{consumption value}} + \underbrace{\kappa_{j_1} IV_t(j_1|m,g,r,z)}_{\text{continuation value}} + \varepsilon_{1t} \tag{9}$$

where $P_{j_1 t}$ is the annual tuition at country j_1 at year t in the destination currency, $\chi_{j_1,t}$ is the INR exchange rate to country j_1 at year t, L_{j_1} is the average program length, $W_t(O)$ is the per-capita GDP of the Indian state, and κ_{j_1} is a discount factor. The tuition cost translates the average total cost of the master's education to Indian rupees, and we allow the sensitivity of the tuition cost to vary with the state GDP.[37] $\varphi(j_1,m,g,r,z)$ is the time-invariant consumption value of the master's education, which depends on individual characteristics and the region r of the individual with the following specification. The detailed specification of φ is described in Appendix G.4. It is worth noting that this term φ not only captures the education taste, but also captures admission selectivity for different tiers of master's

[35] These destinations account for 93.6% of master's education in the data.

[36] Swedish master's programs were basically free before 2010 but charged substantial fees after 2011. Details can be seen in Appendix H.

[37] The detailed construction of country-specific and time-varying tuition fees is written in Appendix H.

programs across countries in a reduced-form way. For example, if the most selective programs are more difficult for Indian students to gain admission to in the United States, this would be reflected in a lower value of $\varphi_{\text{US,STEM Tier 1}}$. Finally, the discount factor κ_{j_1} varies by the destination country. It reflects two sources of variation: it is taken as a calibrated time discount factor $0.95^{L_{j_1}}$ specific to the program length; and it also reflects the difference in variance scales across locations, as we normalize the first-stage utility by the scale of the EVT1 error term σ_1.

5.4.2 Inclusive Values

The inclusive value $IV_t(j_1|m,g,r,z)$ of a master's choice j_1 is the expected value of the best post-graduation choice that the individual will take at the second stage. This term depends on the beliefs about wages and job accessibility in the second stage. We detail the equation and discussion in Appendix G.5.

5.4.3 First-stage Conditional Choice Probabilities

The first-stage choice set is deterministic for all individuals, so the EVT1 idiosyncratic shocks yield the standard logit conditional choice probabilities.

$$P_{1t}(j_1|m,g,r,z) = \frac{\exp v_{1t}(j_1|m,g,r,z)}{\sum_{j_1' \in \mathcal{J}_1} \exp v_{1t}(j_1'|m,g,r,z)} \tag{10}$$

6 Identification and Estimation

This section describes the identification strategy of the model parameters. The key identification challenge arises from the stochastic UK choice set. Because the UK choice set is stochastic, the observed probability that UK masters stay for a high-skill job in the UK equals the product of the job-finding probability and the choice probability conditional on receiving a job offer. Using S_{2t} to denote the observed share of UK master's graduates whose first job is a UK graduate-level job, the problem can be written as:

$$S_{2t}(j_2|j_1 = \text{UK}, M, G, R, Z) = \lambda \cdot P_{2t}(j_2|j_1 = \text{UK}, M, G, R, Z) \tag{11}$$

Since we do not observe each individual's offer arrival, the job-finding probability λ is not directly identified from the data, a standard identification challenge in job search models Flinn and Heckman (1982); Eckstein and Van den Berg (2007). Therefore, the parameters associated with the high-skill UK jobs cannot be identified by the standard inversion of the logit choice functions. To tackle this problem, we leverage the homogeneity of the

wage equation and second-stage cross-country non-wage utility in our model specification to construct student groups whose conditional choice probabilities remain unchanged at both stages between the pre-period and the post-period. These common conditional choice probabilities allow us to take the ratio between the observed first-stage and second-stage choice probabilities and eliminate the variation from the job-finding probabilities. We will elaborate on this identification strategy in detail later in this section.

Our identification strategy proceeds in five steps. For the ease of exposition, we first establish identification in the first four steps under a model without unobserved types, and then show that the unobserved ability type is identified in the fifth step. Each of the first four steps builds on the result of the previous step and identifies a subset of the model parameters. In the first step we show that we can identify the auxiliary wage equation and recover the wage parameters $(\alpha_{j_2,q}, \delta_m^w, \delta^{MA}, \rho)$. The second step takes these wage parameters as known and identifies the second-stage wage utility coefficient β^W and the non-wage utility ξ of all non-stochastic choices. The third step is the most important step in the identification argument. We proceed to recover the non-wage utility of the UK high-skill jobs and the job-finding probabilities in both the pre- and post-reform periods. Finally, in the fourth step, given all the wage and second-stage parameters, we turn to the first stage and identify the tuition sensitivity parameters (β_1, β_2) and the consumption value of education φ. The formal econometric setup and assumptions are written in Appendix I.1. We only present the intuition in Step 3, identification of the amenity and job-finding probabilities of the UK high-skill jobs, in the main text, and defer the remaining steps to Appendix I.2

6.1 Identification Intuition about the Stochastic Second-stage Choices

Sections I.2.1 and I.2.2 have identified everything in the second stage except the non-wage utilities of the stochastic high-skill UK jobs (A, B) and the job-finding probabilities $(\lambda^{Adv.}, \lambda^{Bach.})$. In this section we show the identification of the remaining second-stage parameters in a simpler version of our model and leave the identification of the full model in Appendix I.3. We consider a version with only one stochastic UK job choice H, so the only parameters to be identified are $(\xi_t(\text{UK}, H|j_1 = \text{UK}, m, g, o)$, λ^H_{pre}, $\lambda^H_{post})$. The identification of this simpler model captures the main intuition for the more complicated model, which contains two independent stochastic UK job choices (A, B).

The core idea of the identification is to exploit the exogenous policy reform that exclusively changes the job-finding probabilities and the existence of constant choice odds. The combination of these two features enables us to write the change in the observed choice probabilities (S_{2t}, S_{1t}) between the pre-reform period and the post-reform period, as the product

of the difference in the job-finding probabilities $\Delta\lambda_r$ and some constant term with just one unknown $\xi(UK, H)$ (up to quantities already identified). Therefore, we can take the ratio of the changes in the two choice probabilities and recover the non-wage utility parameters of the UK high-skill job $\xi(UK, H)$. After this, we can back out the job-finding probabilities λ^H_{pre} and λ^H_{post} by plugging in the utility parameters.

This identification result lies in the exclusion assumption we impose. We assume the policy reform only changes the job-finding probability λ of the stochastic job, but not the value of the job once it arrives. This assumption will be violated if the policy also affects the wage or the amenity of working in the U.K. (e.g., a hostile environment against immigrants). Consequently, the pre- and post-reform differences in the choice probabilities in both stages only reflect the change in λ and some known wage variations. By taking the ratio of the two pre- and post-reform changes, we can eliminate the impact of the change in λ and derive a function only of the utility parameters. We will pursue this idea formally in Appendix I.3.

Although restrictive, this exclusion restriction is plausible in this context. It could be violated if, for example, immigrants interpret the reform as a signal of xenophobia, reducing the amenity value of remaining in the U.K. However, the large income gap between the U.K. and India may limit the influence of such concerns. Figures A9 and A10 in the Appendix provide additional evidence consistent with the assumption.

First, if the reform reduced the utility of remaining in the U.K., we would expect higher departure rates among students who had already obtained PSW visas and were living in the U.K. before the reform. Using LinkedIn data, Figure A9 compares three-year retention rates among Indian individuals who began their first U.K. job in 2008 or 2010. The 2010 cohort could respond to the policy announcement by leaving the U.K. before its three-year retention rate was measured in 2013, whereas the three-year retention rate of the 2008 cohort is unaffected by the announcement. Nevertheless, the two cohorts exhibit similar three-year retention rates, suggesting that, conditional on having a job offer, the utility of living in the U.K. did not decline after the reform.

Second, Figure A10 shows that Indian enrollment in postgraduate research programs in the U.K. remained broadly stable around the reform. If the reform made prospective students feel a sense of anti-immigrant sentiment, we would expect enrollment to decline even in programs whose visa pathways were largely unaffected by the policy, i.e., students in research postgraduate programs. Together, these patterns suggest a limited role for the policy's impact on the amenity value of studying and working in the U.K.

6.2 Estimation and Inference

We estimate the model by weighted full-information maximum likelihood, jointly estimating the choice and wage parameters and integrating over unobserved types. Let ϑ denote the parameters governing stage-1 and stage-2 preferences, stochastic choice sets, prior type probabilities, and wages. Conditional on type z, individual i's likelihood contribution is the product of the probability of choosing master's destination j_i, the conditional probability of post-graduation outcome k_i, and, when observed ($d_i = 1$), the conditional density of log wages w_i. Integrating over types yields the log-likelihood:

$$\ell(\vartheta) = \sum_{i=1}^{N} \varpi_i \log \left\{ \sum_{z \in \{H,L\}} \pi_z(m_i, g_i) P_1\big(j_i \mid m_i, g_i, o_i, t_i, z; \vartheta\big) \right. \\ \left. \times P_2\big(k_i \mid j_i, m_i, g_i, o_i, t_i, z; \vartheta\big) \times f\big(w_i \mid k_i, j_i, m_i, g_i, t_i, z; \vartheta\big)^{d_i} \right\}. \tag{12}$$

where P_1 is the first-stage choice probability in equation (10); P_2 is the second-stage choice probability in equations (6) and (7); f is the normal density implied by the wage equation (17).[38]. More detailed discussion regarding the estimation is described in Appendix J.

As discussed in the final paragraph of Section 3.2, each observation receives a weight ϖ_i that adjusts for LinkedIn coverage. This weight equals the inverse coverage rate for the master's destination, normalized to have a mean of one, multiplied by a stay–return correction factor for UK master's graduates. We therefore estimate the model by weighted maximum likelihood, treating the weights as known. Appendix B.3 details their construction.

We report asymptotic standard errors based on the inverse of the observed information matrix, calculated as the negative finite-difference Hessian of the weighted log-likelihood evaluated at the parameter estimates. These standard errors should be interpreted as conditional on $\hat{\rho}$ and the observation weights.

The structural estimation sample is a subset of the full descriptive sample. Due to the tractability of the model, we impose some sample restrictions, such as only keeping the main master's destinations for Indian students. We report the share for each sample selection criterion, and details are provided in Appendix K.

[38] We maximize (12) with L-BFGS using analytic gradients

7 Model Estimates and Model Fit

In this section, we first present and discuss the parameter estimates in Section 7.1, particularly parameters related to the unobserved type. We then evaluate and discuss the model's fit in Section 7.2. The complete set of parameter estimates and standard errors is reported in Table A2.

7.1 Model Estimates

Figure A12 reports the estimated conditional prior probability that student i is the unobserved High type, $\Pr(Z_i = H \mid M_i, G_i)$, by undergraduate institution–major group M_i and undergraduate achievement G_i. As discussed later in Figure A13, the High type is characterized by a larger foreign wage premium, a higher probability of finding a high-skill job in the UK, and a stronger propensity to remain in other destination countries. High-type individuals are concentrated among Tier 1 STEM graduates and, to a lesser extent, non-STEM graduates, whereas their representation among Tier 2 and Tier 3 STEM students is close to zero. Within each institution–major group, students with a LinkedIn profile reporting honors, distinction, or a high GPA have a higher prior probability of being the High type. The estimate on this undergraduate education attainment shifter is positive and statistically significant at the 1 percent level.

Figure A13 and Figure 5 characterize the estimated differences between High- and Low-type students. To isolate the difference of each type-specific parameter, each panel of Figure A13 sets one coefficient to its estimated High-type value while holding all other type-specific parameters at their Low-type values. Thus, the wedge between the High and Low bars only represents the type difference in that coefficient, rather than the compounded effect of all of them. The first panel indicates that High-type students place a lower value on a UK master's relative to other countries at the Stage-1 destination choice, although the estimate is not significant. The second panel shows that they have a higher propensity to remain abroad after completing a master's degree in other destination countries, and the third panel indicates a larger foreign wage premium for the High type. These labor-market advantages are consistent with the interpretation of the High type as capturing greater underlying ability. Figure A11 in the appendix shows that, relative to Low-type students, High-type students are substantially more likely to pursue a master's degree in the United States and less likely to pursue one in the United Kingdom.

A more pronounced distinction between High- and Low-type students emerges in their access to advanced-degree-level jobs. Figure 5 reports the estimated job-offer arrival probabilities for advanced-degree-level and bachelor's-degree occupations, $\lambda^{Adv.}$ and $\lambda^{Bach.}$, re-

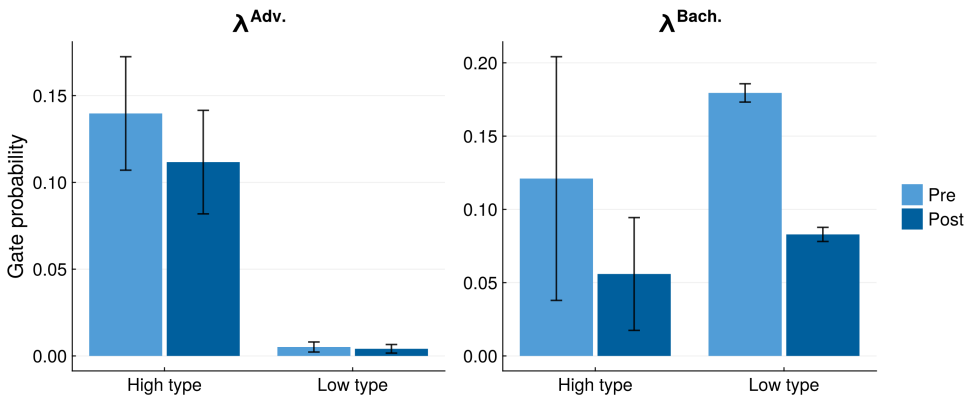


Figure 5: Job Matching Probability λ by Unobserved Types

Notes: The figure plots the estimated probabilities of obtaining an advanced-degree-level job offer in the UK, $\lambda^{Adv.}$ (left panel), and a bachelor-degree-level job offer, $\lambda^{Bach.}$ (right panel), by unobserved type, before and after the reform. Each bar averages the estimated probability over the eight (undergraduate institution–major group $\times$ undergraduate attainment) observed cells within the type, weighted by the model-implied probability of that type among UK master's students. The same weights are used in both pre and post periods, so within each type the pre–post wedge reflects only the change in λ alone. The reform enters as a proportional multiplier common to both types and all observed characteristic cells: $\lambda^g_{post} = \gamma^g \lambda^g_{pre}$. The estimated multiplier for advanced-degree-level jobs is $\hat{\gamma}^{Adv.} = 0.80$ (delta-method standard error 0.06), and $\hat{\gamma}^{Bach.} = 0.46$ (standard error 0.01), significantly below one at the 1 percent level.

spectively, by latent type and policy regime. The light-blue bars correspond to the pre-2011 Post-Study Work (PSW) regime, which allowed international graduates to remain in the UK for two years without employer sponsorship. Under this regime, an advanced-degree-level job enters the realized choice set with a probability of 14 percent for High-type students but with a probability close to zero for Low-type students. By contrast, the corresponding probabilities for bachelor's-degree jobs are 12 percent and 18 percent, respectively. These estimates indicate that access to advanced-degree-level employment, rather than employment more broadly, is a key dimension distinguishing the two latent types.

The dark blue bars in Figure 5 show the set of λ estimates corresponding to the post-reform regime. The abolition of the PSW visa reduced the arrival probability of advanced-degree-level job offers by 20%. By contrast, it reduced the arrival probability of bachelor's-degree job offers by 54 percent, substantially limiting graduates' access to these positions. Consistent with this change, Appendix Figure A14 shows that the model-implied probability of taking a first job in the UK fell from 38 to 22 percent for High-type students and from 26 to 11 percent for Low-type students. For both types, these declines were driven by reductions in the shares entering bachelor's-degree and low-skill jobs, while the share entering advanced-degree-level jobs declined in a much smaller magnitude.

We interpret the decline in $\lambda^{Bach.}$ as reflecting three institutional frictions on high-skill employment introduced unintentionally by the reform. First, the reform shortened the job-

search window, in which the post-graduation grace period fell from 24 months to 4 months. Our auxiliary data pattern is consistent with this time pressure binding: the share of Indian graduates who stayed in the UK for work and started their first job in the year of master's completion rose from 51% pre-reform to 62% post-reform.

Second, the reform made employer sponsorship necessary even for graduate-level jobs. Obtaining a Tier 2 visa requires a Certificate of Sponsorship (CoS) from a licensed employer, and acquiring a sponsor license entails both monetary and administrative costs, which may discourage small and medium-sized firms from participating.[39] Even among licensed employers, they need to bear the administrative costs of hiring a foreign worker that were absent under the PSW regime. For example, they must issue a CoS for each hire and file ongoing compliance reports, notifying the agency of any change affecting a sponsored worker, such as a change in workplace or salary.[40] The abolition of the PSW visa therefore reduced the number of employers willing and able to hire international graduates. These sponsorship costs are likely to be more consequential for bachelor's-degree occupations than for advanced-degree-level occupations. Workers in advanced-degree-level occupations may possess more specialized skills and generate greater match surplus for the firm, making employers more willing to bear the additional cost of sponsorship. In addition, foreign workers in bachelor's-degree occupations may be more easily replaced by domestic labor, whereas the specialized skills required in advanced-degree-level occupations may be relatively scarce in the domestic labor market.

Third, sponsorship requires not only that the occupation fall within the NQF level 6 category but also that the offer meet the minimum salary threshold. The latter had to be at least £20,000 or the occupation-specific rate specified in the relevant code of practice, whichever was higher. This salary requirement was non-trivial relative to prevailing graduate pay. Software-related positions, the most common bachelor's-degree jobs obtained by individuals in our sample, provide a useful benchmark. Contemporary evidence indicates that starting salaries for graduates from IT-related degree programs ranged from £17,200 to £26,000 in 2010, while the expected starting salary for IT professionals averaged approximately £25,500 in 2011.[41] Thus, although the average salary exceeded the Tier 2 floor, the requirement was

[39] Migration Advisory Committee (2012) reports consultation evidence that "Partners also told us that some employers had been put off even attempting to bring migrants in because of the constant change to the system and the bureaucracy involved in obtaining CoS, combined with the responsibility of being a sponsor." Consistent with these fixed costs of sponsorship weighing more heavily on smaller employers, Figure 2 in Migration Observatory (2024) shows a steep firm-size gradient: the share of workers holding a Skilled Worker visa rises from 1% at micro firms (fewer than 10 employees) and 4% at small firms (10–49 employees) to 15% at large firms (250 or more employees) in 2016.

[40] Migration Observatory (2024) reports that small and micro employers voiced concerns that inadvertent non-compliance with monitoring rules could cost them their licences.

[41] See *What Do Graduates Do? 2011*, pp. 22–23.

likely binding for entry-level positions at lower-paying firms or in lower-wage regions, thereby reducing the set of bachelor-degree-level jobs accessible to international graduates after the reform.

7.2 Model Fit

Figure A15 plots the share of students choosing each master's destination across cohorts, comparing the observed shares (solid lines) with their model-implied counterparts (dashed lines). The model reproduces two important features of the data. First, it captures the decline in the UK enrollment share since the 2011 cohort. Second, it matches the partial decline observed for the first post-reform cohort, i.e., the 2011 cohort. The estimated mixing weight governing this transition is $\hat{\tau} = 0.75$ (standard error 0.02). This estimate implies that approximately three-quarters of the 2011 cohort made their enrollment decisions under the beliefs in the pre-reform regime, leaving only the remaining quarter able to adjust immediately in the same application season.

Appendix Figure A16 reports the model's fit to the distribution of post-graduation outcomes by master's destination, separately for the pre-reform cohorts (2010 and earlier; left panel) and the post-reform cohorts (right panel). The blue bars denote observed shares, while the orange bars denote their model-predicted counterparts. As shown in the top row, the model closely reproduces the distribution of outcomes among UK master's graduates. In particular, it captures the two most substantial post-reform changes: the sharp decline in the share entering high-skill employment in the UK and the corresponding increase in the share returning to India.

8 Decomposition and Counterfactuals

In Section 8.1, we quantify the reform's effects on enrollment in master's programs, the supply of high-skill labor in the UK, and the welfare of Indian students. We first decompose the effects on enrollment and labor supply into their underlying mechanisms and assess the contribution of each mechanism along the enrollment and post-graduation employment margins. In particular, we distinguish a direct job-access effect, generated by reductions in the group-specific job-offer arrival probabilities, $\lambda^{Adv.}$ or $\lambda^{Bach.}$, from an indirect enrollment effect that operates through the reform-induced decline and compositional shift in master's enrollment. We then turn to student welfare and monetize the policy-induced losses using the estimated marginal utility of tuition, reporting them as tuition-equivalent compensating variation.

Then in Section 8.2, we exploit the estimated structural model to quantify how changes in labor-market access, wages, and tuition shape the effect of immigration policy on student enrollment. We first conduct a series of comparative-static exercises that separately reduce job-access probabilities, λ, lower wages, or increase tuition, and trace the resulting adjustments in students' education and post-graduation employment choices. We then simulate a policy that conditions PSW eligibility on undergraduate characteristics to assess whether targeted visa policies can yield more desirable outcomes.

8.1 Welfare Analysis and Policy Decomposition

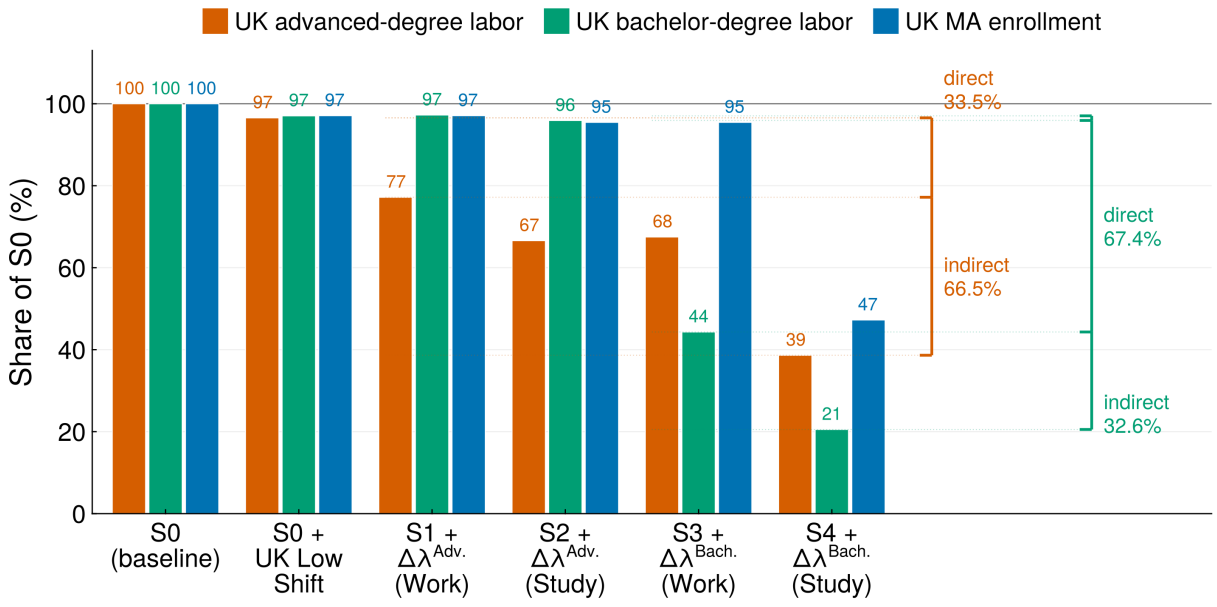


Figure 6: Decomposition

Notes: The figure decomposes the effect of the reform on three UK outcomes, advanced-degree-level employment, bachelor-degree-level employment, and master's enrollment, by activating the components of the reform one at a time. Each scenario is solved in closed form from the estimated model: holding each student's observed characteristics and sampling weight fixed, we compute the stage-1 and stage-2 choice probabilities of the post-reform cohorts (2011–2013) under the indicated combination of components, weight types by their prior probabilities, and aggregate. Each series is indexed to its own no-reform level ($S0 = 100$). Moving left to right, each group adds one component cumulatively: the post-reform shift in the value of low-skill UK employment; the cut in the advanced-degree-level offer probability $\lambda^{Adv.}$, applied first to realized job finding only, with enrollment decisions held fixed (Work), and then also to students' enrollment decision (Study); and the cut in $\lambda^{Bach.}$, split into the same two margins. The braces decompose the decline in advanced-degree-level and bachelor-degree-level high-skill employment into a direct effect—the work-stage step of that series' own change in λ—and an indirect remainder.

Figure 6 decomposes the post-reform declines in skilled labor supply and education enrollment across five sequential counterfactuals, each cumulatively incorporating an additional component of the reform. The baseline, $S0$, evaluates the post-reform cohorts (2011–2013)

under the pre-reform policy regime—using the pre-reform value of λ and retaining the low-skill employment option in the choice set—and normalizes the resulting level of UK high-skill employment to 100%.

The first counterfactual isolates the reform's intended channel by reducing access to low-skill UK jobs while leaving access to high-skill jobs unchanged. In principle, this change could have sizable spillover effects on master's enrollment and subsequent high-skill labor supply, if low-skill employment were an important component of the expected post-graduation value of a UK master's degree and thereby removing this option could induce prospective students to pursue programs elsewhere. The estimated model, however, implies that this channel is quantitatively small: UK advanced-degree-level employment, bachelor's-degree employment, and master's enrollment decline by only around 3% relative to the no-reform baseline. Although Appendix Figure A14 documents a large decline in access to low-skill jobs, these jobs contribute relatively little to the expected value of a UK master's degree. Eliminating access to them therefore generates only modest spillover effects on enrollment and high-skill labor supply.

The second counterfactual incorporates the post-reform change in $\lambda^{Adv.}$ but allows it to affect only stage-2 work choices, holding each student's stage-1 education choice fixed. By construction, this change affects only UK advanced-degree-level labor supply, leaving bachelor's-degree labor supply and master's enrollment unchanged. This change directly reduced advanced-degree-level labor supply by 20%. The third counterfactual then allows stage-1 education choices to respond to the decline in $\lambda^{Adv.}$. Although the lower probability of matching with advanced-degree-level jobs induces only 2% of the total student population to choose other options, it reduces advanced-degree-level labor supply by an additional 10% because those who switch are high-type students.

We next turn to the reform's primary channel: the 54% decline in $\lambda^{Bach.}$. The $S3{+}\Delta\lambda^{Bach.}$ (Work) counterfactual allows this decline to affect stage-2 work choices while holding stage-1 education choices fixed, thereby isolating the direct employment response. With this direct effect, UK bachelor's-degree labor supply falls from 96% to 44% of its baseline level, a decline of more than one-half. The $S4 + \Delta\lambda^{Bach.}$ (Study) counterfactual then allows students' education choices to respond. This further reduces bachelor's-degree labor supply from 44% to 21%. This education response also generates a sizable spillover to advanced-degree-level labor supply, which falls from 67% to 39% of its no-reform level as the pool of UK master's students shrinks. This highlights the importance of bachelor's-degree job access in the continuation value of a UK master's degree, even for high-type students: Figure 5 reports a bachelor-degree-level job appearance probability of approximately 0.12 for these students.

Taken together, the counterfactuals imply that the decline in $\lambda^{Bach.}$ is the primary source of the reform-induced responses along both the education and employment margins, whereas the reductions in access to low-skill and advanced-degree-level jobs play less important roles. The reform reduces UK bachelor's-degree labor supply among Indian students in our sample by approximately 80%. The decomposition attributes 67% of this decline to the direct stage-2 employment response and the remaining 32% to the indirect stage-1 enrollment response. The reform also reduces UK master's enrollment by 53%. Most notably, advanced-degree-level labor supply falls by 61%, and this decrease is mainly attributable to the indirect enrollment response to the decline in $\lambda^{Bach.}$. As discussed above, reduced access to bachelor's-degree jobs shifts high-type students away from the UK, thereby shrinking the pool of graduates eligible for advanced-degree-level employment.

Figure 7 reports the welfare loss from the reform for prospective Indian master's students, measured as the uniform increase in tuition across all master's programs that would generate an equivalent loss in utility. The average loss is largest for the 2011 cohort, at $9,500 (2011 USD), or 26.4% of the enrollment-weighted average program cost. This larger loss reflects the timing of the March 2011 announcement. Approximately 75% of this 2011 enrollment cohort retained the destinations chosen under pre-reform beliefs rather than reoptimizing in response to the policy. By contrast, the 2012 and 2013 cohorts could fully adjust their education choices, and their average welfare losses are approximately $8,000, equivalent to around 22% of average program costs. Across cohorts, the losses for low-type students are roughly twice those for high-type students. Low-type students derive greater education consumption value from UK master's programs, which may reflect the lower admission bars of these programs relative to those in other destinations like the US Consequently, weaker UK employment opportunities induce larger welfare losses on low-type students because shifting away from the UK entails a greater loss in education utility.

8.2 Counterfactual Regimes

Table 3 reports the simulated effects of the PSW abolition and four alternative policy changes. The first column shows that the abolition reduces UK master's enrollment among Indian students by 51.3%, corresponding to an annual loss of £73.9 million in tuition revenue. Because these estimates cover only Indian master's students, the total revenue loss could be larger once accounting further for students from other origin countries and those pursuing other degrees, including undergraduate degrees. The reform also reduces UK high-skill labor supply by around 1,900 workers per year. We apply the upper and lower bounds on the elasticity of annual TFP growth with respect to the foreign STEM employment share

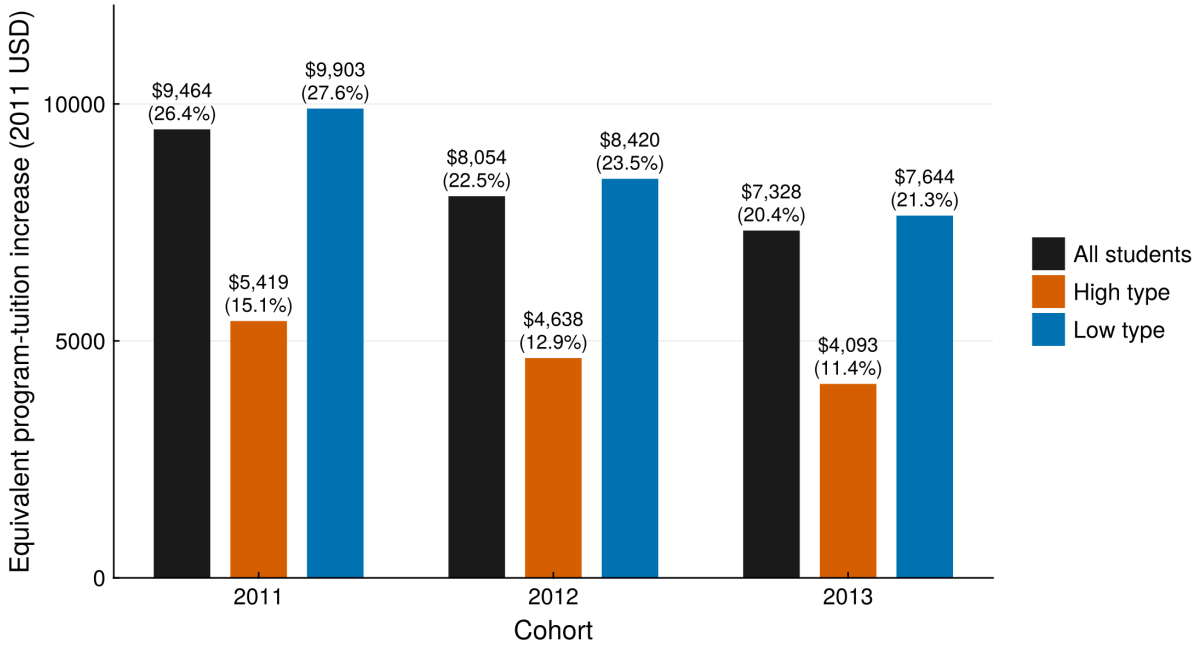


Figure 7: Welfare Decrease

Notes: The figure plots the welfare cost of the UK PSW-abolition reform to students, expressed as the equivalent increase in total master's program tuition across all countries, by cohort and by unobserved type (the three bars in each group). For each student and type we compute the change in ex-ante stage-1 welfare between the estimated post-reform regime and a counterfactual in which the effects of the reform are removed, and convert it to a compensating variation (CV) using the student's own estimated marginal disutility of tuition, $-(\beta_1 + \beta_2 \cdot \mathit{wealth})$; a positive bar is a welfare loss. All amounts are in 2011 US dollars, and the percentage below each dollar amount represents its share relative to the enrollment-weighted average total program cost of the 2011 cohort across destinations. For the 2012 and 2013 cohorts, the welfare change lets students re-optimize their master's destination under post-reform beliefs. For the 2011 cohort, a share $1 - \hat{\tau}$ of students re-optimize, while the remaining $\hat{\tau}$ hold the destination chosen under pre-reform beliefs but face the post-reform regime ($\hat{\tau} = 0.75$, delta-method standard error 0.02); the 2011 bar blends the two groups with weight $\hat{\tau}$. Averages use the sampling weights of the estimation sample and the prior type probabilities $\pi(z \mid M, G)$. Because the model has no outside option at stage 1, the welfare cost is conditional on pursuing a master's degree abroad.

Table 3: Counterfactual

	Counterfactual lever				
	PSW abolition $\lambda^{Adv.}-20\%$, $\lambda^{Bach.}-54\%$	UK wage -10%	$\lambda^{Adv.}$ -10%	$\lambda^{Bach.}$ -10%	UK tuition $+10\%$
Δ UK master's enrollment	−51.33%	−6.93%	−0.84%	−11.50%	−5.38%
Δ UK tuition income (£m/yr)	−73.9	−10.0	−1.2	−16.6	5.8
Δ UK advanced-degree labor	−59.98%	−8.07%	−15.92%	−8.21%	−4.94%
Δ UK bachelor-degree labor	−78.84%	−7.38%	−0.58%	−21.20%	−5.59%
Δ UK high-skill labor (persons/yr)	−1,889	−181	−41	−494	−135
Δ UK GDP (£bn / ‰), 10 yrs, ub	−4.99 (−1.64‰)	−0.48 (−0.16‰)	−0.11 (−0.04‰)	−1.30 (−0.43‰)	−0.36 (−0.12‰)
Δ UK GDP (£bn / ‰), 10 yrs, lb	−2.49 (−0.82‰)	−0.24 (−0.08‰)	−0.05 (−0.02‰)	−0.65 (−0.21‰)	−0.18 (−0.06‰)

Notes: The table reports the effect of the high-skill job-accessibility channel of the 2012 UK PSW-abolition reform (first column) and of four counterfactual policy scenarios (remaining columns). The first column applies the estimated post-reform offer probabilities ($\lambda^{Adv.}$ falls by 20%, $\lambda^{Bach.}$ by 54%); the remaining columns cut the UK wage by 10%, cut $\lambda^{Adv.}$ alone or $\lambda^{Bach.}$ alone by 10%, and raise UK tuition by 10%. All changes are computed from the estimated model for the 2011–2013 cohorts, with sampling weights, relative to a baseline with the reform's low-skill UK value shift active and both offer probabilities at their pre-reform levels. The GDP rows apply the Peri et al. (2015b) elasticity of annual TFP growth with respect to the foreign-STEM employment share (the elasticity $\varepsilon = 0.54$ upper bound, 0.27 lower), following the cumulation of Clemens et al. (2025).

estimated by Peri et al. (2015b), and follow the accumulation method in Clemens et al. (2025) to infer the loss of economic growth. The calculation indicates a UK GDP loss of approximately £2.5–5.0 billion after ten years.[42] For scale, relative to annual UK exports to India of approximately £20 billion, this cumulative loss is equivalent to 12–25% of one year's exports.

The second column in Table 3 considers a 10% decline in UK wages. Lower wages reduce the option value of post-graduation employment in the UK, inducing 6.9% of prospective master's students to choose another destination. The resulting shrinkage in enrollment reduces UK advanced-degree-level and bachelor's-degree foreign labor supply by 8.1% and 7.4%, respectively. The third and fourth columns simulate 10% declines in the job-access probabilities for advanced-degree-level and bachelor's-degree occupations. Because only high-type students are eligible for advanced-degree-level jobs and these students constitute a relatively small share of prospective students, a reduction in $\lambda^{Adv.}$ primarily affects advanced-degree-level labor supply through the direct channel, resulting in a 15.9% drop; master's enrollment and bachelor's-degree labor supply each decline by less than 1%. By contrast, bachelor's-degree jobs contribute to the option value of a UK master's degree for both high- and low-type students. A 10% reduction in $\lambda^{Bach.}$ therefore has broader effects, reducing UK master's enrollment by 11.5%, advanced-degree-level foreign labor supply by 8.2%, and bachelor's-degree foreign labor supply by 21.2%. Finally, a 10% tuition increase

[42] Clemens et al. (2025) estimate that a one-third decline in the number of foreign STEM graduates remaining in the US labor market reduces annual US GDP by $240 billion to $481 billion, an economic loss equivalent to the entire annual output of South Carolina or Wisconsin.

reduces enrollment by 5.4% and advanced- and bachelor's-degree labor supply by 4.9% and 5.6%, respectively. Nevertheless, the higher tuition more than offsets the enrollment decline, increasing annual UK tuition revenue by £5.8 million.

8.3 PSW Visa Eligibility Based on Undergraduate Characteristics

The analysis above suggests that first-best screening is difficult to achieve: abolishing the PSW visa also removes a valuable job-search window and increases students' frictions in accessing high-skill jobs. This trade-off calls for a more targeted design rather than uniform visa abolition for every student. One approach is to differentiate PSW visa eligibility across individuals, making high-ability students more likely to qualify. Indeed, a similar proposal was advanced by a UK think-tank to award graduate visas only to graduates of "the third of universities that require the highest grades to enter."[43] However, such proposals have raised concerns about inequality among both students and institutions. First, students from less affluent backgrounds may struggle to afford tuition at highly ranked UK institutions. Second, lower-ranked institutions outside London may find it even harder to attract international students, widening resource and capacity disparities across UK universities.[44] Moreover, screening based on the master's institution may induce strategic responses from highly ranked institutions, such as expanding enrollment, increasing tuition fees, or reducing educational quality.

By contrast, screening based on students' pre-master's educational backgrounds may address these concerns. Financial constraint plays a relatively minor role in sorting across undergraduate institutions compared to international education, and this approach may mitigate disparities and incentives for strategic responses among UK institutions. Yet, the key question is whether undergraduate characteristics predict subsequent labor market outcomes in the UK. If these characteristics have no predictive power, screening serves little purpose; otherwise, they may provide a useful basis for determining PSW visa eligibility and retaining high-skill workers.

Figure 8 illustrates the informativeness of undergraduate education for subsequent labor market outcomes. The gray bars normalize outcomes under the pre-reform visa regime to 100%, while the black bars show outcomes under the post-reform regime, in which no graduates are eligible for a PSW visa. Under universal abolition, each percentage point reduction in low-skill labor supply comes at the cost of a 0.62 percentage point decline in enrollment, a 0.72 percentage point decline in advanced-degree-level labor supply, and a 0.93

[43] See "Rishi Sunak plans crackdown on student visa salesmen," *Financial Times*, May 17, 2024.

[44] See "Student visa curbs are 'blunt instrument', warns St Andrews vice-chancellor," *Financial Times*, May 21, 2024.

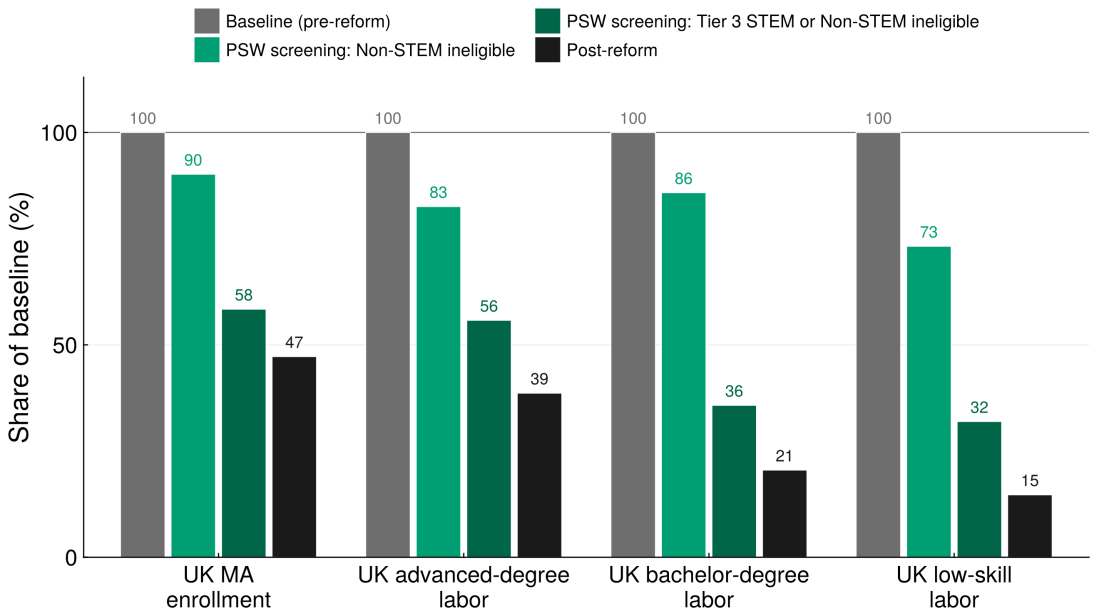


Figure 8: PSW Screening (Observed)

Notes: Each bar is the outcome under the specific regime as a percentage of the pre-reform baseline, pooled over the 2011–2013 cohorts with prior type weights. The four outcomes are UK master's enrollment, the number of UK master's graduates employed in advanced-degree-level, bachelor-degree-level and low-skill jobs. Under PSW screening, graduates in specific groups are ineligible for the post-study work route and face the estimated post-reform regime(λ_{post} and η^{L}_{post}), while every other group keeps the pre-reform regime value of parameters.

percentage point decline in bachelor-degree-level labor supply.

The light green bars represent a counterfactual that removes PSW visa eligibility only for non-STEM students, the group most likely to enter low-skill employment in the UK. Relative to the baseline, this restriction reduces low-skill labor supply by 27%, while reducing UK master's enrollment by only 10%, advanced-degree-level labor supply by 17%, and bachelor-degree-level labor supply by 14%. In other words, each percentage point reduction in low-skill labor supply comes at a smaller cost of 0.37 percentage points in enrollment, 0.63 percentage points in advanced-degree-level labor supply, and 0.52 percentage points in bachelor-degree-level labor supply.

A stricter eligibility rule that excludes both non-STEM students and Tier 3 STEM students further reduces low-skill labor supply to 32% of its baseline level, but entails larger declines of 42% in UK master's enrollment, 44% in advanced-degree-level labor supply, and 64% in bachelor-degree-level labor supply. Under this stricter screening rule, each percentage point reduction in low-skill labor supply comes at the cost of 0.62 percentage points in enrollment, 0.65 percentage points in advanced-degree-level labor supply, and 0.94 percentage points in bachelor-degree-level labor supply. These trade-offs are similar to those under universal abolition, except for a smaller loss in advanced-degree-level labor supply.

Overall, this counterfactual exercise suggests that, when first-best screening is infeasible and screening based on master's education may exacerbate inequality and induce strategic responses, screening PSW eligibility based on undergraduate education can improve policy efficiency, but the gains are concentrated in excluding non-STEM students.

UK government can achieve more efficient screening if it utilizes more information. Figure 9 examines the additional value of screening based on otherwise unobserved types. The results show that information about these types only allows the government to achieve meaningful gains from retaining advanced-degree-level labor. Excluding only low-type students with non-STEM undergraduate degrees reduces advanced-degree-level labor supply by just 1%, compared with 17% when all non-STEM students are excluded. Similarly, excluding only low-type students in the non-STEM and Tier 3 STEM groups leaves advanced-degree-level labor supply at 82% of its baseline level, compared with 56% when both groups are excluded entirely. By contrast, the additional gains in enrollment and bachelor-degree labor supply are small. These limited gains reflect the low prevalence of high types among non-STEM and Tier 3 STEM students. High types account for only 9.2% of non-STEM students without honors and 21.4% of those with honors, while nearly all Tier 3 STEM students are low types (Figure A12). Thus, observed undergraduate characteristics capture much of the information relevant for screening these groups. Because high types are particularly likely to enter advanced-degree-level employment, the main benefit of additional type information

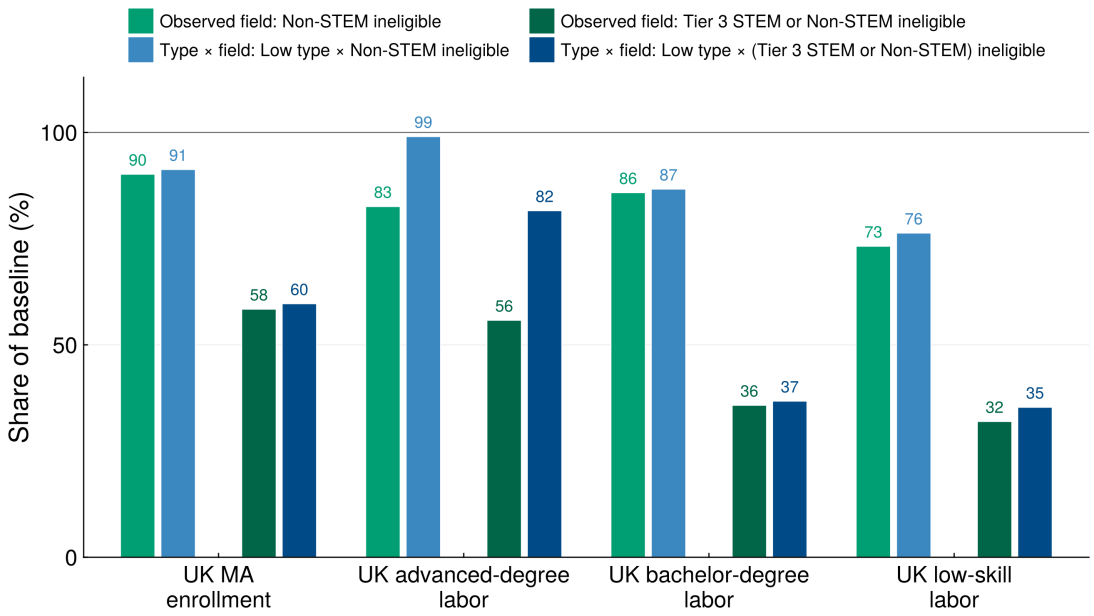


Figure 9: PSW Screening

Notes: Each bar is the outcome under the specific regime as a percentage of the pre-reform baseline, pooled over the 2011–2013 cohorts with prior type weights. The four outcomes are UK master's enrollment, the number of UK master's graduates employed in advanced-degree-level, bachelor-degree-level and low-skill jobs. Under PSW screening, graduates in specific groups are ineligible for the post-study work route and face the estimated post-reform regime(λ_{post} and η^{L}_{post}), while every other group keeps the pre-reform regime value of parameters.

comes from improved retention of talent being qualified for advanced-degree-level jobs.

9 Conclusion

This paper examines the United Kingdom's 2012 abolition of the Post-Study Work visa, a reform intended to restrict low-skill employment pathways for international students. Our analysis shows that the reform unintentionally and substantially increased hiring frictions for high-skill employment. Combining a novel longitudinal dataset of Indian master's students with a two-stage structural model of education and labor-market decisions, we estimate that the reform reduced UK master's enrollment among Indian students by 51%, equivalent to roughly £74 million in forgone annual tuition revenue. In addition, the reform reduced their labor supply in bachelor-degree-level and advanced-degree-level jobs by 80% and 60%, respectively. The decline in bachelor-degree-level labor supply was driven mainly by the direct reduction in job matching probabilities, while the decline in advanced-degree-level labor supply arose predominantly from deterring prospective high-ability students from enrolling. Counterfactual simulations show that screening PSW visa eligibility by undergraduate major reduces low-skill employment at a substantially lower cost to enrollment and high-skill labor supply than universal abolition, while further screening by undergraduate institution tier yields little additional efficiency gain.

Our findings carry broader implications for the design of immigration policy, especially work visas for high-skill immigrants. Reforms that seemingly target visa abuse in low-skill sectors may generate sizable unintended distortions in high-skill matching and retention that counteract stated policy objectives. Flexible post-study work routes facilitate graduates' transition into high-skill employment, which provide a necessary job search window for graduates and serve as low-compliance hiring channels through which employers absorb international talent without bearing the cost of visa sponsorship. This institutional interplay is not specific to the United Kingdom; the mechanisms identified here directly map into key debates surrounding high-skill labor supply in other host countries, most notably the Optional Practical Training (OPT) visa and the H-1B visa in the United States.

In addition, our novel method offers broader empirical implications for dynamic settings in which future options are only *stochastically* available. Without additional exclusion restrictions, the data cannot pin down the trade-off between the upstream benefit and the downstream cost (or vice versa). Specifically, the same baseline choice probabilities can be rationalized by multiple parameter families, depending on how much of the observed behavior is attributed to *constraints* compared to *preferences*. A policy shock that exclusively affects either component can serve as a very useful instrument for identification. This in-

sight is relevant across a wide range of empirical settings in which probabilistic restrictions on downstream choice sets shape upstream investment decisions, such as entrepreneurial entry under acquisition policy reform.

Several questions remain for future work. Our analysis focuses on the first employment spell after master's completion, leaving the medium- and long-run trajectories of international graduates open for further investigation. If low-skill host-country employment serves as a stepping stone into skilled occupations over time, then closing the low-skill entry route may depress long-run skilled retention by more than our model captures; conversely, if post-reform graduates entering skilled jobs later move into lower-skill ones, the reform may have induced occupational mismatch by channeling marginal applicants into ill-fitting skilled positions. Quantifying these lifecycle dynamics is a valuable direction for understanding immigration policy and the global competition for talent.

References

Abarcar, P. and C. Theoharides (2024). Medical worker migration and origin-country human capital: Evidence from us visa policy. *The Review of Economics and Statistics 106*(1), 20–35.

Altonji, J. and Z. Zhu (2025). Returns to specific graduate degrees: Estimates using texas administrative records. Technical report, National Bureau of Economic Research.

Altonji, J. G., P. Arcidiacono, and A. Maurel (2016). The analysis of field choice in college and graduate school: Determinants and wage effects. In *Handbook of the Economics of Education*, Volume 5, pp. 305–396. Elsevier.

Altonji, J. G. and L. Zhong (2021). The labor market returns to advanced degrees. *Journal of Labor Economics 39*(2), 303–360.

Amanzadeh, N., A. Kermani, and T. McQuade (2024). Return migration and human capital flows. Technical report, National Bureau of Economic Research.

Amuedo-Dorantes, C., D. Furtado, and H. Xu (2019). Opt policy changes and foreign born stem talent in the us. *Labour Economics 61*, 101752.

Amuedo-Dorantes, C., K. Shih, and H. Xu (2023). The implications of optional practical training reforms on international student enrollments and quality. *Economic Inquiry 61*(2), 253–281.

Arcidiacono, P., E. Aucejo, A. Maurel, and T. Ransom (2025). College attrition and the dynamics of information revelation. *Journal of Political Economy 133*(1), 53–110.

Arenas, A. (2021). Human capital portability and international student migration. *Journal of Economic Geography 21*(2), 195–229.

Barseghyan, L., F. Molinari, and M. Thirkettle (2021). Discrete choice under risk with limited consideration. *American Economic Review 111*(6), 1972–2006.

Batista, C., D. M. Costa, P. Freitas, G. Lima, and A. B. Reis (2025). What matters for the decision to study abroad? a lab-in-the-field experiment in cape verde. *Journal of Development Economics 173*, 103401.

Beine, M., M. Delogu, and L. Ragot (2020). The role of fees in foreign education: evidence from italy. *Journal of Economic Geography 20*(2), 571–600.

Beine, M., R. Noël, and L. Ragot (2014). Determinants of the international mobility of students. *Economics of Education review 41*, 40–54.

Beine, M., G. Peri, and M. Raux (2023). International college students' impact on the us skilled labor supply. *Journal of Public Economics 223*, 104917.

Beine, M., G. Peri, and M. Raux (2024). The contribution of foreign master's students to us start-ups. Technical report, National Bureau of Economic Research.

Bernstein, S., R. Diamond, A. Jiranaphawiboon, T. McQuade, and B. Pousada (2022). The contribution of high-skilled immigrants to innovation in the united states. Technical report, National Bureau of Economic Research.

Bontemps, C., J.-M. Robin, and G. J. Van den Berg (2000). Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation. *International Economic Review 41*(2), 305–358.

Bound, J., B. Braga, J. M. Golden, and G. Khanna (2015). Recruitment of foreigners in the market for computer scientists in the united states. *Journal of labor economics 33*(S1), S187–S223.

Bound, J., B. Braga, G. Khanna, and S. Turner (2020). A passage to america: University funding and international students. *American Economic Journal: Economic Policy 12*(1), 97–126.

Bound, J., B. Braga, G. Khanna, and S. Turner (2021). The globalization of postsecondary education: The role of international students in the us higher education system. *Journal of Economic Perspectives 35*(1), 163–184.

Bound, J., M. Demirci, G. Khanna, and S. Turner (2015). Finishing degrees and finding jobs: Us higher education and the flow of foreign it workers. *Innovation Policy and the Economy 15*(1), 27–72.

Brinatti, A. and X. Guo (2025). Third-country effects of us immigration policy. Staff Working Paper 2023-60, Bank of Canada. Last updated November 6, 2024.

Callaway, B., A. Goodman-Bacon, and P. H. Sant'Anna (2024). Difference-in-differences with a continuous treatment. Technical report, National Bureau of Economic Research.

Chen, M. (2024). The value of us college education in global labor markets: Experimental evidence from china. *Management Science 70*(2), 1276–1300.

Chen, M., J. Howell, and J. Smith (2023). Best and brightest? the impact of student visa restrictiveness on who attends college in the us. *Labour Economics 84*, 102385.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American economic review 104*(9), 2593–2632.

Clemens, M. A., J. Neufeld, and A. M. Nice (2025). Brain freeze: How international student exclusion will shape the stem workforce and economic growth in the united states.

Compiani, G. and Y. Kitamura (2016). Using mixtures in econometric models: A brief review and some new results. *The Econometrics Journal 19*(3), C95–C127.

Crown, D., A. Faggian, and J. Corcoran (2020). Foreign-born graduates and innovation: Evidence from an australian skilled visa program. *Research Policy 49*(5), 103945.

Demirci, M. (2019). Transition of international science, technology, engineering, and mathematics students to the us labor market: The role of visa policy. *Economic Inquiry 57*(3), 1367–1391.

Demirci, M. (2020). International students and labour market outcomes of us-born workers. *Canadian Journal of Economics/Revue canadienne d'économique 53*(4), 1495–1522.

Dimmock, S. G., J. Huang, and S. J. Weisbenner (2022). Give me your tired, your poor, your high-skilled labor: H-1b lottery outcomes and entrepreneurial success. *Management Science 68*(9), 6950–6970.

Dustmann, C., A. Lindner, U. Schönberg, M. Umkehrer, and P. Vom Berge (2022). Reallocation effects of the minimum wage. *The Quarterly Journal of Economics 137*(1), 267–328.

Eckstein, Z. and G. J. Van den Berg (2007). Empirical labor search: A survey. *Journal of econometrics 136*(2), 531–564.

Flinn, C. and J. Heckman (1982). New methods for analyzing structural models of labor force dynamics. *Journal of econometrics 18*(1), 115–168.

Flynn, R., B. Glennon, R. Murciano-Goroff, and J. Xiao (2024). Building a wall around science: The effect of us-china tensions on international scientific research. Technical report, National Bureau of Economic Research.

French, E. and C. Taber (2011). Identification of models of the labor market. In *Handbook*

of labor economics, Volume 4, pp. 537–617. Elsevier.

Ganguli, I. and M. MacGarvie (2025). International students, immigration policies and implications for innovation. Working Paper 34212, National Bureau of Economic Research.

Glennon, B. (2024). How do restrictions on high-skilled immigration affect offshoring? evidence from the h-1b program. *Management Science 70*(2), 907–930.

Heckman, J. (1990). Varieties of selection bias. *The American Economic Review 80*(2), 313–318.

Hu, S. and J. V. Winters (2025). Growing from the stem? opt classification and international students in economics. Discussion Paper 18336, IZA Institute of Labor Economics.

Hunt, J. (2011). Which immigrants are most innovative and entrepreneurial? distinctions by entry visa. *Journal of Labor Economics 29*(3), 417–457.

Jia, R., G. Khanna, H. Li, and Y. Xu (2025). The ripple effects of china's college expansion on american universities. Working Paper 34391, National Bureau of Economic Research.

Kane, T. J. and D. O. Staiger (2008). Estimating teacher impacts on student achievement: An experimental evaluation. Technical report, National Bureau of Economic Research.

Kasahara, H. and K. Shimotsu (2009). Nonparametric identification of finite mixture models of dynamic discrete choices. *Econometrica 77*(1), 135–175.

Kato, T. and C. Sparber (2013). Quotas and quality: The effect of h-1b visa restrictions on the pool of prospective undergraduate students from abroad. *Review of Economics and Statistics 95*(1), 109–126.

Kerr, S. P., W. R. Kerr, and W. F. Lincoln (2015). Skilled immigration and the employment structures of us firms. *Journal of Labor Economics 33*(S1), S147–S186.

Kerr, W. R. and W. F. Lincoln (2010). The supply side of innovation: H-1b visa reforms and us ethnic invention. *Journal of Labor Economics 28*(3), 473–508.

Khanna, G. (2025). From asia, with skills. Working Paper 34449, National Bureau of Economic Research.

Khanna, G. and N. Morales (2025). The it boom and other unintended consequences of chasing the american dream. Technical Report 460, Center for Global Development. Updated January 2025.

Khanna, G., K. Shih, A. Weinberger, M. Xu, and M. Yu (2025). Trade liberalization and chinese students in us higher education. *Review of Economics and Statistics 107*(5), 1291–1309.

Manzini, P. and M. Mariotti (2014). Stochastic choice and consideration sets. *Econometrica 82*(3), 1153–1176.

McFadden, D. (1972). Conditional logit analysis of qualitative choice behavior.

Migration Advisory Committee (2012, February). Limits on tier 2 (general) for 2012/13 and associated policies. Report, Home Office, United Kingdom. Accessed: 2026-07-07.

Migration Observatory (2024, January). Licenced visa sponsors. Commentary, The Migration Observatory. Accessed 7 July 2026.

Peri, G., K. Shih, and C. Sparber (2015a). Foreign and native skilled workers: What can we learn from the h-1b visa lotteries?

Peri, G., K. Shih, and C. Sparber (2015b). Stem workers, h-1b visas, and productivity in us cities. *Journal of Labor Economics 33*(S1), S225–S255.

Rosenzweig, M. R., D. A. Irwin, and J. G. Williamson (2006). Global wage differences and international student flows [with comments and discussion]. In *Brookings trade forum*, pp.

57–96. JSTOR.

Ruiz, N. G. and A. Budiman (2018). Number of foreign college students staying and working in us after graduation surges. *Pew Research Center*.

Shih, K. (2016). Labor market openness, h-1b visa policy, and the scale of international student enrollment in the united states. *Economic Inquiry 54*(1), 121–138.

Shih, K. (2017). Do international students crowd-out or cross-subsidize americans in higher education? *Journal of Public Economics 156*, 170–184.

Stuen, E. T., A. M. Mobarak, and K. E. Maskus (2012). Skilled immigration and innovation: evidence from enrolment fluctuations in us doctoral programmes. *The Economic Journal 122*(565), 1143–1176.

Teicher, H. (1963). Identifiability of finite mixtures. *The Annals of Mathematical Statistics 34*(4), 1265–1269.

UK Migration Advisory Committee (2018, September). Impact of international students in the UK. Report, Migration Advisory Committee, Home Office, London.

UNESCO Institute for Statistics (2025). UIS data browser: Total outbound internationally mobile tertiary students studying abroad. https://databrowser.uis.unesco.org/.

Zhu, J. L. (2024). Comparative immigration policies and the resource effects of international students in us higher education. Working Paper.

Appendix Tables

Table A1: Summary Statistics (Baseline, Pre-reform, 2008-2010)

	All	Tier 1 STEM	Tier 2 STEM	Tier 3 STEM	Non-STEM
Count	85,038	9,544	8,388	49,538	17,568
Share of sample	100.00%	11.22%	9.86%	58.25%	20.66%
Female	27.47%	21.19%	23.96%	26.54%	35.18%
Honours	3.34%	4.69%	1.60%	1.59%	8.36%
UK Master	31.33%	17.71%	24.79%	28.32%	51.46%
US Master	47.74%	61.15%	54.96%	52.79%	21.32%
ANZ Master	9.18%	4.98%	6.89%	7.94%	16.49%
Canada Master	3.58%	3.83%	3.10%	3.64%	3.52%
Europe Master	8.16%	12.34%	10.26%	7.32%	7.20%
India Work	21.38%	16.51%	20.41%	17.81%	34.53%
Abroad Work	57.59%	58.86%	58.01%	61.53%	45.61%
High-skill Work	69.33%	70.81%	72.19%	71.12%	62.09%
Low-skill Work	9.96%	4.83%	6.56%	8.50%	18.48%
PhD	7.84%	13.56%	9.37%	8.00%	3.55%

Notes: The sample is Indian international master's students whose master's programme starts in 2008–2010, i.e. before the March 2011 announcement and April 2012 abolition of the UK Post-Study Work visa. Columns 2–5 partition the sample by the bachelor type used in the model: three tiers of STEM bachelor institutions (Tier 1 most selective) and a pooled non-STEM category; *Honours* is the indicator for honours, distinction, or a high GPA on the bachelor's degree. The five *Master* rows are destination shares out of the five-group total (master's degrees outside the five groups are excluded), so they sum to 100% within each column; all other rows use the column's own number. The first-job rows describe the first post-master's experience. *India/Abroad Work* and *High- / Low-skill Work* are two splits of the same work spells (79.3% of the sample), each excluding the few spells with no recorded country or mapped NQF level, and *PhD* is a first spell recorded as a doctoral programme; the remainder includes second master's (2.0%) or no first spell observed (8.6%).

Appendix Figures

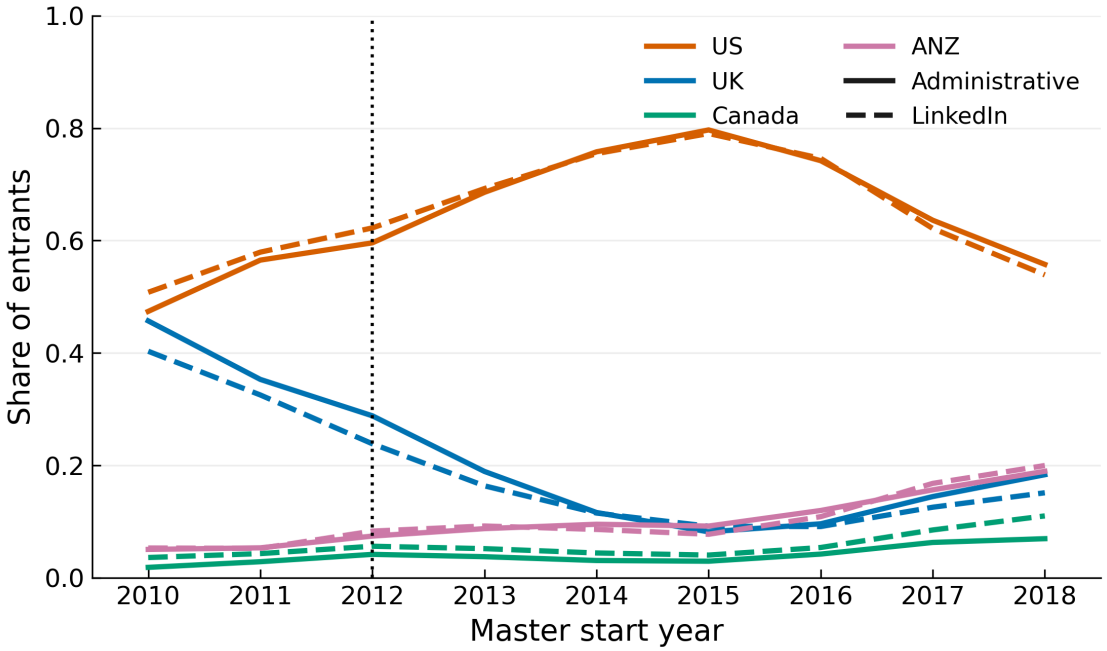


Figure A1: Shares of Indian master's entrants: administrative data v.s. LinkedIn

Notes: Each line is a destination's share of all Indian master's entrants across the four destinations in a given entry year, 2010–2018. Administrative entrant flows are built from the following sources. *United States:* F-1 visas issued by master's programs to Indian citizens in SEVIS (FOIA release, calendar year of program start, one count per individual). *United Kingdom:* India-domiciled postgraduate-taught enrollments at UK higher education providers from HESA (Table 8 to 2013/14, open-data Table 28 from 2014/15). *Canada:* Indian international-student graduates from ISCED-7 (master's) programs in Statistics Canada Table 37-10-0185. *Australia and New Zealand:* Indian master's commencements at public Australian providers among students new to Australia, from the Department of Education's international student data. New Zealand is excluded from both sides. Details of how we harmonize between datasets are described in Appendix Section B.1.1.

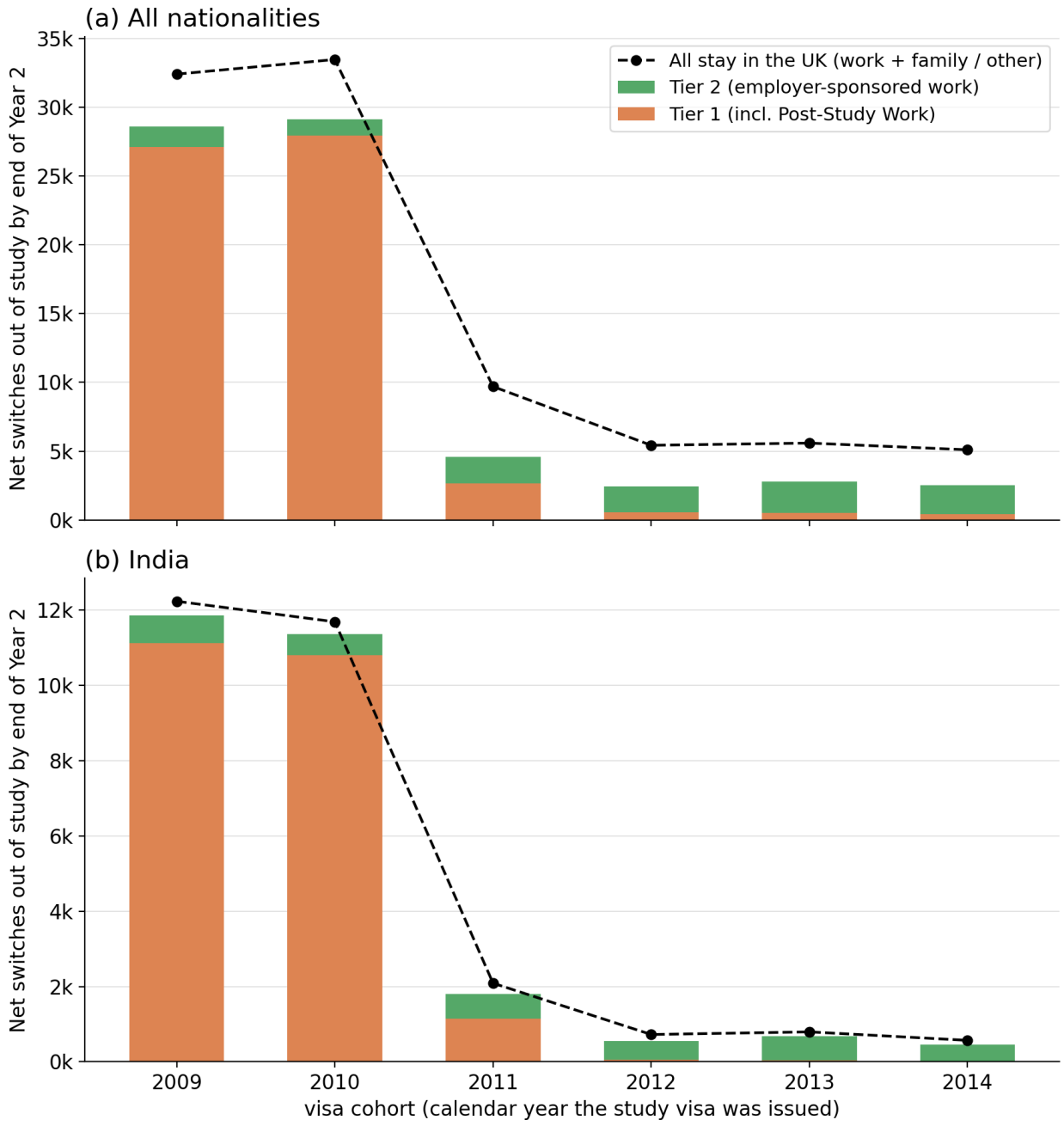


Figure A2: Study-to-Work Visa Transitions in the UK

Notes: Home Office Migrant Journey, 2016 edition, table MJ_03; main applicants only. Bars are the net change in the number holding each work status between the reference date and the end of Year 2; the dashed line is the drop in the number holding study leave minus the number holding expired visa. The dashed line is above the bars because it also includes Tier 5, other work routes, family and settlement leave, and citizenship. The Tier 1 series is the published category, which contains the General, Investor and Entrepreneur routes alongside Post-Study Work. Post-Study Work closed to new applicants on 6 April 2012.

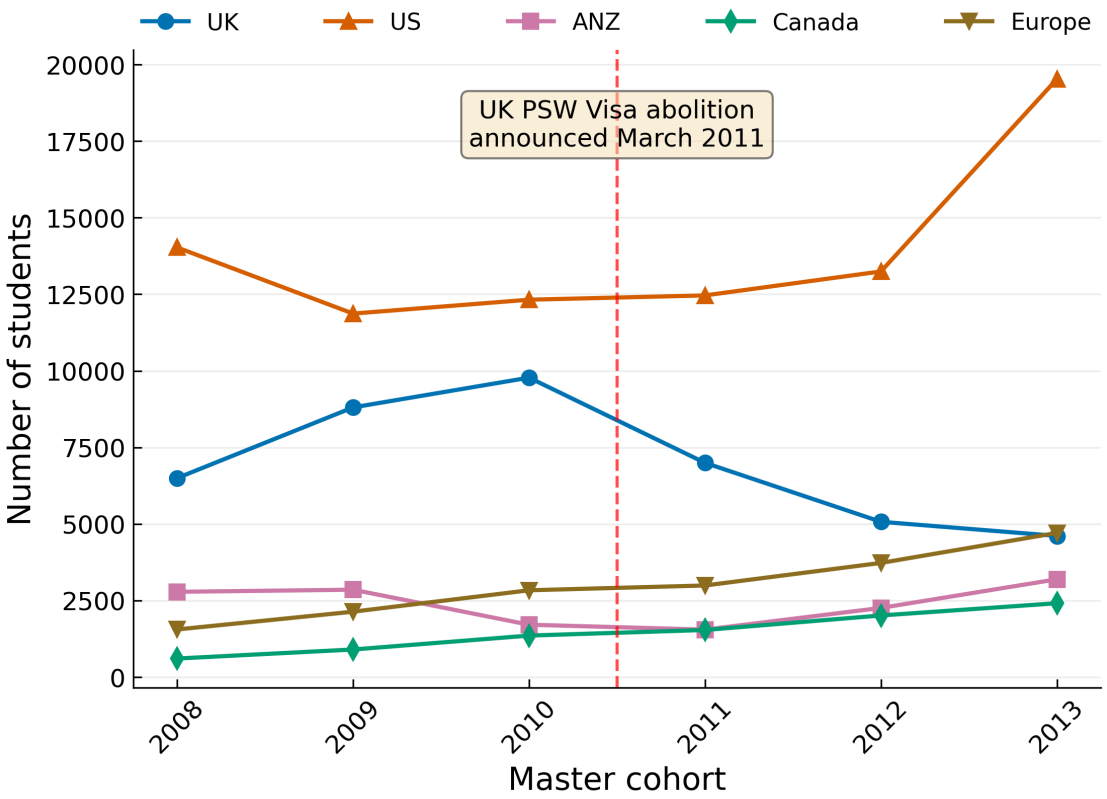


Figure A3: Indian Foreign Master's Students' Destination Choice by Year

Notes: The figure plots the number of Indian foreign master students who begin a foreign master's degree in each cohort year 2008–2013, by destination group. A small share of master students choosing countries outside the five groups (e.g. Singapore) are excluded. The dashed vertical line marks the March 2011 announcement of the abolition of the UK Post-Study Work visa, which took effect in April 2012.

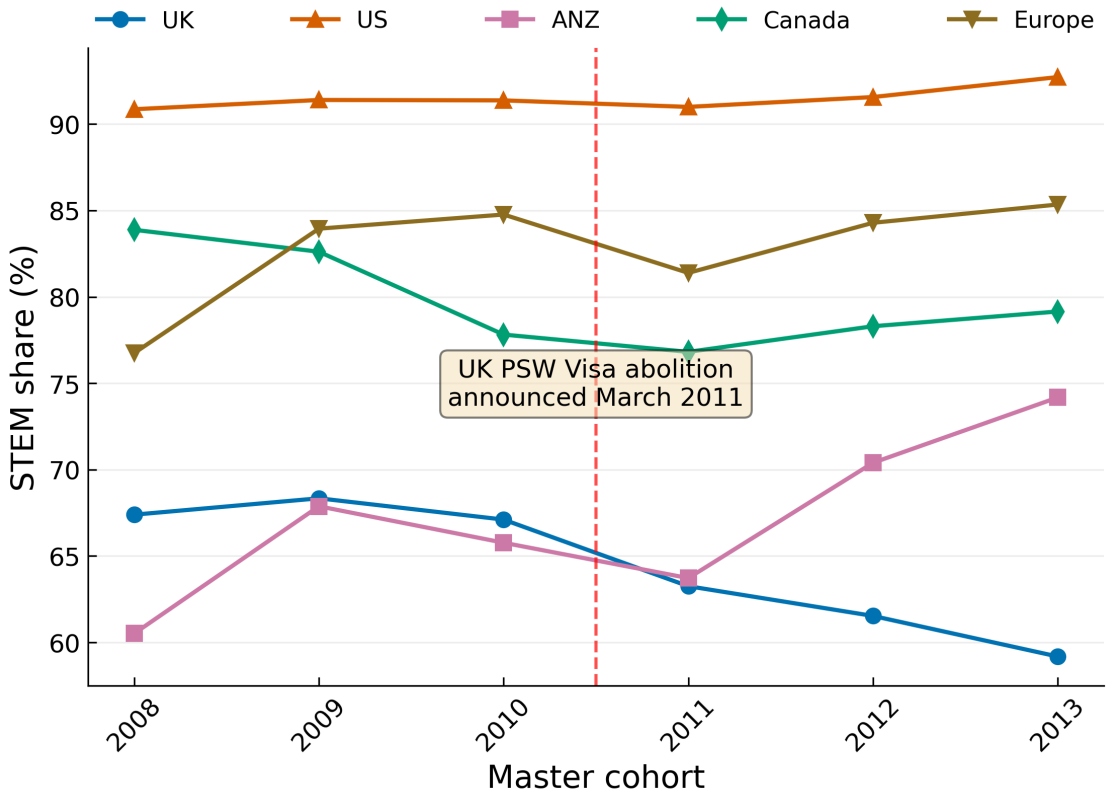


Figure A4: STEM Share By Destination Country

Notes: The figure plots, for each destination group and master start cohort, the share of Indian students whose bachelor's degree is in a STEM field. Master's degrees taken in Singapore or outside the five groups are excluded. The dashed vertical line marks the March 2011 announcement of the abolition of the UK Post-Study Work visa.

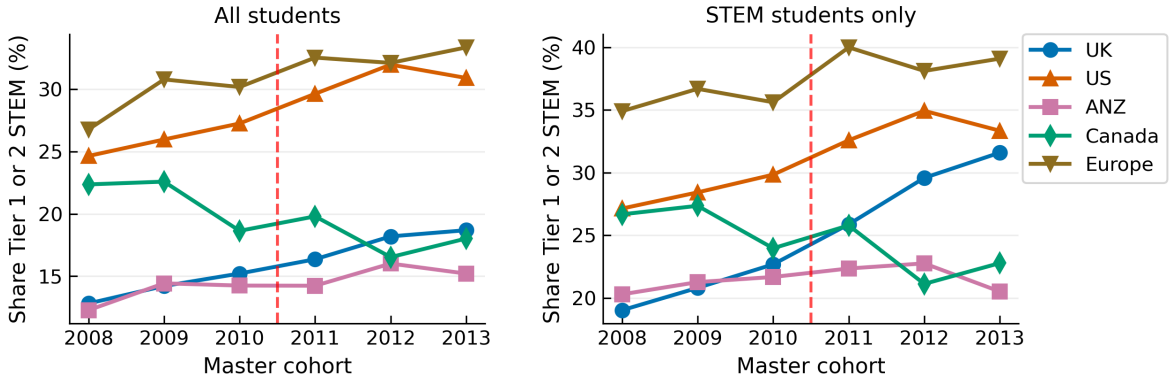


Figure A5: Tier 1 & 2 STEM Share by Destination Country

Notes: Both panels plot the share of students whose bachelor's STEM degrees are obtained in Tier 1 or Tier 2 undergraduate institutions, by destination group and master's start cohort. The two panels differ only in the denominator. The left panel divides by all students in the destination–cohort cell, while the right panel divides by the STEM students alone, so the right panel isolates tiers within STEM.

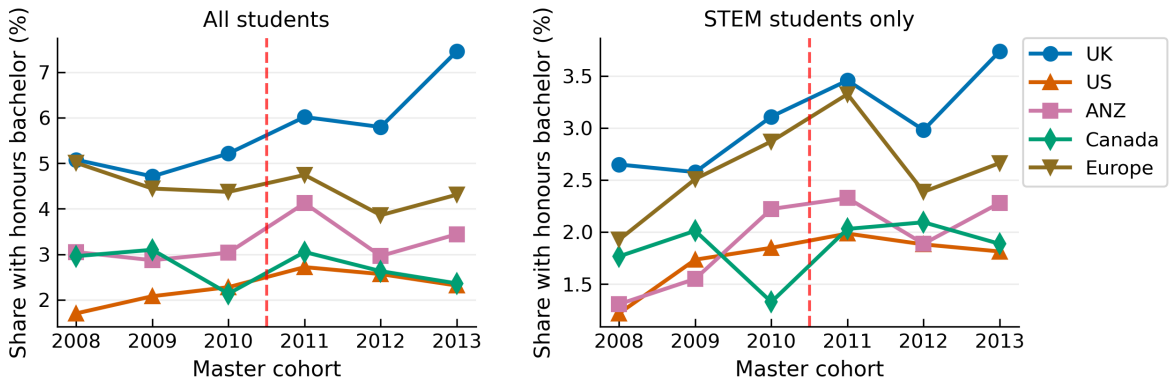


Figure A6: Honour Degree Student Share by Destination Country

Notes: Both panels plot the share of students with an honours, distinction, or high-GPA bachelor's degree, by destination group and master's start cohort. The left panel includes all students as the denominator, while the right only includes STEM students. Honours degrees are concentrated in non-STEM fields, so their share is roughly twice as high in the left panel as in the right.

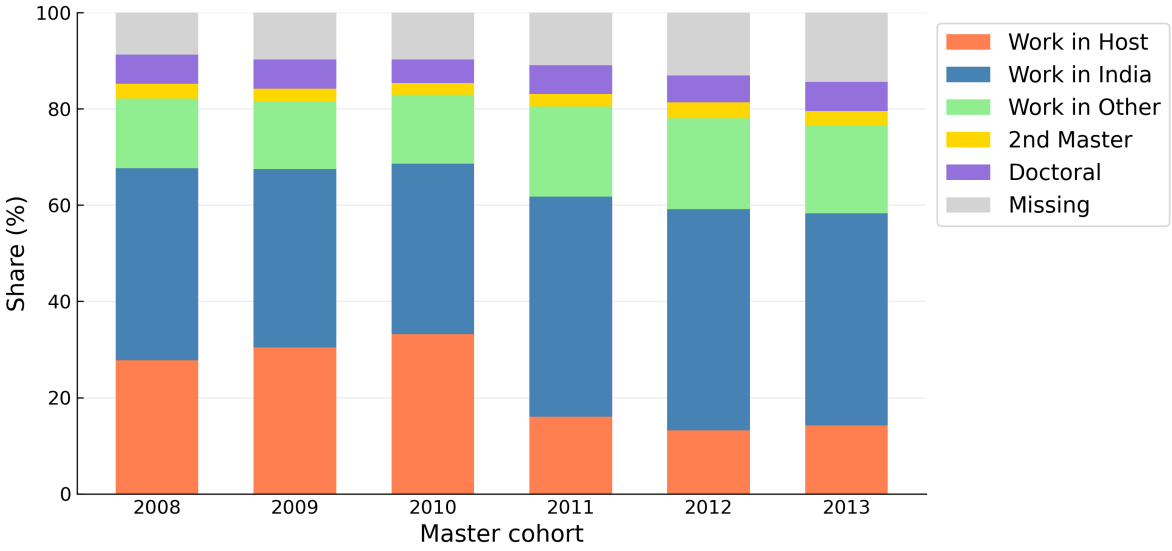


Figure A7: Post-Graduation Placement for UK Master

Notes: The figure decomposes the first post-master's spell of Indian masters studying in the UK into six mutually exclusive and exhaustive categories. *Work in Host* is a first job located in the UK, *Work in Other* a first job in a third country (including work spells with no recorded country), and *Missing* means no first spell is observed.

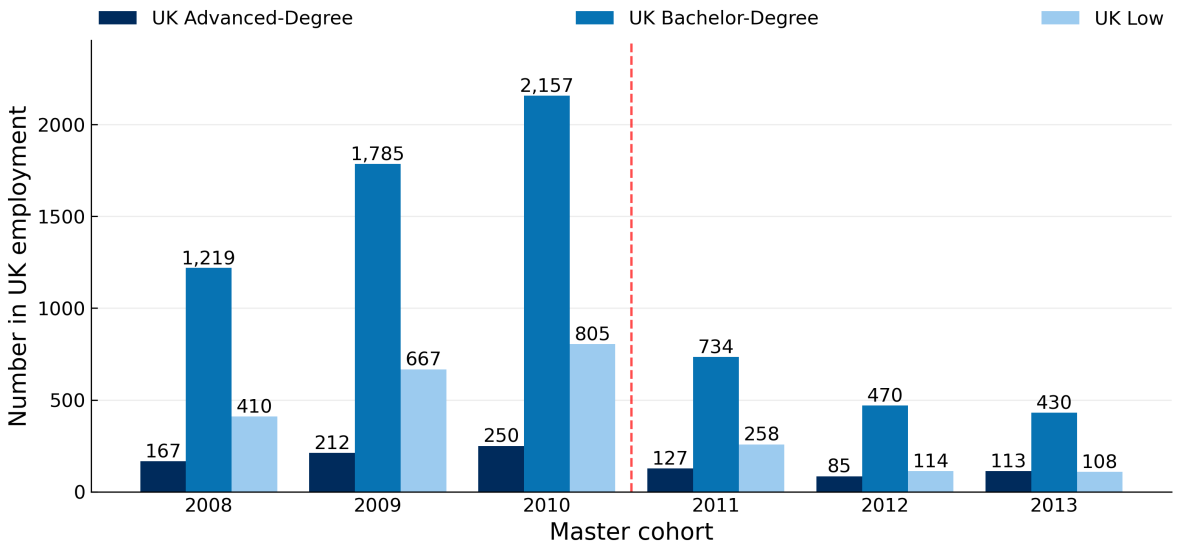


Figure A8: Employment counts among UK "Stayers" by Occupation Groups

Notes: The figure reports the *number* of Indian masters studying in the UK whose first job is in the UK, by master's cohort and by the skill tier of that job. 66 UK jobs with no mapped level are excluded. The dashed line marks the March 2011 announcement of the abolition of the Post-Study Work visa. The contraction between the 2010 and 2013 cohorts is uneven across tiers: bachelor-degree-level employment falls by 80% (2,157 to 430) and Low-skill employment by 87% (805 to 108), while advanced-degree-level employment falls by 55% (250 to 113).

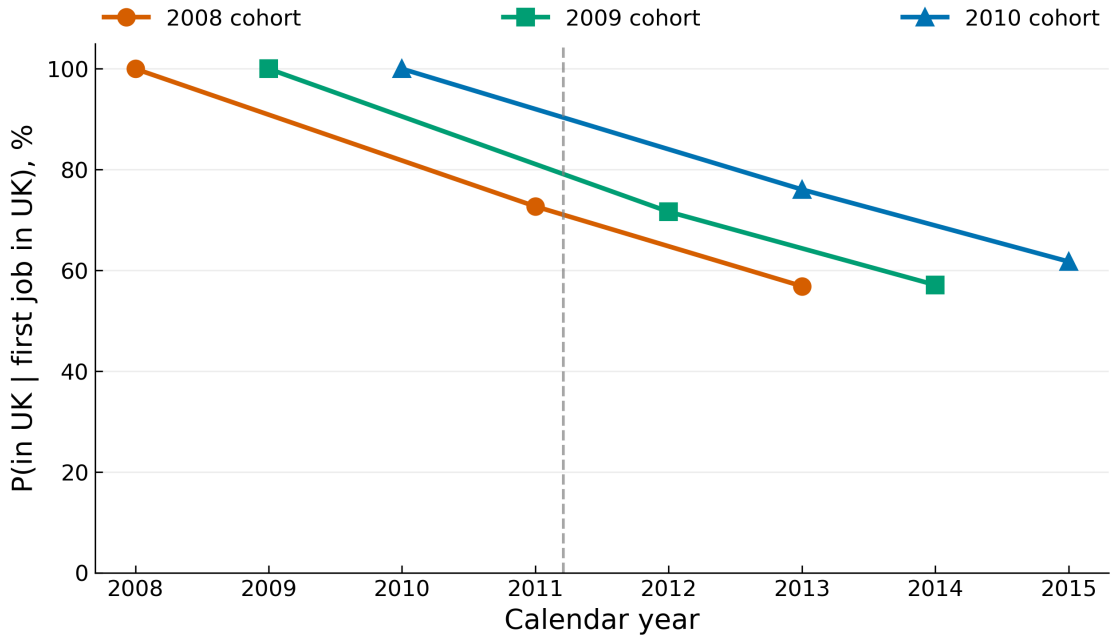


Figure A9: Long-term Stay Rate (UK Master)

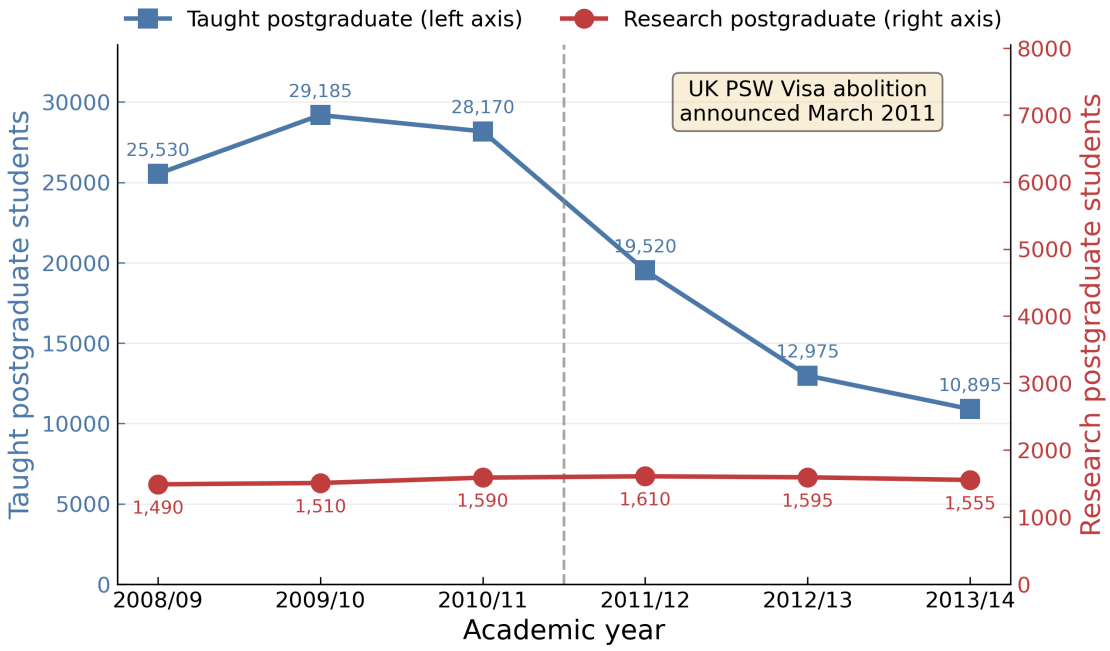


Figure A10: Indian Postgraduate Student Enrollment in the UK

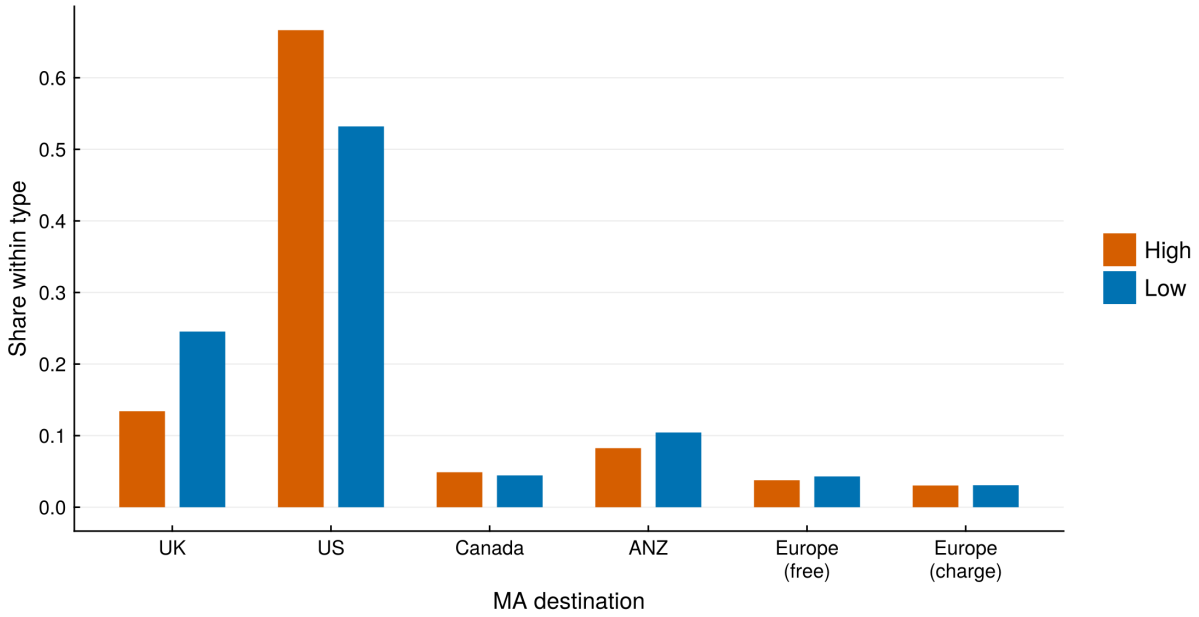


Figure A11: Type Distribution Across Master Destination

Notes: The figure plots the model-implied distribution of master's destinations within each unobserved type; the bars of each type sum to one. Holding each student's observed characteristics and sampling weight fixed, we draw the type from its estimated prior $P(Z \mid M, G)$ and the taste shocks, simulate the stage-1 destination choice under the estimated model for the 2008–2013 cohorts, and average shares over 200 simulation draws.

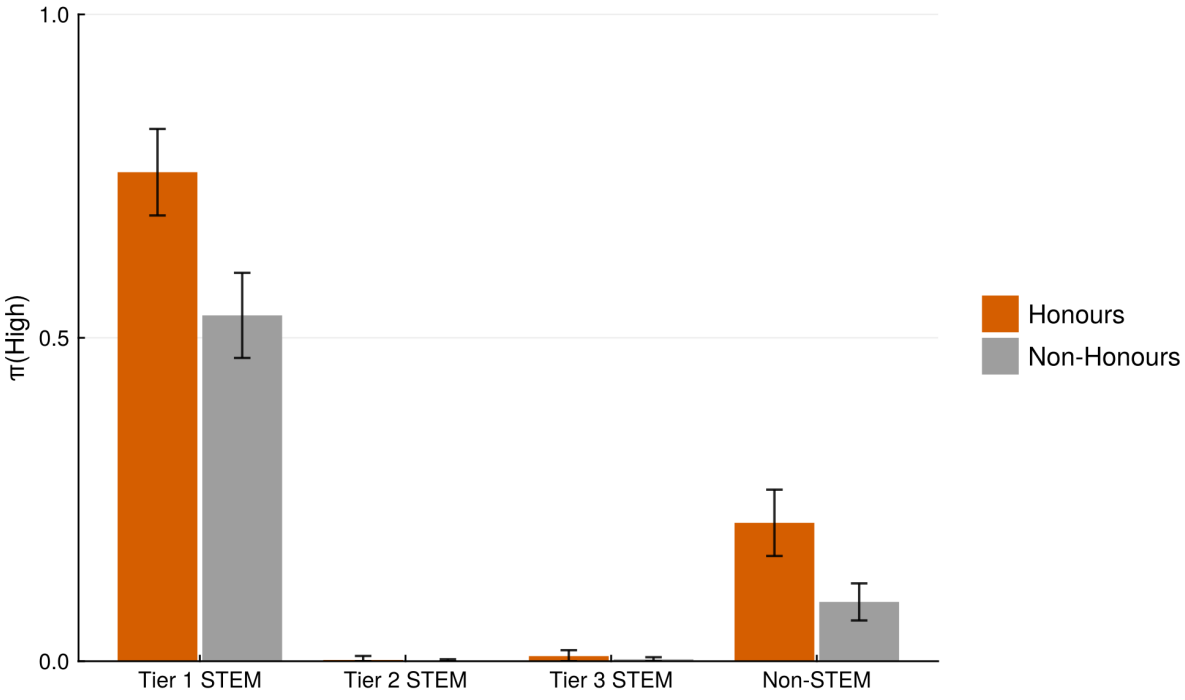


Figure A12: Prior Conditional Type Probabilities

Notes: The figure plots the estimated prior probability that a student is the High type, $\Pr(Z = H \mid M, G)$, by undergraduate tier–major group M and education attainment G. Types are unobserved: the High type is the latent type with the higher wage and the higher elite-sector job-finding rate. The prior is logistic and additive in the two components, $\pi(\text{High} \mid M, G) = \Lambda(\alpha_0^\pi + \gamma_M^\pi + \gamma_G^\pi\, G)$, estimated by maximum likelihood jointly with the remaining model parameters. The estimated honours shifter is $\hat{\gamma}_G^\pi = 0.99$ (standard error 0.14), significant at the 1 percent level. The average unconditional share of high-type students is 9.0% across 2008 to 2013.

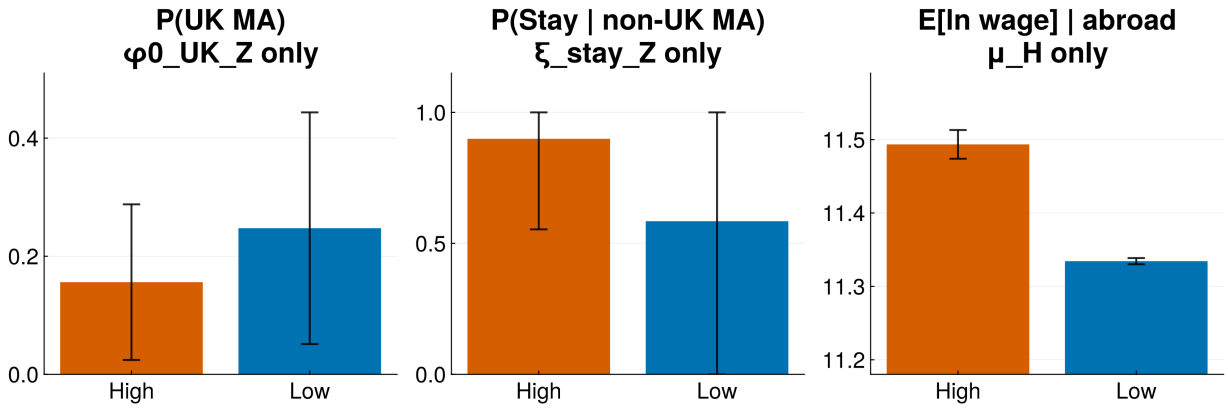


Figure A13: Parameters by Unobserved Types

Notes: The figure isolates the three estimated type-specific coefficients by plotting, for each, the model-implied contrast between the High and the Low type. In each panel exactly one coefficient is set to its estimated High-type value while all remaining type-specific parameters are set at their Low-type values. Thus, the wedge between the High and Low bars only reflects the difference in that parameter. Panel 1 plots the mean probability of choosing a UK master's program for the high and low type, with the UK study-taste shifter $\hat{\phi}_Z^{UK} = -0.63$ and standard error of 0.53. Panel 2 plots the mean probability of remaining in the host country among students observed in a non-UK master's program, with the stay-taste shifter $\hat{\xi}_Z^{stay} = 2.18$ and standard error of 0.29. Panel 3 plots the mean predicted log wage at the observed overseas job, with the high-type wage premium $\hat{\mu}_H = 0.16$ and standard error of 0.01. All means are weighted averages over the estimation sample.

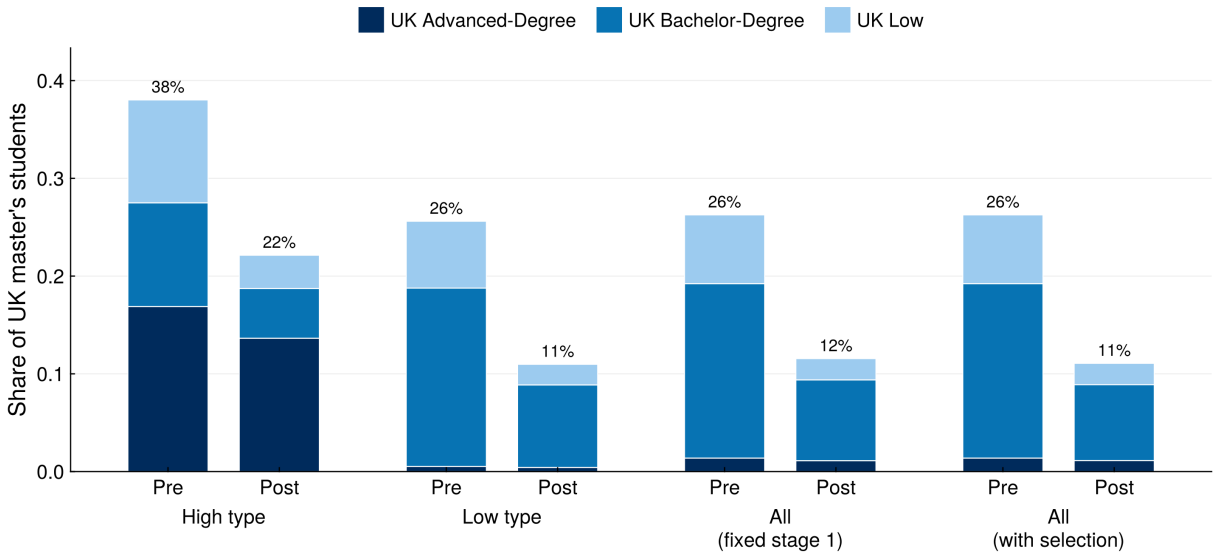


Figure A14: UK Stay Probability

Notes: Each bar is the model-implied share of UK master's students whose first job is in the UK, split by job tier (advanced-degree-level, bachelor-degree-level, low). "Pre" and "Post" represent counterfactual regimes for the same 2011–2013 cohorts. For the High and Low type (the first two groups), the composition of UK master's students is held at its pre-reform level in both bars, so the change isolates the effect of the policy from students' composition change. "All (fixed stage 1)" mixes high and low types, holding enrollment decisions at their pre-reform level and applies the post-reform visa regime to the work choice stage only. "All (with selection)" allows the enrollment stage to respond as well.

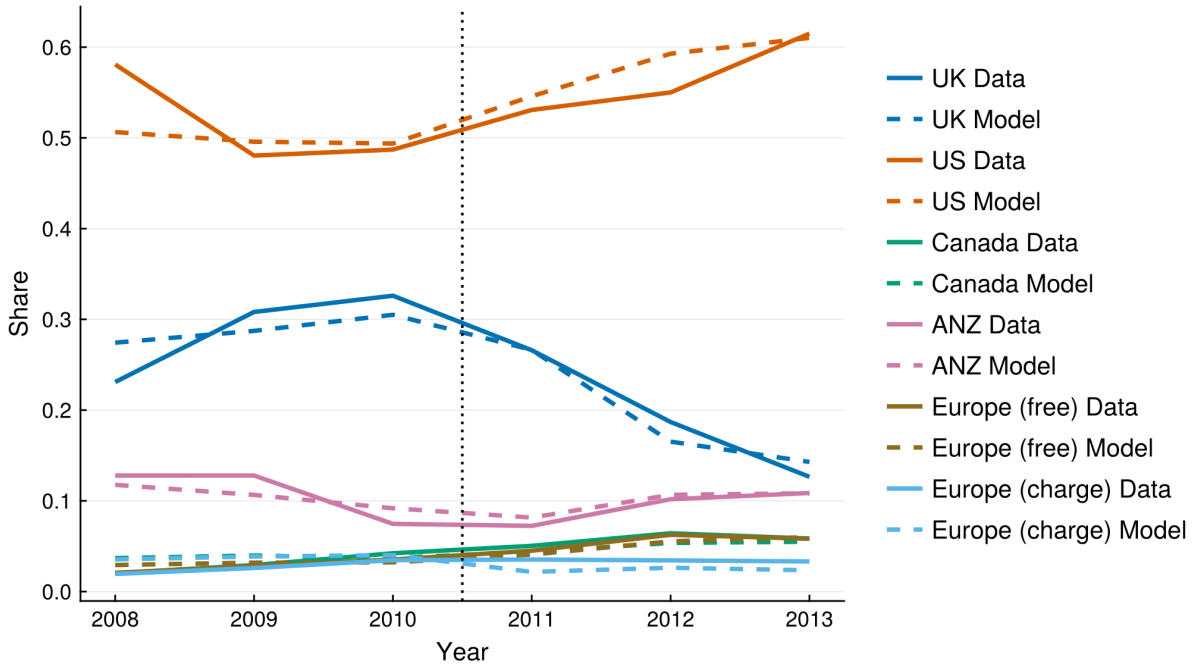


Figure A15: Model Fit – Master Destination Choice

Notes: The figure plots the stage 1 share of each master's cohort enrolling in each of the six destinations—the UK, the US, Canada, Australia–New Zealand, tuition-free continental Europe, and tuition-charging continental Europe—for the 2008–2013 cohorts. Solid lines represent the share calculated from actual data. Dashed lines represent the share calculated from model estimates. This model-based share first holds each student's observed background, origin state, cohort, and sampling weight fixed. Then we draw the unobserved type from its estimated prior $\Pr(Z \mid M, G)$ and the taste shocks 200 times. For each draw, we simulate the stage-1 destination choice from the estimated model, and average the implied shares over 200 simulation draws. The vertical dotted line separates the last pre-reform cohort (2010) from the first post-reform cohort (2011).

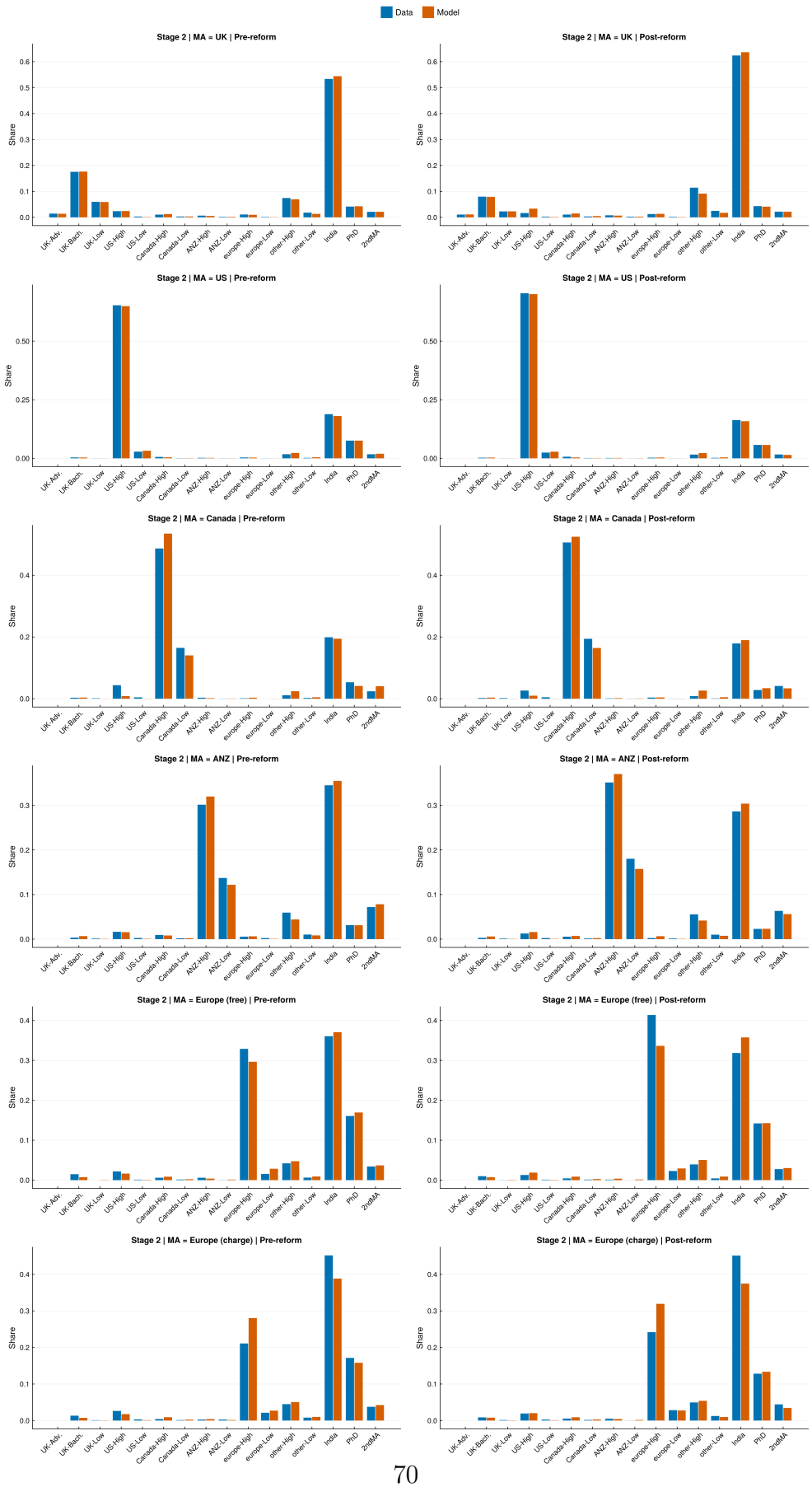


Figure A16: Model Fit –Post-Graduation Placement

A Supplementary Material on Institution Background

A.1 Details about UK Taught Master Program

Taught master's programs in the UK are typically one year in length.[45], compared with the two-year structure that prevails in most US master's programs. Because the academic program was brief and intensive, students had limited time for job searching, networking, or internships during their studies. Under the pre-2012 regime, this limitation was less important, as the two-year Post-Study Work (PSW) visa (described in the next section) provided enough search window. However, this one-year academic structure would later interact with the 2012 policy's four-month search window, imposing severe time constraints on job matching.

Similar to the system in the US, applications to taught master's programs in the United Kingdom were decentralized, with prospective international students generally applying directly to individual institutions. For programs beginning in September or October, applications typically opened in the preceding autumn and were considered on a rolling basis through the following spring or summer, although deadlines varied across institutions and programs. After receiving an unconditional offer from an institution licensed as a Tier 4 sponsor, a non-EU student could be assigned a Confirmation of Acceptance for Studies (CAS) and apply for a Tier 4 (General) student visa. Having guaranteed their visas, most taught master's students began their studies in September or October.

Tuition costs constitute a substantial financial constraint. Taught master's programs charged international students average tuition of approximately £10,431 per year for the 2010 cohort, with substantial variation across institutions and fields. Because UK universities offer little financial aid or subsidies to international master's students, these tuition costs are largely paid out of pocket or financed through education debt. Consequently, tuition represents a highly important factor of the cost-benefit calculation when choosing a destination. In particular, for students from developing countries like India, the option value of the UK labor market access after graduation serves as the fundamental economic justification for their educational investment.

Private institutions could partner with degree-awarding universities, with two common forms of partnership: franchising and validation. Under a franchised-provision arrangement, students were registered with the university but received instruction at a private partner

[45]In the HESA Student Record, the ratio of enrolled stock (Table 11a) to first-year entrants (Table 11b) among full-time non-EU students is 1.24–1.36 for taught master's degrees over 2009/10–2014/15. The modest excess above one reflects programmes that span two HESA reporting years (which run 1 August to 31 July)—taught master's courses typically begin in September and are completed the following autumn—rather than a longer duration.

institution, which delivered a program already approved by the university. Under a validation arrangement, students were normally registered with the private provider and studied a program developed and delivered by that provider but validated by a university. For either pattern, the university could award the degree to the student upon successful completion of the education.

Contemporary investigations documented several serious cases of this collaboration pattern. In 2011, for example, a BBC *Week In Week Out* investigation alleged Rayat London College, a private institution offering University of Wales–validated courses, instructed students on how to cheat in examinations and deceive the UK Border Agency. The alleged scheme would have allowed students to obtain in less than a week a qualification that ordinarily required 15 months and then use it to get an MBA degree with exemptions from approximately two-thirds of the program requirements. Contemporary government evidence also suggests that such weakly regulated private-education sector cases led to immigration-related abuse, identified by parliamentary and government investigations that some institutions provided little genuine instruction and instead facilitated access to student visas and employment in the United Kingdom.[46]

Although no data provides a precise estimate about Indian students' share in bachelor's or taught master's programs, we compare Indian enrollment at private institutions with that at universities, acknowledging that private providers also offer sub-bachelor-degree-level qualifications and English-language courses. However, given that 43% of all non-EU students sponsored by private institutions are enrolled at the higher education level[47] (we expect this proportion to be even higher for Indian students), this comparison is still informative about the share among Indian students in higher education attending private institutions. We exploit two pieces of evidence here. First, a UK Border Agency study of Tier 4 visa holders whose visas began between April and September 2010 found that approximately 85% of Indian students in its sample studied at privately funded institutions, compared with 10% of Chinese students.[48] Second, the CAS data show that non-university sponsors accounted

[46]The House of Commons Home Affairs Committee reported that, among 1,200 education providers inspected, 25% were found not to be genuine. The Committee also cited Home Office concerns that some private institutions provided low-quality instruction to individuals seeking to enter the United Kingdom as students while working illegally, and concluded that most international students with these fraudulently obtained student visas appeared to have entered the country primarily for the work option. See House of Commons Home Affairs Committee (2009), *Bogus Colleges*, Eleventh Report of Session 2008–09, HC 595, paras. 6–8 and 16. Consistent with this evidence, Immigration Minister Damian Green stated in 2011 that numerous colleges offered minimal or no tuition and that some Tier 4 sponsors effectively provided an immigration service rather than an educational one. See Damian Green's QAA Conference Speech, July 7, 2011.

[47]Source: *Overseas students in the immigration system: Types of institution and levels of study*; UK Border Agency Occasional Paper 90

[48]UK Border Agency, *Points-Based System Tier 4: Attendance at Privately Funded Colleges*, March 2011,

for 84.8% of Indian CAS records during the first two quarters of 2010 and 44.2% over the next two quarters.[49] Taken together, these findings suggest that a significant proportion of Indian students were associated with private and non-university sponsors in higher education sector.

A.2 Details about UK PSW Visa

Aggregate evidence suggests that a substantial share of PSW holders did not enter skilled employment. "Home Office Impact Assessment" in 2011 reported that 82% of PSW migrants were employed and that only 51% worked in the top three Standard Occupational Classification major groups, which generally require bachelor-level skills or above. These shares imply that approximately 31% of PSW holders were employed outside those skilled occupational groups, and approximately 18% were unemployed.[50] Another 2010 UK Border Agency operational assessment classified 153 of the 253 (60%) PSW holders in its sample as "unskilled." This estimate should be interpreted cautiously because the sample was restricted to Tier 1 visa holders whose dependents submitted applications in June 2010 and therefore excluded visa holders without dependents. [51]

A.3 Details about UK Tier 2 Visa

Prior to the reform, an applicant for Tier 2 (General) visa was generally required to obtain 70 points across four components: a valid Certificate of Sponsorship issued by a licensed

Figure 1, p. 2. The study draws a random sample of 4,000 Tier 4 visa holders, including 484 Indian students whose leave commenced between April and September 2010. The report's private-institution category includes both private colleges and English-language schools. That April to September window covers the spring and summer intakes, in which non-university sponsors predominate. Approximation from the CAS record data shows that over the full calendar year, the non-university share was approximately 62 percent in 2010 and 55.7 percent in 2011. Chinese students' share was approximately 25% in 2010 and 17.2% in 2011.

[49] Authors' calculations from the Home Office sponsorship dataset (CAS_D02). Four caveats apply. First, coverage of the CAS dataset is incomplete in the first half of 2010, because the Confirmation of Acceptance for Studies became compulsory only in February 2010 and earlier applications used the previous visa system. Second, a CAS assignment is not an arrival: a student holding a CAS may be refused a Tier 4 visa, and refusal rates were substantially higher among applicants sponsored by non-university institutions. Third, "non-university" is broader than private, because it also includes publicly funded non-university institutions, such as art colleges. Fourth, the seasonal difference reflects the concentration of university recruitment in the autumn intake, whereas non-university sponsors operated rolling intakes throughout the year.

[50] Home Office, *Reform of the Points Based Student Immigration System: Final Impact Assessment* (2011), pp. 55–56 and Table A7.2. The top three SOC major groups comprise managers and senior officials, professional occupations, and associate professional and technical occupations. The impact assessment attributes the estimates to the survey underlying Home Office Research Report 49.

[51] UK Border Agency, *Points Based System Tier 1: An Operational Assessment*, Occasional Paper 91 (2010), pp. 1–4. The study's 'unskilled" category included not only individuals apparently employed in unskilled occupations, but also those earning less than £25,000 and those who remained unemployed after having received their visas before 2010.

employer, an appropriate salary above £20,000 per year and the occupation-specific rate, English-language proficiency, and sufficient maintenance funds.[52] More importantly, the sponsored position had to meet the occupational-skill requirement. Between April 2011 and June 2012, eligible occupations generally had to be classified at the National Qualifications Framework (NQF) Level 4 or above; the threshold had previously been Level 3 and was subsequently raised to Level 6 on June 2012.[53] The NQF metrics were used by the immigration authorities to classify the skill requirements of occupations. Levels 1–3 broadly covered secondary-level qualifications; Levels 4 and 5 covered sub-bachelor qualifications; and Levels 6, 7, and 8 corresponded broadly to bachelor's, master's, and doctoral qualifications, respectively.[54]

Tier 2 visa was tied to the sponsored employment. Changing employers generally required a new Certificate of Sponsorship and a new immigration application. Permission was normally granted for up to three years and could subsequently be extended for up to two additional years. After five years of continuous residence in qualifying immigration categories, a worker could apply for indefinite leave to remain (permanent residency), subject to continuing employment, salary, language, and other settlement requirements.[55]

A.4 Details about Education-side Reform

The education-side reforms comprised three main elements. First, on the provider side, the government attempted to remove private "bogus" colleges that were selling immigration rather than education. By April 2012, sponsors were required to obtain Highly Trusted Sponsor (HTS) status and pass rigorous statutory educational inspections. Non-compliant institutions faced strict enrollment caps, license suspensions, or total revocation. By July

[52]Under the rules introduced on 6 April 2011, a valid Certificate of Sponsorship carried 30 points, an appropriate salary 20 points, maintenance funds 10 points, and English-language proficiency at CEFR Level B1 10 points. The £20,000 threshold applied under the Tier 2 rules introduced in April 2011. If some applicable occupation specified a higher salary rate, the higher salary rate would be applied. Salary requirements therefore varied across occupations. See UK Border Agency, *Tier 2 of the Points Based System: Statement of Intent, Transitional Measures and Indefinite Leave to Remain* (2011), especially Appendix B.

[53]See UK Border Agency, *Tier 2 of the Points Based System* (2011), and UK Border Agency, *Statement of Intent and Transitional Measures: Tier 2* (2012). Appendix B in the latter document lists the occupations skilled to NQF 4 that are no longer eligible for Tier 2 visa after 2012, such as "office managers" and "IT operations technicians". Importantly, this was a requirement concerning the skill level of the occupation, not necessarily the applicant's formal educational attainment.

[54]See HM Government, What Qualification Levels Mean." Level 5 is the closest approximate counterpart to a US associate degree. The NQF was subsequently superseded through the development of the Qualifications and Credit Framework and, from October 2015, the Regulated Qualifications Framework (RQF).

[55]See UK Border Agency, *Tier 2 of the Points Based System: Statement of Intent, Transitional Measures and Indefinite Leave to Remain* (2011), sections on changes of employment and indefinite leave to remain and Appendix B. Indefinite leave to remain was not automatic after five years: the applicant had to submit an application and satisfy the settlement rules then in force.

2012, more than 450 colleges had lost the right to sponsor international students.[56] Second, on the student-entry side, the UK Border Agency strengthened its scrutiny of Tier 4 applicants. The reforms raised the English-language requirement, tightened the verification of maintenance funds, and subsequently introduced targeted credibility interviews through which officers could assess whether an applicant was a genuine student.[57] Third, the government restricted the employment and dependant rights attached to student status. Students at most private colleges lost their work rights during study, and the right to bring dependants was restricted to students enrolled in postgraduate university programs lasting longer than 12 months.[58]

A.5 Details on Tier 2 Visa Sponsorship Post-Reform

In 2012, the monetary cost for acquiring a Tier 2 sponsorship license was £500 for a small employer and £1,500 for a large employer. Additionally, sponsoring each worker required a Certificate of Sponsorship costing £179 for firms, while workers also needed to pay the visa application fees starting from £561.[59] Employers using legal assistance for applications could also incur fees of approximately £500–£1,500 per worker.[60]

Apart from costs in the application stage, Tier 2 visa sponsors also need to bear significant compliance costs. Failing to do so might induce the revocation of the sponsorship license. Reports from the House of Commons Home Affairs Committee indicate that during the first three quarters of 2012 alone, 469 Tier 2 sponsoring employers were suspended and 252 had their licenses permanently revoked, out of approximately 22,000 licensed sponsors — a revocation rate of roughly 3.3%, compared with approximately 2% during the first two years of the Tier 2 system (2008–2010). This increase reflected an expansion in post-license compliance visits, which reached 4,901 in the first three quarters of 2012, of which approximately 40% were unannounced — compared with only approximately 15% for new applicants before 2011. These compliance costs deterred sponsorship at the margin, particularly among smaller and less specialized employers.

[56] See Home Office, "Government Outlines Major Overhaul of Student Visas" (22 March 2011), and "Student Visa Interviews and Genuine Student Rule" (9 July 2012).

[57] See Home Office, *Final Impact Assessment: Reform of the Points Based Student Immigration System* (2011), pp. 10–11; and Home Office, "Student Visa Interviews and Genuine Student Rule" (9 July 2012). The latter introduced a targeted interview program and a formal genuine-student test from 30 July 2012.

[58] See Home Office, *Final Impact Assessment: Reform of the Points Based Student Immigration System* (2011), pp. 10–11; and Theresa May, "Student Visas," House of Commons Debates, 22 March 2011, vol. 525. These restrictions took effect in stages beginning in 2011.

[59] These fees took effect on April 6, 2012. See Home Office, "Immigration and Nationality Services," (February 9, 2012).

[60] In evidence to the House of Commons Public Accounts Committee on March 28, 2011, immigration solicitor Tony Haque reported this amount. See "Immigration—Points Based System," oral evidence, Q27.

B The LinkedIn Profile Data

B.1 Coverage for Indian International Master Across Education Markets

B.1.1 Harmonization across different administrative data counting Indian master's students

This appendix documents how we construct an administrative count of Indian master's entrants that is comparable to our LinkedIn sample for four destinations. Throughout, the LinkedIn count is the number of profiles with a bachelor's degree from an Indian institution and a first foreign master's degree starting in a given calendar year; the administrative count is the number of Indian students starting a master's program in that destination in the same year. Three aspects differ across sources and require harmonization: whether the source reports entrants, graduates, or an enrollment stock; whether it counts persons or course enrollments; and which credentials and providers it covers. We describe each destination in turn.

United States. We use record-level SEVIS data on F-1 students obtained under the Freedom of Information Act. Each fiscal-year file lists every F-1 record active during that year, with the program start date and education level. We count master's-level program starts by Indian citizens in each calendar year, taking January–September from the file of the same fiscal year and October–December from the precedent one, and we count each individual once so that transfers between schools are not double counted. This is a direct entrant flow and needs no stock-to-flow conversion. We apply no selection on either side: SEVIS covers all F-1 master's students regardless of institution, and Indian students from a US bachelor's degree, who are outside the LinkedIn universe, only account for approximately 1% of entrants.

United Kingdom. We use HESA counts of India-domiciled students enrolled in postgraduate taught programs at UK providers: the published Table 8 series for 2006/07–2013/14 and the open-data Table 28 series from 2014/15 onward, scaled at the overlap year to the later series, which additionally covers designated courses at alternative providers. HESA reports a December census *stock* rather than *entrants*. Because a one-year master's that starts in September and ends after 1 August may appear in two reporting years, the stock exceeds the first-year flow by a factor of approximately 1.3, a scaling parameter we obtain from comparing enrollment stock with HESA's first-year tables for non-EU students. (The

first-year tables for Indian students are unavailable.) Thus, we recover entrants from stocks by netting out the carry-over of the previous cohort.

$$\text{UK Entrants}_t = \text{HESA Stock}_t - 0.30 \times \text{UK Entrants}_{t-1}, \tag{13}$$

We apply no selection on either side. LinkedIn coverage of UK entrants is nonetheless lower than other destinations, particularly before 2012 and after 2018, because a large portion of Indian taught masters were enrolled in private franchised providers; HESA registers these students at the awarding university, and their LinkedIn presence is low (See Section 2.1). On the contrary, if we focus on Russell Group universities for 2014–2019, during which we have finer data, coverage rate for this group of Indian master students from matching LinkedIn records is approximately 79 percent. However, Open data do not separate franchised students from the awarding university's own campuses before 2014, so we retain all providers and acknowledge the difference as an "underestimated" coverage rate.

Canada. Statistics Canada publishes no entrant series by citizenship and level. We use Table 37-10-0185, graduates of ISCED-7 (master's) programs among international students with Indian citizenship, and convert *graduates* to *entrants* by distributing each graduating cohort back over start years using the distribution of program length observed in the LinkedIn sample, which is concentrated at one and two years. This treats completion as certain and therefore understates entrants slightly. Two harmonizations concern the population. First, permanent residents are excluded from the administrative count because we cannot identify immigration status on LinkedIn; the resulting coverage rate is biased upward, most in the early years. Second, 58% of LinkedIn records listing a Canadian "master's" are in fact one-year *college postgraduate diplomas* or *graduate certificates*, mostly at Ontario colleges. Statistics Canada classifies these credentials below ISCED 7, so they are absent from the denominator, and only 13–22 percent of such students appear on LinkedIn against approximately 79.4 percent for university master's students. We exclude these records from the LinkedIn count.

Australia and New Zealand. We use the Department of Education's international student data, which report *commencements* by nationality and, from 2019, by level of study, provider type, and whether the student is new to Australia. Commencements are an entrant flow, but they count *course enrollments* rather than *persons*, so a student who packages a graduate certificate with a master's is counted twice. We apply three harmonizations. First, private providers account for 26–32 percent of Indian master's commencements in the administrative data but under 5 percent of the LinkedIn sample; we restrict both sides to public

universities. Second, we exclude students who were not new to Australia, who typically hold an Australian prior qualification and are therefore outside the LinkedIn universe of Indian bachelor's graduates. Third, we deflate enrollments to persons using the ratio of the Department's enrollment counts to its student headcounts for Indian students, approximately 1.04–1.12 depending on the year. Before 2019 the data carry no level or provider detail, and we apply the 2019 share of restricted master's commencements in all higher-education commencements to earlier years. The administrative data cover Australia only, so we exclude New Zealand institutions, 12% of the LinkedIn ANZ sample, from the LinkedIn count.

B.1.2 Validation of LinkedIn dataset

Figure A17 compares pooled destination shares for Indian master's students in LinkedIn and administrative data over 2010–2020 (Panel a) and 2013–2017 (Panel b). Consistent with Figure A1, the shares are broadly consistent between the administrative data and the LinkedIn data, with the largest discrepancy occurring for the UK. This gap reflects HESA's inclusion of university enrollment in franchised programs delivered by private providers. The closer alignment in Panel (b), covering a period when such programs were less prevalent and subject to tighter government scrutiny, is consistent with this interpretation.

Figure A18 reports LinkedIn coverage rates, defined as the ratio of Indian master's entrants observed on LinkedIn to the corresponding administrative counts for each destination and entry year. Coverage is substantially lower for the UK than for other destinations, although the gap narrows during 2013–2017. This pattern is again consistent with the explanation discussed above: HESA includes enrollment in franchised programs delivered by private providers, which were more prevalent before 2012 and after 2018 and less represented on LinkedIn.

Figure A19 compares administrative counts of Indian master's entrants with LinkedIn counts scaled by the inverse of the destination-specific pooled coverage rates reported in Figure A18. We use rates pooled over all years, except for the UK, where we use 2013–2017 to limit the contamination of franchised programs. The scaled LinkedIn counts closely track administrative trends for the United States, Canada, and Australia and New Zealand, indicating that the Linkedin data captures the trend well and consistent with relatively stable coverage. For the UK, the two series align closely during 2012–2018. The divergence outside this period is consistent with the greater prevalence of franchised programs discussed above. In model estimation, we use the inverse pooled coverage rates as individual weights, as discussed in Appendix Section B.3.

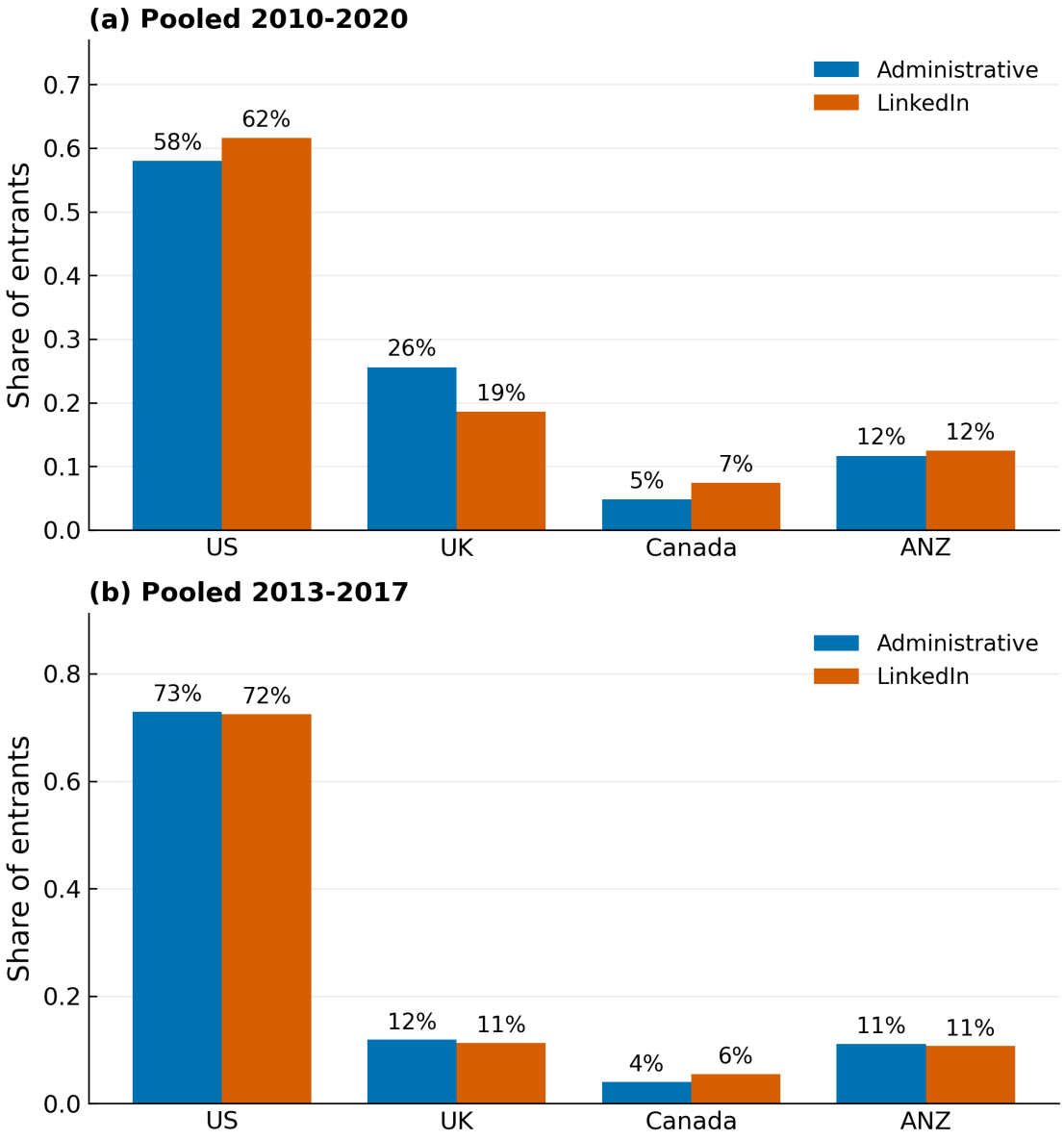


Figure A17: Destination shares of Indian master's entrants across years

Notes: Bars show each destination's share of all Indian master's entrants across the four destinations, pooled over 2010–2020 in panel (a) and over 2013–2017 in panel (b). Blue bars use administrative entrant counts; orange bars use our LinkedIn sample. The administrative constructs and the harmonization of the two populations are described in Appendix Section B.1.1.

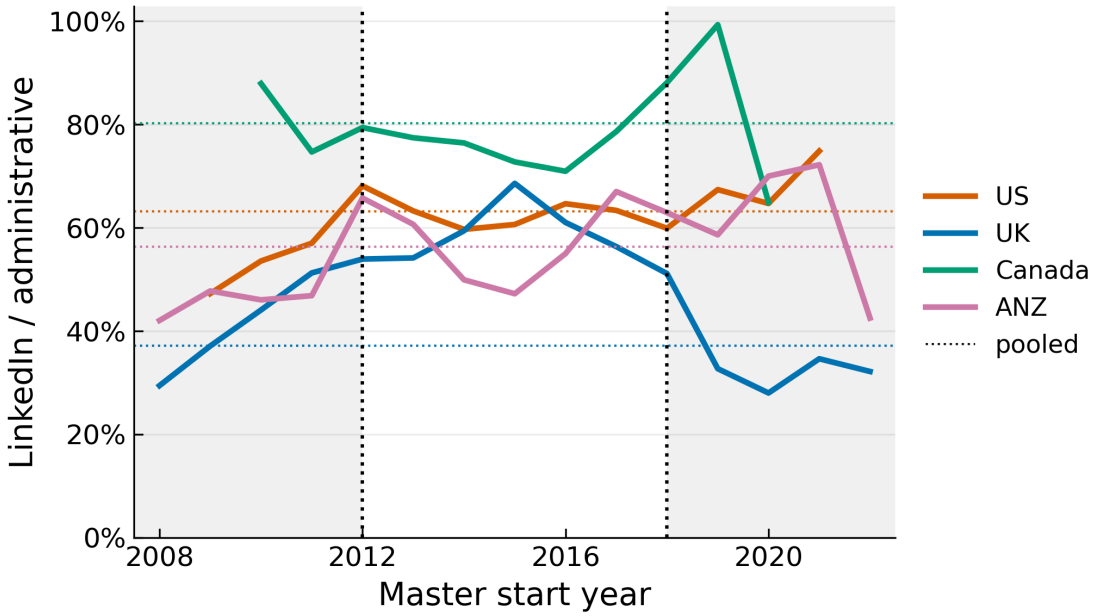


Figure A18: LinkedIn coverage of Indian master's entrants by destination and year
Notes: Each solid line is the ratio of Indian master's entrants in our LinkedIn sample to the administrative entrant count for that destination and entry year; dotted horizontal lines are the ratio pooled over all years with administrative data. Administrative constructs and population adjustments follow Appendix Section B.1.1.

B.2 Coverage for Indian International Master Across Labor Markets

In this section, we describe how we use the *Migration Journey* dataset to construct an appropriate denominator for the coverage rate of our LinkedIn sample of Indian taught master's students who stayed in the UK for the first job (i.e. "stayers"). *Migration Journey* tracks changes in the visa status of Tier 4 student migrants in the years following entry, separately for each entry cohort. For example, it reports the number of students in each cohort who, by the end of the first, second or third year, transitioned to Tier 1 or Tier 2 work visas, remained on student visas, or left the UK (i.e. student visa expires).

One limitation is that *Migration Journey* does not disaggregate Tier 4 entrants by level of study, so we cannot restrict the population to taught master's students and must instead use entrants at all education levels. Nevertheless, these data remain informative about the representation of taught master's students in our sample. Bachelor's and postgraduate research students typically require more than two years to complete their studies, so few of them are likely to appear as transitions to work at the end of the first or second year. Students on certificate and diploma courses, and students from lower-quality master's programs such as franchised provision, are also unlikely to possess the skills required for jobs meeting the NQF

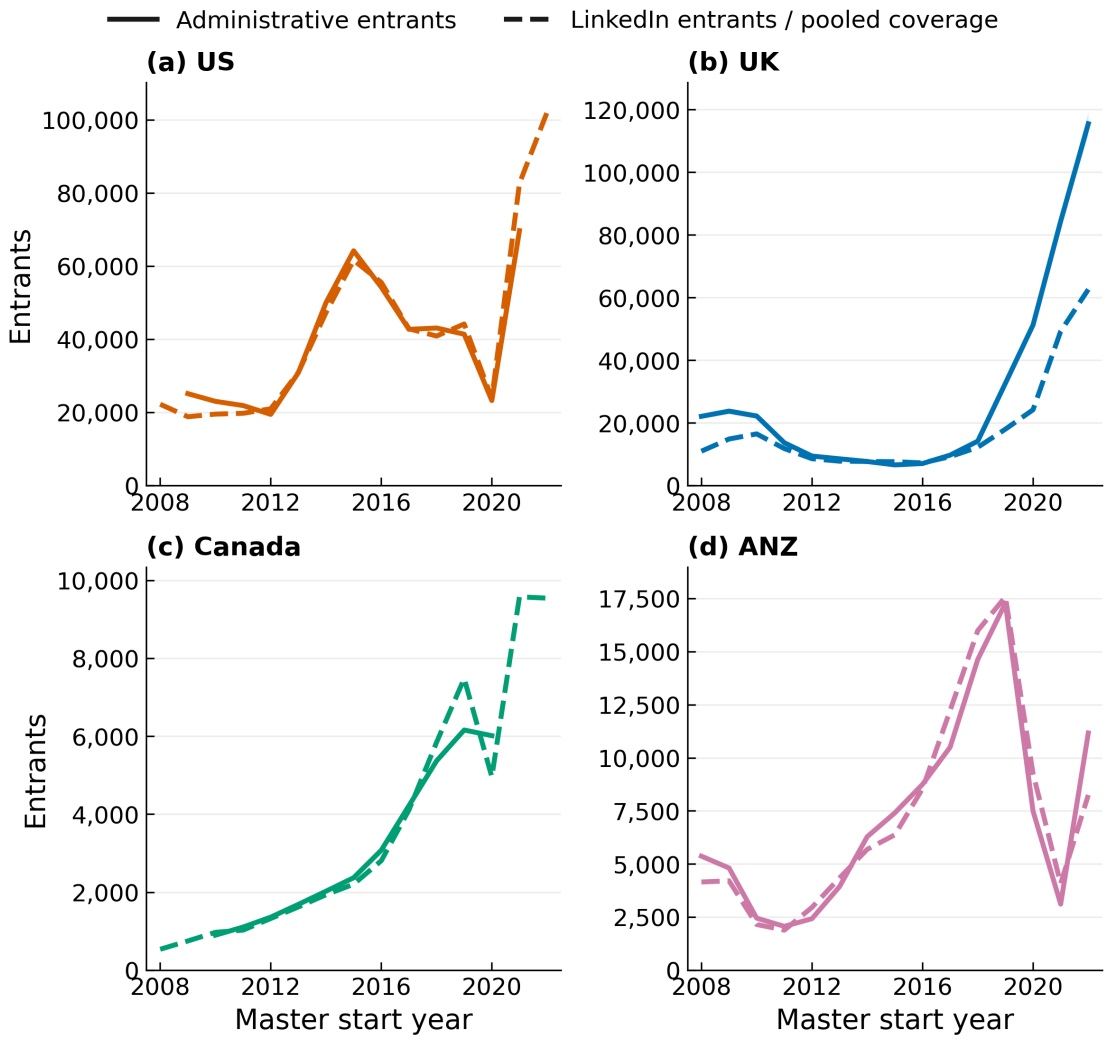


Figure A19: Indian master's entrants: administrative v.s. coverage-scaled LinkedIn

Notes: Solid lines are administrative counts of Indian master's entrants by entry year; dashed lines are LinkedIn counts divided by the destination's pooled coverage rate over years. The pooled coverage rate is the ratio of LinkedIn to administrative entrants pooled over all years with administrative data for the United States (63.1%), Canada (80.3%), and Australia and New Zealand (56.4%), and over 2013–2017 for the United Kingdom (59.4%) derived from Figure A18. Both series are on the same scale and the comparison is of trends rather than levels. Perfectly overlapping lines across years indicate that LinkedIn coverage is constant over time.

level 6 skill threshold that Tier 2 visa imposes after 2012. Post-reform stayers may therefore provide a reasonable approximation to the population corresponding to our LinkedIn sample of taught master's students.

In panel (b) of Figure A2, the dashed line plots the total number of Indian students in each cohort who remain in the UK after their studies, measured as transitions to work visas, family visas, and other forms of leave, summed over years one and two. The green bars show the subset transitioning to Tier 2 work visas. The predominance of Tier 2 transitions post-reform is consistent with our LinkedIn data pattern, in which most stayers work in high-skill occupations. We use total transitions rather than Tier 2 transitions alone to account for LinkedIn users who may remain on family visas or other forms of leave, consistent with the small share of LinkedIn records reporting employment in low-skill occupations even after the reform. Combining total transitions for the 2012 and 2013 cohorts yields 1,518 individuals, which we use as the denominator, compared with 1,320 individuals whom we observe on LinkedIn and whose first job is in the UK, which we use as the numerator. Thus we obtain the coverage rate of 87% for "UK stayers".

For UK leavers, we infer the coverage rate using the coverage rate for all Indian taught master's students in the UK and the coverage rate for UK stayers calculated above. Specifically, the total number of stayers and leavers observed on LinkedIn, divided by their combined population in the administrative data, should equal the overall coverage rate of Indian taught master's students.[61] For the 2012 and 2013 cohorts, we observe 9,683 Indian students starting a UK master's on LinkedIn, against 14,972 taught-master entrants at UK higher education institutions in the administrative data, an overall coverage rate of 64.7%. Subtracting the 1,320 stayers leaves 8,363 LinkedIn individuals who did not stay in the UK for work, including those who returned to India, moved to a third country, continued their education, or have no observed first spell. The corresponding administrative population is 13,454. The implied coverage rate for "UK leavers" is therefore approximately 62%.

[61] This calculation abstracts from students who continue their education after completing a master's program, such as those pursuing a second master's degree or a PhD. They are excluded in both the denominator and the numerator. This group is small, and its exclusion has little effect on the results.

B.3 Individual Weight in the estimation

Each observation enters the likelihood with a weight

$$\varpi_i = \frac{1}{C} \cdot \frac{1}{c_{j_1(i)}} \cdot s_i, \qquad s_i = \begin{cases} 0.74 & \text{UK master's stayed in the UK (work, PhD, second master's),} \\ 1.03 & \text{UK master's, left the UK,} \\ 1 & \text{otherwise,} \end{cases} \tag{14}$$

where $c_{j_1(i)}$ is the LinkedIn coverage rate for master's destination j_1 and C normalizes the weights to have a mean of one. As Figure A18 shows, LinkedIn coverage differs across master's destinations. Thus, we use inverse-coverage weights to recover the population's destination composition, following a similar rationale to Amanzadeh et al. (2024). For destinations other than the UK, we calculate c_j by pooling observations over 2008–2013. For the UK, we instead use 2013–2017 to limit contamination from franchised programs. The coverage rates are reported in the note to Figure A18. The second factor adjusts for the higher LinkedIn coverage of UK stayers relative to UK leavers. We calculate the two correction factors using the stayer coverage rate imputed in the previous section and the overall UK coverage rate, normalizing them to an average of one within the UK sample.

C Indian Bachelor Ranking Data

C.1 Institutional Structure and Admissions in Indian Higher Education

Indian higher education institutions can be classified into four legal categories based on their founding authority and statutory status: Central Universities, State Universities, Private Universities, and Deemed to be Universities. Central Universities are established by the central government, State Universities by state governments, and Private Universities by private entities under state legislation, while Deemed to be Universities are institutions granted university status through statutory authorization. Notably, these legal categories do not directly correspond to a quality hierarchy, and substantial heterogeneity exists within each category.

A distinct statutory designation within the higher education system is that of Institutions of National Importance (INIs), which are established through parliamentary legislation. INIs include selected public professional institutions, primarily in engineering, medicine, and management, and are consistently ranked among the top institutions in their respective fields.

They operate under separate legal frameworks and receive direct funding from the central government, with relatively centralized governance structures. In constructing our institutional quality measures, INI status is explicitly incorporated into the imputation procedure.

Undergraduate admissions in India are organized primarily by field of study. Major professional disciplines, including engineering, medicine, law, and management, rely on separate national or multi-state entrance examination systems. Candidates take examinations corresponding to their chosen field and participate in centralized seat allocation based on national rank and submitted preferences. The admission process is administered through designated examination authorities and centralized allocation mechanisms. As a result, admission systems operate independently across fields, and no single entrance examination covers all institutions and programs.

The combination of multiple institutional categories and discipline-specific admission systems implies the absence of a unified quality evaluation framework. Admission procedures and information disclosure practices vary across institution types and fields, leading to fragmented observable data. Some institutions appear in national ranking systems, others are observable only through field-specific examination outcomes, and some lack publicly reported ranking or admission information. Measuring institutional quality therefore requires integrating information from multiple sources. The next subsection describes the two sources of quality measures used in this study.

C.2 Quality Measures

Evaluating institutional quality in the Indian higher education system is complicated by discipline-specific admissions mechanisms and incomplete coverage of ranking data. To construct consistent measures of institutional quality, we rely on two complementary sources: (i) official national rankings, and (ii) discipline-specific entrance examination cutoffs.

C.2.1 National Institutional Ranking Framework (NIRF)

The National Institutional Ranking Framework (NIRF) was launched in 2015 by the Ministry of Education (MOE) as a standardized framework for ranking higher education institutions in India. It provides an annual ranking based on a uniform evaluation methodology.

NIRF publishes rankings across multiple categories, including an overall category as well as discipline-based categories such as engineering, law, management, and medicine. Institutions are evaluated and ranked within their respective categories. The ranking is determined by a composite score constructed from five dimensions: Teaching, Learning and Resources; Research and Professional Practice; Graduation Outcomes; Outreach and Inclusivity; and

Perception. Each dimension aggregates multiple quantitative indicators, and the final score is calculated as a weighted average of these components.

Participation in NIRF requires institutional data submission, and published rankings typically include only a limited number of institutions within each category. In our construction of institutional quality measures, we implement a score imputation procedure to combine information across categories. However, due to the limited coverage of NIRF, a substantial number of institutions do not receive an observed ranking score.

C.2.2 Entrance Examination Cutoffs

Admission to undergraduate engineering and law programs in India is conducted through centralized national entrance examinations followed by centralized seat allocation procedures.

In engineering, admissions are based on the Joint Entrance Examination (JEE). The examination is administered by the National Testing Agency (NTA), which conducts the test, evaluates performance, and publishes standardized All-India Rank lists. Seat allocation is conducted separately through a centralized counseling mechanism under which candidates submit ordered program preferences and seats are assigned based on rank and availability. Program-level opening and closing ranks are publicly reported for each institution and admission round.

In law, admissions to National Law Universities are conducted through the Common Law Admission Test (CLAT). The examination generates a national rank list, and seat allocation is conducted centrally among participating institutions. Opening and closing ranks are publicly reported for each institution and admission round.

We use the publicly reported opening and closing ranks from these examinations to construct rankings for institutions that admit students through these examination systems, which serve as a supplementary measure of institutional quality.

D The Classification Rule of Job Types

O*NET Job Zones are a five-category classification of occupations by the education, experience, and training they typically require, ranging from Zone 1 (little or no preparation) to Zone 5 (extensive preparation). We classify Zone 5 occupations, most of which typically require graduate school education as noted by O*NET, as "advanced-degree-level" jobs. Note that the definition of Zone 5 occupations is broadly consistent with that of NQF Level 7 and NQF Level 8 jobs classified by the UK administration. Among Indian workers working in the UK, the most common occupations in this category are "*investment fund managers*" and "*financial quantitative analysts*".

We draw the boundary between the remaining two job levels at the skill threshold that the policy itself applies. Post-reform Tier 2 (General) sponsorship required the occupation to sit at UK NQF Level 6 or above. Therefore, we classify all Zone 4 occupations, together with some Zone 3 occupations that map to Level 6 or above on the NQF occupation list, as bachelor-degree-level occupations. Some of the most common job titles among Indian workers in the UK in this tier are "*software developers*" and "*mechanical engineers*". The remaining Zone 1 and Zone 2 occupations, together with Zone 3 occupations below NQF Level 6, are classified as low-skill occupations.

E More Discussion on Descriptive Evidence

Table A1 shows that Indian master's destination choices differ a lot across fields and bachelor institution quality. Over half of STEM students chose the United States, while nearly half of non-STEM students chose the United Kingdom as their master's destination, and they also favored Australia and New Zealand relative to STEM students. STEM Students from higher tiers of undergraduate institutions were also more likely to choose master's programs in the US and Europe relative to STEM students from the lower tiers, while the latter group preferred UK and ANZ master's programs more.

Regarding post-graduation placement, 57.59% of students secured employment outside India, while 21.38% returned to India for work. There is also substantial heterogeneity by major along this margin: non-STEM students were more likely to return to India for work after completing their master's programs by 14 to 18 percentage points. In terms of occupational skill requirements, non-STEM workers were also 10.0%–13.7% more likely to work in low-skill jobs (i.e., jobs requiring skills below the bachelor-degree level). STEM students, especially Tier 1 STEM students, were significantly more likely to enter doctoral programs after completing their master's degrees.

F Quasi-Experimental Design

F.1 Econometrics Framework and Identification

Let $Y_{c,t}(0)$ denote the potential number of UK master's entrants from undergraduate institution cluster c in cohort t under the pre-reform visa regime, and let $Y_{c,t}(1, d)$ denote the corresponding potential outcome under the reform at exposure level d. Let D_c denote cluster c's observed exposure, measured by its pre-reform UK stay rate over 2001–2007, and let $s(c)$ denote its state. Assuming no anticipation before the 2011 announcement, observed

enrollment satisfies

$$Y_{c,t} = \begin{cases} Y_{c,t}(0), & t \leq 2010 \\ Y_{c,t}(1, D_c), & t \geq 2012 \end{cases} \tag{15}$$

Let $\mathcal{T}_{\text{pre}} = 2008, 2009, 2010$ and $\mathcal{T}_{\text{post}} = 2012, 2013$ denote the cohorts included in the DiD estimation.

Assumption 1 (Log-Linear Response with a Common Slope)**.** *For every cluster c, post-reform cohort $t \in \mathcal{T}_{\text{post}}$, and exposure level d in the observed exposure range,*

$$\log Y_{c,t}(1, d) - \log Y_{c,t}(0) = \beta_{0,t} + \beta_1 d.$$

The reform's effect on log enrollment has a common slope across clusters and post-reform cohorts. The intercept $\beta_{0,t}$ allows for a common policy effect that may vary across cohorts. Under this assumption, increasing exposure from d' to d changes log enrollment under the reform by $\beta_1(d - d')$. Combined with the parallel trends assumption below, this restriction allows comparisons between differently exposed clusters to identify β_1 even though all clusters are "treated".

The assumption restricts both the functional form of the response and its heterogeneity across clusters. This distinction matters because parallel trends in untreated outcomes alone does not generally permit cross-exposure comparisons to be interpreted as causal responses to changes in exposure (Callaway et al., 2024). Importantly, we do not extrapolate to zero exposure throughout the analysis: the discussion regarding magnitudes written in Section 4 compares clusters at different points of the observed exposure distribution, such as its top and bottom deciles.

Assumption 2 (Parallel Trends in Untreated Unit)**.** *For any clusters with exposure d and d' in the same state and any cohorts $t, t' \in \mathcal{T}_{\text{pre}} \cup \mathcal{T}_{\text{post}}$,*

$$\begin{aligned} &E[\log Y_{\cdot,t}(0) \mid D_c = d, S_c = s] - E[\log Y_{\cdot,t'}(0) \mid D_c = d, S_c = s] \\ &\quad = E[\log Y_{\cdot,t}(0) \mid D_c = d', S_c = s] - E[\log Y_{\cdot,t'}(0) \mid D_c = d', S_c = s]. \end{aligned}$$

This assumption implies that, in the absence of the reform, expected enrollment would have grown at the same proportional rate across institution clusters within a state. This condition allows enrollment levels to differ across clusters and growth rates to differ across states and cohorts. To see this more clearly, the untreated outcome can be written as

$$Y_{c,t}(0) = \exp\{\alpha_c + \kappa_{s(c),t}\},$$

where α_c is a cluster fixed effect and $\kappa_{s(c),t}$ is a state-by-cohort fixed effect, which absorbs shocks that have a common proportional effect on enrollment across clusters within a state. The assumption rules out differential untreated enrollment growth correlated with exposure within a state. The pre-reform event-study coefficients provide some suggestive evidence: the 2008 and 2009 coefficients are individually and jointly statistically indistinguishable from zero (joint $p = 0.24$). These results provide no evidence of differential pre-reform trends, although they do not directly establish the identifying assumption.

To establish identification, define observed conditional mean enrollment as $m_{d,t,s} = \mathbb{E}[\log Y_{\cdot,t} \mid D_c = d, S_c = s]$. Consider two groups of institution clusters in the same state, with exposures $D_c = d$ and $D_{c'} = d'$, where $d \neq d'$. For any $t \in \mathcal{T}_{\text{post}}$ and $t' \in \mathcal{T}_{\text{pre}}$, Assumptions 1 and 2, together with the observed-outcome relationship above, imply

$$\big(m_{d,t,s} - m_{d,t',s}\big) - \big(m_{d',t,s} - m_{d',t',s}\big) = \beta_1(d - d').$$

Thus, β_1 gets identified. The coefficient β_1 represents the common semi-elasticity of expected enrollment with respect to reform exposure; or a one-percentage-point increase in exposure corresponds to a proportional enrollment change of $\exp(0.01\beta_1) - 1$. The common policy component $\beta_{0,t}$, and hence the reform's total effect at a given exposure level, remains unidentified.

F.2 Construction of Institution Clusters

We identify each student's undergraduate institution from the LinkedIn education record matched to the All India Survey on Higher Education (AISHE) register. Affiliated colleges are assigned to their affiliating universities, so an 'institution" generally corresponds to a university together with its affiliated colleges. Among Indian students who began a master's degree abroad in 2008–2013, this matching identifies 2,843 distinct institutions. However, most are too small to support a reliable institution-specific pre-policy stay rate. We therefore construct institution clusters in three ordered steps using only information from the 2001–2007 cohorts. First, an institution with at least 30 UK master's students in these cohorts forms its own cluster. Second, we pool the remaining institutions within Indian state and bachelor's-institution tier—the three STEM tiers defined in Section 3 and the non-STEM group—into a single "other institutions" cluster for each state and tier. Third, if such a pool still contains fewer than 30 pre-reform UK master's students, we merge it into a single residual cluster for the state. Pools that remain below the threshold after this third step are excluded. Because cluster membership is determined entirely by pre-2008 information, it cannot respond to the reform or to post-reform enrollment.

Using the threshold of 30, the procedure yields 143 clusters across 24 states: 98 standalone institutions, 35 state-by-tier pools, and 10 state-level pools. The retained clusters account for 95% of outbound master's students in the four fields during 2008–2013; the excluded students come from four small states with thin within-state pools. Every cluster has at least 30 pre-reform UK master's students, with a median of 56, and only 4 of the 858 cluster-by-cohort cells from 2008 to 2013 have zero UK entrants. Approximately half of the variation in exposure across clusters is within states, providing the within-state variation needed to include state-by-cohort fixed effects in our main specification. The threshold involves a trade-off between sampling noise in the stay rate and the number of clusters retained. Specifically, a higher threshold improves the reliability of the exposure measure but reduces the number of clusters and the available variation in exposure.

F.3 Empirical-Bayes Shrinkage of the Institution-Level Stay Rate

The raw exposure measure for institution-cluster c is the share $\hat{p}_c$ of its n_c UK master's students in the 2001–2007 cohorts whose first job was in the United Kingdom. As a sample proportion, $\hat{p}_c$ measures the underlying stay propensity p_c with sampling error: $\hat{p}_c = p_c + e_c$, where e_c has variance $s_c^2 = p_c(1-p_c)/n_c$, which we estimate by substituting $\hat{p}_c$ for p_c. With a median of 56 students per cluster, this sampling variance is of the same order as the dispersion in p_c across clusters, implying attenuation bias in a regression using $\hat{p}_c$ in the usual way. In the Empirical Bayes specification, we treat p_c as drawn from a distribution with mean μ_s and variance σ^2 and replace $\hat{p}_c$ with its posterior mean,

$$\tilde{p}_c \;=\; \lambda_c\,\hat{p}_c + (1-\lambda_c)\,\mu_s, \qquad \lambda_c \;=\; \frac{\sigma^2}{\sigma^2 + s_c^2}, \tag{16}$$

where μ_s is the stay rate among pre-reform UK master's students from clusters in state s, and σ^2 is estimated by the method of moments as the cross-cluster variance of $\hat{p}_c$ net of the average sampling variance, $\hat{\sigma}^2 = \mathrm{Var}_c(\hat{p}_c) - \overline{s_c^2}$, with $\hat{\sigma}^2$ bounded below by a small positive number. The weight λ_c captures the reliability of the cluster's own estimate: it approaches one for large clusters (n_c) with small s_c^2, for which the sample proportion is measured precisely, and approaches zero for small clusters, for which the sample proportion is dominated by noise and is therefore shrunk toward the state-level prior. At the threshold of 30 pre-reform students, λ_c ranges from approximately 0.37 for the smallest clusters to 1.0 for the largest, with a mean of 0.61; the pooled reliability, $\hat{\sigma}^2/\mathrm{Var}_c(\hat{p}_c)$, is 0.57.

Shrinkage corrects attenuation bias because $\tilde{p}_c$ is the best linear predictor of p_c given $\hat{p}_c$, so the prediction error $p_c - \tilde{p}_c$ is orthogonal to $\tilde{p}_c$ by construction. A regression on $\tilde{p}_c$ therefore

recovers the coefficient on the true stay propensity, whereas a regression on $\hat{p}_c$ recovers that coefficient scaled by the reliability of the measure (Kane and Staiger, 2008; Chetty et al., 2014). Consistent with this, the estimate in column 4 of Table 1 is approximately 1.5 times the raw estimate in column 2, close to the reciprocal of the estimated reliability. Two limitations apply. First, we estimate a single σ^2 for all states, so the dispersion of true stay propensities is assumed to be common across states. Second, the standard errors in column 5 treat $\tilde{p}_c$ as known and do not account for the estimation of the shrunken measure, following the convention in the value-added literature we adopt. For the decomposition by job tier in Table 2, we apply equation (16) to each tier's share separately, using tier-specific state priors and signal variances; consequently, the three shrunken components need not sum to the shrunken total.

G Structural Model Specification Details

G.1 Definitions of Individuals' Observed Characteristics

The tier-major type M_i indexes an individual's undergraduate institution and field of study. STEM graduates are assigned to one of three institutional quality tiers, and non-STEM graduates form a single residual category, as defined in section 3.3. $G_i \in \{0, 1\}$ indicates undergraduate academic attainment, equal to one if the student's bachelor's education profile reports an honors degree, a first-class distinction, or a high GPA. The undergraduate school state O_i is the Indian state in which the individual completed their Bachelor's degree. We further define the macroeconomic region $R_i = R(O_i) \in \mathcal{R} = \{1, 2, 3, 4, 5\}$(Southern, Western, Northern, Central, Eastern; with Southern being $r = 1$), following the India's administrative Zonal Councils classification.[62]. We will use the less granular macroeconomic region R_i in most of the specifications to avoid the small-sample problem for some worker type cells. The master's enrollment year $t_i \in \{2008, 2009, ..., 2013\}$ is the cohort year of the individual's master's enrollment.

These individual observed states enter the individual's decisions in multiple ways. The tier-major type M_i, Indian state of undergraduate school O_i, and the unobserved ability type z_i affect the value of both the first-stage and second-stage choices. Specifically, the tier-major type M_i and the unobserved ability type z_i affect the expected wage returns associated with each labor-market outcome. Meanwhile, the Indian state of the undergraduate school O_i is a proxy for the individual's state of origin and enters the model through the state-level GDP. It reflects the financial resources available to the individual for their master's education. In

[62]See *Zonal Council-Ministry of Home Affairs*

addition, we allow heterogeneous individual preferences among different (M_i, O_i, z_i) triples. Last but not least, the enrollment cohort t_i directly reflects the policy environment (pre-reform for $t_i \leq 2010$ and post-reform for $t_i \geq 2011$) and the macroeconomic conditions in which the individual makes her decision.

G.2 Wage Equations and Beliefs

The realized log wage of a worker i with bachelor's tier-major type m, attainment g, unobserved ability type z, who took a master's in destination $j_1 \in \mathcal{J}_1$ and works in (j_2, q) in calendar period t is

$$\ln w_i = \underbrace{\alpha_{j_2,q}}_{\text{Country-Job Type}} + \underbrace{u_{j_2,q,t}}_{\text{Macro}} + \underbrace{\delta^w_m}_{\text{UG Tier-Major}} + \underbrace{\delta^w_g}_{\text{UG Attainment}} + \underbrace{\delta^{MA}\mathbf{1}\{j_2 = j_1\}}_{\text{Match Premium}} + \underbrace{\mu_z \mathbf{1}\{z = H, j_2 \neq \text{IN}\}}_{\text{Ability Premium}} + \nu_i, \tag{17}$$

with $\nu_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2_w)$ independent of all choice shocks,[63] and $\sum_t u_{j_2,q,t} = 0$. Here $\sum_t u_{j_2,q,t} = 0$ is a mechanical result since $\alpha_{j_2,q}$ absorbs its expectation across time.

This specification corresponds to the realized wage directly observed in the data, and it permits an additively separable structure for multiple sources of variation in realized wages. The mean wage gap between different groups defined by (m, g, z, j_1, j_2, q) is governed by the fixed effects. $\alpha_{j_2,q}$ is a location-skill fixed effect, which captures the average log wage for jobs of skill level q in country j_2. δ^w_m is a bachelor's tier-major shifter that measures the wage premium of the individual's bachelor institution's tier and bachelor major relative to some reference type (Tier 3 STEM). δ^w_g is a bachelor's education attainment shifter that measures the wage premium for individuals with good education performance in bachelor's studies. δ^{MA} is the match wage premium for students who stay in their master's country. μ_z is the abroad-work wage premium, relative to working in India, for the unobserved high-ability type. In contrast, the variation in realized wages across time is governed by the macro deviations $u_{j_2,q,t}$. $u_{j_2,q,t}$ is a macro shock at the country-skill cell that aggregates business-cycle conditions and other macroeconomic shifters unobserved to the econometrician. We assume $u_{j_2,q,t}$ is observed by students in each period t. Finally, the within-group variation among individuals comes from an i.i.d. idiosyncratic shock ν_i with mean zero and constant variance σ^2_w.[64]

[63] We implicitly assume that any correlation between the idiosyncratic unobserved wage residual and choice probabilities is captured by the unobserved types Z.

[64] This additively separable structure is helpful for identifying certain important parameters, as we will show in later sections.

As said before, we assume the macro deviation follows an AR(1) process,

$$u_{j_2,q,t+1} = \rho u_{j_2,q,t} + \eta_{j_2,q,t+1}, \quad \eta \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_e^2) \tag{18}$$

where $\rho \in (-1, 1)$ is the persistence parameter of macroeconomic shocks and σ_e^2 is the innovation variance.

The AR(1) process determines students' beliefs about their wages upon graduation. Students need to form two beliefs at the two stages using the most recent macroeconomic shock they observe.

For the first belief in the education choice stage, we assume students hold rational expectations over the labor market condition in the future, based on the information set they have before the master's choice. The forecast horizon is determined by the program length: a student in cohort t graduates at $t+1$ or $t+2$, and we assign the horizon $L_{j_1} \in \{1, 2\}$ by rounding the average program length in the data. In particular, the UK master's program averages 1.3 years, so we round its horizon to one period. The horizon of all other destinations is rounded to two periods. Thus, at the first stage, the student projects forward the macro deviation one or two periods later.

Regarding the second belief upon master's graduation and entering the labor market, the student has already observed the macro deviation $u_{\cdot,t}$, so she only needs to take the expectation over the mean-zero idiosyncratic shock ν_i. Writing $(\bar{w}^{(1)}, \bar{w}^{(2)})$ for the beliefs at the first and the second stages, respectively,

$$\bar{w}_t^{(1)}(j_2, q; j_1, m, g, z) = \alpha_{j_2,q} + \delta_m^w + \delta_g^w + \delta^{MA}\mathbf{1}\{j_2 = j_1\} + \mu_z \mathbf{1}\{z = H, j_2 \neq \text{IN}\} + \rho^{L_{j_1}} u_{j_2,q,t-L_{j_1}} \tag{19}$$

$$\bar{w}_t^{(2)}(j_2, q; j_1, m, g, z) = \alpha_{j_2,q} + \delta_m^w + \delta_g^w + \delta^{MA}\mathbf{1}\{j_2 = j_1\} + \mu_z \mathbf{1}\{z = H, j_2 \neq \text{IN}\} + u_{j_2,q,t} \tag{20}$$

It is worth noting that our specification of the belief structure implicitly assumes that students could not forecast changes in future immigration policies or job-finding regimes. This is supported by the quasi-experimental evidence in Figure 3 that we do not see students from different policy exposure react differently prior to the policy. In addition, we assume away the potential general equilibrium effects that the immigration policy might operate through on workers' wages. We leave this issue to future research.

G.3 Non-wage Utility Term

The parametrization and interpretation of this term ξ_t vary across countries. For UK jobs, since we explicitly model the stochastic job arrival, this non-wage utility can be interpreted as an **amenity term**, which includes factors such as career progression, immigration en-

vironment, and home preference relative to the origin. For non-UK jobs, since we do not model job arrival and assume jobs appear in the choice set with probability one, the ξ_t term should be interpreted as a comprehensive **non-wage utility**, which consists of both amenities and job-search frictions. Taking India as the reference alternative, we assume the following piecewise specification of ξ_t:

$$\xi_t(k|j_1, m, g, r, z) = \begin{cases} 0, & k = \text{IN} \\ U_H^{UK} + \alpha^{UK}(r), & k \in \{(\text{UK}, A), (\text{UK}, B)\},\ j_1 = \text{UK} \\ U_L^{UK} + \delta_L^{post}\mathbf{1}\{t \geq t^*\} + \alpha^{UK}(r), & k = (\text{UK}, L),\ j_1 = \text{UK} \\ \xi_{j_1,m}^{stay} + \xi_g^{stay} + \xi_r^{stay} + & \\ \delta_{c(j_1)}^{stay,post}\mathbf{1}\{t \geq T_{c(j_1)}\} + \xi_Z^{stay}\mathbf{1}\{Z = H\}, & k \in \{(j_2, L), (j_2, H)\},\ j_2 = j_1 \neq \text{UK} \\ \xi_{j_2}^{cross}, & \text{all other } k = (j_2, q) \end{cases} \tag{21}$$

Here $\delta_c^{stay,post}$ for $c \in \{\text{US, ANZ}\}$ is an additive post-reform stay shifter at known cohorts T_{US} (2013 OPT STEM extension) and T_{ANZ} (2012 subclass-485 extension).[65] This term is included to capture the effect of immigration policy reforms that happened in non-UK countries contemporaneously.

Several assumptions of this specification of ξ_t are worth discussing. First, we assume the advanced-degree-level level job A and the bachelor-degree-level level job B share the same amenity. Their utilities differ only through $\beta^W \bar{w}^{(2)}$. Second, we assume cross-country non-wage utility only depends on the inflow country, as most of the mobility from the master's country to a third country depends on the specific immigration policies of the latter country. Third, the unobserved ability type Z does not enter the non-wage utility of UK stayers. Since the unobserved type reflects the underlying ability of individuals, it should only explicitly affect the job finding probabilities instead of amenities. In contrast, the unobserved ability type enters the non-wage utilities of the stayers in all other countries, since these non-wage utilities also reflect the job matching probability, which depends on individual's ability.

G.4 Education Consumption Value

We take the UK master's program as the reference group, and assume the location of the bachelor's education and the individual's unobserved ability would affect the education taste for UK master's education. Specifically,

[65] Policy Details TBD here.

$$\varphi(j_1, m, g, r, z) = \begin{cases} \varphi_r^{\text{UK}} + \varphi_z^{\text{UK}} \mathbf{1}\{Z = H\}, & j = \text{UK}, \quad \varphi_1^{\text{UK}} \equiv 0 \\ \varphi_{j_1,m} + \varphi_{j_1,g}, & j \neq \text{UK} \end{cases} \tag{22}$$

G.5 Inclusive Values

Since the individual is forming the expectation of the second stage, we need to specify the individual's belief at the first stage. Recall that the deterministic component of the second-stage utility from (4):

$$v_{2t}(k|j_1, m, g, r, Z) = \beta^W \bar{w}_t^{(2)}(j_2, q; j_1, m, g, Z) + \xi_t(k|j_1, m, g, r, Z) \tag{23}$$

Since both components of v_{2t} are time varying, we need to plug in the individual's first-stage belief to construct their expectation. As defined in (19), $\bar{w}_t^{(1)}$ is the individual's first-stage belief about her wage at time t, which is not affected by the policy reform. In contrast, the expectation of the other time-varying parameters, $(\xi_t(\text{defined in (21) }), \lambda^{Adv.}, \lambda^{Bach.})$, is affected by whether the individual is aware of the reform at the first stage. As discussed in Section 5.2, the pre-2010 cohorts would not expect the policy reform, while the post-2011 cohorts would be fully aware of it. Therefore, the pre-2010 cohorts would expect $(\lambda_{pre}, \xi_{t<t^*})$ while the post-2011 cohorts would expect $(\lambda_{post}, \xi_{t>t^*})$. The 2011 transition cohort is a mixture between the two cases, with probabilities $(\tau, 1-\tau)$ for the pre- and post- beliefs respectively.

Using $\tilde{v}_{2t}$ to denote the first-stage expectation of the deterministic second-stage utility, we write the inclusive values of the master's destinations for the pre-reform and post-reform cohorts. As discussed in Section 5.1, the second-stage choice sets of non-UK masters are deterministic, so the EVT1 shocks give the standard representation of the expected maximum utility.

$$IV_t(j_1|m, g, r, Z) = \log \sum_{k \in \mathcal{J}_2(j_1)} \exp \tilde{v}_{2t}(k|j_1, m, g, r, Z), \quad j_1 \neq \text{UK} \tag{24}$$

For the UK masters, the choice set is stochastic, so the inclusive value is the expectation of the conditional inclusive value taken over the distribution of the stochastic choice set:

$$IV_t(j_1 = \text{UK}|m, g, r, Z) = \sum_{\mathcal{J}_2 \in supp(\pi_J)} \pi_J(\mathcal{J}_2; \lambda) \log \sum_{k \in \mathcal{J}_2} \exp \tilde{v}_{2t}(k|j_1, m, g, r, Z), \quad j_1 = \text{UK} \tag{25}$$

where $\lambda = (\lambda^{Adv.}, \lambda^{Bach.})$ is the job finding probabilities of (A, B) jobs in the UK, and that $\lambda = \lambda_{pre}$ for the pre-reform cohorts and $\lambda = \lambda_{post}$ for the post-reform cohorts.

Finally, since the transition cohort $(t = 2011)$ is a $(\tau, 1-\tau)$ mixture of the pre-reform and

post-reform beliefs, their inclusive value is also a mixture of the pre-reform and post-reform inclusive values,

$$IV_{t^*}(j_1|m,g,r,,Z) = \tau \cdot IV_{t^*}^{pre}(j_1|m,g,r,Z) + (1-\tau) \cdot IV_{t^*}^{post}(j_1|m,g,r,Z) \tag{26}$$

where $(IV_{t^*}^{pre}(j_1|m,g,r,Z), IV_{t^*}^{post}(j_1|m,g,r,Z))$ are the inclusive values that the transition cohort would have if they hold pre-reform and post-reform beliefs, respectively. Note that the inclusive value is the only object for which the belief of the transition cohort matters, since the true regime is fully revealed when the transition cohort graduates and enters the second stage.

H Construction of Master's Program Tuition

This appendix describes how we construct $P_{j_1,t}$, the average annual tuition charged to international master's students at destination $j_1 \in \mathcal{J}_1$ for cohort t. Each series is indexed by the academic year in which the program begins and is expressed in the destination's own currency; we convert these amounts to contemporary Indian rupees using the exchange rate $\chi_{j_1,t}$ obtained from Indian official data.[66] Three of the master's destinations have publicly available official statistics on international master's tuition. For the United Kingdom, we take the overseas classroom-based taught-postgraduate fee reported by each institution in the Reddin Survey of University Tuition Fees and average across institutions within each academic year, using the midpoint when a fee is reported as a range.[67] For the United States, we take the out-of-state tuition for full-time graduate students reported by each institution to IPEDS and average across institutions, weighting by the number of nonresident-alien graduate students.[68] For Canada, we use Statistics Canada's national average tuition for international graduate students.[69]

For the remaining destinations, we cannot find publicly available national tuition statistics online, so we construct tuition series from fee schedules at leading master's destination institutions in our sample. For ANZ, we use international tuition fees for a computing (STEM) and a commerce or economics (non-STEM) master's program at the University of Melbourne and the University of Sydney, weighting the four institution–field cells by their shares among pre-2008 students in our sample.

For Europe (High), we collect non-EU master's tuition fees for an engineering or computer

[66] Source: Handbook of Statistics on the Indian Economy, Table 139
[67] Source: Reddin Survey (Complete University Guide)
[68] Source: IPEDS (National Center for Education Statistics)
[69] Source: Statistics Canada, Table 37-10-0045-01

science program at thirteen institutions in Sweden (KTH, Uppsala, Lund), Ireland (Trinity College Dublin, Limerick, Dublin Business School), the Netherlands (Delft, Amsterdam, Maastricht), and Italy (Politecnico di Milano, Politecnico di Torino, Siena, Trento). Because the current websites do not report historical tuition, we recover period-specific fees from snapshots archived by the Internet Archive's Wayback Machine. For missing years, we interpolate linearly between the nearest documented years and hold fees constant outside the observed range. For Sweden, we assign the full fee in 2011, the first year of tuition charges, rather than interpolate between zero tuition and the 2012 fee. After collecting tuition for each program, we then average tuition fees across institutions within each country and aggregate across countries using pre-2008 sample shares as weights (Sweden 0.54, Ireland 0.19, the Netherlands 0.17, and Italy 0.10). The resulting series rises from approximately €4,300 in 2008–2010 to €14,100 in 2011, reflecting Sweden's introduction of tuition fees for students from outside the EEA and Switzerland in autumn 2011.[70]

For Europe (Low), we assign a constant annual tuition fee of €500. German public universities generally charged no tuition, although five states hosting more than half of international students levied fees of up to €1,000 per year, while French public universities charged annual registration fees of only €226–€254.

I Identification Details

I.1 Observables and Assumptions

Before proceeding to the identification arguments, we will first define the observed population objects and describe the set of technical assumptions that we maintain in addition to the model specification. Throughout the whole section, we define the following observed population objects:

Definition 1 (Population Objects)**.** *For each observed cell* $(M = m, G = g, O = o, T = t)$*, the first-stage master's education choice probability* $S_{1t}(m, g, o)$*, second-stage work location choice probability* $S_{2t}(m, g, o)$*, and wage distribution* F_{wt} *are observed.*

$$S_{1t}(j|m, g, o) := \mathbb{P}(j_1 = j|M = m, G = g, O = o), \quad j \in \mathcal{J}_1 \tag{27}$$

$$S_{2t}(k|j, m, g, o) := \mathbb{P}(j_2 = k|j_1 = j, M = m, G = g, O = o), \quad k \in \mathcal{J}_W(j_1) \tag{28}$$

$$\log w_t|j_2, j_1, m, g \sim F_{wt}(\cdot|j_2, j_1, M = m, G = g) \tag{29}$$

[70] See Uppsala University, "Fees for students from outside Europe established", July 12, 2010.

Meanwhile, we maintain the following technical assumptions:

Assumption 3 (Logit Errors). *ε_{1t} and ε_{2t} are i.i.d. standard Type-I Extreme Value shocks, mutually independent, and independent of all the other variables.*

Assumption 4 (Independent and Excluded Job Arrival). *The arrivals of the two UK high-skill jobs A and B are mutually independent and independent of (ε, ν) conditional on individual states (M, G, O, Z, T).*

Assumption 5 (Policy Exclusion). *The policy reform only changes the job finding probabilities λ and the non-wage utility of L jobs in the UK. All the other parameters are constant across t.*

Assumption 6 (Proportional Policy Effect). *$\lambda^q_{post}(M, G, Z) = \gamma_q \lambda^q_{pre}(M, G, Z)$ with $\gamma_q \in (0, 1)$ not depending on (M, G, Z, O), for $q \in \{A, B\}$.*

Assumption 7 (Interior Support). *$\lambda^q_{pre}(M, G, Z), \lambda^q_{post}(M, G, Z) \in (0, 1)$, $\pi_z(M, G) \in (0, 1)$, $\tau \in (0, 1)$, $\sigma_w > 0$, and every population probability in Definition 1 is strictly positive on its support.*

Assumption 8 (Sufficient Variation). *The observed cells include, for each m, at least one (o, t) with $t \in \mathcal{T}^{pre} \backslash \{t^*\}$ and at least one with $t \in \mathcal{T}^{post} \backslash \{t^*\}$; at least two distinct regions r; and, for the tuition parameters, variation in $P_{j_1,t} \chi_{j,t} L_{j_1}$ across t within $j_1 \in \mathcal{J}_1$ and variation in $W_t(o)$ across o within region r.*

Assumption 9 (Parallel Wage Trend). *For each cell (m, g, o, z), there exists $t_1 \in \mathcal{T}^{pre} \backslash \{t^*\}$ and $t_2 \in \mathcal{T}^{post} \backslash \{t^*\}$ such that*

$$\mathbb{E} w_{t_2}(k | UK, \cdot) - \mathbb{E} w_{t_1}(k | UK, \cdot) = \mathbb{E} w_{t_2}(k' | UK, \cdot) - \mathbb{E} w_{t_1}(k' | UK, \cdot), \quad \forall k, k' \in \mathcal{J}_2(UK) \tag{30}$$

The average observed wage across workers with a UK master's has a common slope across job types from t_1 to t_2.

Under Assumptions 3-9, we establish the identification results in Sections I.2.1-I.2.4 for a version of the model *without* the unobserved ability type, and Section I.2.5 reintroduces the unobserved ability type Z.

I.2 Identification of Different Steps

I.2.1 Step 1: Wage Equations

Without the unobserved ability type Z, the wage equation (17) becomes

$$\ln w_t(k | m, g, j_1) = \alpha_{j_2,q} + u_{j_2,q,t} + \delta^w_m + \delta^w_g + \delta^{MA} \mathbf{1}\{j_2 = j_1\} + \nu_t, \tag{31}$$

with $\nu_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_w^2)$ independent of all choice shocks, and $\sum_t u_{j_2,q,t} = 0$.

The identification of this wage equation can be established by standard panel model techniques. First we take the average of equation (31) across periods to get

$$\frac{1}{T}\sum_t \ln w_t(k|m,g,j_1) = \alpha_{j_2,q} + \delta_m^w + \delta_g^w + \delta^{MA}\mathbf{1}\{j_2 = j_1\} + \frac{1}{T}\sum_t \nu_t \tag{32}$$

Since $\{\nu_t\}_t$ are i.i.d. and independent of all the other variables, $(\alpha_{j_2,q}, \delta_m^w, \delta_g^w, \delta^{MA})$ can be identified by an OLS regression of $\frac{1}{T}\sum_t \ln w_t(k|m,g,j_1)$ on a vector of indicators on the right hand side of (32) given our Assumption 8 Sufficient Variation. To identify the AR(1) process of $u_{j_2,q,t}$, consider the cross-individual expectation of (31) within the same cell (k, m, g, j_1, t):

$$\mathbb{E}\ln w_t(k|m,g,j_1) = \alpha_{j_2,q} + u_{j_2,q,t} + \delta_m^w + \delta_g^w + \delta^{MA}\mathbf{1}\{j_2 = j_1\}, \tag{33}$$

Given the identified $(\alpha_{j_2,q}, \delta_m^w, \delta_g^w, \delta^{MA})$, we can recover $u_{j_2,q,t}$ by

$$u_{j_2,q,t} = \mathbb{E}\ln w_t(k|m,g,j_1) - \alpha_{j_2,q} - \delta_m^w - \delta_g^w - \delta^{MA}\mathbf{1}\{j_2 = j_1\} \tag{34}$$

Finally, since $\eta_{j_2,q,t}$ i.i.d. given (j_2, q) and independent of $u_{j_2,q,t-1}$, we can recover the persistence parameter ρ in (18) again through an OLS regression of $u_{j_2,q,t}$ on $u_{j_2,q,t-1}$.

The variance (σ_w^2, σ_e^2) of $(\nu_t, \eta_{j_2,q})$ are identified through the second moment of the recovered residual terms. From (32) we can recover $\frac{1}{T}\sum_t \nu_t$, and by its i.i.d. distribution, we have

$$\sigma_w^2 = T\cdot var(\frac{1}{T}\sum_t \nu_t) = T\cdot var(\frac{1}{T}\sum_t \ln w_t(k|m,g,j_1) - \alpha_{j_2,q} - \delta_m^w - \delta_g^w - \delta^{MA}\mathbf{1}\{j_2 = j_1\}) \tag{35}$$

Similarly, by the AR(1) representation of $u_{j_2,q,t}$, we have

$$var(u_{j_2,q,t}) = \frac{\sigma_e^2}{1-\rho^2} \Rightarrow \sigma_e^2 = (1-\rho^2)var(u_{j_2,q,t}) \tag{36}$$

I.2.2 Step 2: The Deterministic Second-stage Choices

In this section, we show that the wage coefficient β^W, the non-wage utility ξ_t for the deterministic choices, and the variance of the first-stage EVT1 shocks σ_1^2 are all identified. First we establish Lemma 1 that maps the observed conditional choice probabilities S_{2t} to the deterministic second-stage utilities v_{2t}.

Lemma 1 (IIA of Always-available Choices)**.** *Fix a cell (m, g, o, t) and the first-stage choice*

j_1, for any $k, k' \in \mathcal{J}_2^0(j_1)$ and any $(\lambda^{Adv.}, \lambda^{Bach.}) \in (0,1)^2$,

$$\log \frac{S_{2t}(k|j_1, m, g, o)}{S_{2t}(k'|j_1, m, g, o)} = v_{2t}(k|j_1, m, g, o) - v_{2t}(k'|j_1, m, g, o) \tag{37}$$

That is, the odds between two always-available alternatives obey IIA, as if there were no stochastic choice set at all.

Proof. See Appendix I.3.1 □

Lemma 1 allows us to recover the deterministic second-stage utility v_{2t} as in standard logit conditional choice probabilities (McFadden, 1972). This result, combined with our specification of a common cross-country non-wage utility, helps us identify the wage coefficient β^W.

Proposition 1 (Identification of β^W)**.** *Suppose there exists a first-stage choice j_1, a work country $j_2 \neq j_1, IN$, and a period t with*

$$\bar{w}_t^{(2)}(j_2, H; j_1, m, g) \neq \bar{w}_t^{(2)}(j_2, L; j_1, m, g) \tag{38}$$

then β^W is identified and

$$\beta^W = \frac{\log S_{2t}(j_2, H|j_1, m, g, o) - \log S_{2t}(j_2, L|j_1, m, g, o)}{\bar{w}_t^{(2)}(j_2, H; j_1, m, g) - \bar{w}_t^{(2)}(j_2, L; j_1, m, g)} \tag{39}$$

Proof. See Appendix I.3.2 □

With the wage coefficient β^W, we can recover the non-wage utilities of the deterministic choices using the always-available job in India as the reference choice. Note that in (21) we normalized $\xi_t(\text{IN}|j_1, m, g, o, z) = 0$.

Proposition 2 (Identification of ξ_t for deterministic choices)**.** *Under Assumptions 3-9 and the normalization $\xi_t(IN|j_1, m, g, o, z) = 0$, the second-stage non-wage utility parameters $\xi_t(k|j_1, m, g, o)$ are identified for any always-available choice $k \in \mathcal{J}_2^0(j_1)$, $j_1 \in \mathcal{J}_1$.*

Proof. See Appendix I.3.3 □

After we recover the wage coefficient β^W and the non-wage utilities ξ_t for deterministic second-stage choices, we can recover σ_1^2, the variance of the first-stage EVT1 shocks from the first-stage choice probabilities.

Proposition 3 (Identification of σ_1^2)**.** *Under the conditions of Proposition 1 and Proposition 2, given the calibrated $\kappa_{j_1} = 0.95^{L_{j_1}}/\sigma_1$, σ_1 is identified.*

Proof. See Appendix I.3.4

□

I.2.3 Step 3: The Stochastic Second-stage Choices

From now on, we suppress the dependence of the values on the worker characteristics (m, g, o) and the first-stage choice $j_1 = \text{UK}$. For more convenient notations, we first define a few key objects:

Definition 2 (Single-gated Model)**.** *For the first-stage choice $j_1 = UK$, define the total exponential value of the second-stage always-available set $\mathcal{J}_2^0(j_1 = UK)$ as*

$$B_t := \sum_{k \in \mathcal{J}_2^0(j_1 = UK)} \exp v_{2t}(k) \tag{40}$$

The exponential value of the UK stochastic job H is

$$a_t := \exp v_{2t}(UK, H) \tag{41}$$

The second-stage odds

$$\theta_t := a_t / B_t \tag{42}$$

The first-stage counterparts of (a_t, B_t, θ_t) are $(\tilde{a}_t, \tilde{B}_t, \tilde{\theta}_t)$, which are expected second-stage objects constructed using the first-stage beliefs.

$$\begin{aligned} \tilde{a}_t &:= \exp \tilde{v}_{2t}(UK, H) \\ \tilde{B}_t &:= \sum_{k \in \mathcal{J}_2^0(UK)} \exp \tilde{v}_{2t}(k) \\ \tilde{\theta}_t &:= \tilde{a}_t / \tilde{B}_t \end{aligned} \tag{43}$$

The cross-stage odds ratio

$$\varkappa_t := \tilde{\theta}_t / \theta_t \tag{44}$$

Note that B_t and $\tilde{B}_t$ are functions of observed wages and parameters already identified, so they are both known quantities. Besides, $\varkappa_t$ is also a known positive given the wage equations since it reflects the wage belief adjustment.

Our model specification implies a nice invariance property of $(\theta_t, \tilde{\theta}_t)$ before and after the visa reform, which is stated formally in Lemma 2.

Lemma 2. *There are cohorts $t_1 \in \mathcal{T}^{pre} \backslash \{t^*\}$ and $t_2 \in \mathcal{T}^{post} \backslash \{t^*\}$ at which $\lambda_{pre} \neq \lambda_{post}, \theta_{t_1} = \theta_{t_2}$ and $\tilde{\theta}_{t_1} = \tilde{\theta}_{t_2}$.*

Proof. See Appendix I.3.5 □

With Lemma 2, we provide a constructive identification of $\xi_t(\text{UK}, B)$ in Proposition 4 below.

Proposition 4 (Identification of ξ_t for the UK stochastic job)**.** *Consider the model with only one stochastic UK job in Definition 2 under Assumptions 3-9. Define the observed response ratio*

$$\mathcal{R} := \frac{\kappa_{UK^{-1}}\mathcal{D}(t_1,t_2) - \Delta_{t_1}^{t_2}\log\tilde{B}_t}{S_{2t_2}(H|UK) - S_{2t_1}(H|UK)} \tag{45}$$

where

$$\mathcal{D}(t_1,t_2) := \Delta_{t_1}^{t_2}\left[\log\frac{S_{1t}(UK)}{S_{1t}(j_0)}\right] - \Delta_{t_1}^{t_2}\left[\left(\beta_1 + \beta_2 W_t(o)\right)\left(T_{UK,t} - T_{j_0,t}\right)\right] + \kappa_{j_0}\Delta_{t_1}^{t_2}\left[IV_t(j_0)\right] \tag{46}$$

for any reference destination $j_0 \neq UK$, with $\Delta_{t_1}^{t_2}[f(t)] := f(t_2) - f(t_1)$ differencing the bracketed expression across the two cohorts. Then:

1. *the model implies a testable restriction $\mathcal{R} > \varkappa$;*
2. *θ is uniquely determined, and so is the non-wage utility of the UK stochastic job $\xi_t(UK, H)$;*
3. *the job finding probabilities $(\lambda_{pre}, \lambda_{post})$ are recovered.*

Proof. See Appendix I.3.6 □

I.2.4 Step 4: First-stage Preferences and Transition Weight

Given all the identified second-stage parameters, β_1, β_2, and κ_{j_1} for all $j_1 \in \mathcal{J}_1$, the first-stage parameters are relatively easy to identify. Recall the first-stage conditional choice probability:

$$P_{1t}(j_1|m,g,o,z) = \frac{\exp v_{1t}(j_1|m,g,o,z)}{\sum_{j_1'\in\mathcal{J}_1}\exp v_{1t}(j_1'|m,g,o,z)} \tag{47}$$

Therefore, taking $j_1 = \text{UK}$ and $r(o) = 1$ as the reference level with $\varphi_1^{\text{UK}} = 0$, we have

$$\begin{aligned}
&\log\frac{P_{1t}(j_1|m,g,o,z)}{P_{1t}(\text{UK}|m,g,r(o)=1,z)} \\
&= (\beta_1 + \beta_2 W_t(o))P_{j_1,t}\chi_{j_1,t}L_{j_1} + \varphi(j_1,m,g,r,z) + \kappa_{j_1}IV_t(j_1|m,g,o,z) \\
&- (\beta_1 + \beta_2 W_t(r(o)=1))P_{\text{UK},t}\chi_{\text{UK},t}L_{\text{UK}} - \underbrace{\varphi(\text{UK},m,g,r=1,z)}_{=0} + \kappa_{\text{UK}}IV_t(\text{UK}|m,g,o,z)
\end{aligned} \tag{48}$$

Note that everything above except $\varphi(j_1, m, g, r, z)$ is known, so $\varphi(j_1, m, g, r, z)$ is identified by simply rearranging the equation.

Finally, given all the parameters identified, the transition weight τ for the $t^* = 2011$ cohort is identified by the following proposition.

Proposition 5 (The transition weight τ)**.** *Given all the identified parameters, τ is identified as long as $IV_{t^*}^{pre}(UK/m,\ g,\ o) \neq IV_{t^*}^{post}(UK/m,\ g,\ o)$ for some cell.*

Proof. See Appendix I.3.11 □

I.2.5 Step 5: The Unobserved Type

Steps 1–4 treat every state variable as observed. Once we reintroduce the ability type $Z \in \{H, L\}$, the choice probabilities observed in a cell (m, g, o, t) are mixtures of the two type-specific choice probabilities, weighted by the type composition of the cell. Arguments related to direct logit inversions in Sections I.2.1–I.2.4 do not apply to the mixture.

Our strategy is therefore to recover the type-specific population objects first. Once they are known, Z plays the role of an additional observed state, and proofs in Sections I.2.1–I.2.4 can be applied. In the discussion below, Steps 1–4 identify the model type by type, and the type shifters $(\mu_z, \xi_Z^{stay}, \eta_Z^q, \varphi_z^{UK})$ are the differences between the two sets of type-specific parameters.

In particular, the wage distribution is what allows us to recover the type composition. By Equation (17), the high type earns a premium μ_z in every job outside India, and conditional on the observed states and the type, the wage shock ν_i is independent of the choice shocks. Selection into a choice cell therefore affects the wage distribution of the cell only through Z: within a cell, the log wage is a **two-component normal mixture**, and the mixing weight is the share of high types among the students who made those choices (i.e., posterior). We recover this share from the wage distribution in Step 5.1 and then carry it through the rest of the decision tree with Bayes' rule, first in the second stage (Step 5.2) and then in the first stage (Step 5.3). Step 5.4 recovers the type shifters.

Throughout this section, we fix a cell of observed states (m, g, o, t) and suppress it from the notation except where the dependence matters; every result below holds cell by cell. For any object f that depends on the type, define the type difference

$$\Delta_Z\big[f(z)\big] := f(H) - f(L), \tag{49}$$

Definition 3 (Type-specific Population Objects)**.** *For each type $z \in \{H, L\}$, first-stage choice $j_1 \in \mathcal{J}_1$, and second-stage alternative $k \in \mathcal{J}_2(j_1)$, the type-specific first-stage and*

second-stage choice probabilities are

$$S_{1t}(j_1 \mid z) := \Pr(j_1 \mid Z = z), \tag{50}$$

$$S_{2t}(k \mid j_1, z) := \Pr(j_2 = k \mid j_1, Z = z). \tag{51}$$

The posterior type shares $\Pr(Z = H \mid j_1)$ *and* $\Pr(Z = H \mid j_1, k)$ *are the shares of high types among students who choose* j_1 *in the first stage, and among those who choose* j_1 *and then* k*, respectively.*

The type enters the deterministic second-stage utility v_{2t} in (23) through two terms only: the wage premium μ_z on jobs outside India through the wage belief (20), and the stay shifter ξ_Z^{stay} in (21). The non-wage utilities of the India job, the cross-country jobs, and the UK jobs in (21) do not depend on the type, and neither do the utilities of continued education in (5). Hence, for every job $k = (j_2, q) \in \mathcal{J}_2(j_1)$,

$$\Delta_Z\big[v_{2t}(k \mid j_1, z)\big] = \beta^W \mu_z \, \mathbf{1}\{j_2 \neq \text{IN}\} + \xi_Z^{stay} \, \mathbf{1}\{j_2 = j_1 \neq UK\}, \tag{52}$$

and $\Delta_Z\big[v_{2t}(k \mid j_1, z)\big] = 0$ for $k \in \{\text{PhD}, \text{2nd Master}\}$.

Step 5.1: Type composition of the wage cells. Consider the students who choose j_1 in the first stage and a job $k = (j_2, q)$ in the second stage. By Equation (17), a low type in this cell earns log wage $a(j_1, k) + \nu_i$, where

$$a(j_1, k) := \alpha_{j_2,q} + u_{j_2,q,t} + \delta_m^w + \delta_g^w + \delta^{MA} \mathbf{1}\{j_2 = j_1\}, \tag{53}$$

and a high type earns $a(j_1, k) + \mu_z \mathbf{1}\{j_2 \neq \text{IN}\} + \nu_i$. Because ν_i is independent of the choice shocks given Z, the observed log-wage distribution of a cell (j_1, k) with $j_2 \neq \text{IN}$ is a two-component normal mixture with a common variance:

$$\begin{aligned} F_{wt}(\cdot \mid k, j_1) = \big[1 - \Pr(Z = H \mid j_1, k)\big] \, \mathcal{N}\big(a(j_1, k), \sigma_w^2\big) \\ + \Pr(Z = H \mid j_1, k) \, \mathcal{N}\big(a(j_1, k) + \mu_z, \sigma_w^2\big). \end{aligned} \tag{54}$$

A finite mixture of normal distributions is identified up to the labeling of its components (Teicher, 1963); Compiani and Kitamura (2016) review this and related identification results for mixtures in econometric models. Lemma 3 applies the result to (54).

Lemma 3 (Wage Mixture)**.** *Fix* $j_1 \in \mathcal{J}_1$ *and a job* $k = (j_2, q) \in \mathcal{J}_2(j_1)$ *with* $j_2 \neq \text{IN}$*. Suppose Assumption 7 holds and* $\mu_z \neq 0$*, and label the types so that* $\mu_z > 0$*. Then* $\Pr(Z = H \mid j_1, k)$*,* $a(j_1, k)$*,* μ_z*, and* σ_w^2 *are identified from* $F_{wt}(\cdot \mid k, j_1)$*.*

Proof. See Appendix I.3.7. □

Two features of Lemma 3 deserve discussion. First, the condition $\mu_z \neq 0$ is the substantive one. If the two types earned the same wage abroad, the wage would carry no information about the type, and Step 5.1 would fail. Second, the lemma is silent on the alternatives (India and continued education) whose wage carries no information about the type. Step 5.2 recovers the type composition of these alternatives from the choice data.

Lemma 3 also restores Step 1. The low-type mean $a(j_1, k)$ takes the place of the cell mean in Section I.2.1, and in the India cells, where the log wage is a single normal, the cell mean equals $a(j_1, k)$ directly. Applying the argument of Section I.2.1 to $a(j_1, k)$ identifies $(\alpha_{j_2,q}, \delta^w_m, \delta^w_g, \delta^{MA}, \rho, \sigma^2_e)$.

Step 5.2: Bayes' rule in the second stage. Lemma 3 gives the type composition of every alternative $\Pr(Z = H|j_1, k)$ with an informative wage. To recover the type-specific choice probabilities, we combine Bayes' rule with the logit structure of the second stage. Lemma 4 provides the link: across two always-available alternatives, the posterior log odds of being a high type differ by exactly the type gap in their relative utility.

Lemma 4 (Posterior Log Odds)**.** *Fix $j_1 \in \mathcal{J}_1$. For any two always-available alternatives $k, k' \in \mathcal{J}^0_2(j_1)$,*

$$\Lambda^{-1}\big(\Pr(Z = H \mid j_1, k)\big) - \Lambda^{-1}\big(\Pr(Z = H \mid j_1, k')\big) = \Delta_Z\big[v_{2t}(k \mid j_1, z)\big] - \Delta_Z\big[v_{2t}(k' \mid j_1, z)\big], \tag{55}$$

where $\Lambda^{-1}(x) = \log\big(x/(1-x)\big)$ is the inverse of the logistic cdf in (3).

Proof. See Appendix I.3.8. □

The proof applies Lemma 1 within each type. Conditional on the type, the odds between two always-available alternatives follow IIA even when the UK choice set is stochastic, so the job-finding probabilities of either type do not enter (55). Combined with (52), Lemma 4 turns the posteriors from Step 5.1 into statements about the type shifters in Stage 2.

Proposition 6 (Type Composition in the Second Stage)**.** *Suppose Assumptions 3–9 hold and $\mu_z \neq 0$. Then for every $j_1 \in \mathcal{J}_1$:*

1. *The posterior $\Pr(Z = H \mid j_1, k)$ is the same for every cross-country job $k = (j_2, q)$ with $j_2 \notin \{j_1, \text{IN}\}$ and, when $j_1 = UK$, for the UK low-skill job (UK, L). Proposition 1 continues to hold, and β^W is identified.*

2. *For $j_1 \neq UK$, the type shifter ξ_Z^{stay} in (21) is identified by*

$$\xi_Z^{stay} = \Lambda^{-1}\big(\Pr(Z = H \mid j_1, (j_1, q))\big) - \Lambda^{-1}\big(\Pr(Z = H \mid j_1, k')\big) \tag{56}$$

for any stay job (j_1, q) and any cross-country job k'.

3. *For the alternatives without a type-revealing wage, $k \in \{\text{IN}, \textit{PhD}, \textit{2nd Master}\}$,*

$$\Lambda^{-1}\big(\Pr(Z = H \mid j_1, k)\big) = \Lambda^{-1}\big(\Pr(Z = H \mid j_1, k')\big) - \beta^W \mu_z \tag{57}$$

for any cross-country job k', so $\Pr(Z = H \mid j_1, k)$ is identified.

4. *The type composition of each first-stage choice, $\Pr(Z = z \mid j_1) = \sum_{k' \in \mathcal{J}_2(j_1)} \Pr(Z = z \mid j_1, k')\, S_{2t}(k' \mid j_1)$, is identified, and so are the type-specific second-stage choice probabilities*

$$S_{2t}(k \mid j_1, z) = \frac{\Pr(Z = z \mid j_1, k)\, S_{2t}(k \mid j_1)}{\sum_{k' \in \mathcal{J}_2(j_1)} \Pr(Z = z \mid j_1, k')\, S_{2t}(k' \mid j_1)}, \qquad k \in \mathcal{J}_2(j_1). \tag{58}$$

Proof. See Appendix I.3.9. □

Each part of Proposition 6 has a simple reading. Part 1 says that high types do not sort across third countries. Every cross-country job pays the premium μ_z and carries no other type shifter, so the type composition is the same across these jobs and, for UK graduates, also in the UK low-skill job. Proposition 1 survives for the same reason: it compares two jobs in the same third country, where the premium cancels and β^W is identified. Part 2 identifies the stay shifter from the over-representation of high types among stayers relative to third-country movers. For example, suppose high types make up a larger share of the Canadian master's graduates who stay in Canada than of those who move to a third country. Since both groups earn the premium μ_z, the difference must come from the non-wage value of staying, reflecting the job search friction faced by students with different ability types. Part 3 applies the same logic to returnees. Returning to India forgoes the premium, so in log odds, high types are under-represented among returnees relative to movers by exactly $\beta^W \mu_z$. Continued education carries no type shifter either, so the same holds for PhD and second-master's students.

Part 4 is Bayes' rule. Its denominator requires the posterior for every alternative in $\mathcal{J}_2(j_1)$, including those without an informative wage, which is why Part 3 must come first. Equation (58) also holds when $j_1 = UK$. Bayes' rule does not depend on how the choice set

is formed, and the posteriors of the UK high-skill jobs come directly from Lemma 3 because these jobs pay the premium μ_z.

Step 5.3: Bayes' rule in the first stage. The denominator in Part 4 of Proposition 6 gives the type composition $\Pr(Z = H \mid j_1)$ of every master's destination. The type share of the cell and the type-specific first-stage choice probabilities follow from the law of total probability and Bayes' rule:

$$\pi_z(m, g) = \sum_{j' \in \mathcal{J}_1} \Pr(Z = H \mid j')\, S_{1t}(j'), \tag{59}$$

$$S_{1t}(j_1 \mid z) = \frac{\Pr(Z = z \mid j_1)\, S_{1t}(j_1)}{\sum_{j' \in \mathcal{J}_1} \Pr(Z = z \mid j')\, S_{1t}(j')}, \qquad j_1 \in \mathcal{J}_1. \tag{60}$$

The right-hand side of (59) is computed separately for each origin state o and cohort t, while the type share $\pi_z(m, g)$ varies with neither. This is one of the over-identifying restrictions we collect at the end of this section.

Step 5.4: The type-specific parameters. After Steps 5.1–5.3, the type-specific population objects in Definition 3 are identified for every cell, and so is the wage distribution of each type, given by the two components of the normal mixture distribution in Equation (54). Conditional on $Z = z$, the model has the structure of the model without types in Sections I.2.1–I.2.4, with z added to the cell. Steps 1–4 therefore apply to each type separately, and the type shifters are the type differences of the recovered parameters. The wage premium μ_z comes from Lemma 3, and the stay shifter ξ_Z^{stay} from Proposition 6. Proposition 4, applied to each type, also recovers ξ_Z^{stay} as the type difference of the stay utilities, and Proposition 3, applied to each type, recovers $(\beta_1, \beta_2, \sigma_1)$. The two remaining type shifters, the job-finding premium and education taste, require more work.

The job-finding premium η_Z^q. The type-specific share of UK graduates who take a UK high-skill job is the product of a job-finding probability and a conditional choice probability, so by itself it does not reveal η_Z^q. We separate the two as in Step 3. By Equation (21), the type does not enter the non-wage utility of the UK high-skill jobs. Applying Proposition 8 to the type-specific choice probabilities of either type recovers this common non-wage utility, together with that type's job-finding probabilities. Given the non-wage utility, the odds of the UK high-skill jobs are known for both types, and Lemma 2 recovers the job-finding probabilities of the other type from its second-stage shares alone. By Equation (3), the

job-finding probabilities of the two types differ only through η_Z^q on the logit scale, so

$$\eta_Z^q = \Lambda^{-1}\big(\lambda_{pre}^q(m,g,H)\big) - \Lambda^{-1}\big(\lambda_{pre}^q(m,g,L)\big), \qquad q \in \{\text{Adv.}, \text{Bach.}\}. \tag{61}$$

The education taste shifter φ_z^{UK}. The first stage has a deterministic choice set and logit shocks, so the argument of Lemma 4 carries over. Bayes' rule gives

$$\Lambda^{-1}\big(\Pr(Z = H \mid j_1)\big) = \Lambda^{-1}\big(\pi_z(m,g)\big) + \Delta_Z\big[\log S_{1t}(j_1 \mid z)\big], \qquad j_1 \in \mathcal{J}_1. \tag{62}$$

Differencing Equation (62) between $j_1 = UK$ and a reference destination $j_0 \neq UK$ and substituting the first-stage utility (9) and the education consumption value (22) gives

$$\begin{aligned} &\Lambda^{-1}\big(\Pr(Z = H \mid UK)\big) - \Lambda^{-1}\big(\Pr(Z = H \mid j_0)\big) \\ &\qquad = \varphi_z^{UK} + \kappa_{UK}\,\Delta_Z\big[IV_t(UK \mid z)\big] - \kappa_{j_0}\,\Delta_Z\big[IV_t(j_0 \mid z)\big]. \end{aligned} \tag{63}$$

The tuition terms and the education taste for j_0 cancel because neither depends on the type. The left-hand side is known from Proposition 6. Once the second-stage parameters of both types are known, so are the inclusive values in (24)–(25), and (63) identifies φ_z^{UK}. The equation also shows why the type gap in UK enrollment does not by itself measure a taste for UK education. High types face a different post-graduation choice set: they earn μ_z abroad, receive UK job offers at a different rate, and value staying in other destinations differently. The inclusive-value terms remove these differences, and φ_z^{UK} is what remains.

Finally, the inclusive values of the transition cohort t^* are the τ-mixtures in (26), and Proposition 5 identifies τ for each type. Proposition 7 summarizes the results of this section.

Proposition 7 (Identification with the Unobserved Type)**.** *Suppose Assumptions 3–9 hold, $\mu_z \neq 0$, and the common non-wage utility of the UK high-skill jobs is identified by Proposition 8 for at least one type. Then the type share $\pi_z(m,g)$, the type shifters $(\mu_z, \xi_Z^{stay}, \eta_Z^q, \varphi_z^{UK})$, and all type-specific parameters in Steps 1–4 are identified.*

Proof. The result follows from Lemma 3, Proposition 6, equations (59)–(60), and the type-by-type application of Steps 1–4 in Step 5.4. □

Sources of identification. Step 5.1 is the only place where we use the normality of the wage shock. Steps 5.2–5.4 rely on the logit structure and on the restrictions on where the type enters the model, which Steps 1–4 already maintain. Without a parametric outcome equation, nonparametric identification of a finite mixture typically requires several conditionally independent measurements of the type for each individual (Kasahara and Shimotsu, 2009;

Compiani and Kitamura, 2016). Our data provide two choices, the second of which depends on the first, and one wage per individual, so we rely on the parametric form of the wage shock instead. The additive structure of (17) provides a check on Step 5.1 that does not use normality. Consider two students in the same job $k = (j_2, q)$ abroad who took their master's in two different countries $j_1, j_1' \neq j_2$. By (53), they share the same low-type mean, so the gap between the mean log wages of their cells is $\mu_z\big[\Pr(Z = H \mid j_1, k) - \Pr(Z = H \mid j_1', k)\big]$.

Over-identifying restrictions. With the observed states restored, the model implies more restrictions than identification requires:

1. μ_z and σ_w^2 are common to every wage cell (j_1, k, m, g, t).
2. Within each j_1, the posterior type share is equal across all cross-country jobs and, for UK graduates, the UK low-skill job (Part 1 of Proposition 6).
3. ξ_Z^{stay} is the same for every $j_1 \neq UK$, every pair of a stay job and a cross-country job, and every cell.
4. $\pi_z(m, g)$ does not vary with (o, t).
5. η_Z^q does not vary with (m, g), the policy multiplier γ_q is common to both types, and the non-wage utility of the UK high-skill jobs recovered from the two types coincides.
6. φ_z^{UK} is the same for every reference destination j_0, cohort, and cell, and τ is the same for both types.

The estimator in Section 6.2 imposes these restrictions jointly, which disciplines the type composition recovered from the wage in Step 5.1.

I.3 Other Identification Proofs

I.3.1 Proof of Lemma 1

Proof. First consider the simpler case $j_1 \neq$ UK whose second-stage choice set is deterministic. From the second-stage choice probability representation (6),

$$S_{2t}(k|j_1, m, o) = \frac{\exp v_{2t}(k|j_1, m, o)}{\sum_{\tilde{k} \in \mathcal{J}_2(j_1)} \exp v_{2t}(\tilde{k}|j_1, m, o)} \tag{64}$$

Therefore, taking the log ratio cancels the denominator and leads to (37)

Then consider the case $j_1 = \text{UK}$, whose second-stage choice set is stochastic. By (7) the choice probability of any always-available choice $k \in \mathcal{J}_2^0(\text{UK})$ is

$$\begin{aligned} S_{2t}(k|j_1 = \text{UK}, m, g, o) &= \sum_{\mathcal{J}_2 \in supp(\pi_J)} \pi_J(\mathcal{J}_2; \lambda^{Adv.}, \lambda^{Bach.}) P_{2t}(k|j_1 = \text{UK}, m, g, o, \mathcal{J}_2) \\ &= \sum_{\mathcal{J}_2 \in supp(\pi_J)} \pi_J(\mathcal{J}_2; \lambda^{Adv.}, \lambda^{Bach.}) \frac{\exp v_{2t}(k|\text{UK}, m, g, o)}{\sum_{\tilde{k} \in \mathcal{J}_2(\text{UK})} \exp v_{2t}(\tilde{k}|\text{UK}, m, g, o)} \\ &= \exp v_{2t}(k|\text{UK}, m, g, o) \cdot \sum_{\mathcal{J}_2 \in supp(\pi_J)} \frac{\pi_J(\mathcal{J}_2; \lambda^{Adv.}, \lambda^{Bach.})}{\sum_{\tilde{k} \in \mathcal{J}_2(\text{UK})} \exp v_{2t}(\tilde{k}|\text{UK}, m, g, o)} \end{aligned} \tag{65}$$

Note that $\sum_{\mathcal{J}_2 \in supp(\pi_J)} \frac{\pi_J(\mathcal{J}_2; \lambda^{Adv.}, \lambda^{Bach.})}{\sum_{\tilde{k} \in \mathcal{J}_2(\text{UK})} \exp v_{2t}(\tilde{k}|\text{UK}, m, g, o)}$ does not depend on the always-available choice k, so taking the log ratio again cancels it and leads to (37).

□

I.3.2 Proof of Proposition 1

Proof. Both (j_2, H) and (j_2, L) lie in the deterministic always-available set $\mathcal{J}_2^0$ regardless of the first-stage choice j_1. Note that $j_1 \neq j_2$ and $j_2 \neq \text{IN}$, so according to (21) the second-stage non-wage utility is $\xi_{j_2}^{cross}$ for both choices. By Lemma 1,

$$\begin{aligned} \log \frac{S_{2t}(j_2, H|j_1, m, g, o)}{S_{2t}(j_2, L|j_1, m, g, o)} &= v_{2t}(j_2, H|j_1, m, g, o) - v_{2t}(j_2, L|j_1, m, g, o) \\ &= \beta^W \left(\bar{w}_t^{(2)}(j_2, H; j_1, m, g) - \bar{w}_t^{(2)}(j_2, L; j_1, m, g) \right) + \underbrace{\xi_{j_2}^{cross} - \xi_{j_2}^{cross}}_{=0} \end{aligned} \tag{66}$$

Using (38) and rearranging (66) gives β^W. □

I.3.3 Proof of Proposition 2

Proof. WLOG fix the cell (m, g, o, t) and consider any first-stage choice $j_1 \neq \text{UK}$, whose second-stage choice set is deterministic and always available. The observed second-stage choice probability of a job $(j_2, q) \neq \text{IN}$ is

$$S_{2t}(j_2, q|j_1, m, g, o, z) = \frac{\exp v_{2t}(j_2, q|j_1, m, g, o, z)}{\sum_{k \in \mathcal{J}_2^0(j_1)} \exp v_{2t}(k|j_1, m, g, o, z)} \tag{67}$$

Similarly, the observed second-stage choice probability of the Indian jobs is

$$S_{2t}(\text{IN}|j_1,m,g,o,z)=\frac{\exp v_{2t}(\text{IN}|j_1,m,g,o,z)}{\sum_{k\in\mathcal{J}_2^0(j_1)}\exp v_{2t}(k|j_1,m,g,o,z)} \tag{68}$$

Taking the difference in the log choice probabilities gives

$$\log\frac{S_{2t}(j_2,q|j_1,m,g,o,z)}{S_{2t}(\text{IN}|j_1,m,g,o,z)}=v_{2t}(j_2,q|j_1,m,g,o,z)-v_{2t}(\text{IN}|j_1,m,g,o,z) \tag{69}$$

From (4) we know

$$\begin{aligned}&v_{2t}(j_2,q|j_1,m,g,o,z)-v_{2t}(\text{IN}|j_1,m,g,o,z)\\&=\beta^W(\bar{w}_t^{(2)}(j_2,q;j_1,m,g,z)-\bar{w}_t^{(2)}(\text{IN};j_1,m,g,z))+\xi_t(j_2,q|j_1,m,g,o,z)-\underbrace{\xi_t(\text{IN}|j_1,m,g,o,z)}_{=0}\end{aligned} \tag{70}$$

Rearranging (69) and (70) gives

$$\xi_t(j_2,q|j_1,m,g,o,z)=\log\frac{S_{2t}(j_2,q|j_1,m,g,o,z)}{S_{2t}(\text{IN}|j_1,m,g,o,z)}-\beta^W(\bar{w}_t^{(2)}(j_2,q;j_1,m,g,z)-\bar{w}_t^{(2)}(\text{IN};j_1,m,g,z)) \tag{71}$$

The RHS of (71) only contains observed quantities and already identified parameters, so $\xi_t(j_2,q|j_1,m,g,o,z)$ is identified. □

I.3.4 Proof of Proposition 3

Proof. WLOG fix the cell (m,g,o) and consider two first-stage choices $j_1,j_1'\neq\text{UK}$ and two time periods $t_1\neq t_2$. By (10), the first-stage choice probability is:

$$S_t(j_1|m,g,o,z)=\frac{\exp v_{1t}(j_1|m,g,o,z)}{\sum_{j\in\mathcal{J}_1}\exp v_{1t}(j|m,g,o,z)} \tag{72}$$

Taking the difference in log $S_t(j_1|m,g,o,z)$ and $S_t(j_1'|m,g,o,z)$ gives

$$\begin{aligned}&\log\frac{S_t(j_1|m,g,o,z)}{S_t(j_1'|m,g,o,z)}=v_{1t}(j_1|m,g,o,z)-v_{1t}(j_1'|m,g,o,z)\\&=\big(\beta_1+\beta_2W_t(o)\big)\Big(P_{j_1,t}\chi_{j_1,t}L_{j_1}-P_{j_1',t}\chi_{j_1',t}L_{j_1'}\Big)+\varphi(j_1,m,g,r,z)-\varphi(j_1',m,g,r,z)\\&+\kappa_{j_1}IV_t(j_1|m,g,r,z)-\kappa_{j_1'}IV_t(j_1'|m,g,r,z)\end{aligned} \tag{73}$$

Since the first-stage education consumption values $\big(\varphi(j_1,m,g,r,z),\varphi(j_1',m,g,r,z)\big)$ are time-invariant, they are canceled in the difference of (73) between t_1 and t_2. Besides, we calibrate

$\kappa_j = 0.95^{L_j}/\sigma_1$, so we have

$$\begin{aligned}
&\log \frac{S_{t_2}(j_1|m,g,o,z)}{S_{t_2}(j_1'|m,g,o,z)} - \log \frac{S_{t_1}(j_1|m,g,o,z)}{S_{t_1}(j_1'|m,g,o,z)} \\
&= \left(\beta_1 + \beta_2 W_{t_2}(o)\right)\left(P_{j_1,t_2}\chi_{j_1,t_2}L_{j_1} - P_{j_1',t_2}\chi_{j_1',t_2}L_{j_1'}\right) - \left(\beta_1 + \beta_2 W_{t_1}(o)\right)\left(P_{j_1,t_1}\chi_{j_1,t_1}L_{j_1} - P_{j_1',t_1}\chi_{j_1',t_1}L_{j_1'}\right) \\
&+ \frac{0.95^{L_{j_1}}}{\sigma_1}\left(IV_{t_2}(j_1|m,g,r,z) - IV_{t_1}(j_1|m,g,r,z)\right) - \frac{0.95^{L_{j_1'}}}{\sigma_1}\left(IV_{t_2}(j_1'|m,g,r,z) - IV_{t_1}(j_1'|m,g,r,z)\right)
\end{aligned} \tag{74}$$

By Assumption 8, we have sufficient variation in our observed data, so $\beta_1, \beta_2, \sigma_1$ are jointly identified by the above regression.

□

I.3.5 Proof of Lemma 2

Proof. We first explicitly write down the components of the second-stage utility $v_{2t}(\text{UK}, B)$

$$v_{2t}(\text{UK}, B) = \beta^W \bar{w}_t^{(2)}(\text{UK}, B) + \xi_t(\text{UK}, B) \tag{75}$$

Plugging in $\bar{w}_t^{(2)}(\text{UK}, B)$ from (20) and noting $\xi_t(\text{UK}, B) = \xi(\text{UK}, B)$ is time-invariant, we have

$$v_{2t}(\text{UK}, B) = \beta^W \left[\alpha_{\text{UK},B} + \delta_m^w + \delta^{MA} + u_{\text{UK},B,t}\right] + \xi(\text{UK}, B) \tag{76}$$

Now consider $t_1 \in \mathcal{T}^{pre}\backslash\{t^*\}$ and $t_2 \in \mathcal{T}^{post}\backslash\{t^*\}$. the difference between their second-stage utilities can be written as

$$v_{2t_2}(\text{UK}, B) - v_{2t_1}(\text{UK}, B) = \beta^W \left(u_{\text{UK},B,t_2} - u_{\text{UK},B,t_1}\right) \tag{77}$$

This difference in second-stage utilities comes from the difference in wages only, so by Assumption 9, this difference is common across any job type $q \in \{A, B, L\}$ in the UK between t_1 and t_2.

$$v_{2t_2}(\text{UK}, q) - v_{2t_1}(\text{UK}, q) = \beta^W \left(u_{\text{UK},B,t_2} - u_{\text{UK},B,t_1}\right) = \beta^W \Delta_{t_1}^{t_2} u_{\text{UK}} \tag{78}$$

Then consider $(\theta_{t_1}, \theta_{t_2})$. Plugging in a_t and B_t, we have

$$\begin{aligned}
\theta_{t_2} &= \frac{a_{t_2}}{B_{t_2}} = \frac{\exp v_{2t_2}(\text{UK}, B)}{\sum_{k \in \mathcal{J}_2^0(\text{UK})} \exp v_{2t_2}(k)} \\
&= \frac{\exp\left(v_{2t_1}(\text{UK}, B) + \beta^W \Delta_{t_1}^{t_2} u_{\text{UK}}\right)}{\sum_{k \in \mathcal{J}_2^0(\text{UK})} \exp\left(v_{2t_1}(k) + \beta^W \Delta_{t_1}^{t_2} u_{\text{UK}}\right)} \\
&= \frac{\exp\left(\beta^W \Delta_{t_1}^{t_2} u_{\text{UK}}\right) \exp v_{2t_1}(\text{UK}, B)}{\exp\left(\beta^W \Delta_{t_1}^{t_2} u_{\text{UK}}\right) \sum_{k \in \mathcal{J}_2^0(\text{UK})} \exp v_{2t_1}(k)} \\
&= \frac{\exp v_{2t_1}(\text{UK}, R)}{\sum_{k \in \mathcal{J}_2^0(\text{UK})} \exp v_{2t_1}(k)} \\
&= \theta_{t_1}
\end{aligned} \tag{79}$$

Repeating the same steps with $\tilde{v}_{2t}$ and $\bar{w}_t^{(1)}(\text{UK}, R)$ would give $\tilde{\theta}_{t_1} = \tilde{\theta}_{t_2}$.

Finally, consider $\tilde{B}_{t_1}$ and $\tilde{B}_{t_2}$. Again we write down explicitly the components of $\tilde{B}_t$.

$$\begin{aligned}
\tilde{B}_t &= \sum_{k \in \mathcal{J}_2^0(\text{UK})} \exp \tilde{v}_{2t}(k) \\
&= \sum_{k \in \mathcal{J}_2^0(\text{UK})} \exp\left(\beta^W \bar{w}_t^{(2)}(k) + \xi(k)\right) \\
&= \sum_{k \in \mathcal{J}_2^0(\text{UK})} \exp\left(\beta^W (\alpha_k + \delta_m^w + \delta^{MA} + \rho^{L_{\text{UK}}} u_{k,t}) + \xi(k)\right)
\end{aligned} \tag{80}$$

□

I.3.6 Proof of Proposition 4

We provide a constructive proof of Proposition 4 in this section.

Proof. The proof proceeds in five steps.

Step 1: the second-stage response. The UK choice set is stochastic: $\mathcal{J}^1 = \mathcal{J}_2^0(\text{UK}) \cup \{B\}$ occurs with probability λ and $\mathcal{J}^2 = \mathcal{J}_2^0(\text{UK})$ occurs with probability $1 - \lambda$. Therefore, the observed second-stage choice probability of the high-skill job B is

$$S_{2t}(B) = \lambda \cdot P_{2t}(B) + (1 - \lambda) \cdot 0 = \lambda \cdot \frac{a_t}{a_t + B_t} \tag{81}$$

Define $\pi^*(\theta) = \frac{\theta}{1+\theta}$, then the observed second-stage choice probability can be written as

$$S_{2t}(B) = \lambda \pi^*(\theta_t) \tag{82}$$

By Lemma 2, $\theta_{t_1} = \theta_{t_2} = \theta$, so the difference of the observed second-stage choice probabilities between t_1 and t_2 is proportional to the difference in the job finding probability.

$$\Delta S_2 := S_{2t_2}(B) - S_{2t_1}(B) = (\lambda^{post} - \lambda^{pre})\pi^*(\theta) = \Delta\lambda \pi^*(\theta) \tag{83}$$

By Assumption 6, $\Delta\lambda = \lambda^{post} - \lambda^{pre} \neq 0$.

Step 2: the first-stage response. From (25), the inclusive values can be written as

$$IV_t(UK) = \lambda \log(\tilde{a}_t + \tilde{B}_t) + (1-\lambda)\log \tilde{B}_t = \log \tilde{B}_t + \lambda \log(1+\tilde{\theta}_t) \tag{84}$$

Again by Lemma 2, $\tilde{\theta}_{t_1} = \tilde{\theta}_{t_2} = \tilde{\theta}$, so the difference of the inclusive values between t_1 and t_2 is

$$\Delta IV(\text{UK}) := IV_{t_2}(UK) - IV_{t_1}(UK) = \log \tilde{B}_{t_2} - \log \tilde{B}_{t_1} + \Delta\lambda \log(1+\tilde{\theta}) \tag{85}$$

Step 3: $\mathcal{D}$ measures $\kappa_{\text{UK}}\Delta IV(\text{UK})$. From the first-stage choice probability representation (10), for any $j_0 \in \mathcal{J}_1, j_0 \neq \text{UK}$

$$\log \frac{S_{1t}(\text{UK})}{S_{1t}(j_0)} = \left(\beta_1 + \beta_2 W_t(o)\right)\left(T_{\text{UK},t} - T_{j_0,t}\right) + \varphi(\text{UK}) - \varphi(j_0) + \kappa_{\text{UK}} IV_t(\text{UK}) - \kappa_{j_0} IV_t(j_0) \tag{86}$$

Applying $\Delta_{t_1}^{t_2}$, the education consumption values $(\varphi(\text{UK}), \varphi(j_0))$ are time-invariant and canceled. Rearranging gives exactly $\mathcal{D}(t_1, t_2) = \kappa_{\text{UK}}\Delta IV(\text{UK})$. Note that $\mathcal{D}(t_1, t_2)$ is a function of observed quantities and parameters already identified, so it is known. So is κ_{UK}.

Step 4: the ratio. Combining all the results above, the ratio can be written as

$$\begin{aligned}
\mathcal{R} &:= \frac{\kappa_{\text{UK}^{-1}}\mathcal{D}(t_1,t_2) - \Delta_{t_1}^{t_2}\log \tilde{B}_t}{S_{2t_2}(R|UK) - S_{2t_1}(R|UK)} \\
&= \frac{\Delta IV(\text{UK}) - \Delta_{t_1}^{t_2}\log \tilde{B}_t}{\Delta\lambda\pi^*(\theta)} \\
&= \frac{\Delta\lambda \log(1+\tilde{\theta})}{\Delta\lambda\pi^*(\theta)} \\
&= \frac{\log(1+\tilde{\theta})}{\theta/(1+\theta)} \\
&= \frac{(1+\theta)\log(1+\varkappa\theta)}{\theta}
\end{aligned} \tag{87}$$

Step 5: inversion and recovery. Define the function

$$\Phi_{\varkappa}(\theta) = \frac{(1+\theta)\log(1+\varkappa\theta)}{\theta} \tag{88}$$

It is easy to show that $\Phi_{\varkappa}(\theta)$ is continuous and strictly increasing on $(0,\infty)$ for $\varkappa \in (0,2]$, and it is a bijection from $(0,\infty)$ to $(\varkappa,\infty)$. Therefore, θ is identified as the unique solution to the equation

$$\mathcal{R} = \Phi_{\varkappa}(\theta). \tag{89}$$

With $\theta_{t_1} = \theta_{t_1} = \theta$, for $t = t_1$ or $t = t_2$, we can recover $\xi(\text{UK})$ through

$$\begin{aligned} v_{2t}(\text{UK}) &= \log a_t = \log\left(\theta \cdot B_t\right) \\ \xi(\text{UK}, R) &= v_{2t}(\text{UK}, R) - \beta^W \bar{w}_t^{(2)}(\text{UK}, R) \end{aligned} \tag{90}$$

Given θ, the latent conditional choice probability $\pi^*(\theta)$ is known, so λ is identified by taking the ratio between the observed choice probability and the latent conditional choice probability.

$$\lambda^{pre} = \frac{S_{2t_1}(\text{UK, q})}{\pi^*(\theta)} \tag{91}$$

$$\lambda^{post} = \frac{S_{2t_2}(\text{UK, q})}{\pi^*(\theta)} \tag{92}$$

□

I.3.7 Proof of Lemma 3

Proof. Fix j_1 and a job $k = (j_2, q)$ with $j_2 \neq \text{IN}$, and write $p := \Pr(Z = H \mid j_1, k)$ and $a := a(j_1, k)$. By (17), conditional on (j_1, k) and $Z = z$, the log wage equals $a + \mu_z \mathbf{1}\{z = H\} + \nu_i$. The choices (j_1, k) depend only on the observed states, the type, the choice shocks, and the job arrivals. Given the observed states and Z, the wage shock ν_i is independent of the choice shocks by (17) and of the job arrivals by Assumption 4, so conditioning on (j_1, k) leaves its distribution at $\mathcal{N}(0, \sigma_w^2)$. Integrating over Z given (j_1, k) yields (54). Under Assumption 7, $\pi_z(m, g) \in (0, 1)$ and every type-specific choice probability is strictly positive, so $p \in (0, 1)$.

The moment generating function of the log wage in the cell,

$$M(s) := E\big[\exp(s \log w) \mid j_1, k\big] = \exp\left(as + \tfrac{1}{2}\sigma_w^2 s^2\right)\big[1 - p + p\exp(\mu_z s)\big], \qquad s \in \mathbb{R}, \tag{93}$$

is finite for every s and known, since $F_{wt}(\cdot \mid k, j_1)$ is observed. Let $G(s) := \log M(s)$. The

last term of $G(s)$ grows at most linearly in $|s|$, so

$$\sigma_w^2 = 2 \lim_{s\to\infty} \frac{G(s)}{s^2}. \tag{94}$$

Given σ_w^2, define $g(s) := G(s) - \frac{1}{2}\sigma_w^2 s^2 = as + \log\big(1 - p + p\exp(\mu_z s)\big)$. With $\mu_z > 0$,

$$\lim_{s\to+\infty} \big[g(s) - (a+\mu_z)s\big] = \log p, \qquad \lim_{s\to-\infty} \big[g(s) - as\big] = \log(1-p). \tag{95}$$

The function g therefore has two linear asymptotes. Their slopes identify $a+\mu_z$ and a, and hence a and μ_z; either intercept identifies p. If the types were labeled the other way, so that $\mu_z < 0$, the two asymptotes would switch roles. This is why the lemma needs a labeling convention but no further restriction. The lemma is a special case of the identification result in Teicher (1963); the argument above gives the parameters in closed form. □

I.3.8 Proof of Lemma 4

Proof. By Bayes' rule, for $z \in \{H, L\}$ and any $k \in \mathcal{J}_2(j_1)$,

$$\Pr(Z = z \mid j_1, k) = \frac{\Pr(Z = z \mid j_1)\, S_{2t}(k \mid j_1, z)}{S_{2t}(k \mid j_1)}. \tag{96}$$

Taking the log ratio between $z = H$ and $z = L$ gives

$$\Lambda^{-1}\big(\Pr(Z = H \mid j_1, k)\big) = \Lambda^{-1}\big(\Pr(Z = H \mid j_1)\big) + \Delta_Z\big[\log S_{2t}(k \mid j_1, z)\big]. \tag{97}$$

The first term on the right-hand side does not depend on k, so differencing (97) between k and k' gives

$$\Lambda^{-1}\big(\Pr(Z = H \mid j_1, k)\big) - \Lambda^{-1}\big(\Pr(Z = H \mid j_1, k')\big) = \Delta_Z\left[\log \frac{S_{2t}(k \mid j_1, z)}{S_{2t}(k' \mid j_1, z)}\right]. \tag{98}$$

Conditional on $Z = z$, the second stage is the model without unobserved types, i.e., with z added to the cell, and the proof of Lemma 1 goes through unchanged with the job-finding probabilities of type z. Hence, for $k, k' \in \mathcal{J}_2^0(j_1)$, the right hand side equals to:

$$\log \frac{S_{2t}(k \mid j_1, z)}{S_{2t}(k' \mid j_1, z)} = v_{2t}(k \mid j_1, z) - v_{2t}(k' \mid j_1, z), \qquad z \in \{H, L\}. \tag{99}$$

Substituting into (98) gives (55). □

I.3.9 Proof of Proposition 6

Proof. Fix j_1. Every cross-country job lies in the always-available set $\mathcal{J}_2^0(j_1)$, and so do the India job, the two continued-education options and, when $j_1 = UK$, the UK low-skill job. We take the four parts in turn.

Part 1. Let $k = (j_2, q)$ and $k' = (j_2', q')$ be two cross-country jobs, with $j_2, j_2' \notin \{j_1, \text{IN}\}$. By (52), $\Delta_Z[v_{2t}(k \mid j_1, z)] = \Delta_Z[v_{2t}(k' \mid j_1, z)] = \beta^W \mu_z$, so Lemma 4 implies $\Pr(Z = H \mid j_1, k) = \Pr(Z = H \mid j_1, k')$. The same holds for $k' = (UK, L)$ when $j_1 = UK$, since the UK low-skill job pays the premium μ_z and its non-wage utility in (21) does not depend on the type. Now take two jobs $k = (j_2, q)$ and $k' = (j_2, q')$ in the same third country j_2, as in Proposition 1. Because their posteriors are equal, (96) gives

$$\frac{S_{2t}(k \mid j_1, z)}{S_{2t}(k' \mid j_1, z)} = \frac{\Pr(Z = z \mid j_1, k)}{\Pr(Z = z \mid j_1, k')} \cdot \frac{S_{2t}(k \mid j_1)}{S_{2t}(k' \mid j_1)} = \frac{S_{2t}(k \mid j_1)}{S_{2t}(k' \mid j_1)}, \qquad z \in \{H, L\}, \tag{100}$$

so the observed odds equal the type-specific odds. By Lemma 1 applied to either type, the log of these odds equals $\beta^W \left[\bar{w}_t^{(2)}(j_2, q; j_1, z) - \bar{w}_t^{(2)}(j_2, q'; j_1, z)\right] = \beta^W \left[a(j_1, k) - a(j_1, k')\right]$, where the premium μ_z and the cross-country utility $\xi_{j_2}^{cross}$ cancel. The low-type means are identified by Lemma 3. Therefore, Equation (39) holds with the wage difference replaced by $a(j_1, k) - a(j_1, k')$, and β^W is identified whenever this difference is nonzero.

Part 2. Let $j_1 \neq UK$, let $k = (j_1, q)$ be a stay job, and let k' be a cross-country job. By (52), $\Delta_Z[v_{2t}(k \mid j_1, z)] = \beta^W \mu_z + \xi_Z^{stay}$ and $\Delta_Z[v_{2t}(k' \mid j_1, z)] = \beta^W \mu_z$. Lemma 4 then gives (56). Neither job is in India, so both posteriors on the right-hand side of (56) are identified by Lemma 3.

Part 3. Let $k \in \{\text{IN}, \text{PhD}, \text{2nd Master}\}$ and let k' be a cross-country job. By (52), $\Delta_Z[v_{2t}(k \mid j_1, z)] = 0$ and $\Delta_Z[v_{2t}(k' \mid j_1, z)] = \beta^W \mu_z$. Lemma 4 then gives (57). The posterior of k' and the premium μ_z are identified by Lemma 3, and β^W by Part 1.

Part 4. Every alternative in $\mathcal{J}_2(j_1)$ is either a job outside India, whose posterior is identified by Lemma 3, or one of the alternatives in Part 3. When $j_1 = UK$, the first group includes the stochastic UK high-skill jobs, since they pay the premium μ_z. The law of total probability then gives $\Pr(Z = z \mid j_1) = \sum_{k' \in \mathcal{J}_2(j_1)} \Pr(Z = z \mid j_1, k')\, S_{2t}(k' \mid j_1)$, and rearranging (96) gives (58). □

I.3.10 Identification of Two Stochastic Choices

In this section, we extend the identification argument in Section I.2.3 from a single stochastic choice to the full specification with two independently arriving UK jobs (A, B). As in the first four steps of the identification section, we consider the model without the unobserved

ability type. We take the wage parameters, β^W, the non-wage utilities of the always-available choices, and $(\beta_1, \beta_2, \kappa_{j_1})$ as known from Sections I.2.1 and I.2.2.

The main idea is the same as in Section I.2.3. We use the first-stage response to the reform to identify η_H. The common non-wage utility of the two jobs $\{A, B\}$ in (21) is important: although there are two stochastic choices, their preferences contain only one remaining unknown parameter η_H. The first-stage response therefore gives an equation in a single unknown. Unlike the single-choice argument, the construction below allows the odds and the always-available values to vary across cohorts. In particular, it accounts for the already identified change Δ_L in the non-wage utility of the UK low-skill job.

Step 1: the common non-wage utility. Fix a cell (m, o) and suppress the dependence on (m, o) and the first-stage choice $j_1 = \text{UK}$. For $q \in \{A, B\}$, define the exponential values and odds as

$$\begin{aligned} a_{qt} &:= \exp v_{2t}(\text{UK}, q), \qquad & \theta_{qt} &:= \frac{a_{qt}}{B_t}, \\ \tilde{a}_{qt} &:= \exp \tilde{v}_{2t}(\text{UK}, q), \qquad & \tilde{\theta}_{qt} &:= \frac{\tilde{a}_{qt}}{\tilde{B}_t}, \end{aligned} \tag{101}$$

where B_t and $\tilde{B}_t$ are the sums over the always-available set $\mathcal{J}_2^0(\text{UK})$ defined in Definition 2. Both quantities are known. Since the two stochastic jobs share the non-wage utility $\eta_H + \omega_{r(o)}$, their odds can be written as

$$\begin{aligned} \theta_{qt}(\eta_H) &= e^{\eta_H} \zeta_{qt}, \qquad & \zeta_{qt} &:= \frac{\exp\left\{\beta^W \bar{w}_t^{(2)}(\text{UK}, q) + \omega_{r(o)}\right\}}{B_t}, \\ \tilde{\theta}_{qt}(\eta_H) &= e^{\eta_H} \tilde{\zeta}_{qt}, \qquad & \tilde{\zeta}_{qt} &:= \frac{\exp\left\{\beta^W \bar{w}_t^{(1)}(\text{UK}, q) + \omega_{r(o)}\right\}}{\tilde{B}_t}. \end{aligned} \tag{102}$$

The quantities $(\zeta_{qt}, \tilde{\zeta}_{qt})$ are known and strictly positive. In particular,

$$\frac{\theta_{At}}{\theta_{Bt}} = \exp\left\{\beta^W \left[\bar{w}_t^{(2)}(\text{UK}, A) - \bar{w}_t^{(2)}(\text{UK}, B)\right]\right\} \tag{103}$$

is known, and the same holds for the ratio of the first-stage odds. However, the levels of these odds are not yet identified. They become known only conditional on a candidate η_H.

Step 2: the second-stage response. Write $\lambda_{A,t}$ and $\lambda_{B,t}$ for the job finding probabilities applicable to cohort t, and abbreviate $S_{2t}(q) := S_{2t}(\text{UK}, q|\text{UK}, m, o)$. Summing over the

four possible choice sets gives

$$\begin{aligned} S_{2t}(A) &= \lambda_{A,t}(1-\lambda_{B,t})\frac{\theta_{At}}{1+\theta_{At}} + \lambda_{A,t}\lambda_{B,t}\frac{\theta_{At}}{1+\theta_{At}+\theta_{Bt}}, \\ S_{2t}(B) &= \lambda_{B,t}(1-\lambda_{A,t})\frac{\theta_{Bt}}{1+\theta_{Bt}} + \lambda_{A,t}\lambda_{B,t}\frac{\theta_{Bt}}{1+\theta_{Bt}+\theta_{At}}. \end{aligned} \tag{104}$$

For more convenient notation, define

$$p_{qt} := \frac{\theta_{qt}}{1+\theta_{qt}}, \qquad c_{qt} := p_{qt} - \frac{\theta_{qt}}{1+\theta_{At}+\theta_{Bt}} > 0, \quad q \in \{A, B\}. \tag{105}$$

Then (104) becomes

$$\begin{aligned} S_{2t}(A) &= \lambda_{A,t}\left(p_{At} - \lambda_{B,t}c_{At}\right), \\ S_{2t}(B) &= \lambda_{B,t}\left(p_{Bt} - \lambda_{A,t}c_{Bt}\right). \end{aligned} \tag{106}$$

The following lemma shows that, conditional on η_H, these two equations uniquely recover the job finding probabilities whenever an interior solution exists.

Lemma 5 (Conditional identification of the job finding probabilities)**.** *Fix a cohort t and a candidate η_H. The system (106) has at most one solution $(\lambda_{A,t}, \lambda_{B,t}) \in (0,1)^2$. If such a solution exists, $\lambda_{B,t}$ is the unique root in $(0,1)$ of*

$$\begin{aligned} p_{Bt}c_{At}\lambda_{B,t}^2 - \left[p_{At}p_{Bt} - S_{2t}(A)c_{Bt} + S_{2t}(B)c_{At}\right]\lambda_{B,t} \\ + S_{2t}(B)p_{At} = 0, \end{aligned} \tag{107}$$

and

$$\lambda_{A,t} = \frac{S_{2t}(A)}{p_{At} - \lambda_{B,t}c_{At}}. \tag{108}$$

Proof. For any $\lambda_{A,t} \in [0,1]$, the denominator in (108) is strictly positive since

$$p_{At} - \lambda_{B,t}c_{At} \geq p_{At} - c_{At} = \frac{\theta_{At}}{1+\theta_{At}+\theta_{Bt}} > 0. \tag{109}$$

Therefore, the first equation in (106) gives (108). Substituting it into the second equation and collecting terms gives (107).

Now suppose $(\lambda_{A,t}, \lambda_{B,t}) \in (0,1)^2$ is a solution. By the product of the roots of a quadratic,

the other root of (107) is

$$\begin{aligned}\lambda'_{B,t} &= \frac{S_{2t}(B)p_{At}}{p_{Bt}c_{At}\lambda_{B,t}} = \frac{p_{At}}{c_{At}}\left(1-\lambda_{A,t}\frac{c_{Bt}}{p_{Bt}}\right) \\ &= \frac{1+\theta_{At}+\theta_{Bt}-\lambda_{A,t}\theta_{At}}{\theta_{Bt}} \\ &= \frac{1+\theta_{Bt}+(1-\lambda_{A,t})\theta_{At}}{\theta_{Bt}} > 1.\end{aligned} \tag{110}$$

where the second equation follows from equation (106); the third equation follows from equation (105). Hence the other root is greater than one and cannot be a probability. The admissible $\lambda_{B,t}$ is unique, and (108) uniquely determines $\lambda_{A,t}$. □

For any candidate η_H admitting an interior solution, denote the recovered probabilities as $(\hat{\lambda}_{A,t}(\eta_H), \hat{\lambda}_{B,t}(\eta_H))$. They can be calculated directly by taking the smaller root of (107) and then applying (108). This step alone does not identify η_H: different candidate non-wage utilities may be matched by different job finding probabilities.

Step 3: the first-stage response. Fix $t_1 \in \mathcal{T}^{pre}\backslash\{t^*\}$ and $t_2 \in \mathcal{T}^{post}\backslash\{t^*\}$. At these two cohorts, students' first-stage beliefs about job arrival agree with the applicable policy regime. Write $\hat{\lambda}_q^{pre}(\eta_H) := \hat{\lambda}_{q,t_1}(\eta_H)$ and $\hat{\lambda}_q^{post}(\eta_H) := \hat{\lambda}_{q,t_2}(\eta_H)$ for $q \in \{A, B\}$. Define the admissible set

$$\begin{aligned}\mathcal{E} := \big\{\eta_H \in \mathbb{R} : &\text{ both cohort inversions exist in } (0,1)^2, \\ &0 < \hat{\lambda}_q^{post}(\eta_H) < \hat{\lambda}_q^{pre}(\eta_H) < 1, \quad q \in \{A, B\}\big\}.\end{aligned} \tag{111}$$

The second restriction incorporates the proportional policy effect in Assumption 6. For this fixed m, it requires the implied policy multipliers to lie in $(0, 1)$; their equality across different m will provide additional restrictions below.

For each candidate $\eta_H \in \mathcal{E}$, define

$$\begin{aligned}\ell_{qt}(\eta_H) &:= \log\left(1+\tilde{\theta}_{qt}(\eta_H)\right), \qquad q \in \{A, B\}, \\ \Delta_{AB,t}(\eta_H) &:= \log\frac{1+\tilde{\theta}_{At}(\eta_H)+\tilde{\theta}_{Bt}(\eta_H)}{\left(1+\tilde{\theta}_{At}(\eta_H)\right)\left(1+\tilde{\theta}_{Bt}(\eta_H)\right)} < 0.\end{aligned} \tag{112}$$

The implied UK inclusive value is

$$\begin{aligned}\widehat{IV}_t(\mathrm{UK};\eta_H) = \log\tilde{B}_t &+ \hat{\lambda}_{A,t}(\eta_H)\ell_{At}(\eta_H) + \hat{\lambda}_{B,t}(\eta_H)\ell_{Bt}(\eta_H) \\ &+ \hat{\lambda}_{A,t}(\eta_H)\hat{\lambda}_{B,t}(\eta_H)\Delta_{AB,t}(\eta_H).\end{aligned} \tag{113}$$

The last term accounts for the fact that the student chooses only one job when both arrive. Consequently, the inclusive value gain from having both jobs available is smaller than the sum of the gains from having each job available separately.

As in Proposition 4, take any reference master's destination $j_0 \neq \text{UK}$ and define the observed first-stage response net of the known tuition and reference-country changes:

$$\begin{aligned}\mathcal{D}(t_1, t_2) := \Delta_{t_1}^{t_2}\left[\log \frac{S_{1t}(\text{UK})}{S_{1t}(j_0)}\right] \\ - \Delta_{t_1}^{t_2}\left[\left(\beta_1 + \beta_2 W_t(o)\right)\left(T_{\text{UK},t} - T_{j_0,t}\right)\right] \\ + \kappa_{j_0}\Delta_{t_1}^{t_2}\left[IV_t(j_0)\right],\end{aligned} \tag{114}$$

where $T_{j,t} := P_{j,t}\chi_{j,t}L_j$. Every quantity on the RHS is observed or already identified. The remaining equation for η_H is

$$\begin{aligned}H(\eta_H) := \kappa_{\text{UK}}\left[\widehat{IV}_{t_2}(\text{UK}; \eta_H) - \widehat{IV}_{t_1}(\text{UK}; \eta_H)\right] \\ - \mathcal{D}(t_1, t_2) = 0.\end{aligned} \tag{115}$$

Proposition 8 (Identification of the two stochastic UK choices)**.** *Maintain the model specification and Assumptions 3-8, and suppose the parameters taken as known at the beginning of this subsubsection have been identified. Let η_H^0 denote the true common non-wage utility parameter. Then:*

1. *$\eta_H^0 \in \mathcal{E}$ and $H(\eta_H^0) = 0$.*
2. *If $H'(\eta_H^0) \neq 0$, then η_H^0 and the four job finding probabilities for this m are locally identified.*
3. *If H has a unique zero on $\mathcal{E}$, then η_H^0 is globally identified. The job finding probabilities for every observed type m and the policy multipliers (γ_A, γ_B) are then recovered by the second-stage inversions.*

Proof. The proof proceeds in four steps.

Step 1: recovery at the true parameter. At η_H^0, the true job finding probabilities satisfy (106) in each cohort. They are interior by Assumption 7, so Lemma 5 implies that they are exactly the probabilities recovered by the inversion. Assumption 6 gives $\lambda_q^{post} < \lambda_q^{pre}$ for $q \in \{A, B\}$, so $\eta_H^0 \in \mathcal{E}$.

Step 2: the inclusive value response. Averaging the conditional log sums of expo-

nential utilities over the four choice sets gives, with the candidate arguments suppressed,

$$\begin{aligned}\widehat{IV}_t(\text{UK}) = &(1-\hat{\lambda}_{A,t})(1-\hat{\lambda}_{B,t})\log\tilde{B}_t \\ &+\hat{\lambda}_{A,t}(1-\hat{\lambda}_{B,t})\log(\tilde{B}_t+\tilde{a}_{At}) \\ &+(1-\hat{\lambda}_{A,t})\hat{\lambda}_{B,t}\log(\tilde{B}_t+\tilde{a}_{Bt}) \\ &+\hat{\lambda}_{A,t}\hat{\lambda}_{B,t}\log(\tilde{B}_t+\tilde{a}_{At}+\tilde{a}_{Bt}).\end{aligned} \tag{116}$$

Factoring out $\tilde{B}_t$ inside each logarithm and collecting terms gives (113). At the true parameter, this is the true inclusive value. Meanwhile, taking the difference of the first-stage log odds between t_1 and t_2 cancels the time-invariant education consumption values. Rearranging gives

$$\mathcal{D}(t_1,t_2) = \kappa_{\text{UK}}\left[IV_{t_2}(\text{UK}) - IV_{t_1}(\text{UK})\right]. \tag{117}$$

Combining these results gives $H(\eta_H^0)=0$.

Step 3: local identification. Let F_t denote the vector of the two expressions on the RHS of (106). Its Jacobian with respect to the job-finding probabilities is

$$J_t := \frac{\partial F_t}{\partial(\lambda_{A,t},\lambda_{B,t})} = \begin{pmatrix} p_{At}-\lambda_{B,t}c_{At} & -\lambda_{A,t}c_{At} \\ -\lambda_{B,t}c_{Bt} & p_{Bt}-\lambda_{A,t}c_{Bt}\end{pmatrix}. \tag{118}$$

For $(\lambda_{A,t},\lambda_{B,t})\in[0,1]^2$, its determinant satisfies

$$\begin{aligned}\det J_t &= p_{At}p_{Bt} - p_{At}\lambda_{A,t}c_{Bt} - p_{Bt}\lambda_{A,t}c_{At} \\ &\geq p_{At}p_{Bt} - p_{At}c_{Bt} - p_{Bt}c_{At} \\ &= \frac{\theta_{At}\theta_{Bt}}{(1+\theta_{At})(1+\theta_{Bt})(1+\theta_{At}+\theta_{Bt})} > 0.\end{aligned} \tag{119}$$

Therefore, the implicit function theorem implies that the inverted probabilities are continuously differentiable in η_H near the true parameter. The same is true of H. If $H'(\eta_H^0)\neq 0$, then η_H^0 is the unique zero of H in a neighborhood of the true parameter. This establishes local identification of η_H^0, and Lemma 5 then establishes local identification of the four job finding probabilities.

Step 4: global identification and recovery. Every parameter vector consistent with the model and the observed data must generate a candidate in $\mathcal{E}$ satisfying $H(\eta_H)=0$. If this zero is unique, it must be η_H^0. The non-wage utilities are then recovered as

$$\xi_t(\text{UK},A) = \xi_t(\text{UK},B) = \eta_H^0 + \omega_{r(o)}. \tag{120}$$

Applying Lemma 5 in a pre-reform and a post-reform cell for each m recovers both job finding probabilities in each regime. Finally, Assumption 6 gives

$$\gamma_q = \frac{\lambda_q^{post}(m)}{\lambda_q^{pre}(m)}, \qquad q \in \{A, B\}. \tag{121}$$

In the model without the unobserved ability type, $\lambda_q^{pre}(m) = \Lambda(\varrho_m^q)$, so the baseline job arrival parameters are also recovered:

$$\varrho_m^q = \log \frac{\lambda_q^{pre}(m)}{1 - \lambda_q^{pre}(m)}, \qquad q \in \{A, B\}. \tag{122}$$

□

The nonzero derivative in Proposition 8 is a sufficient condition for local identification. A zero derivative does not by itself establish non-identification, since a higher-order zero can still be isolated. Similarly, global identification requires uniqueness over the admissible set; the second-stage inversion alone does not establish this uniqueness. Neither result requires the constant-odds property in Lemma 2 or the parallel wage trend in Assumption 9.

Step 5: additional identifying variation. The construction above uses only one (m, o) cell and one pre- and post-reform cohort pair. The model provides additional restrictions that may distinguish candidates satisfying this single response equation.

1. *Common job finding probabilities within a policy regime.* For fixed m, job finding probabilities do not depend on o or on t within the same regime. Restoring these arguments, the inversions must therefore satisfy

$$\hat{\lambda}_q(\eta_H; m, o, t) = \hat{\lambda}_q(\eta_H; m, o', t'), \qquad q \in \{A, B\}, \tag{123}$$

whenever t and t' belong to the same regime. The odds may nevertheless differ through wages, $\omega_{r(o)}$, and B_t. This restriction locally identifies η_H if, for some such pair and some q,

$$\frac{\partial}{\partial \eta_H} \left[\hat{\lambda}_q(\eta_H; m, o, t) - \hat{\lambda}_q(\eta_H; m, o', t') \right] \Big|_{\eta_H = \eta_H^0} \neq 0. \tag{124}$$

This source of identification relies on the restriction that job arrival is common across these cells, together with the logit specification.

2. *Common proportional policy effects across worker types.* By Assumption 6, the policy multipliers do not depend on m. Hence, after imposing the within-regime restrictions,

for any $m \neq m'$,

$$\frac{\hat{\lambda}_q^{post}(\eta_H; m)}{\hat{\lambda}_q^{pre}(\eta_H; m)} = \frac{\hat{\lambda}_q^{post}(\eta_H; m')}{\hat{\lambda}_q^{pre}(\eta_H; m')}, \qquad q \in \{A, B\}. \tag{125}$$

These restrictions use only the second-stage shares and the parameters already identified.

3. *Multiple cells and cohort comparisons.* Equation (115) holds for every observed (m, o) and every pre- and post-reform cohort pair excluding t^*. Since η_H is common across these comparisons, all the resulting response equations must hold at the same candidate value.

Thus, multiple zeros of one response equation do not by themselves imply non-identification in the full model. If the response equations and the additional restrictions jointly admit only one candidate η_H with admissible probabilities in every cell, the common non-wage utility and all job finding probabilities are globally identified. If more than one candidate survives these restrictions, further identifying information is needed before claiming point identification from this construction.

I.3.11 Proof of Proposition 5

Proof. Through all the identified parameters and the first-stage choice probability of cohort t^*, we can recover $IV_{t^*}(\text{UK}|\text{m, g, o})$. Besides, we can also calculate $IV_{t^*}^{pre}(\text{UK}|\text{m, g, o})$ and $IV_{t^*}^{post}(\text{UK}|\text{m, g, o})$. Note that

$$IV_{t^*}(\text{UK}|\text{m, g, o}) = \tau IV_{t^*}^{pre}(\text{UK}|\text{m, g, o}) + (1-\tau) IV_{t^*}^{post}(\text{UK}|\text{m, g, o}), \tag{126}$$

so τ is identified as the solution to this equation as long as $IV_{t^*}^{pre}(\text{UK}|\text{m, g, o}) \neq IV_{t^*}^{post}(\text{UK}|\text{m, g, o})$. □

J Estimation Detail Discussion

Two features of the estimator deserve discussion. First, we estimate the wage and choice parameters jointly rather than estimating the wage equation in a preliminary step and treating its parameters as fixed in the choice likelihood. Unobserved type enters both components: the premium μ_z shifts observed wages and affects choice probabilities through expected wages, while each individual's posterior type probability depends jointly on her choices and

observed wage. The likelihood is separable across the two blocks only conditional on unobserved z, so we need to integrate over unobserved types to connect the wage and choice components of the likelihood.

Second, we do not estimate the persistence parameter ρ as an unrestricted parameter in the likelihood maximization. It enters choice probabilities only through stage-1 wage beliefs, while the cell-by-year wage deviations are otherwise unrestricted. We instead impose consistency between the persistence underlying wage beliefs and the dynamics of the estimated wage deviations. Specifically, we maximize ℓ holding $\hat{\rho}^{(x)}$ fixed, fit an AR(1) process to the estimated cell-by-year deviations using equation (18), update $\hat{\rho}^{(x+1)}$ to the fitted persistence coefficient, and recompute wage beliefs using the updated $\hat{\rho}^{(x+1)}$. We iterate until the fitted coefficient equals the value implied by the choice likelihood. The resulting fixed point is $\hat{\rho} = 0.466$.

K Sample Selection in the Model

1. Undergraduate Major Based on the undergraduate majors reported on LinkedIn, we exclude students in five specialized professional and applied fields: pharmacy, medicine, law, architecture, and agriculture. Among the 2008–2013 cohorts, this restriction removes 11,864 individuals (6.7%), leaving 165,819 individuals. These fields may involve distinct professional training and employment pathways, including access to occupation-specific immigration visas, which may reduce graduates' negative exposure to the PSW abolition reform.[71]

2. Master's Destination We first restrict the estimation sample to international master's students whose destination is the UK, the US, Canada, ANZ (Australia and New Zealand), or Europe. This restriction reduces the number of destination-specific parameters to be estimated in stage 1 and preserves statistical power while retaining the major master's destinations. It excludes 10,578 individuals (6.4% of the initial sample), reducing the sample from 165,819 to 155,241. Singapore is the largest excluded destination, accounting for approximately 28% of the exclusions.

3. Indian States with too few observations We further exclude students who obtained their bachelor's degree in an Indian state or union territory with fewer than 100 observations, as these samples are too small to support estimation of the state-level component of

[71] For example, eligible overseas doctors could undertake temporary training in the UK through the Medical Training Initiative under a Tier 5 (Government Authorized Exchange) visa. See the House of Lords debate, "NHS: Medical Training Initiative,", March 11, 2010

the model: Meghalaya (39), Tripura (9), Arunachal Pradesh (8), Manipur (8), Dadra and Nagar Haveli and Daman and Diu (6), Nagaland (4), Andaman and Nicobar Islands (3), Mizoram (2). This restriction excludes 79 individuals (0.05% of the remaining 155,241), leaving 155,162 individuals from 24 states and union territories.

4. Missing post-graduation outcomes We retain the 15,536 individuals with missing post-graduation outcomes and the 1,084 individuals who report a job but no location, classifying both groups as returnees to India. This classification rests on the assumption that these individuals most likely returned to India and stopped updating their LinkedIn profiles, given LinkedIn's lower coverage in India than in the destination countries. In the estimation, these individuals contribute only to the choice likelihood, not to the wage likelihood.

L Full Estimates of the Model

Table A2: Parameter Estimates (Full)

Estimate (SE)		Estimate (SE)		Estimate (SE)	
Panel A: Wage return and migration-stay tastes					
$\hat{\beta}^{w}$	5.486 (0.092)	ξ^{stay}[europe, a_3=2]	-10.009 (0.182)	ξ^{stay}[europe, a_3=4]	-11.397 (0.183)
ξ^{stay}[US, a_3=1]	-10.536 (0.204)	ξ^{stay}[US, a_3=3]	-12.206 (0.313)	ξ^{stay}_{Xh}	-0.375 (0.059)
ξ^{stay}[Canada, a_3=1]	-7.881 (0.160)	ξ^{stay}[Canada, a_3=3]	-9.739 (0.296)	ξ^{stay}_{region}[Western]	0.186 (0.016)
ξ^{stay}[ANZ, a_3=1]	-9.353 (0.161)	ξ^{stay}[ANZ, a_3=3]	-11.371 (0.299)	ξ^{stay}_{region}[Northern]	-0.031 (0.021)
ξ^{stay}[europe, a_3=1]	-9.948 (0.177)	ξ^{stay}[europe, a_3=3]	-11.752 (0.302)	ξ^{stay}_{region}[Central]	-0.209 (0.030)
ξ^{stay}[US, a_3=2]	-10.639 (0.205)	ξ^{stay}[US, a_3=4]	-12.047 (0.202)	ξ^{stay}_{region}[Eastern]	-0.572 (0.034)
ξ^{stay}[Canada, a_3=2]	-8.040 (0.178)	ξ^{stay}[Canada, a_3=4]	-8.444 (0.163)	$\delta^{stay,post}$[US]	0.182 (0.021)
ξ^{stay}[ANZ, a_3=2]	-9.686 (0.169)	ξ^{stay}[ANZ, a_3=4]	-9.803 (0.159)	$\delta^{stay,post}$[ANZ]	0.375 (0.036)
Panel B: UK stay amenity					
U^{UK}_{H}	2.214 (0.471)	α^{UK}[Western]	-0.071 (0.063)	α^{UK}[Eastern]	-0.149 (0.130)
U^{UK}_{L}	-9.217 (0.152)	α^{UK}[Northern]	0.001 (0.072)		
δ^{L}_{post}	-1.360 (0.147)	α^{UK}[Central]	-0.219 (0.117)		
Panel C: Cross-country work, PhD and second master's					
ξ^{cross}[UK]	-13.810 (0.175)	ξ_{PhD}[Canada, STEM]	49.917 (0.861)	$\xi_{\text{PhD},Xh}$	0.685 (0.068)
ξ^{cross}[US]	-15.465 (0.213)	ξ_{PhD}[ANZ, STEM]	49.079 (0.859)	ξ_{2ndMA}[UK]	47.783 (0.857)
ξ^{cross}[Canada]	-13.143 (0.164)	ξ_{PhD}[europe, STEM]	50.638 (0.858)	ξ_{2ndMA}[US]	48.921 (0.857)
ξ^{cross}[ANZ]	-14.350 (0.173)	ξ_{PhD}[UK, NSTEM]	47.487 (0.856)	ξ_{2ndMA}[Canada]	49.526 (0.859)
ξ^{cross}[europe]	-14.308 (0.185)	ξ_{PhD}[US, NSTEM]	48.640 (0.857)	ξ_{2ndMA}[ANZ]	49.518 (0.857)
ξ^{cross}[other]	-10.994 (0.152)	ξ_{PhD}[Canada, NSTEM]	47.930 (0.885)	ξ_{2ndMA}[europe]	48.831 (0.858)
ξ_{PhD}[UK, STEM]	48.959 (0.858)	ξ_{PhD}[ANZ, NSTEM]	47.548 (0.862)		
ξ_{PhD}[US, STEM]	50.524 (0.858)	ξ_{PhD}[europe, NSTEM]	48.374 (0.861)		
Panel D: Stage-1 destination preferences and cost					
φ[US, a_3=1]	3.457 (0.083)	φ[Canada, a_3=3]	-0.034 (0.152)	φ_{Xh}[ANZ]	-0.289 (0.097)
φ[Canada, a_3=1]	-0.081 (0.082)	φ[ANZ, a_3=3]	1.637 (0.151)	φ_{Xh}[europe_free]	0.057 (0.107)
φ[ANZ, a_3=1]	1.971 (0.084)	φ[europe_free, a_3=3]	1.965 (0.150)	φ_{Xh}[europe_charge]	0.055 (0.112)
φ[europe_free, a_3=1]	1.505 (0.090)	φ[europe_charge, a_3=3]	1.209 (0.148)	φ^{UK}_{0}[Western]	-0.277 (0.023)
φ[europe_charge, a_3=1]	0.615 (0.083)	φ[US, a_3=4]	1.573 (0.063)	φ^{UK}_{0}[Northern]	-0.346 (0.027)
φ[US, a_3=2]	3.459 (0.092)	φ[Canada, a_3=4]	-1.506 (0.069)	φ^{UK}_{0}[Central]	-0.121 (0.038)
φ[Canada, a_3=2]	-0.439 (0.107)	φ[ANZ, a_3=4]	0.979 (0.064)	φ^{UK}_{0}[Eastern]	-0.434 (0.043)
φ[ANZ, a_3=2]	1.732 (0.097)	φ[europe_free, a_3=4]	-0.246 (0.078)	β_1	-0.801 (0.022)
φ[europe_free, a_3=2]	1.777 (0.101)	φ[europe_charge, a_3=4]	-0.399 (0.066)	β_2	0.424 (0.031)
φ[europe_charge, a_3=2]	0.665 (0.099)	φ_{Xh}[US]	-0.328 (0.092)		
φ[US, a_3=3]	3.788 (0.146)	φ_{Xh}[Canada]	-0.407 (0.116)		
Panel E: UK job-access gates and transition					
$\rho^{Adv.}$[a_3=1]	-4.936 (0.340)	$\rho^{Bach.}$[a_3=1]	-1.295 (0.021)	$\rho^{Adv.}_{\gamma}$	1.382 (0.356)
$\rho^{Adv.}$[a_3=2]	-4.464 (0.353)	$\rho^{Bach.}$[a_3=2]	-1.488 (0.042)	$\rho^{Bach.}_{\gamma}$	-0.152 (0.058)
$\rho^{Adv.}$[a_3=3]	-7.539 (0.523)	$\rho^{Bach.}$[a_3=3]	-1.519 (0.109)	ρ_{τ}	1.108 (0.093)
$\rho^{Adv.}$[a_3=4]	-6.817 (0.485)	$\rho^{Bach.}$[a_3=4]	-1.965 (0.036)		
$\rho^{Adv.}_{Xh}$	-0.068 (0.116)	$\rho^{Bach.}_{Xh}$	0.198 (0.063)		
Panel F: Unobserved type					
$\eta^{Adv.}_{Z}$	5.251 (0.518)	φ^{UK}_{Z}	-0.628 (0.527)	γ^{π_Z}[a_3=4]	3.537 (0.595)
$\eta^{Bach.}_{Z}$	-0.298 (0.462)	$\alpha^{\pi_Z}_{0}$	-5.830 (0.567)	$\gamma^{\pi_Z}_{Xh}$	0.992 (0.141)
ξ^{stay}_{Z}	2.182 (0.286)	γ^{π_Z}[a_3=2]	-1.388 (1.657)		
μ^{abroad}_{Z}	0.159 (0.011)	γ^{π_Z}[a_3=3]	5.969 (0.578)		
Panel G: Wage equation (auxiliary)					
δ^{w}[a_3=2]	0.010 (0.004)	δ^{w}[a_3=4]	-0.118 (0.004)	δ^{w}_{MA}	-0.064 (0.008)
δ^{w}[a_3=3]	-0.038 (0.005)	δ^{w}_{Xh}	-0.001 (0.008)	$\log \sigma_w$	-0.597 (0.002)

Notes: Joint MLE (ρ=0.466). Standard errors (in parentheses) are square roots of the diagonal of H^{-1} from the joint observed information in the 197-dimensional additive parameter space. The 84 wage cell–year fixed effects $\alpha_{jw,q,t}$ complete the space and are reported in the companion table (param_est_wageFE). Normalizations: a_3=1, Southern region, India work, and the Low type carry zero shifts; Xh denotes the Honours additive shifter.

M Counterfactual: Signals to Reduce Job Prospect Uncertainty

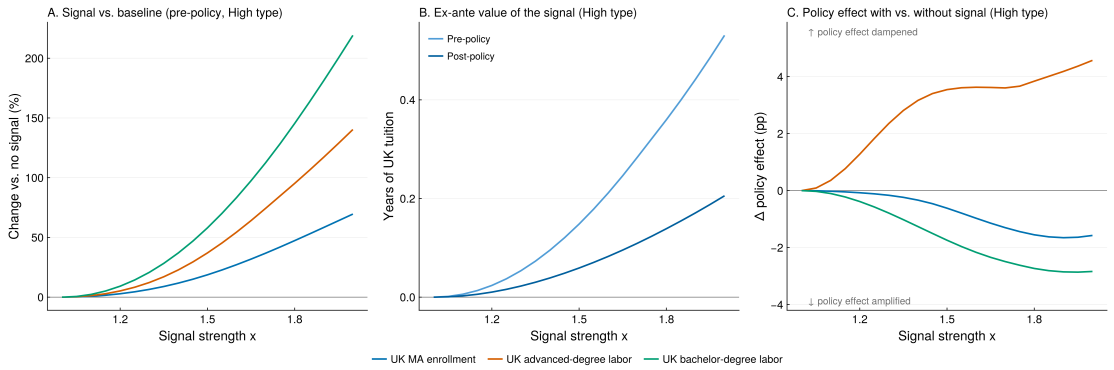


Figure A20: Effects of Employment Prospect Signal

Notes: This figure shows responses for High-type students only. Before enrolling, each student observes a binary signal with probability a half for each about their post-graduation UK job gates: $\lambda^g_{\text{good}} = \min(x\lambda^g, 1)$ and $\lambda^g_{\text{bad}} = 2\lambda^g - \lambda^g_{\text{good}}$, $g \in \{Adv., Bach.\}$. Students act on the signal at both the enrollment and work stages. Outcomes pool the 2011–2013 cohorts with prior type weights. Panel A shows percentage change in UK master's enrollment and UK advanced- and bachelor-degree-level employment relative to no signal. Panel B shows the expected gain in the enrollment-stage inclusive value, converted at the marginal disutility of tuition, in years of UK master's tuition. Panel C shows the policy effect (percentage change from the pre- to the post-policy regime) with the signal minus without it, in percentage points; since the effect is a decline, positive values mean the signal dampens it.

As discussed in the decomposition above, uncertainty about employment prospects plays an important role in students' education decisions. Policymakers seeking to attract high-quality students could therefore provide high-ability students with more information about their chances of finding high-skill employment and remaining in the UK after completing a master's degree. In practice, this could involve greater transparency about post-graduation outcomes, including employment statistics and profiles of graduates who remain in the UK. Alternatively, a points-based visa system could make eligibility depend on pre-master's characteristics, thus providing more certainty about UK job prospects during the education choice stage.[72] Figure A20 illustrates the value of this information and the resulting reduction in uncertainty.

Specifically, we allow for additional heterogeneity within the unobserved high type: some students have higher probabilities of receiving high-skill offers than others, reflected in different values of λ. In the baseline model, students know their type but are unaware of this

[72]For example, Hong Kong's young talent system states: "Applicants who have obtained a degree at bachelor level or higher from an institution listed among the top 100 universities/institutions in any of the above four tables, or in the top 30 colleges in the National Liberal Arts Colleges Rankings published by the US News and World Report will be awarded an additional 30 points."

additional heterogeneity. In the counterfactual, the UK government provides high-type students with a signal about their employment prospects before they decide where to pursue a master's degree. Half receive a positive signal indicating better prospects for both bachelor-level and advanced-level employment in the UK, while the other half receive a negative signal indicating worse prospects. The signals leave aggregate employment probabilities unchanged, satisfying signal neutrality. Signal strength determines the magnitude of students' positive or negative posterior updates. We make comparisons on two metrics: outcomes with and without signals under the pre-policy baseline, and the PSW abolition policy effects with and without signals. The outcomes are master's enrollment, bachelor-level labor supply, and advanced-level labor supply.

Figure A20 shows that providing signals increases master's enrollment and skilled labor supply in the UK (left panel): the "pull-in" effect among students receiving positive signals dominates the "pull-out" effect among those receiving negative signals. These results suggest that providing more information can help the government attract talent. Information also substantially increases students' welfare (middle panel). Its implications for the effects of the visa reform depend on two opposing channels (right panel). First, students receiving a positive signal may still pursue a UK master's degree because they expect to obtain an advanced-level job despite reduced access to bachelor-level jobs, dampening the policy effect. Second, because a positive signal raises the perceived probability of obtaining a bachelor-level job, restricting access to these jobs can cause a larger decline in the expected value of a UK master's degree, amplifying the policy effect. The first channel dominates for advanced-level labor supply, so information dampens the policy-induced decline. By contrast, the second channel dominates for master's enrollment and bachelor-level labor supply, amplifying their declines. Students receiving positive signals place greater value on the bachelor-level employment opportunities lost under the reform. Consequently, some shift away from UK master's education even though their advanced-level employment prospects remain favorable and unchanged by the reform.