The IT Boom and Other Unintended Consequences of Chasing the American Dream*

Gaurav Khanna†
Nicolas Morales‡

Abstract

We study how US immigration policy and the Internet boom affected not just the US, but also led to a tech boom in India. Students and workers in India acquired computer science skills to join the rapidly growing US IT industry. As the number of US visas was capped, many remained in India, enabling the growth of an Indian IT sector that eventually surpassed the US in IT exports. We leverage variation in immigration quotas and US demand across occupations to show that India experienced a 'brain gain' when the probability of migrating to the US was higher. Changes in the US H-1B cap induced changes in fields of study, and occupation choice in India. We then build and estimate a quantitative model incorporating trade, innovation, and dynamic occupation choice in both countries. We find that high-skill migration raised the average welfare of workers in each country, but had distributional consequences. The H-1B program induced Indians to switch to computer science occupations, and helped drive the shift in IT production from the US to India. We show that accounting for endogenous skill acquisition is key for quantifying the gains from migration.

JEL: F16, F22, J61

Keywords: High-skill migration, H-1Bs, India, computer scientists, IT sector, brain gain

*We thank Ashish Arora, Dominick Bartelme, John Bound, Michael Clemens, Mark Colas, Gordon Hanson, Andrei Levchenko, Anna Maria Mayda, David McKenzie, Justin Sandefur, Jagadeesh Sivadasan, Sebastian Sotelo and seminar participants at the NBER Summer Institute, Michigan, Purdue, Montevideo, Riverside, Raleigh (Society of Labor Economists), Western Ontario, Claremont Mckenna, Hawaii, Mt Holyoke, Center for Global Development, Clermont-Ferrand, Bonn (IZA), Yale, Delhi (ISI), Upjohn Institute, Bank of Italy, and the ASSA (Atlanta) for comments. We are grateful to the Alfred Sloan Foundation and the NBER Fellowship on High-Skill Immigration for generous research support. The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Board of Governors.

†Assistant Professor, University of California – San Diego, gakhanna@ucsd.edu, econgaurav.com
‡Federal Reserve Bank of Richmond, nicolas.morales@rich.frb.org, sites.google.com/view/nicolasmorales

1 Introduction

Migration policy, and the skills of migrants, have been at the forefront of elections, policy debates, and academic discourse throughout the world. The effects of high-skill migration, as exemplified by the high-profile US H-1B program, are theoretically ambiguous for both sending and receiving countries. For instance, immigration can attract global talent and lead to production growth in receiving countries, but migration-induced technological catch-up may contrarily shift production to the sending country. The sending country may experience brain drain as human capital departs, or experience brain gain as the opportunity to migrate induces skill acquisition. We resolve these ambiguities by modeling and estimating the long-term welfare consequences of high-skill migration to the US, and examine how the combination of the US IT boom and US immigration policies led to a boom halfway across the world in India.

We evaluate the role played by US immigration policy in the shift of IT production from the US to India. We start by presenting a series of facts that describe how changes in US immigration policy and shocks to US labor demand affected the occupational choices of Indian college graduates. We build and estimate quantitative model that incorporates trade, and dynamic occupational choices of forward-looking workers and students in both countries. In counterfactual exercises that restrict the migration of Indian workers, our results indicate that the H-1B program and the tech boom had a powerful impact on IT sectors in both countries. US high-skill immigration policy facilitated the spread of the US-led boom to India, and by the mid-2000s India became the major exporter of software. Despite various distributional effects, we find that, on average, workers in each country are better off due to immigration.

As a first step, in Section 2, we use descriptive trends and background information to describe our narrative and ground our model. Innovation rapidly caused an expansion in the US IT sector in the early 1990s (Bound et al., 2015; Kerr, 2013a), and almost a decade later the IT sector in India quickly grew from 1.2% of GDP in 1998 to 7.5% in 2012 (NASSCOM, 2012). We argue that Indian workers and students responded to these booms and migration opportunities by accumulating computer science skills valuable both at home and abroad. While a fraction of these workers entered the US labor market via the restricted supply of H-1B visas, many joined the rapidly growing IT sector in India, helping production shift from the US to the lower-wage destination of India. This educated workforce in India enabled the Indian IT sector to grow rapidly, with new firms joining the race and older firms expanding. Over time, India became a major producer of software, eroding the US dominance in IT exports.

To reinforce our quantitative model, we show well-identified evidence supporting the underlying driving force behind our hypothesis: labor demand shocks in the US affected human capital accumulation in India. In Section 3, we combine data from both countries, and estimate specifications with major and occupational choice in India as the outcome, and probability of migrating to the US as the explanatory variable. Since the probability of migration to the US is likely correlated with unobserved features of the Indian labor market, we derive variation in the probability of migrating from changes to the H-1B visa cap, and occupation-specific migration propensities. Changes to the cap were plausibly independent from unobserved shocks to supply-side worker preferences and occupation-specific demand shocks from local industry in India. We re-weight the immigration policy shocks with the baseline propensities of Indian migrants to work in certain occupations. Our instrument is strongly predictive of the number of recent high-skilled migrants from India working in the US.

We estimate a positive, quantitatively significant response to US demand shocks in both the choice of major and the choice of occupation in India. This represents a 'brain gain' driven by immigration prospects. A 1% increase in the emigration probability of a specific occupation, results in a 0.74% increase in employment in India. This effect is larger for younger workers. We use our estimated labor-response elasticity as a targeted moment when estimating our general equilibrium (GE) model.

We examine threats to identification, and show robustness to additional specifications, controls, falsification tests, and tests of pre-trends. For instance, our results are similar when leveraging alternative instruments that rely on the differential occupation-specific demand from the US over time, captured by the annual flows of migrants from *other* countries. We obtain data from other destination countries and highlight the lack of occupation-specific correlated demand for Indian migrants. We further leverage variation across regions in India, and show these responses are larger in areas that had a higher propensity to produce certain occupations. A variety of different sources of variation and specifications consistently point to our posited 'brain gain' hypothesis.

Motivated by these facts, the primary substance of our paper is the GE model outlined in Section 4, which contains three crucial features. First, we model dynamic labor supply decisions and how they respond to changes in immigration policy. Students in both countries have heterogeneous preferences, and make forward-looking decisions on choosing their college major given their expected future earnings in different occupations. After graduation, workers (also with heterogeneous preferences), choose every year to either continue working in their current occupation or switch occupations (paying a cost), given the labor demand shocks and their expected future benefits in each occupation.

Indian computer scientists (henceforth, CS) incorporate the possibility of migration when making their occupational and major decisions. As such, workers choose their occupation considering the *expected* value of working as CS, observing both the wages in India and the US. While choosing occupations, workers have uncertainty over the migration cap, as well as whether they will be selected to migrate to the US. Such uncertainty opens up the possibility of brain gain in India driven by the prospect of migration.

Second, we model how firms hire both local and foreign workers, across different occupation-skill groups. As migration increases the size of the US CS workforce, firms demand more workers in complementary occupations, such as managerial positions. At the same time, skill-biased technical change shifts labor demand in favor of high-skill occupations, while computer scientists, both domestic and foreign, are innovators and increase the overall IT productivity in both countries (Kerr and Lincoln, 2010). With return migration some Indian CS working in the US return to India with specialized skills that are complementary to local CS skills in India.

Our final feature captures trade between India, the US and the rest of the world. The IT sector produces a continuum of varieties, the productivities of which differ across countries. The products are traded between countries as in the canonical Eaton and Kortum (2002) framework. Restricting immigration, or rapid growth in the Indian IT sector can shift production across borders. Since the final output uses software as an intermediate input in production, an expansion in IT raises productivity in the rest of the economy. Consumers benefit from lower prices, while both countries compete for global markets. The potential to trade helps drive India's growth, as workers switch to the high-innovation IT sector. At the same time, in both countries, immigration-wage impacts are muted by trade as resources shift across sectors (Ventura, 1997), and tradable sectors can absorb labor supply shocks (Burstein et al., 2020).

Our model generates countervailing forces, making the theoretical impacts of the H-1B program ambiguous. For instance, the effects of brain drain from India, compete with brain gain as more Indians acquire skills valued in the US and return migrants also improve local production. Similarly, the impact on the US IT sector is theoretically unknown: on the one hand, an influx of computer scientists helps the US IT sector grow, but on the other hand, the H-1B program spurs growth in the competing Indian IT sector, eroding the US's market share.

To discipline the model and resolve these theoretical ambiguities, we use a combination of estimation and calibration techniques. The parameters in our major and occupational choice model jointly determine a key elasticity for our results: the dynamic labor supply elasticity. This plays a crucial role in the occupational response of Indian and US college graduates when immigration policy in the US changes. We estimate this elasticity by minimizing the distance between a series of targeted moments and their empirical counterparts using data from 1995-2010. One important targeted moment is the occupational response in India to labor demand shocks in the US, causally estimated in Section 3. Given the dynamic nature of labor supply decisions, even though the short-run labor supply curve is relatively inelastic, the long-run labor supply curve is fairly elastic (as students choose different majors over time), implying that reduced-form estimates of contemporaneous responses to immigration changes in some other work may not pick up the entirety of the labor supply adjustments.

Intuitively, the production side helps us determine the exogenous innovation shocks that shift the labor demand curve for computer scientists, allowing us to trace out the dynamic labor supply curve. We rely on methods from the trade literature to estimate trade costs and technology parameters. We use instrumental variables, implicitly leveraging variation in US immigration policy and migrant-supply shocks, to measure the innovation response (captured by patenting) to industry-level CS flows, and find results consistent with previous work (Kerr and Lincoln, 2010; Khanna and Lee, 2018; Peri et al., 2015).

Methodologically, our approach is similar to recent developments in the migration literature that combine reduced-form identification with structural models to make meaningful statements about welfare (Lagakos et al., 2018), and general equilibrium effects (Colas, 2019; Llull, 2018). Given the complexity of the model, it is important to do validation exercises. In Appendix G, we show that our model does a good job of matching both levels and trends in wages, employment, and IT sector output in out-of-sample tests for both countries.

In Section 6, we conduct counterfactual exercises to study the impact of a more restrictive immigration policy on both the US and Indian IT sectors by restricting H-1B migration to only half the number of Indian migrants over the 1995-2010 period. Our results indicate that US immigration policy did play a significant role in the spread of the IT boom from the US to India. The possibility of migrating to the US under the H-1B program incentivized students and workers in India to choose CS degrees and occupations, increasing the size of the non-migrant Indian CS workforce. However, the migration led native US CS workers to switch to non-CS occupations, which led to a 5.87% drop in the native US CS workforce in 2010.

An increase in the size of the Indian CS workforce due to the H-1B program also led to an increase in India's IT sector productivity. Under the H-1B program, IT production grows more in India than in the US – US IT output is 0.11% higher, and Indian IT output 35.5% higher in 2010. The production shift to India, however, hurts some US workers, most notably US-born computer scientists. World IT output increases, the US-India combined welfare is higher by 0.14%, and the average worker in each country is better off in a world with skilled migration. The net H-1B welfare gains to US workers are about $2,227 per migrant, and gains to Indian non-migrants are about $2,364 per migrant.

We highlight important mechanisms and show the quantitative relevance of our modeling decisions. We show that dynamic endogenous labor supply decisions are key for a quantitative exercise on migration. Ignoring India's occupational choice would predict welfare gains of immigration for the US that are 30% larger than our baseline and predicts India would experience welfare losses driven by brain drain. We describe the consequences of not allowing for brain drain, brain gain and return migration, to understand the contributions of each.

We explore a few extensions to our baseline model. First, in addition to the major-choice, we model the decision to go to college in the first place. Second, we explore how our conclusions change when considering immigrants as imperfect substitutes to native US CS. Third, we allow for heterogeneity in abilities and allow for migrants to be positively selected. Finally, we unpack the importance of trade and remittances, vary the cap sizes across other values, and time the migration restrictions to begin in different periods to better understand the dynamics.

Our paper is innovative in four main ways. First, we contribute to the literature that uses quantitative models to understand the general equilibrium consequences of immigration (Burstein et al., 2019; Caliendo et al., 2020; Colas, 2019; Desmet et al., 2018; di Giovanni et al., 2015; Llull, 2018; Monras, 2020; Morales, 2019). We add to this discussion by considering how migration-driven incentives to invest in human capital facilitate the corresponding growth in production for sending and receiving countries. We show that incorporating uncertainty in migration when making occupational choices substantially alters the global gains from migration.

Second, we address the debate between 'brain drain' and 'brain gain' (Abarcar and Theoharides, 2020; Agrawal et al., 2011; Beine et al., 2001; Dinkleman and Mariotti, 2016; Easterly and Nyarko, 2009; Shrestha, 2016; Stark et al., 1997). While many worry about the large number of educated Indians emigrating, we show how better paid jobs abroad incentivize students to choose certain majors and supply a skilled workforce to Indian firms. Migrants who return with newly acquired skills and technical knowhow help develop the IT sector at home, contributing to the brain gain, and facilitating a growth in exports (Bahar and Rapoport, 2018).

Third, we delve deeper into the role of migration and trade between developed and developing economies to either help or hinder structural transformation. On the one hand, North-South trade may hinder structural transformation as developing economies specialize in less productive sectors (Matsuyama, 1992) and migration may help developed economies maintain its comparative advantage by attracting global talent (Freeman, 2006b; Rybczynski, 1955). On the other hand, technological diffusion through trade and migration can help developing countries catch up with developed ones (Acemoglu et al., 2015; Davis and Weinstein, 2002; Krugman, 1979). By proposing a new mechanism through which immigration affects production, we provide novel evidence to understand major 'big push' sectoral transformations to high-skill production in emerging economies (Lagakos and Waugh, 2013; Lewis, 1954; Murphy et al., 1989).

Lastly, our paper addresses crucial issues raised by the labor literature on the impacts of high-skill immigrants on the US economy. Several papers quantify costs and benefits of high-skill immigration for the US (Borjas, 1999; Bound, Khanna, and Morales, 2016; Doran, Gelber, and Isen, 2017; Freeman, 2006b; Hunt and Gauthier-Loiselle, 2010; Kerr and Lincoln, 2010) but abstract away from the role played by other countries. In our current paper, we incorporate the occupational responses in India and the growth of the Indian tech sector, which greatly affects incomes in the US. To study the linkages across the countries and the feedback into the US, we model what happened on both sides of the world.

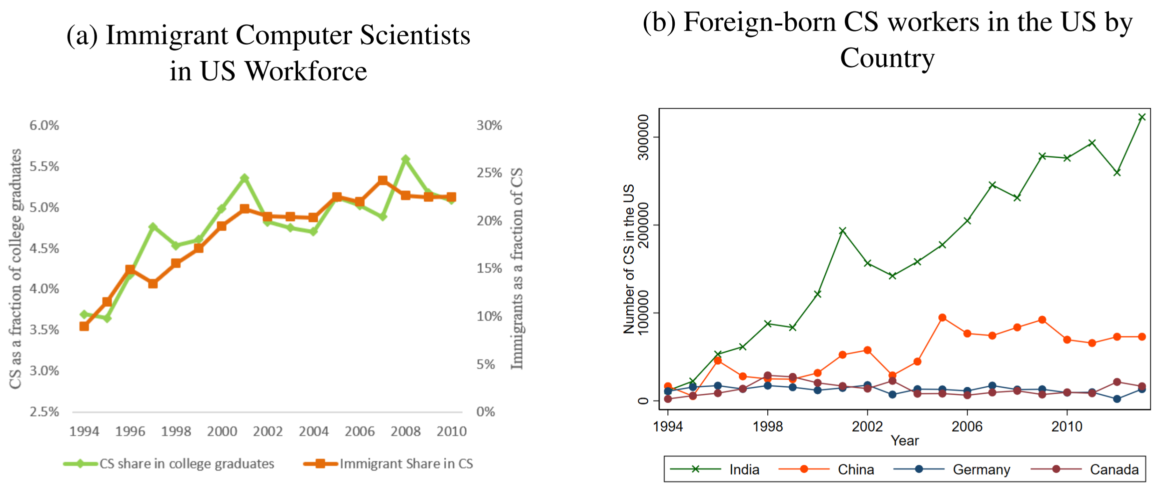
2 The Tech Boom in the US and India

2.1 The Internet Boom in the US and the H-1B Visa

Starting in the mid-1990s, innovation in the IT sector led to an increase in demand for computer scientists (CS) in the US. As Figure 1a shows, CS employment, as a share of the college-educated workforce, rose dramatically in the second half of the 1990s. By the turn of the century the CS share among college graduates increased by 43%.

CS occupations were the fastest growing occupations in the second half of the 1990s, and were

Figure 1: High-Skill Immigration and the IT Boom



Source: March Current Population Survey (CPS). Immigrants defined as foreign born who migrated after the age of 18. In the left panel, we plot the number of CS as a fraction of US college graduates (left axis) and the number of immigrant CS as a fraction of total CS (right axis). In the right panel (March CPS) we restrict the sample to foreign-born workers in CS who immigrated after the age of 18. Data details are in Appendix A.1.

expected to remain the fastest over the next decade (BLS, 1996). This growth was, in part, fueled by foreign-born workers (Figure 1a). In 1994, the share of foreign-born among CS was 8.9%. But given the dramatic growth in the second half of the 1990s, by 2010 foreigners accounted for more than 20% of the CS workforce. The Immigration Act of 1990, which established the H-1B visa program for temporary workers in specialty occupations (requiring *"theoretical and practical application of a body of highly specialized knowledge in a field of human endeavor"*), played an important role in this growth.

By the time the IT boom was starting in the mid-1990s, the 65,000 H-1B cap started binding and the allocation was filled on a first come, first served basis. According to the USINS (2000) about two-thirds of all H-1B visas were awarded to computer-related occupations in 1999 and U.S. Department of Commerce (2000) estimated that during the late 1990s, 28% of all programmer jobs in the US went to H-1B visa holders. H-1B visas, therefore, became an important source of labor for the technology sector. As shown in Figure 2a, the H-1B cap changed overtime and remains until today the main source of college educated immigrants.

2.2 The Spread of the IT Boom to India

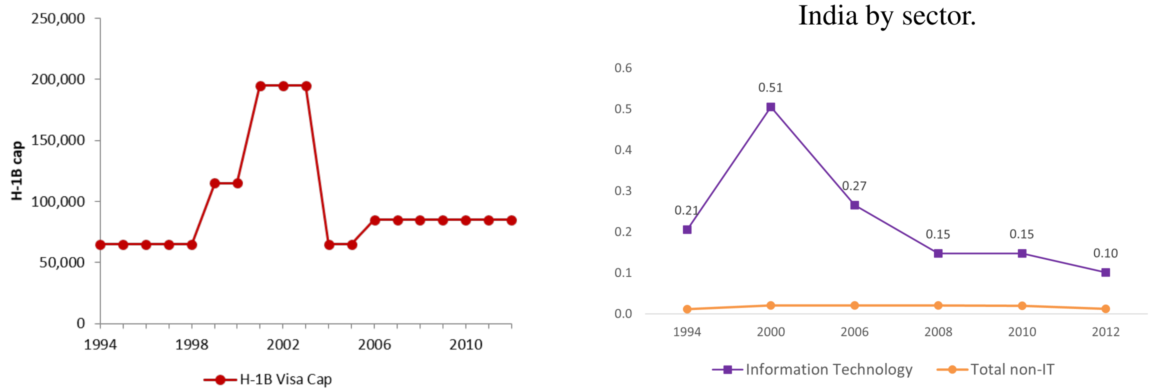
Our paper quantitatively confirms qualitative work, discussed in Appendix H, on why this spread of the tech boom was limited to certain countries like India, and how a large part of the success of Indian firms was attributable to high-skill Indian immigrants in the US. Other countries (for e.g., Israel, Ireland, and Germany) that eventually became important IT exporters, did not have trajectories that were linked to migration to the US.

India sent top engineers during the earlier hardware-boom of the 1970s and 1980s (Appendix Figure H1). Yet, as immigration was restricted, the numbers then were lower. This diaspora helped

Figure 2: Migration Prospects and Changes in the H-1B Cap

(a) Change in H-1B Cap (1994-2012)

(b) Ratio of Indian workers in US relative to India by sector.



Source: The left panel shows authors' calculations based on USCIS reports. The right panel calculates the ratio of Indian college graduates working in the US relative to the number of Indian college graduates working in India. We plot the ratio separately for the Information Technology (IT) sector and all other sectors except IT. The data comes from the American Community Survey for US and National Sample Survey for India. Data details are in Appendix A.1.

establish strong connections and a reputation for well-trained workers (Arora et al., 2001; Saxenian, 1999). Bhatnagar (2006) notes that Indian professionals in Silicon Valley "*built personal networks and valuable reputations and used their growing influence within US companies to help Indian companies get a foot in the door*" in the expanding IT sector. With the advent of the H-1B program facilitating migration, it is then in these occupations with strong connections that saw large emigration responses, as these tech workers "*migrate to better paid jobs in other countries*" (Kumar, 2006). The US had a large market for client services and software development, along with capital, technology, and industrial agglomeration.

As the cap was binding in most years, changes in the cap (Figure 2a) often mirrored the flow of computer scientists from India (Figure 1b). By 2014, more than 70% of all H-1B visas (USCIS, 2014), and 86% of all CS H-1B visas were awarded to Indians, while only 5% were awarded to Chinese nationals (Computerworld, 2015). The top H-1B firms were in IT, and many were based in India (USCIS, 2014).

Even as the US IT sector was growing rapidly, the Indian IT sector was comparatively small in the early years. At the turn of the century, for every two IT workers in India, there was one Indian-born migrant in the US IT sector (Figure 2b). Indeed, even in later years, the lack of a domestic market in India meant that 90% of software revenues came from the US (NASSCOM, 2003). The large relative demand from US IT implied that migration prospects potentially influenced college graduates in India to choose CS occupations to increase their chances of migrating. Given the large CS wage differentials between the US and India and the non-trivial probability of migrating to the US, the US IT boom raised the expected returns to working in an engineering/CS occupation in India (Figure 3a).

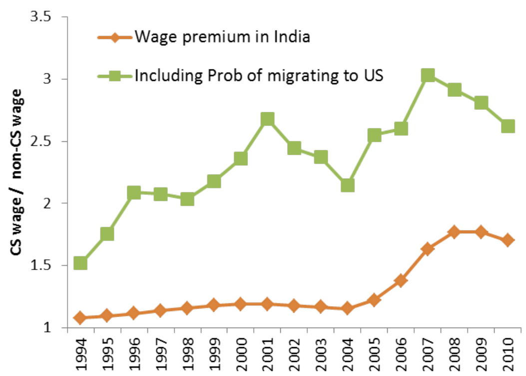
These higher (expected) returns to CS, affected the education sector in India. Bhatnagar (2006) notes since Indian programmers were "*in a domestic environment with few job opportunities, growth* (*in training and degrees*) *was also driven by larger salaries in the IT industry abroad*." To meet the rising demand, engineering schools introduced CS-oriented degrees (Figure 3b), and companies started training, building technical skills for the industry. Indian software professionals have engineering degrees, and engineering colleges increased their emphasis on IT (Arora and Athreye, 2002).

These aspiration are reflected in a survey of those who topped the school-leaving exam between 1996-2015: 40% are in the US, 56% have an Engineering degree, and 32% of workers work in IT (Indian Express, 2020). Figure 3b suggests that this enrollment response in India began soon after 1994 when the H-1B program and US tech boom started; even though the Indian IT industry was underdeveloped and local computer science premiums were low, but the expected returns to the degree were rising because of the H-1B program (Figure 3a).

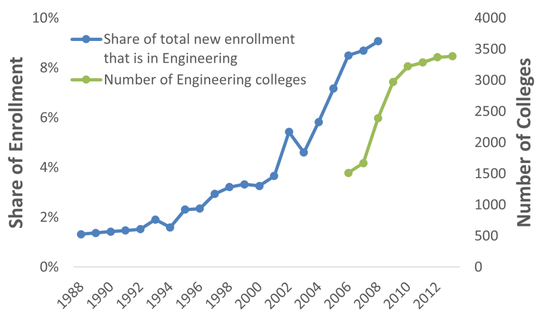
Even as Indians acquired skills that were valuable abroad, the H-1B visa was capped and many were unable to migrate to the US. As H-1B visas are awarded for 3 to 6 years and green card queues for Indians are long, many migrants return to India after their H-1Bs expire. This meant that the highly skilled immigrant workforce was not in the US in the longer term and many joined the Indian workforce with specialized CS skills. Indian firms tapped into this skilled workforce, which partly lead to an IT boom later in India (Figure 4a).

Figure 3: Returns to Skill and Growth in College Degrees in India

(a) CS Wage Premium in India with Prob of Migration



(b) Growth in Degrees and Colleges in India



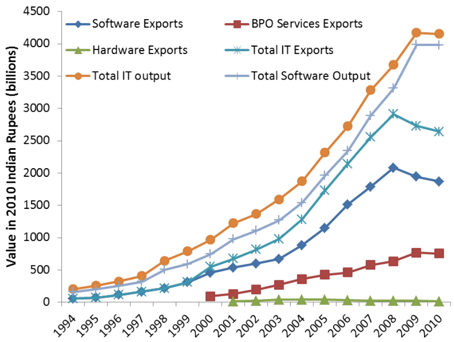
Source: March Current Population Survey (CPS), National Sample Survey (NSS), Ministry of Human Resources and Development and the All India Council for Technical Education. In the left panel we restrict the sample to college graduates, and compare the CS and non-CS wage in India (orange) to the expected wage, taking into account the probability of migrating and the wage differential in the US. Data details are in Appendix A.1.

The growth in India affected the US's dominance in IT exports, as production shifted to the other side of the world. The US was historically the largest exporter of software: by 1997, they accounted for 58% of all export revenues. In the mid-2000s, however, India overtook the US as the major exporter of IT products (Figure 4b). Most of the early growth was export-led: in 1995, software was only 2% of all exports, but by the turn of the century, it was 26%. Indeed, Indian IT firms were export oriented catering to a consumer base abroad (Figure 4a).

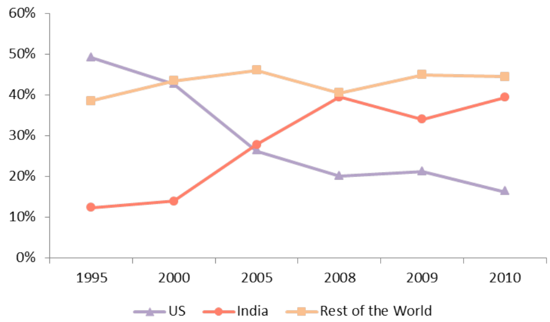
In Appendix H, we explore other details on why this boom missed other countries but settled on India. India has not only had high-quality engineering schools that train potentially lower-wage, English-speaking workers, but also developed strong networks with the US during the earlier hardware boom (Bhatnagar, 2006). Our hypothesis is that this spread of the IT boom from the US to India was partly driven by the H-1B program, and we analyze how important US immigration policy was for a structural transformation half-way across the world.

Figure 4: The IT Boom Spreads to India

(a) Growth in the Indian IT sector



(b) Share of IT Exports



Source: National Association of Software and Service Companies (NASSCOM), and OECD Trade in Value Added Statistics for industry C72: Computer and Related Activities. Data details are in Appendix A.1.

3 Educational and Occupational Choice in India

Despite these concurrent trends, one may contend that such human capital investments in India were independent of demand shocks from the US. To support the quantitative exercise in the main part of our analysis, we first establish that labor demand shocks in the US did indeed lead to education and labor supply responses in India.

3.1 Education responses to US demand.

We begin by studying whether changes in US immigration policy impacted the education choices of Indian students. The goal of this analysis, is to quantify how the choice of field of study in India responds to immigration incentives from the US. Yet, a simple correlation between migrating to the US and education decisions in India may be confounded by various factors. For instance, local demand for certain skills in India may be correlated with migrant flows to the US, and this local demand may be driving education decisions. Or, individual preferences for certain majors in India may drive major choice. Subsequently, an increase in individuals with these skills will lead to more emigration to the US of such individuals. This would be a case of reverse causality.

To credibly examine the relationship between migration probabilities and education decisions in India, we obtain two new databases: First, we obtain college-by-degree-by-field-by-year level government reports on end-of-year examinations, for registered colleges and universities in India. We digitize these data for 437 colleges across 10 fields of study to create a new database of examinations, and complement this with state-by-degree-by- field-year-level examinations data. We use the number of students who appear for examination, rather than those who pass the exams. The former reflects enrollment better, as some students may appear for an exam but fail to get the passing score. These examinations reflect the number of people who appear at all levels/years of the degree (whether they are first-year or final-year students). We take a two-year lead of this variable, as enrollment decisions may take time to affect the number of students appearing for examinations. We show robustness to alternative lead-structures.

Our second new database is data from the universe of LinkedIn profiles for Indian migrants to the US. These profiles report school, field of study, and graduation year. We use the universe of LinkedIn profiles of users who graduated from an Indian university before 2000 and eventually migrated to the US to construct school-by-field level migration exposure. While not all migrants have a LinkedIn profile, we show in Appendix A.2 that we cover over 60% of those studying computer science and engineering. In Appendix A.2, we describe these datasets in detail and compare the LinkedIn data with nationally representative data from the American Community Survey (ACS).

We estimate the following specification using these data:

$$Ln(\mathrm{N}_{srft+2}) = \beta_1 \left(Log(\text{H-1B cap}_t) \times (\text{Mig exposure}_{srf}) \right) + \delta_{srf} + \delta_{srt} + \delta_{frt} + \varepsilon_{srft} \tag{1}$$

Here, $N_{srf,t+2}$ is the number of students (who are eligible for examinations) in school s, located in region (state) r, field of study f, and year $t+2$. The H-1B cap_t is a variable that calculates the policy-predicted H-1B stock solely from changes in the cap. As such, changes in the stock of immigrants are only driven by policy changes. The changes in the cap are likely to have bigger impacts on schools-majors with relatively more alumni in the US. These are schools-by-majors that will be known to US employers, and with the help of alumni network, facilitate the hiring of workers to the US. As a result, we may expect that changes in the probability of migrating would have a bigger impact on those schools-by-majors. We interact the changes in the cap with a measure of $\text{Mig exposure}_{srf} = \frac{\text{N grads from srf before 2000 in US}}{\text{Enrollment in srf in 2001}}$.

Despite the fact that we derive variation from the cap changes, we iteratively control for a slew of fixed effects, to account for various possible confounders. These include school-by-field fixed effects δ_{srf}, school-by-time fixed effects δ_{srt}, and field-by-region-by-time fixed effects δ_{frt}. This is a fully saturated model, that accounts for changes in local demand for certain majors over time, baseline links at the school-field level, and growth in enrollment at specific universities, among other things.

Table 1 shows that, even with a stringent fixed-effects specification, increases in the H-1B cap, lead to differential increases in enrollment in schools-by-fields that had stronger links to the US.

In Appendix tables we test for robustness. First, in Table B1 we show that using a 1-year lead (instead of a 2-year lead) in examinations produces similar patterns. In Table B2 we use an alternative source of variation: the number of H-1B migrants from India (rather than changes in the cap). Table B3 uses the number of H-1B migrants from other countries (excluding India). Both these two new shifters capture actual migration propensities under the H-1B program, and so reflect the likelihood of migrating to the US. Finally, in Table B4 we use the state-level reports on examinations, that cover

Table 1: The Effect of the H-1B cap on Enrollment in Majors

	Dependent variable: Log(Exam Eligible$_{srf,t+2}$)			
Migration Exposure × Log(H1B Stock)	0.098***	0.075**	0.096*	0.065*
	(0.034)	(0.038)	(0.049)	(0.036)
School-Field FE	Y	Y	Y	Y
Time FE	Y			
Field-Time FE		Y		
State-Time FE		Y		
Field-State-Time FE			Y	Y
School-Time FE				Y
N	9513	9509	8773	8655
R-sq	0.900	0.908	0.917	0.947

Dependent variable is the number of students eligible to appear for an examination in school s, located in state r, studying in field f, and year $t+2$. Migration exposure is defined as the number of pre-2000 graduates from school s and field f working in the US, divided by the total pre-2000 enrollment at school s and field f. The H-1B stock is the cap-predicted number of H-1B workers in the US. Data for the migration exposure are from the universe of LinkedIn users that studied in India, and work in the US. Data on examinations are from annual Indian Government reports. Data details are in Appendix A.2. Standard errors clustered at the school-field level. $*p<0.1$, $**p<0.05$, $***p<0.01$.

a longer time-frame.

All these tables show that as the probability of migrating to the US increases, there is an increase in enrollment in colleges and fields that are more strongly connected to the US. Next, we explore whether this major choice decision also mirrors changes in occupation choices in India.

3.2 Occupational responses to US demand.

As a second step, we study whether the emigration of Indians to the US in particular occupations induced Indians who had not yet migrated into choosing such occupations. In our baseline model, workers in India choose occupations based on their preferences, labor demand in India and labor demand from the US, as in equation 2:

$$Ln(\mathrm{N}_{ort}) = \gamma_1 \text{Demand India}_{ort} + \gamma_2 \text{Demand from the US}_{ort} + \gamma_3 \text{Supply \& Preferences}_{ort}\ , \qquad (2)$$

where N_{ort} is the Indian college-educated workforce in occupation o, region (state) r, and time t. Our primary independent variable of interest is Demand from the US$_{ort}$. This derives the US demand for Indian migrants to work in occupation o, from both migration policy and demand shocks in the US. Demand India$_{ort}$ is the demand for work by occupation, region and time, from within the Indian economy. Supply and Preferences$_{ort}$ capture worker preferences for certain occupations across regions that may also vary over time.

A standard immigration model that ignores endogenous occupational choice driven by immigration would predict pure brain drain in India, $\gamma_2 < 0$, such that a stronger US demand for migrant workers in occupation o would reduce the number of workers in India in o. Yet, with endogenous occupation choice, more demand in the US for occupation o can potentially increase the supply of occupation o in India, as workers choose occupations that may help them migrate in the future, $\gamma_2 > 0$. Our goal is to estimate the elasticity of local occupational choice with respect to US demand, γ_2. Yet, not adequately accounting for the other components will bias our estimates of γ_2.

Estimating γ_2 in equation 2 generates two main challenges. First, the increase in occupation-specific migration may be the result of either demand from the US or an increase in supply from India (say, driven by unrelated investments in universities or changes in preferences of Indian college graduates). As such, there may be an increase in the number of workers in occupation o and region r over time, if worker preferences or education investments (that are unrelated to demand shocks) happen to change worker composition. The second challenge is that demand shocks in the US may be correlated with demand shocks in India or other parts of the world, confounding estimates of γ_2.

To tackle these issues we take multiple steps. We propose an instrumental variables strategy that leverages plausibly exogenous changes in the H-1B cap over time, which are likely to be uncorrelated with unobserved confounders. We then discuss other concerns such as global shocks that affect occupational choice in India, and perform a series of checks with alternative instruments, specifications, and falsification tests that reinforce the validity of our approach.

First, in Appendix B.2 we describe in detail the different possible components of Demand India$_{ort}$ and Supply & Preferences$_{ort}$, and how they may affect occupation choice in India. To account for domestic demand, we construct a shift-share control using the initial industrial composition in a given region (number of workers in occupation o, region r as a share of college employment in India) interacted with national level growth by industry as in equation 3:

$$Local_{ort} = Ln\left(\sum_{ind} \frac{N^{ind}_{or,1994}}{\sum_{ind} N_{or,1994}} \times \text{Total employment in } ind \text{ and } t\right) \tag{3}$$

We use variation across Indian states, and occupation-time variation to separately identify labor demand shocks experienced in India.[1] The term $Local_{ort}$ also controls for aggregate demand for occupations across all regions. If the overall demand for occupation o' from firms based in India increases, $Local_{o'rt}$ will increase across all regions. In this way, we also control for any demand-shocks from firms in India that are correlated with demand from other countries.

The (Demand from US)$_{ort}$ component primarily rests on the changing probability of migration driven by evolving immigration policies and demand shocks for migrant workers. While our primary analysis surrounds the changing probability of migration, in Appendix B.6 we do examine variations of this exercise that include the changing wage structure as a component of demand from abroad. In our main specification, our independent variable of interest is the probability of migrating, which falls

[1] We define "Total employment in *ind* and t" as the total number of college graduates in a given industry in t excluding own occupation-region employment in t.

as non-US workers submit more H-1B applications:

$$Ln\left(\text{Migration prob}_{ort}\right) \equiv Ln\left(\frac{\text{Indians in US}_{ort}}{\text{Total US Migration Applicants}_t}\right) \tag{4}$$

The key variable $(\text{Indians in US})_{ort}$ is defined as the number of Indian college graduates from region r working in occupation o in the US, who migrated in the past 5 years. $(\text{Total US Migration Applicants})_t$ is the total number of H-1B applicants over the same 5-year period. In our structural model we match the coefficient γ_2 that corresponds to this migration probability.

Here, we are careful about measuring Indians in US_{ort}. There are no data on the precise origin region r of migrants in the US, and variation in the demand from abroad (which we leverage) is more likely driven at the occupation-year level. For these reasons, we start by assuming the US migration shock affects all Indian regions alike and estimate:[2]

$$Ln(\text{N}_{ort}) = \gamma_0 + \gamma_1 \text{Local}_{ort} + \gamma_2 Ln(\text{Indians in US}_{ot}) + \delta_{or} + \delta_{rt} + \varepsilon_{ort} \tag{5}$$

As we describe in Appendix B.2, the set of fixed effects help account for meaningful components of the Demand India_{ort} and Supply & Preferences_{ort}. These include preferences that vary across region and occupation, that evolve across regions over time, and employment growth driven by regional development over time. As we describe in Appendix B.2, the error term ε_{ort} contains the unexplained component of supply-side preferences ε^S_{ort}, and of local demand shocks ε^D_{ort}, after accounting for these controls. Table A1 presents summary statistics for our main variables.

In Section 3.2.4, we define our main variable $(\text{Indians in US})_{ort}$ at the region-occupation-time level where we no longer assume the US shock is equally affecting all Indian regions. We also do the entire analysis without the region dimension (all dependent and control variables), producing similar estimates of γ_2 (Appendix Table B9). Yet, the region dimension allows us to comprehensively control for local demand shocks, and account for meaningful variation in occupational choice across regions. Region-time fixed effects imply that we can interpret our estimates as an occupation-choice elasticity in response to changes in opportunities abroad. Similarly, our local demand control helps capture occupation-region specific local shocks that may otherwise confound the aggregate specification.[3]

Endogeneity concerns in OLS arise, for instance, as unobserved preferences (the error term ε^S_{ort}) affect both occupational choice and migration probability by occupation. Additionally, if local demand in India contains a component ε^D_{ort} that is yet unaccounted for, but somehow correlated with the migration probability for occupation o, it would bias γ_2.

The direction of the OLS biases is ambiguous. On the one hand, it is possible that Indians have strong preferences for certain occupations in India that happen to be demanded by the US

[2] In Appendix B.2, we explain step-by-step how we derive equation 5 from equation 2.

[3] For instance, the state of Karnataka is a known tech hub in India. If IT local demand increases only in Karnataka and not in other states, the local disaggregation would help the local demand control better capture the increase in demand for CS than an aggregate measure.

$\left(Cov(\varepsilon^S_{ort}, \text{Indians in US}_{ot}) > 0\right)$, driving an upward bias. On the other hand, it is possible that Indians choose to migrate in occupations where job opportunities at home are low $\left(Cov(\varepsilon^D_{ort}, \text{Indians in US}_{ot}) < 0\right)$. For instance, Indians that have a high preference for being space engineers would be more likely to migrate than than those interested in textile manufacturing. This is a case of reverse causality: when local job opportunities are low, those with high preferences for specialized occupations migrate, biasing down the OLS coefficient. Our empirical specifications using a variety of instruments and controls, consistently suggest a downward OLS bias. To account for these possible OLS biases, we use an instrumental variables strategy, leveraging changes in US immigration policy to identify the supply response in India.

3.2.1 Instrumental Variables

To isolate the demand-from-abroad channel, we construct different instruments that leverage changes in US visa policy, demand for different types of workers, and occupation-specific migration propensities. Our aim is to derive variation from either US policy changes, or from US business cycle fluctuations, that are not correlated with supply-side changes in India or demand from other destinations.

Our first instrument leverages changes to the H-1B cap over time shown in Figure 2a, where we reweight these time-varying exogenous changes with baseline exposure-weights. Rather than using the actual H-1B stock, we create a predicted stock relying solely on the cap. That is, we measure the immigrant stock by adding the H-1B cap every year since 1994, whereby changes in the stock of immigrants are only driven by policy changes and not by endogenous employment decisions. The instrument using the H-1B variation as 'shifter' is described by equation 6:

$$\text{H-1B Instrument}_{ot} = \underbrace{\frac{(\text{N Indians in US})_{o,1990}}{(\text{N Indians in US})_{1990}}}_{\text{Initial share}} \times \underbrace{(\text{Policy-predicted H-1B stock})_t}_{\text{time shifter}} \tag{6}$$

We assume that the exogenous variation in cap changes by the US Congress, are unlikely to be driven by occupation-specific shocks to the Indian labor market and worker preferences. Since our specification includes occupation-region, and region-time fixed effects, the remaining variation is primarily coming from the changes in the cap over time, reweighted by the baseline occupation migration propensities. Indeed, since the cap is not occupation specific, we believe such variation to be more exogenous to underlying changes in occupation-specific demand. The shares capture the propensity of Indians to be differentially represented in certain occupations at baseline, so these occupations are more likely to respond when faced with time-varying shocks to the propensity to migrate.

Figure 2a shows the changes in caps that occurred over our period of analysis, likely reflecting political bargaining in the US Congress. The cap was raised substantially with the help of a series of Acts between 1998 and 2001, signed by President Clinton. It was lowered in 2004 by President Bush, with separate provisions for those with graduate degrees from the US. Unlike other work that leverages changes in the cap to study outcomes in the US (Kerr et al., 2015; Kerr and Lincoln, 2010), we need weaker identification assumptions as our outcomes are in India. We do not expect cap changes to be associated with either shocks to supply-side preferences in India, nor with demand in India that is not from the US. As such, we believe that H-1B Instrument$_{ort}$ is uncorrelated with ε^S_{ort} and ε^D_{ort}. The added advantage of leveraging changes over time is that we can conduct tests of pre-trends and baseline correlations which help support our identification assumptions, or control for occupation-trends and time-varying wages in other formulations.

Table 2: US Demand Affects Supply in India

	OLS		**2SLS**	
	Log Employment All	Log Employment Young	Log Employment All	Log Employment Young
Log Indians in US migrated in past 5 years	0.204***	0.169***	0.740***	1.065***
	(0.0455)	(0.0495)	(0.262)	(0.311)
Demand Control	0.654***	0.716***	0.724***	0.833***
	(0.214)	(0.239)	(0.225)	(0.259)
First Stage				
Instrument			0.00293***	0.00293***
			(0.000356)	(0.000356)
Demand Control			-0.165*	-0.165*
			(0.0983)	(0.0983)
N	3,114	3,114	3,114	3,114
1st stage F-stat			67.64	67.64

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$ All regressions include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Columns 1 and 3 show OLS and 2SLS regressions for the dependent variable of log employment of all college graduates in India while columns 2 and 4 count only college graduates who are between 25 and 40. Main explanatory variable is the log number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o. Demand Control is defined in equation 3. Sample is restricted to occupations that are high-skill intensive. Final sample includes 38 occupations, 30 Indian states and 5 periods. We only include occupation-region pairs where workers are found in at least one year. The instrument exploits variation in the H-1B cap as a time shifter. Data for India, from the National Sample Surveys (1994, 2000, 2005, 2010 and 2012), and data for the US from the US Census (1990, 2000), and the American Community Survey (2005, 2010, 2012). The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3.

We describe the results of our exercise in Table 2, showing first the OLS estimate followed by the IV. Across columns we show results by age group, distinguishing between all cohorts, and the young (those below 40 years). Our instrument has a strong first stage. The coefficient for the sample of all cohorts suggests that the labor-response elasticity is 0.74: for a 1% increase in migration probabilities from India in occupation o, there is a 0.74% increase in employment in India. On net, this suggest a brain gain driven by immigration.

When restricting the analysis to young college graduates below the age of 40, the migration elasticity is higher, suggesting that youth are more responsive to demand shocks from abroad. Indeed, this is a nuance we capture with our modeling exercise, whereby we model not just the choice of switching occupations (that varies by age), but also the choice of college major. As youth choose different majors in response to shocks, their labor response is more elastic.[4]

In Table 2, the OLS coefficient suggests a downward bias with respect to the IV. As discussed earlier, one possible reason is that where demand in India is low, workers might put additional effort to

[4]This implies that the long run labor response is more elastic than the short run response, as over time more cohorts major in different fields. We also estimate an occupational switching cost that increases with age.

migrate such that we see high migration rates for occupations with low demand in India. As changes in the H-1B cap are plausibly uncorrelated with worker preferences and demand conditions in India, we can isolate the labor-response elasticity.

3.2.2 Alternative Instruments and Correlated Demand Shocks

We consider other instrumental variables that instead of relying on the H-1B cap, leverage variation in the overall demand for migrants from US firms. In alternative formulations of our instrument, we incorporate shifters that vary at the occupation-by-time level (the demand for all other migrants by occupation), and overall demand for migrants from other countries. That is, to isolate the demand-from-US channel, we replace the time-shifter in equation 6 to be the total number of college-educated migrants working in the US, from all other countries in the world $(\text{Migrants from Other Countries})_t$. We also calculate an alternative version of this instrument, where we allow the time-shifter in equation 7 to vary across occupations, capturing occupation-specific demand $(\text{Migrants from Other Countries})_{ot}$.

$$\text{Demand from US IV}_{ot} = \underbrace{\frac{(\text{N Indians in US})_{o,1990}}{(\text{N Indians in US})_{1990}}}_{\text{Initial share}} \times \underbrace{(\text{Migrants from Other Countries})_{ot}}_{\text{time shifter}} \tag{7}$$

Once again, these instruments capture the demand for migrants from the US, and are unlikely to be correlated with supply-side changes in India. As we show in Panel A of Table 3, these instruments yield similar results both in terms of magnitude and significance.

One may question whether the H-1B cap changes are themselves correlated with shocks to the US economy. This does not raise identification concerns *per se*, as we do want variation from the US economy (whether business cycle changes or policy shocks, or a combination), to drive the demand from the US. Figure 2a suggests that sometimes these forces may move in opposite directions – for instance, the cap increase coincided with the 2001 dot-com bust.

Even if the H-1B IV was not correlated with US demand shocks, the IV proposed in equation 7 (and Table 3) might be. While much of the variation in equation 7 likely still stems from the binding H-1B cap, some part may be due to fluctuations in US economic cycles. This is only a concern if this demand for workers from the US (that is not driven by H-1B policy changes) is also correlated with demand for workers from other destinations (such as, the UK or Canada), altering the interpretation of the estimated elasticity. We examine this possibility in detail in Appendix B.5, where we combine micro data from other non-US destinations, and show that the demand from firms in other parts of the world play a negligible role. In Figures B1a and B1b, it is clear that even in 2012, when other countries have had time to develop their IT sectors, the US swamps overall demand when compared to the UK or Canada. As expected, in 2000, when only the US IT sector was developed, the scale of the difference in the demand for IT from the US was many magnitudes larger than the UK (Figure B1c).

Together, these instruments allow us to estimate the supply-side responses in India to demand from the US. A final point of examination would be if the US demand shocks (say, from the H-1B cap) were correlated with changes in India for *other* reasons. It is unlikely that the political wrangling in the US on the H-1B cap changes resulted from college major-choice in India. Indeed, as we intimate in Section 2.2 and Appendix H, the fledgling Indian IT sector was still nascent and local CS wages low, when the US innovation boom happened and drove up-skilling in India (Figure 4).

Yet, this is why we control for local demand from within India, as in equation 3. If demand from firms in India are correlated with demand from other countries for occupation o', it will increase the term $Local_{o'rt}$. Indeed, in Appendix Table B5, we find that in removing the control, our estimates do not change. Our measure of local demand shocks is itself meaningful across all tables, and as expected, has a direct effect on occupational choice. But it does not affect the relationship between the probability of migration and occupational choice, suggesting that it is unlikely to be associated with H-1B cap changes. This control not only accounts for local industrial growth but also other nationwide shocks. For instance, if there was a nationwide demand shock for IT, the IT component of the shift-share will increase, and those region-occupation pairs more intensive in IT will have larger overall labor demand than less intensive region-occupation pairs. This way, the control also accounts for local demand changes for each occupation over time.

In Appendix B.5 we describe the role played by an additional demand control derived from the industry-level flow of exports from India to the US. This captures correlations in product-level demand for goods that may affect the demand for labor. In Table B5 we find that in controlling for these product demand shocks, our estimated elasticities remain unaffected.

3.2.3 Falsification Tests and Alternative Specifications

We consider alternative specifications and falsification exercises to challenge our conclusions. In all such specifications we consistently show results for the full sample and young workers separately. First, in Appendix B.4, we start by modifying the H-1B cap instrument to depend on the length of the visas, and the probability of return migration. We then consider the importance of various controls and specification decisions, for instance, where we exclude local demand controls, and use the inverse hyperbolic sine to include occupation-time observations with zero shares. Appendix B.5 shows these robustness checks for the alternative instruments.

Second, in Appendix B.6 we explore the role of the changing earnings for migrants in the US. While our primary framework best reflects the specification consistent with our structural model, here we also consider an alternative specification. That is, we reformulate our demand shock from abroad (Demand from the US$_{ort}$), to take into account not only the probability of migration but also changes in the occupation-specific wages in the US. This augments our primary specification above, which leveraged only the change in the probability of migrating. We also discuss in detail, alternative measures of the endogenous variable, whereby we use the wage-bill (rather than employment) of recent Indian migrants in the US. The results in Appendix B.6, suggest that accounting for wages does not affect our primary elasticity estimates of the labor-response with respect to the probability of migrating.

Table 3: Alternative Instrumental Variables

Panel A: Alternative Instruments	Non Indian Migrants IV Log Employment All	 Log Employment Young	Non Indian Migrants by Occupation Log Employment All	 Log Employment Young
Log Indians in US	0.629***	0.770***	0.582**	0.915***
migrated in past 5 years	(0.118)	(0.138)	(0.240)	(0.291)
Demand Control	0.710***	0.795***	0.703***	0.814***
	(0.222)	(0.250)	(0.221)	(0.254)
N	3,114	3,114	3,114	3,114
1st stage F-stat	178.1	178.1	66.71	66.71

Panel B: Regional Re-weighting	H-1B cap IV Log Employment All	 Log Employment Young	Non Indian Migrants IV Log Employment All	 Log Employment Young
Log Indians in US	0.925**	1.245**	0.810***	0.923***
migrated in past 5 years	(0.457)	(0.533)	(0.281)	(0.346)
Demand Control	0.725**	0.822**	0.728***	0.828***
	(0.298)	(0.370)	(0.280)	(0.316)
N	3,087	3,087	3,087	3,087
1st stage F-stat	9.429	9.429	12.10	12.10

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$ All regressions include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Columns 1 and 3 show 2SLS regressions for the dependent variable of log employment of all college graduates in India while columns 2 and 4 count only college graduates who are between 25 and 40. Main explanatory variable is the log number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o. Demand Control is defined in equation 3. Sample is restricted to occupations that are high-skill intensive. Final sample includes 38 occupations, 30 Indian states and 5 periods. We only include occupation-region pairs where workers are found in at least one year. The instrument exploits variation in the H-1B cap as a time shifter. Data for India, from the National Sample Surveys (1994, 2000, 2005, 2010 and 2012), and data for the US from the US Census (1990, 2000), and the American Community Survey (2005, 2010, 2012). The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3. Panel A shows the primary specifications with two alternative instruments. In columns 1 and 2, as in equation 7, we capture the demand for migrants by looking at the flow of college-educated workers from all other countries to the US. In columns 3 and 4, we modify the instrument to be the number of migrants from other countries *by occupation*. In panel B, we reformulate the explanatory variable and the instrumental variable to be reweighted by the baseline region-share of Indians in a specific occupation, as in equations 9 and 8. Columns 1 and 2 shows results where the time-shifter in the instrument is the H-1B stock. Columns 3 and 4 shows results where the time-shifter in the instrument is the number of non-Indian migrants to the US.

Third, in Appendix B.7, we conduct various specification tests. We show our results using levels instead of logs, and controlling for occupation-by-region specific trends. We document the sensitivity to various sub-samples, for instance, by dropping each occupation, and then each region one at a time. Importantly, we conduct a pre-trends analysis to show that future changes to shocks do not affect past labor supply. We further show that the baseline shares are not correlated with other observable characteristics of the labor market and education sector prior to the IT boom. We further show that the instrument itself is not predictive of other changes in the local labor market (like the unemployment rate). Such tests of pre-trends, and baseline correlations help support our identification assumptions.

Finally, in Appendix B.8 we conduct a set of falsification and specification tests that have also been reiterated by recent work on shift-share instruments (Adao et al., 2019; Borusyak et al., 2020; Goldsmith-Pinkham et al., 2020). We confirm that our specification favorably fits certain guidelines and recommendations suggested by such recent work. In incorporating occupation-by-region fixed effects, we already control for the baseline shares, and rely on the time-varying (H-1B policy) shocks for identification. And in controlling for local demand shocks, we control for local confounders. Given the lack of correlation between changes over time and baseline shares, we find that inference is unaffected (Adao et al., 2019). Unlike other work on migration, (Card, 2001), our outcomes and explanatory variables are based on datasets from different parts of the world, preventing mechanical biases, and necessitating weaker identification assumptions (Jaeger et al., 2018). The lack of pre-trends, solidify the identification assumptions. Balance with baseline characteristics, robustness to dropping each occupation and region one at a time, and the inability to explain the variation in baseline shares from a long list of covariates, while not necessary for identification (Borusyak et al., 2020), do help support our design (Goldsmith-Pinkham et al., 2020)

3.2.4 Dimensions of the Data, and Regional Variation

So far most of the variation we leverage is at the occupation-time level. We now describe specifications that incorporate the regional dimension in our instrument, reweighting the baseline propensities of each region in India to produce workers of a certain occupation. Then, we also conduct all our analyses solely at the occupation-time level. While this last specification limits the ability to control for changes driven by region-specific shocks, we show that our estimated elasticities are similar to before.

First, we consider specifications where we reweight (Indians in US)$_{ot}$ with baseline propensities of each state in India to produce workers of occupation o. Our independent variable is now $Ln\,(\text{Indians in US}_{ot} \times \text{Region share}_{ort})$. In our main specification, we had weighted each state equally (i.e. $\text{Region share}_{ort} = 1$). Now we define the region share to be the share of Indians in an occupation, from a specific Indian state r. The baseline propensities not only capture the presence of local universities and local jobs in the state, but also potential connections to migrants abroad, amplifying demand shocks from abroad. We lag this region-share measure by five years (see Appendix B.3 for data details) to capture the baseline shares that will be affected by US emigration:

$$\text{Region share}_{rot} = \frac{\text{N Indians}_{ro,t-5}}{\text{N Indians}_{o,t-5}} \tag{8}$$

As demand for occupation o increases, regions with higher baseline propensities to produce workers of occupation o would be particularly affected. We also define our instrument at the region-occupation-time level by interacting the initial distribution of occupations across regions with the initial distribution of occupations of Indians in the US as shown in equation 9:

$$\text{H-1B IV'}_{ort} = \underbrace{\frac{(\text{N in India})_{o,r,1991}}{(\text{N in India})_{o,1991}} \times \frac{(\text{N Indians in US})_{o,1990}}{(\text{N Indians in US})_{1990}}}_{\text{Initial share}} \times \underbrace{(\text{Policy-predicted H-1B stock})_{t}}_{\text{time shifter}} \tag{9}$$

Panel B of Table 3 shows these results for the instrumental variables that use the H-1B stock and the number of non-Indian migrants to the US as time shifters. In the Appendix Table B8, we show the results for the remaining instruments described in Section 3.2.3. Across specifications, once again, we find similar estimates showing that as demand for certain occupations from the US increase, there is a detectable labor supply response in India.

Importantly, we also conduct the entire analysis without the region dimension altogether (Appendix Table B9). That is, all our dependent, independent, instrumental and control variables are all at the occupation-time level, once again producing similar estimates of γ_2. We show robustness to using logs or levels, across each type of instrument, controlling or not controlling for occupation trends, showing results by age groups, and finally (in Panel D) normalizing by the US workforce to account for heteroskedasticity.

Yet, the regional dimension allows us to comprehensively control for local demand shocks, and account for meaningful variation across regions and over time. As certain industries (like IT) are concentrated in few parts of the country, the region dimension allow us to control for these baseline differences across regions with δ_{or}. The region-by-time fixed effects δ_{rt} account for the fact that certain regions may grow for reasons unrelated to demand from abroad.

4 Model

Our model consists of two main parts: in Section 4.1, we model the labor supply decisions of college graduates in both the US and India, and in Section 4.2 we discuss how goods are produced and sold to consumers in each country, and the rest of the world. In the product market, firms and consumers make static decisions each period conditional on the parameters of the model, and the availability of each type of labor in the economy. The college labor market has a dynamic horizon: since human capital investments and career choices have long-term payoffs, workers in both countries are allowed to choose their fields of study and occupations based on the information they have today and their expected payoffs in the future. Workers in India are uncertain on whether they will get a slot to migrate when making their occupational choices, which opens up the possibility of brain-gain. In Section 4.3, we describe the equilibrium.

4.1 The Supply of Workers in India and the US

Workers in India and the US work for a finite number of periods N and are either high (college graduates) or low skilled (non-college graduates), denoted by $H_{t,k}$ and $L_{t,k}$ respectively, where k represents country and t time. In our baseline framework, we assume that the decision to attend college is made outside the model, but an extension models this decision as well.[5]

College graduates in both countries have the choice to become computer scientists ($CS_{t,k}$) or work in some other college-graduate occupation ($G_{t,k}$). We allow for two type of decisions for college graduates. First, prior to joining the labor market, students choose whether to enroll in CS or in a non-CS field of study, which influences their initial occupation after graduation. Second, workers choose every period between working as a computer scientist and working in another occupation. We group non-CS occupations together despite the heterogeneity across jobs: we capture this heterogeneity in tastes and proclivities as individual, period-specific taste shocks of studying each field and working

[5]We consider the marginal computer scientist's main choice to be between CS and other college-graduate occupations instead of between CS and not going to college. In section 7.4 we show that quantitatively, our results change very little when adding an endogenous college decision.

in each occupation.[6]

Some Indian CS have the opportunity to migrate every period to work in the US as computer scientists ($CS^m_{t,k}$). The total number of Indians migrating depends on the US migration policy cap each period (cap_t). Once in the US, migrants face an exogenous probability of return, where if selected, they have to go back to India and work as return computer scientists ($CS^R_{t,k}$) until they retire. Each period, every Indian CS who never migrated knows they will participate in a lottery that can give them the chance to migrate, but are uncertain on whether they will get one of the capped slots when making their occupation decisions. At the same time, workers in India and the US face uncertainty over the actual value of the cap, and future wages. Figure C1 summarizes the timing of the decisions each period.

We present the equations related to the decisions of workers in India. The decisions for the US-born are a subset of Indian workers' decisions, since the US-born only choose between CS and non-CS, but do not have the opportunity to migrate. Individual i's decision problem, before joining the labor market, is summarized by equation 10:

$$\max\{\beta \mathbb{E}_t V^{cs}_{t+1} + \bar{F} + \sigma \eta^{cs}_{i,t},\ \beta \mathbb{E}_t V^{g}_{t+1} + \sigma \eta^{g}_{i,t}\} \tag{10}$$

Individuals compare the expected future payoffs of joining the labor force next period with a CS degree V^{cs} with the future payoff of joining with some other college degree V^g. Such choices are also affected by a fixed education cost for studying CS, $\bar{F}_{cs}$, which can be positive or negative, and idiosyncratic taste shocks for studying each field: η^{cs}_i and η^g_i.[7] We assume that η^{cs}_i and η^g_i are independently and identically distributed as a standard Type I Extreme Value distribution (Rust, 1987). The parameter σ controls the sensitivity of major and occupation choices to preference shocks. A smaller σ implies small changes in career prospects produce big variations in the number of students graduating with CS degrees. Enrolling in a CS major allows individuals to join the labor force as CS, while non-CS majors join non-CS occupations.

Once they join the labor market, at the start of each period, individuals choose to work in CS or another occupation to maximize the expected present value of their lifetime utility. We denote occupational choice as $o = \{cs, g\}$ in equation 11:

$$V^o_{t,a} = \max_o \left\{ \underbrace{w^o_t}_{\text{current wage}} + \underbrace{\chi(a) \times \mathbb{1}(o_t \neq o_{t-1})}_{\text{switching cost}} + \underbrace{\zeta \times \mathbb{1}(o_t = g)}_{\text{distaste for CS}} + \underbrace{\beta \mathbb{E}_t [V^o_{t+1,a+1}]}_{\text{future payoffs}} + \underbrace{\sigma \eta^o_{i,t}}_{\text{preferences}} \right\}, \tag{11}$$

where $V^o_{t,a,k}$ is the value of starting in occupation o in period t at age a. At the beginning of each period, individuals learn their period-specific preference shock and decide whether to switch occupations, taking into account their expected lifetime stream of income. ζ is the taste attractiveness parameter

[6] Appendix C.2 shows how grouping multiple occupations in the CES framework preserves the CES elasticity.

[7] The fixed cost for getting a CS major $\bar{F}$ captures the taste frictions faced by college goers that potentially deter them from choosing CS degrees despite CS having higher relative wages.

for not working as a computer scientist, and $\chi(a)$ is the age-dependent monetary cost of switching occupations. For simplicity, we assume switching costs are linear in age: $\chi_0 + (\chi_1 \times age)$, to capture that switching can become harder as workers get older. In the model, all workers retire after working N periods.

After each period of CS work in India, workers participate in a lottery to migrate to the US and work as computer scientists from the next period onward. Since the outcome of the lottery is uncertain when making occupation decisions, the future payoff of choosing a CS occupation becomes a weighted average between the future payoffs of working in India and the future payoffs of working in the US as shown in equation 12:

$$\mathbb{E}_t[V^{cs}_{t+1,a+1}] = \mathbb{E}_t[p_{t+1} \times V^{us}_{t+1,a+1} \ + \ (1-p_{t+1}) \times V^{in}_{t+1,a+1}] \,, \tag{12}$$

where p_{t+1} is the migration probability, which depends on the US migration cap in period $t+1$ (cap_{t+1}) and the total number of CS in India in period t as in equation 13:[8]

$$p_{t+1} = \frac{cap_{t+1}}{CS_{t,in}} \tag{13}$$

Once in the US, Indian migrants either stay working as CS in the US until they retire or return to India, based on an exogenous return probability, ρ. The future payoff of migrating is:

$$V^{us}_{t+1,a+1} = \kappa w^{cs}_{us,t+1} \ + \ \beta \mathbb{E}_{t+1}[(1-\rho)V^{us}_{t+2,a+2} + \rho V^{in,R}_{t+2,a+2}] \,, \tag{14}$$

where $V^{in,R}_{t+2,a+2}$ is the future payoff of returning to India and working as a returned CS until retirement. Parameter κ captures any utility cost/benefit of each dollar earned in the US relative to India. For instance, if moving to the US also improves access to higher quality services and amenities, κ would be larger than 1 and workers would respond more to changes in the US wage than the Indian wage. A key simplification of our model is that both the decision to migrate and the decision to return are deterministic for the worker. If they win the lottery they will always migrate, and if they get drawn to return they will return. As long as there is a large wage premium in the US, and the H-1B cap is small relative to the total number of CS in India, it is a reasonable simplification to assume that there will always be enough workers who want to migrate.[9] We assume all workers have identical abilities and CS workers in India have the same probability of migration, both of which we relax in Section 7.5.

For US workers, the structure of the model is identical to the one presented above with the exception of emigration. All parameters in the labor supply model are different for US and Indian workers. The labor supply parameters jointly determine the dynamic elasticity of occupational choices with respect to wage: the short-run labor supply curve may be inelastic, but as more students choose majors,

[8]We assume that all Indian CS in period t are potential applicants to migrate to the US in period $t+1$. Since Indian CS are the only workers that can migrate in the model, the overall probability is the value of the cap divided by the total number of Indian CS in India each period.

[9]Similarly, given the costly green card sponsorship process, and the long green card queues for Indian nationals, we would expect that return decisions are mostly driven by the firm rather than the worker.

the long-run elasticity is higher.

4.1.1 Wage and Cap expectations

When choosing occupations, workers in both countries form expectations on the path of future wages and immigration caps. We assume worker expectations follow an AR(1) process:

$$x_{k,t+1} = \phi_0^{x,k} + \phi_1^{x,k} x_{k,t} + \varepsilon_{k,t+1}^x \,, \tag{15}$$

where x_k stands for future wages $\{w_k^g, w_k^{cs}\}$ or migration caps $\{cap\}$. The x,k superscript indicates that there is a different set of parameters for each country-outcome pair. $\varepsilon_{k,t+1}^x$ are forecasting errors assumed to be *iid* over time. As in Dix-Carneiro (2014), we assume this forecasting rule is consistent with the equilibrium outcomes of the model, such that workers have rational expectations over wages and the migration cap. This structure indicates there are two dimensions of uncertainty when it comes to immigration. First, workers do not know what the total number of new migrants each period will be when they make their migration decisions. Second, even if they knew the migration cap, they do not know with certainty whether they will get a slot given the lottery. Both of these features open the possibility of brain gain in our model. If the migration cap increases today, many Indian workers will go into CS to try to migrate. However, many of those will not win the lottery and will stay working as CS in India.

4.2 Product Market

4.2.1 The Household Problem

We close the model by specifying how consumption, production, and trade occurs. Consumers in each economy supply one unit of labor, and have the same preferences over final good Y, which has Constant Elasticity of Substitution (CES) form over different varieties $v \in [0,1]$.[10]

$$Y = \left(\int_0^1 y_v^{\frac{\iota-1}{\iota}} dv \right)^{\frac{\iota}{\iota-1}} \,, \tag{16}$$

where ι is the elasticity of substitution between the varieties of the final good. These varieties may be produced in other parts of the world and imported. Access to more varieties at lower prices (say, as IT production expands) raises consumer welfare.

A consumer's labor income is spent on these goods as there are no savings. Consumers maximize utility subject to a budget constraint, where expenditure equals wage income. While consumers have identical consumption preferences, they do not receive the same labor income as they work in three different occupations (CS, non-CS graduates, and non-graduates).

[10]Since the product market is static and that the structure is the same across countries, we omit country and time subscripts k and t for convenience, but all endogenous variables and parameters are time and country specific. The production function elasticities are constant and common for both countries.

4.2.2 Final Goods Production

Each firm producing variety v in the final goods sector (subscript y) has Cobb Douglas constant returns to scale technology over intermediate inputs from the IT sector $C_{v,y}$, and a labor aggregate, with Cobb Douglas parameter γ. Each variety is produced with productivity $z_{v,y}$:

$$y_v = z_{v,y} C_{v,y}^{\gamma} \left(\left[\alpha (\ell_{v,y})^{\frac{\tau-1}{\tau}} + (1-\alpha)(h_{v,y})^{\frac{\tau-1}{\tau}} \right]^{\frac{\tau}{\tau-1}} \right)^{1-\gamma} \tag{17}$$

The IT good is an input in final goods production: importantly, this implies that innovation in IT can increase productivity in other sectors of the economy.[11] Following the framework introduced by Eaton and Kortum (2002), each producer has a different level of efficiency in producing each variety, denoted by $z_{v,y}$. The final goods sector employs low-skilled non-college graduates $\ell_{v,y}$, and an aggregate of high-skilled college graduates $h_{v,y}$, with an elasticity of substitution τ between them, and a distributional parameter α.

Using a nested CES format, the aggregate of college graduates $h_{v,y}$ can be represented by equation 18, where $cs_{v,y}$ is the number of CS hired in the final goods sector, and $g_{v,y}$ is the number of non-CS graduates hired in the final goods sector. This complementarity ensures that as the US hires more CS workers, it raises the demand for non-CS occupations (like managers), tending to raise the non-CS wage. Native and foreign-born CS are perfect substitutes in production.[12] The elasticity of substitution between CS and non-CS college graduates is λ.

$$h_{v,y} = \left[\delta \left(cs_{v,y}\right)^{\frac{\lambda-1}{\lambda}} + (1-\delta)\left(g_{v,y}\right)^{\frac{\lambda-1}{\lambda}} \right]^{\frac{\lambda}{\lambda-1}} \tag{18}$$

As immigration increases the size of the CS workforce, demand will rise for workers in complementary occupations, raising their wages. This may induce native CS workers to switch to other occupations, mitigating negative wage impacts. At the same time, skill-biased technical change towards CS workers δ, or towards college graduates $(1-\alpha)$, shift over time with the innovation boom. In India, we assume CS workers are a CES aggregate of native CS (cs^n) and return CS (cs^r) as in equation 19:

$$cs_{v,y} = \left[\left(cs_{v,y}^{n}\right)^{\frac{\varepsilon-1}{\varepsilon}} + \left(cs_{v,y}^{r}\right)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}} \tag{19}$$

Native and return-migrant CS workers are not perfect substitutes, as return migrants have different sets of skills given their work history abroad. ε is the elasticity of substitution between the native CS workers and return migrants.

The first-order conditions determine the demand for intermediate IT inputs and the different types of labor in the final goods sector. Together with the demand for labor from the IT sector we derive the aggregate labor demand for each worker.

[11] A major component of US productivity growth is attributable to industries that use IT as an input (Jorgenson et al., 2016), such as financial services, motor-vehicle manufacturing, and scientific production.

[12] In Section 7.3, we allow foreign CS to be imperfect substitutes for natives CS.

4.2.3 Production in the IT Sector

For each IT variety j we assume that there are infinitely small firms with constant returns to scale technology willing to produce the good. Firms in the final goods sector have preferences over different types of IT goods c_j, such that:

$$C_y = \left(\int_0^1 c_j^{\frac{\iota-1}{\iota}} dj \right)^{\frac{\iota}{\iota-1}} \tag{20}$$

IT firms have CES technology in the labor aggregate (equation 21), where $cs_{j,c}$ is the number of CS and $g_{j,c}$ non-CS college graduates employed by IT firm j. Here λ is the elasticity of substitution between CS and non-CS college graduates, and $\delta + \Delta$ is the distributional CES parameter. $\Delta > 0$ as IT is more intensive in CS than the final goods sector. For India, $cs_{j,c}$ is the CES composite between native and return computer scientists as in equation 19.

$$c_j = z_{j,c} \left[(\delta + \Delta) \left(cs_{j,c} \right)^{\frac{\lambda-1}{\lambda}} + (1 - \delta - \Delta) \left(g_{j,c} \right)^{\frac{\lambda-1}{\lambda}} \right]^{\frac{\lambda}{\lambda-1}} \tag{21}$$

4.2.4 International Trade

We model the world economy as a set of three regions: the United States, India, and the rest of the world (RoW), with preferences and production as described in sub-sections 4.2-4.2.3. Although we focus on India and the US, we incorporate the RoW to capture how India and the US compete in the world market and how they have the option of buying and selling products to a third region. While workers in the RoW produce and consume both final goods and IT goods, we simplify the analysis by assuming they do not receive or send CS migrants.[13]

All three regions trade both goods (final and IT) following the standard framework of Eaton and Kortum (2002), where each region has a comparative advantage in producing some of the varieties of each good. We assume that country k's efficiency in producing good j in sector s is the realization of the random variable Z_k^s, drawn independently for each j from a distribution $F_{s,k}(z)$, where productivity $z_{j,s,k}$ is from a Frechet (Type II extreme value) distribution:

$$F_{s,k}(z) = e^{-T_{s,k} z^{-\theta}} \tag{22}$$

Here $\theta > 1$ governs the dispersion of the productivity draws across varieties. Higher $T_{s,k}$ increases the likelihood of drawing higher efficiencies for good j, and is the technology level for each country-sector pair. If the US has higher $T_{s,k}$, the US is more efficient at producing more varieties in sector s on average, even as India and the RoW will be efficient at producing certain varieties in the sector. Innovation by CS workers shifts out the distribution, raising $T_{s,k}$.

Consumers in each country buy each variety from the lowest-price producer. If a consumer in b buys from k, they pay an iceberg-trade cost $d_{b,k}$. All else equal, a country becomes a more attractive

[13] Figure B1 describes the relatively negligible skilled migration from India to non-US destinations.

provider of the good whenever their prices are lower as a consequences of one of three things: an increase in the technology (that allows for better draws of $z_{s,k}$), a decrease in trade costs $d^s_{b,k}$, or a decrease in labor costs $(w_{\ell,k}, w_{g,k}, w_{cs,k})$.

Such features of our model will be relevant to capture the empirical patterns shown in Figure 4b: while the US was the predominant exporter of IT goods for most of the 1990s, India takes over soon thereafter as technology in India increases. Since varieties may be produced in any part of the world and imported, restricting immigration to the US may affect growth in the US IT sector and lead to certain varieties being produced in other countries. At the same time, more migration raises the prospect of migrating from India; which increases the size of the Indian CS workforce, potentially shifting some production from the US to India.

We capture the possibility of firms from country b outsourcing production to k, reflected by more exports from k to b. US-owned firms producing and exporting from India count towards Indian production and exports. The same is true for Indian-owned firms in the US.[14]

Importantly, we model directed technological change (Acemoglu, 1998). Since production in IT is heavily reliant on technology, this is an important driver of how technology spreads to India. Computer scientists in both countries are innovators and increase the technological productivity in the IT sector (Kerr and Lincoln, 2010). This can potentially raise wages on average and can mitigate the depression in CS wage growth due to immigration. Since IT output is an intermediate input into the final goods sector, technological advances can increase the productivity of other downstream sectors of the economy as well.

Innovation depends on the number of CS workers in the IT sector. The 'brain drain' of CS workers to the US is countered by the 'brain gain' of workers acquiring CS skills with the prospect of migrating. As we show later the brain gain makes the Indian IT sector more productive and over time the leading exporter of IT.[15] We parametrize IT sector productivity in country k to be a function of the total number of CS IT workers in country k:

$$T_{c,k} = \bar{T}_{c,k} CS^{\beta}_{c,k} \;\; for\; k = \{us, in\} \;, \tag{23}$$

where $\bar{T}_{c,k}$ is baseline IT technology in country k, and $CS^{\beta}_{c,k}$ is the endogenous technology component. β is the elasticity of technology level to changes in CS IT workers in country k.

4.3 Equilibrium

Equilibrium in each period is a set of prices and wages ($P_{t,c,k}$, $P_{t,y,k}$, $w^{\ell}_{t,k}$, $w^{g}_{t,k}$, $w^{cs}_{t,k}$), quantities of output and labor ($C_{t,y,k}$, $Y_{t,k}$, $L_{t,k}$, $G_{t,k}$, $CS_{t,k}$), and level of technology ($T^s_{t,k}$) such that: (1) consumers in the US, India, and the rest of the world, maximize utility by choosing $Y_{t,k}$ taking prices as given, (2) college graduates in the US and India choose their field of major and occupations, taking wages

[14] Morales (2019) explores the implications of US high-skill immigration policy on multinational activity.

[15] Here we address a growing literature on technological diffusion and directed technological change within the Ricardian framework (Alvarez et al., 2013; Dasgupta, 2012; Kerr, 2013b; Perla et al., 2015; Somale, 2014).

as given, and form expectations, (3) firms in both the IT and the final goods sector maximize profits taking wages and prices as given, (4) trade between the three regions is balanced, and (5) output and labor markets clear. In Appendix C.3 we describe in detail the equilibrium equations regarding trade, occupation choice and the price index.

5 Empirically Determining the Model's Parameters

Here we describe how we use data to determine the parameters of the model for the period 1995 to 2010. We map each period in the model as a 5-year period in the data (1995, 2000, 2005, 2010). We assume students choose their majors at age 20 and join the labor force at age 25. Then they work for 8 periods until age 65, when they retire.[16]

To make the exposition transparent, we separate this process into three building blocks of parameters: 1) product market elasticities: τ, λ, ε, ι, θ and β; 2) time-varying product demand parameters: $\delta_{k,t}$, $\alpha_{t,k}$, $T^s_{k,t}$, $d^s_{k,b,t}$, $\gamma_{k,t}$, $\Delta_{t,k}$, the labor quantities and the migration cap; and finally, 3) the elasticity of labor supply of college graduates in the US and India determined by labor supply parameters: $\Theta = \{\sigma_k, \zeta_k, \chi_{0,k}, \chi_{1,k}, \bar{F}_k, \kappa, \phi_0^{x,k}, \phi_0^{x,k}\}$ for $k = \{us, in\}$.

The solution algorithm consists of the following steps: first, we estimate product market elasticities, and set some to values estimated in the literature (Block 1). Second, we guess the labor supply elasticity parameters $\hat{\Theta}_{guess}$ and discipline the time-varying product demand parameters to match trends in wages and productivity that reflect the skill- / sector-biased technological change that occurred between 1995-2010 (Block 2). Once all product market parameters are determined, we solve for equilibrium, conditional on the labor supply parameters being $\hat{\Theta}_{guess}$. Finally, we construct a series of targeted data moments related to the labor supply. Our algorithm repeats the process iteratively, searching over the labor supply parameters until the distance between predicted and observed data moments is minimized (Block 3). In Section 5.1 - 5.3, we discuss the estimation of each of these blocks in detail.

5.1 Product Market Elasticities

As a first step, we determine the product market elasticities. We set the elasticity of substitution between college and non-college graduates, $\tau = 1.7$, based on papers the estimate that parameter (Card and Lemieux, 2001; Goldin and Katz, 2007; Katz and Murphy, 1992), and we explicitly estimate it using data from India.[17] For elasticities of substitution between CS and non-CS college graduates, we

[16] The 5-year period decision is mainly due to the NSS in India being available roughly every 5 years until 2005 and every 2 years after that. We describe in detail the US and India datasets in Appendix A.

[17] We replicate Card and Lemieux (2001) using the India data and estimate an elasticity of complementarity of 0.55 (see Table D1). This corresponds to $\tau = 1.8$, and is statistically indistinguishable from 1.7. These papers estimate the overall substitution between college and non-college graduates, while our parameter is sector specific. However, when calculating the overall substitution between college and non college graduates our estimates are indistinguishable from our assigned value of τ.

set $\lambda = 2$, which is within the estimates of Ryoo and Rosen (2004) and Burstein et al. (2019).[18] To determine the substitution elasticity between CS who never emigrated, and those who return from the US, we follow the literature on return migration. We may expect this elasticity to be greater than the elasticity between CS and non-CS graduates. In our steady-state year, we match the average premium of 15% across papers in this literature, corresponding to a value of $\varepsilon = 30$.[19] For the substitution between varieties in each country, we follow Bernard et al. (2003) who estimate the elasticity of substitution across US plants to be 3.79 and set $\iota = 4$. For the trade elasticity we use $\theta = 8.28$, proposed by Eaton and Kortum (2002). We show in Table E3 that our results are robust to variations in these elasticities.

Finally, for the elasticity of technology with respect to the total number of CS working in IT, we estimate $\beta = 0.228$. As we elaborate on in Appendix D.2, we use an instrumental variable strategy together with variation in patenting across industries in the US to identify this parameter. We use a shift-share instrument that interacts the industry-specific dependence on immigrant CS workers at baseline with the total number of immigrant CS in the US each year. Our instrument leverages variation in US immigration policy (like changes to the H-1B cap), sending country shocks, and the fact that immigrants are more likely to be CS. Importantly, our estimated elasticity is very close to other work.[20]

5.2 Time-varying Product Market Parameters

For a given guess of the labor supply parameters $\hat{\Theta}_{guess}$, we discipline the time-varying product demand parameters such that we match observed trends in relative wages, trade flows and production shares. A detailed summary of our data sources can be found in Appendix A. The time-varying parameters γ_k, $\alpha_{t,k}$, $\delta_{t,k}$, $\Delta_{t,k}$, $T^s_{t,k}$ and $d^s_{t,k,b}$ are calibrated separately for 1995, 2000, 2005 and 2010 for each country. Here we just mention what trends we are explicitly matching and in Appendix D.3 we document details on the implementation.

To calibrate the Cobb Douglas parameters γ_k we match the share of income from the final goods sector spent on varieties of the IT sector for each country. The share parameter of non graduates in the production function, $\alpha_{t,k}$, is determined in both India and the US such that it matches the observed share of expenditures from the final goods sector in non-college graduates. The distributional parameter between CS and non-CS college graduates $\delta_{t,k}$ is calibrated to match within-country relative wages

[18] Ryoo and Rosen (2004) estimate the elasticity of substitution between engineers and other graduates to be 1.2 - 2.2. Burstein et al. (2019) estimate an occupational elasticity of substitution between 1.81-2.1.

[19] The literature on return migrants find that in other contexts, those who emigrated for work and return home earn a wage premium relative to those who never migrate (Barrett and O'Connell, 2001; Hazans, 2008; Reinhold and Thom, 2013).

[20] Peri et al. (2015) estimate a 1% increase in the US STEM workforce increased TFP by 0.27%, whereas Kerr and Lincoln (2010) find patenting elasticities that lie between 0.1 and 0.4. In our earlier work, Bound, Khanna, and Morales (2016), we estimate an elasticity of 0.23 from how the price of IT goods change with changes in the CS workforce. In recent work, Khanna and Lee (2018) we find an elasticity of 0.2 using measures of innovation derived from Schumpeterian growth. In Section 7.2, we discuss how our results change for different values of β, including when there is no spillover in India.

between CS and non-CS college graduates in the data. The additional distributional parameter in the IT sector $\Delta_{t,k}$ captures the extra intensity of CS in the IT sector. We calibrate $\Delta_{t,k}$ to be proportional to $\delta_{t,k}$ every period such that it matches the within-country relative share of CS between the IT and non-IT sector in 1995. To estimate the productivity levels ($T^{s}_{t,k}$) and bilateral trade costs ($d^{s}_{t,k,b}$) for each country-sector pair we use trade data and match the observed trade flows every year. Finally, for relative technology in the non-IT sector between India and the US, we calibrate $\left(\frac{T^{y}_{t,in}}{T^{y}_{t,us}}\right)$ such that we match relative wages of non graduates between India and the US.

We also calibrate the total quantity of college and non-college workers in each country using population statistics. We use information on new-Indian immigrants to the US to calibrate the migration cap under the real scenario. Finally, for the return migration rate ρ we set it to $\rho = 23.5\%$ for every period. According to Lee (2016), the OECD estimates that 23.5% of high-skill immigrants in the US return to their home countries after a six-year period. We use the American Community Survey to follow cohorts of Indian CS in the US over time and estimate the return rate to be indeed close to 23.5%.

5.3 Dynamic Labor Supply Identification

We are now ready to estimate the labor supply parameters $\hat{\Theta}$, which determine the dynamic elasticity of labor supply. Every year, the labor demand curve for the US and India shifts due to changes in technology and production function parameters as in Section 5.2. Such exogenous innovation shocks capture skill-biased and sector-biased technological progress that shift out the relative demand curve for CS. These exogenous shifts in labor demand allow us to trace out the labor supply curve, and identify the underlying labor supply parameters.

We use a minimum distance estimation technique (McFadden, 1984) where we identify the labor supply parameters $\hat{\Theta}$ jointly using specific moments of the data. Let the product market parameters in Sections 5.1 and 5.2 be $\hat{\Omega}$. We calculate a vector of targeted moments predicted by the model, $m(\hat{\Omega}, \hat{\Theta}_{guess})$, using parameters $\hat{\Omega}$ and labor supply guess $\hat{\Theta}_{guess}$. The algorithm searches over $\hat{\Theta}$ such that it minimizes the distance between the targeted moments predicted by the model and their empirical counterparts as in equation 24:

$$\hat{\Theta}^{*} = \min_{\Theta} \left(m(\hat{\Omega}, \Theta) - m(Data)\right)' W \left(m(\hat{\Omega}, \Theta) - m(Data)\right) \,, \tag{24}$$

where $m(Data)$ are the empirical counterparts of the targeted moments and W is the weighting matrix. $\hat{\Theta}$ is composed of eleven parameters: the taste dispersion parameters σ_k, the mean tastes for non-CS occupations ζ_k, the occupation switching costs $\chi_{0,k}$, $\chi_{1,k}$, the education costs for CS $\bar{F}_k$, and the utility weight on US wages κ.

To separately identify each parameter, we choose eleven moments that are differentially affected by each of the parameters, such that the solution of equation 24 yields parameters that minimize the distance between the simulated and data moments. For both India and US we choose the following

Table 4: Empirical vs. Simulated moments

	US		India	
	Data Moments	Simulated Moments	Data Moments	Simulated Moments
Share CS 1995	2.94%	2.92%	0.12%	0.11%
	[2.71%-3.17%]		[0.06%-0.19%]	
Share CS 2010	3.90%	3.68%	2.76%	2.82%
	[3.66%-4.15%]		[2.28%-3.24%]	
Transition rate 95-00	1.39%	1.20%	0.19%	0.70%
	[0.10%-0.28%]		[0.1%-0.28%]	
Ratio CS Share [25-30]/[31-60] 2010	1.00	1.16	3.96	3.27
	[0.83-1.17]		[2.45-5.47]	
Ratio CS share [45-60]/[31-60] 2010	0.87	0.81	0.19	0.15
	[0.80-0.94]		[0.05-0.33]	
Supply response to migration			0.74	0.76
			[0.25-1.28]	

Simulated method of moments results comparing empirical moments to data moments. 'Supply response to migration' is based on parameter defined in Section 3 and estimated in Table 2. 'Transition rate' is defined as the net occupational switching rate between CS and non-CS occupations. 'Ratio CS Share' is the share of CS workers in age group 25-30 and age group 31-60. 'Share CS' is the share of the college graduate workforce that is in CS. The parameters in Table 5 are simultaneously estimated by matching all these moments across both countries. 95% confidence intervals for empirical moments are in parenthesis.

moments: the share of workers in CS in 1995 and 2010; the ratio between the CS share among those between 25-30 years relative to 31-60 years old in 2010; the net occupation switching rate between 1995 and 2000, and the ratio between the CS aged 45-60 years relative to 31-60 years old in 2010. The final moment we target is the elasticity of occupational choice with respect to migration probability estimated in Section 3. According to our reduced-form estimates, a 1% increase in migrants in a given occupation in the US, causes a 0.74% increase in the labor supply of that occupation in India (Table 2).

In Table D3, we corroborate the intuition on which moments identify each parameter by calculating elasticities of the targeted moments with respect to changes in the supply-side parameters as in Lagakos et al. (2018). While our system uses all empirical moments together to simultaneously identify the parameters, there is strong intuition behind the identification of each parameter. The CS share in 1995 and 2010 help identify both, the mean taste for non-CS ζ_k and the dispersion parameter σ_k. A higher ζ_k will make CS less desirable and lower the average CS share. The change in the CS share between 1995 and 2010 helps identify σ_k, since a higher σ_k means that individuals assign high weights to the idiosyncratic preference shocks they receive every period, making them less responsive to changes in the relative wage. The ratio of the CS share between those aged 25-30 relative to those 31-60 helps identify the education costs $\bar{F}_k$, since a higher ratio implies that it is easier to join CS occupations at the initial period (paying the education cost) than later by switching occupations. The net switching rate will primarily identify the switching costs $\chi_{0,k}$, since high switching costs are expected to decrease switching both in and out of CS. The ratio of the CS share between those aged 45-60 relative to those 31-60 will help identify the age specific component of the switching rate captured by parameter $\chi_{1,k}$. Finally, the employment elasticity with respect to migration will help identify the utility weight parameter κ. The larger the weight on US wages, the larger is the occupational response

Table 5: Labor supply parameter estimates

	US	India
Mean taste for non-CS occupations	0.47	0.45
	[0.42,0.51]	[0.33,0.58]
Preference dispersion parameter	0.46	0.26
	[0.40,0.52]	[0.18,0.32]
Education cost for CS degrees	-0.09	-0.02
	[-0.10,-0.08]	[-0.02,-0.01]
Occupation switching costs	-0.65	-1.09
	[-0.83,-0.48]	[-1.33,-0.85]
Age-specific switching cost	-0.29	-0.01
	[-0.34,-0.24]	[-0.01,-0.005]
Utility weight on US wage		2.72
		[1.85,3.58]

Estimated labor supply parameters based on the simulated method of moments exercise. These parameters jointly determine the short-run and long-run labor supply elasticities. 'Mean taste for non-CS' is the average preference for non-CS occupations. 'Preference dispersion' is the variance in tastes for occupations. 'Switching costs' are occupation switching costs between CS and non-CS occupations. 'Education cost for CS' is the average cost / non-preference for CS degrees. All costs and tastes in units of the numeraire (consumption basket). 'Utility weight on US wage' is the non-pecuniary premium on US wages (relative to India) for Indian workers. Confidence intervals for parameters are the 5th and the 95th percentiles of the bootstrapped (with replacement) sample.

to increased migration probabilities.

We construct the eleven data moments using the CPS and the NSS for the US and India, respectively. We consider the CS share as the total number of CS workers relative to the total number of college graduates in each country, excluding the Indian migrants from both India and US shares. For the switching rate, we use net flows by cohort, calculated by adding the absolute value of net flows across each cohort and taking the ratio with respect to the sum of the net stayers across each cohort. Finally, for the employment elasticity with respect to migration, we use the estimated elasticity of 0.74 from Section 3, Table 2.

The equation 24 routine minimizes the distance between simulated and empirical moments. The model is perfectly identified (eleven data moments to identify eleven parameters). Table 4 shows the predicted moments match the empirical data moments closely.

The estimated labor supply parameters are in Table 5. As expected, there is a positive mean taste for non-CS occupations in both countries which explains the positive CS wage premium in both countries. Switching costs in India are higher than in the US, indicating that occupational choices are less responsive to wage changes in India. Age specific switching costs, however, are higher in the US. A negative education cost implies that it is costlier to study CS than other majors as fewer students enroll in CS despite higher wages. Finally, the utility weight on US wage parameter is larger than one, which means that each dollar earned in the US provides additional non-pecuniary utility with respect to a dollar earned in India.

Finally, we estimate the expectation parameters in equation 15, $\phi_0^{x,k}$ and $\phi_1^{x,k}$, where k represents country and x stands for each of the college-graduate wages $\{w_k^g, w_k^{cs}\}$ and the migration cap $\{cap\}$. Since workers have rational expectations for each x this implies there exists a fixed point, in which the wage and cap expectations workers have, when making occupation decisions, is on average the same as the equilibrium outcomes. Hence, parameters $\phi_0^{x,k}$ and $\phi_1^{x,k}$ are estimated as a function of the rest of the parameters of the model.

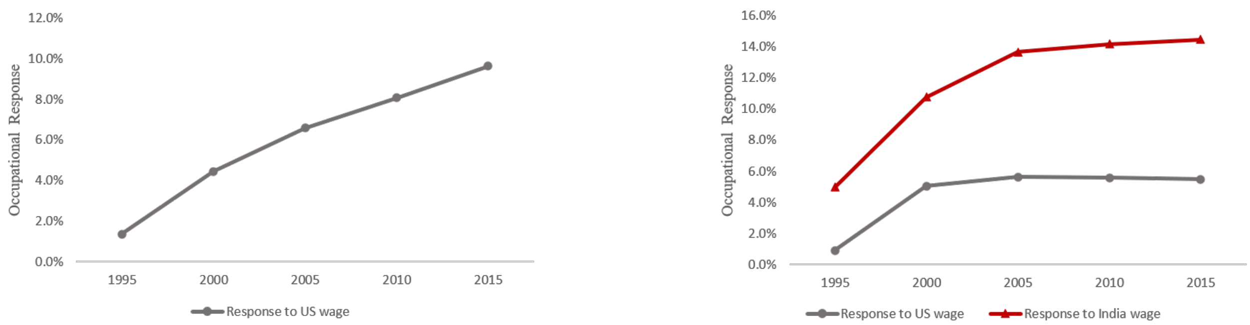
Together, the parameters estimated in this section determine the dynamic labor supply elasticity in each country. We represent this elasticity in Figure 5, where we artificially raise the CS wage in the US by 1% starting in 1995, and estimate the labor supply response over time. We increase the Indian wage by the same magnitude as the 1% increase in the US wage. In the short run, the labor supply curve is inelastic, as few shift into CS. But over the longer run, as more and more students major in CS, and more workers switch into CS, the labor supply curve becomes elastic. Estimating this dynamic labor supply elasticity is an important contribution of our work: previous reduced-form estimates in some of the literature are often estimating a static elasticity, and as such, providing a limited picture of the long run labor-supply response.

Short-run labor supply elasticity in the US is 1.37 while in the long-run, we estimate the US occupational elasticity to be 9.63. Ryoo and Rosen (2004) estimate the supply of engineers relative to other graduates to be between 2.5 and 4.5, within our short and long term elasticities. Traiberman (2019) finds that a permanent decline of 1% in an occupation skill price, leads to a 17% higher exit rate from that occupation. We would expect our elasticity to be higher than the one for engineers but lower than when looking at all occupations. As shown in Figure 5, Indian college graduates respond to both changes in the US and Indian wages. Overall labor supply in India is more elastic than in the US, partially due to a lower preference dispersion parameter in India, which makes Indians more sensitive to wage changes.

Figure 5: Dynamic occupation choice in response to wages: permanent increase in CS wage

(a) United States

(b) India



Graphs show the dynamic labor supply response to a 1% permanent increase in the US CS wages since 1995. The short- and long-run response depend on the labor supply parameters estimated in Table 5. For the response to changes in the India wage, instead of changing the wage by 1%, we change it by the same magnitude as the US wage change to make it comparable.

In order to evaluate the fit of our model, we compare our simulated results with features from the data as out-of-sample tests. In the estimation exercise, we explicitly match certain data points or trends, whereas here we discuss how well our model matches the data on items we do not explicitly use to discipline the model. In Appendix G, we show that we match fairly well some of the key aspects that we are trying to capture, such as the relative wages across countries and the observed changes in IT prices and output.

6 Counterfactual Exercises

Policy makers often debate changing the H-1B cap, and recent political debates focus on restricting the number of H-1B visas. To evaluate the impact of the H-1B program on US and Indian economies, we conduct counterfactual exercises where we first reduce the migration cap by 50% every year since 1995.[21] In Appendix F.5, we examine extensions where we vary the cap across different sizes, and start the cap restrictions in different time periods. With a lower cap, some workers who may have been granted visas to the US are now forced to work in India. We look at the changes in going from the counterfactual scenario with limited migration (50% of the cap) to the real scenario where migration is as observed in the data. Hence, we refer to the impacts as a result of 'increased migration,' or simply migration under the H-1B program.

Our attempt is to empirically resolve theoretical ambiguities. Some of these ambiguities include whether the effects of brain drain outweigh brain gain, whether wage gains due to innovation overcome wage depression due to an influx of workers, and whether immigration allows the US IT sector to grow or instead facilitates the shift in production to India.

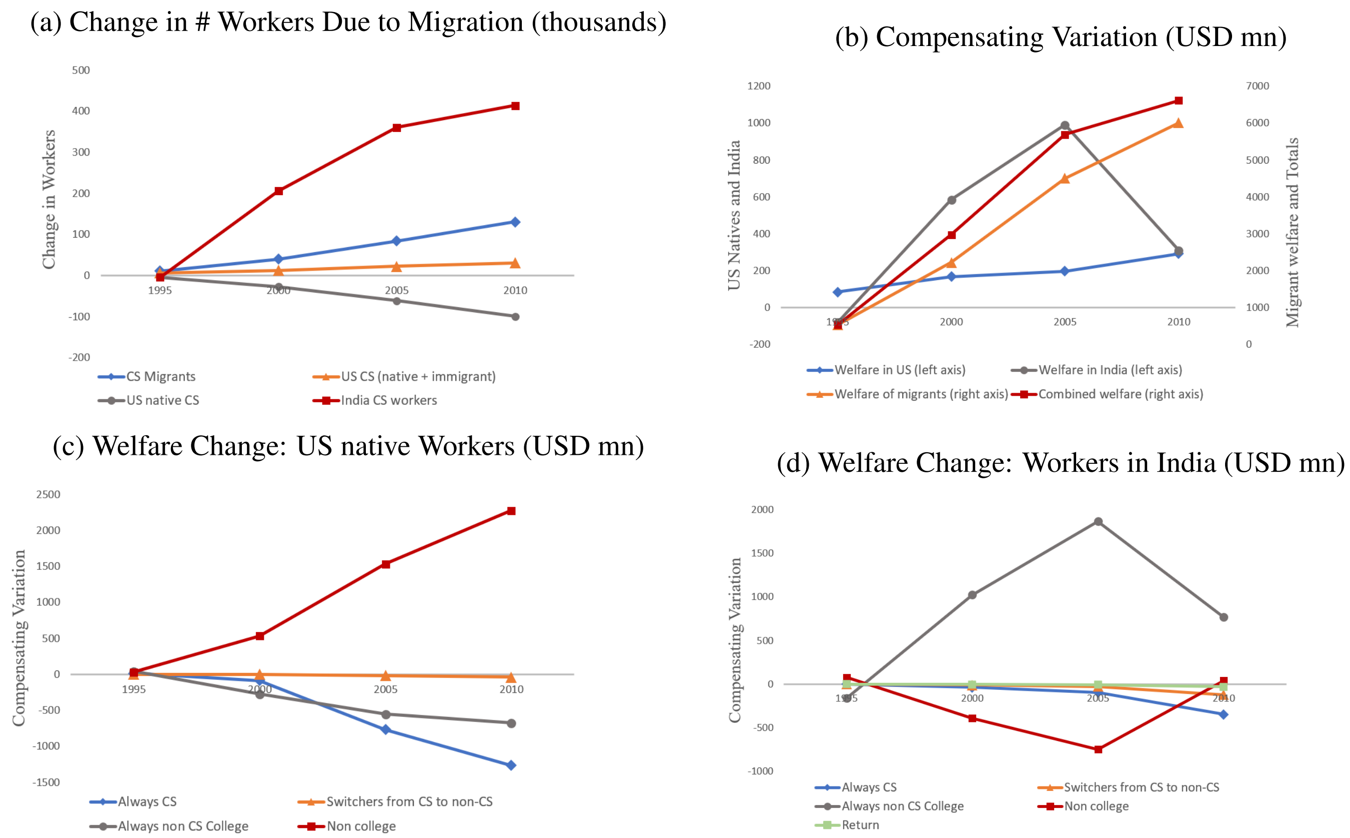
6.1 Baseline results

As we show in Figure 6a, when going from the counterfactual to the real scenario, the total number of migrants (mechanically) increases by almost 11,000 in 1995 and by over 130,000 in 2010. The increase in migration of CS drives down CS wages in the US, leading more than 99,000 US-born CS to switch to non-CS occupations. In India, college graduates respond to the migration incentives by switching to CS occupations. While some migrate many others stay in India, increasing the number of CS in India by more than 400,000 workers by 2010.

Both the US and India increase the total number of CS, an occupation complementary to non-CS college graduates and high-school graduates. In addition, CS workers have a spillover effect that benefits all workers, making total net welfare gains positive for both countries, as shown in Figure 6b. Welfare for US natives increases by $292 million by 2010 while Indian natives who stay in India gain $310 million when migration to the US increases. In Appendix Table E1, we summarize the distributional gains and losses per migrant for each group of workers. In 2010, there is a net welfare gain to US workers of $2,227 per migrant, and of $2,364 per migrant to workers in India. If we also include gains to migrants themselves, overall net welfare increases by $50,400 per migrant. The hump-shaped pattern for India's compensating variation shown in Figure 6b can be explained by changes in magnitudes of two opposing effects. In the earlier periods, Indians motivated by migration start selecting into CS and working in India, increasing productivity domestically. As time goes by, getting a lottery slot gets harder given the large numbers of local CS, which makes migration play less of a role on welfare.

[21] In our model we only allow for migration of Indian workers to the US. Hence, we calibrate the cap with the number of Indians that came to the US under the H-1B program every period.

Figure 6: Effect of immigration on occupational choice and welfare



Graphs show the consequences of increase migration. The top left panel shows 'CS Migrants' as the difference in migration between the real and the counterfactual (restricted lower cap) scenario. The occupational choices in the top left panel are responses to this increased migration. The remaining panels show the compensating variation for agents due to a restriction in migration. Compensating variation is defined as the amount of USD that must be provided to agents in a world with restricted migration to provide them with the same welfare as in a world with H-1B migration.

Even though natives from both countries benefit from migration, there are distributional gains and losses, which we describe in Figures 6c-6d and Appendix Table E1. Since the total number of CS in both countries increases, those already working in CS lose from migration (despite gains from innovation) as their wages decline with increased competition. The effect on those who are non-CS college graduates differs by country. In the US, as the number of migrants increases, some college graduates working in CS switch to non-CS occupations, lowering the non-CS college graduate wage (despite the gains from complementarity) and hence, their overall welfare. In India, the opposite happens: some college graduates switch into CS, leaving non-CS with fewer workers and increasing their wages and welfare. Non-college graduates in the US benefit from migration as prices are lower, and they are complements in production. In India, non-college graduates face two counteracting effects: They gain from the additional productivity from local CS but lose since the total number of graduates decreases in India (due to emigration). Overall they lose in the earlier periods and gain in the last period when the stock of CS in India is sufficiently high.

Accounting for the occupational response of Indian college graduates to changes in migration policy in the US is crucial to quantifying the gains and losses from migration. To see this, in Table 6 we compare the results between our baseline model (column 1), and alternative models that shut down the supply response in India. Column 2 shows the results for a model where individuals in both

Table 6: Effect of Migration in 2010: Baseline vs No Occupational Choice

	Baseline	No occupational choice	
		In both countries	**In India only**
Wages			
US CS workers	-0.97%	-3.28%	-0.80%
India CS workers	-11.5%	0.60%	0.9%
Occupational Choice			
US CS (native plus immigrant)	1.65%	7.50%	1.95%
US CS workers	-5.87%	-	-5.56%
India CS workers	56.6%	-	-
IT production			
US IT output	0.11%	4.20%	1.30%
India IT output	35.55%	-10.2%	-9.80%
Welfare			
Welfare of US natives	0.007%	0.04%	0.009%
Welfare in India	0.13%	-0.05%	-0.05%

Percent difference in main outcomes when we go from a scenario with restricted migration (counterfactual) to a scenario with full migration (real). In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995. 'No occupational choice in both countries' is a model where occupational choice is prohibited. 'No occupational choice in India only' is a model where occupational choice is only prevented in India.

countries cannot switch occupations, thus ignoring endogenous occupational choice in each country. Finally, column 3 shows the results for an alternative model where workers in the US are allowed to switch occupations but workers in India are not (and thereby, cannot respond to the increased prospect of migrating to higher wages abroad).

In our baseline model, 5.87% of US graduates in CS switch to other occupations when migration is increased, driven by a -0.97% reduction in the CS wage. In India, the total supply of CS increases (despite emigration) when migration is allowed. This increases the size of the Indian IT sector by 35.55%. The US IT sector increases by only 0.11% as natives switch away from CS, and there is increased competition from India for the world market. Welfare is higher by 0.007% for US natives, and by 0.13% for those working in India. The combined increase in welfare, including the welfare of migrants, is higher by 0.14% (Appendix Table E2).

In an analysis that did not allow for endogenous occupation choice in India (column 3), India would only experience brain drain as their CS leave the country, lowering IT output by 9.80%. As a consequence, the increase in total CS in the US would increase US IT output by 1.30%, a larger effect than in the baseline model, as the US captures the market that was satisfied by India. A model that ignores the supply response only in India has significant consequences for welfare. The total welfare of US natives due to immigration is 30% higher than in the baseline, while in India welfare *decreases* by 0.05% as there cannot be brain gain when supply is fixed. When restricting occupational choice in both countries (column 2), total welfare in the US is six times higher than in the baseline, as US CS cannot change occupations and the US captures all the spillover created by CS innovation. The results in Table 6 highlight the importance of the main mechanism introduced in this paper: taking into account supply responses in sending countries would predict an opposite effect for the sending countries like India, and drastically mute the welfare effects of migration for the US.

Appendix Table E2 shows how different variables change over time across scenarios. In Appendix Table E3, we study the sensitivity to varying the four elasticities we take from the literature (λ, τ, θ and ε). Our results are qualitatively similar to the baseline.

7 Mechanisms and Alternative Specifications

Our model is comprehensive enough to capture the main channels through which migration affects both India and the US. We run several extensions to our model in order to tease out the mechanisms underlying our base specification, each time recalibrating the entire model to match the moments. First, in Section 7.1, we separate out the mechanisms underlying brain drain and brain gain. We additionally evaluate how our quantitative results change when we do not include the possibility of return migration. In Section 7.2, we explore the role played by endogenous technology by varying the technology elasticity, and in Section 7.3, we look into the possibility of migrant workers in the US being imperfect substitutes for native CS. Section 7.4 incorporates an endogenous college decision where workers choose to either go to college or join the labor market as high-school graduates. Section 7.5 discusses an alternative model that has heterogeneity in abilities and positive selection into migration. Finally, in Appendix F, we discuss alternative models that include lower trade costs, remittances, different migration caps, and starting immigration restrictions in later periods.

7.1 Brain Drain, Brain Gain and Brain Circulation

The baseline results in Section 6.1 include a combination of brain drain and brain gain for India. On the one hand, some CS leave for the US, driving brain drain. On the other, some graduates choose CS with the prospects of migrating but end up staying, generating a spillover effect and brain gain. To disentangle such effects, we compare the baseline with two alternative scenarios. First, we shut down India's occupational response to migration opportunities. While workers are still allowed to migrate, they do not take into account the possibility of migrating when choosing occupations, such that India experiences no brain gain. As shown in Table 7 column 2, if Indians do not respond to migration incentives, the CS workforce in India decreases by 10.21% since some CS migrate and not many join the workforce. This makes the IT sector in India *shrink* when migration to the US increases, and the US IT sector increases by 1.30%. Net welfare gains from migration are larger in the US, but welfare falls in India, as India only experiences brain drain, and no brain gain, when migration is increased.

As a second exercise in Table 7 column 3, we remove the possibility of brain drain by shutting down migration but still allowing workers in India to choose occupations in response to the prospect of migrating (i.e. allowing brain gain). The CS workforce in India grows rapidly, and IT output expands by 36.97%. Such increases makes the US IT sector shrink by 0.79%. When there is no brain drain, total welfare in India increases by 0.15% while decreasing in the US by 0.002%. In Appendix Table E4 we show how the brain drain and brain gain channels affect each of the different types of workers in the US and India.

Table 7: Brain Drain vs Brain Gain: Main Outcomes

	Baseline	No occupational choice in India	No migration but reallocation	No return migration
Occupational Choice				
US CS (native plus immigrant)	1.65%	1.95%	-0.26%	2.51%
US CS native	-5.87%	-5.56%	-0.27%	-16.80%
India CS	56.64%	-10.21%	64.14%	40.97%
IT production				
US IT output	0.11%	1.30%	-0.79%	0.62%
India IT output	35.55%	-9.80%	36.97%	28.46%
Total Welfare				
Welfare of US natives	0.007%	0.009%	-0.002%	0.012%
Welfare in India	0.130%	-0.047%	0.149%	0.106%

Percent difference in main outcomes when we go from a scenario with restricted migration (counterfactual) to a scenario with full migration (real). In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995. The 'no occupational choice in India' scenario restricts occupational choice and inhibits 'brain gain' in response to migration opportunities but allows for emigration ('brain drain'). 'No migration but reallocation' restricts migration (no 'brain drain') but allows for responses to migration opportunities ('brain gain'). 'No return migration' sets the return rate ρ to zero.

Last, we examine the importance of return migration. In our modification, Indian CS migrants stay in the US until the end of their careers. As shown in Table 7 column 4, the US experiences larger welfare gains when there is no return migration. This is driven by two channels. First, the stock of migrants in the US increases more as they no longer return to India. The additional CS create a larger innovation spillover benefiting all US workers. Second, India does not have the additional productivity effect provided by return migrants, which stymies the shift in IT production from the US to India. India experiences lower welfare increases than in the baseline, as now they have only one type of (never-emigrated) CS in the country.

7.2 Endogenous Technology and Innovation

The endogenous technology elasticity that we estimate in Section D.2 using an instrumental variables strategy has a value of 0.226, and is consistent with other values in the literature (Kerr and Lincoln, 2010; Khanna and Lee, 2018; Peri et al., 2015). In Table F1, we show our main results as we vary the parameter value all the way down to 0 (no endogenous technology). We also try a specification where technological spillovers are positive in the US and nonexistent in India, in order to capture the fact that the US may do more R&D than India.

Migration affects US welfare in similar magnitudes as we vary this parameter. Yet, India is adversely affected when there are no innovation spillovers from CS workers (column 2). In such a scenario, many Indian graduates switched to CS with the prospect of migrating, but were unable to migrate due to the cap. There is, therefore, a suboptimal oversupply of CS workers, but no gains from innovation, lowering welfare. For the same reason, when spillovers are lower (columns 2-4) than the baseline scenario, there is less of a shift in IT production from the US to India under H-1B migration, as CS in both countries contribute less to innovation.

The technology spillover in India is key for our result that India experiences a brain gain. As noted by Stark (2004), when productivity depends on the average level of human capital in the economy, individual workers tend to under invest in skills. The possibility of migrating to the US, motivates Indians to acquire CS skills and create a larger spillover that increase IT production and welfare in India. From the innovation literature (Byrne et al., 2013; Jorgenson et al., 2016) we expect a positive spillover generated by the IT sector. Column 4 shows that even with a small level of spillover, India experiences positive welfare gains from migration.

7.3 Imperfect Substitution Between Immigrants and Natives

The literature on migration has considered that migrants may have different skills than natives, and so instead of being perfect substitutes in production, there may be some degree of complementarity. Peri and Sparber (2011) shows that among high-skill workers, migrants tend to select into more quantitative tasks while natives relocate towards communication-intensive tasks. As we are looking at natives and migrants within the narrow and specific occupation of computer science, the assumption of perfect substitution is more reasonable than if we were looking at college graduates as a whole. Nevertheless, we explore how our results change if we consider natives and migrants to be imperfect substitutes. Details on the implementation of this alternative modeling specification can be found in Appendix F.2. As shown in Table F2 column 2, US welfare gains from increasing migration would be higher (due to worker complementarities) than the baseline of perfect substitutes. The US increases IT production more than in the baseline model while India's IT sector increases sightly less than in the baseline.

7.4 Endogenous College Decision

So far we consider the pool of college graduates to be the same between real and counterfactual scenarios. We relax that assumption by allowing high-school graduates to choose whether to join the labor force without going to college, or go to college and join the labor force in the next period. The details of this implementation can be found on Appendix F.3. As shown in Table F2, when increasing immigration, India reduces their high-school graduates by 0.18% since workers are motivated by immigration prospects to get into college. The US on the other hand, increases their high-school graduates by 0.03% since some natives decide not to go to college given the lower wages for college graduates when high-skilled immigration is larger. In terms of welfare however, our results are quantitatively almost identical to the baseline model with slightly larger gains for both countries under the endogenous college specification.

7.5 Heterogeneity in Abilities

Next we allow for heterogeneity in abilities and positive selection into migration. If more able individuals are more likely to migrate, we would expect welfare and IT production to increase more than in our baseline in the US, and a more significant brain drain effect on India.

We assume college graduates in the US and India draw an ability for each occupation, $\phi_{cs,i}$, $\phi_{g,i}$ from independent Normal distributions with mean 0 (a normalization) and variance $\sigma_{\phi,cs,k}$ and

$\sigma_{\phi,g,k}$.[22] The wage that worker i receives for working in occupations o is:

$$w_{i,t,o,k} = exp(\phi_{o,i}) \times w_{t,o,k} \ , \tag{25}$$

where $exp(\phi_{o,i})$ is a parametrization of worker i's human capital in occupation o, and $w_{t,o,k}$ is the wage per effective unit paid for occupation o in country k at time t. This heterogeneity extension is a simplified version of the labor market presented in Dix-Carneiro (2014). To identify the additional parameters $\sigma_{\phi,cs,k}$ and $\sigma_{\phi,g,k}$ for both countries, we include them in the joint estimation (of Section 5.3) by further matching the wage dispersion for CS and non-CS college graduates in the US and India, as four additional moments in our Simulated Method of Moments (SMM) procedure.

The ability draws are constant throughout individual careers while idiosyncratic preference shocks are drawn every period. Hence, heterogeneity in abilities will act as an additional friction, reducing occupation switching over time, while heterogeneity in preferences will induce more two-way switching after individuals take a new draw every period. The precise empirical moments we additionally use are the wage dispersion within each country-occupation pair, relative to the mean wage for that country-occupation pair. Finally, we assume migration into the US is more likely if the ability draw in CS is higher, as in equation 26:

$$p_{t+1,i} = \bar{p}_{t+1} \times \frac{cap_{t+1}}{CS_{t,in}} \times F(\phi_{cs,i}|\sigma_{\phi,cs,in}) \ , \tag{26}$$

where the baseline probability of migration from equation 13, $\frac{cap_{t+1}}{CS_{t,in}}$ is multiplied by the Normal distribution function $F(\phi_{cs,i}|\sigma_{\phi,cs,in})$ evaluated at individual i's ability draw in CS, and re-weighted by parameter $\bar{p}_{t+1}$. The sum of all the probabilities of migration among those who choose CS adds up to the total number allowed by the migration cap at time $t+1$.[23]

The model is overall very similar to the baseline model presented in Section 4.1, but computationally intensive, as we now simulate the occupational choices for each individual, as each worker has different ability draws and continuation values. To simplify the computation of the equilibrium and estimation, we assume individuals have perfect foresight on the future migration cap and wages instead of rational expectations. We also re-estimate the baseline model with perfect foresight to compare it to the heterogeneous model. Overall, the quantitative results do not change much when considering perfect foresight instead of rational expectations.

As shown in Table 8, results are qualitatively similar to the case with homogeneous abilities but there are some interesting quantitative nuances. With heterogeneous abilities, the US IT sector grows more than in the baseline, as we increase migration. This is partly driven by the most able CS immigrating, which drives IT production in the US. Welfare for US natives is higher than in the baseline, as the US receives more effective units of CS, generating a larger spillover. India's welfare

[22] We normalize the mean ability to 0 as it cannot be separately identified from the taste parameters η_k.

[23] Intuitively, $F(\phi_{cs,i})$ makes more able individuals have a higher probability of migration. $\bar{p}_{t+1}$ is needed such that when we add all probabilities of migration across individuals we get the current migration cap, cap_{t+1}.

Table 8: Heterogeneity in Abilities

	Baseline		Heterogeneity in Abilities	
	US	India	US	India
IT output	0.24%	36.22%	1.83%	13.79%
Welfare				
Always CS	-1.14%	-12.93%	-2.34%	-3.40%
Switchers	-0.58%	-11.73%	-0.01%	-5.57%
Always non-CS graduates	-0.02%	0.93%	0.06%	0.46%
Non college	0.12%	0.02%	0.14%	-0.04%
Total Natives	0.005%	0.13%	0.08%	0.06%
Migrants	41.43%		50.14%	

Percent difference on main outcomes when we go from a scenario with restricted migration (counterfactual) to a scenario with full migration (real). In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995. For computational reasons we estimate the baseline and the model with heterogeneous abilities assuming perfect foresight instead of rational expectations. The model with heterogeneous abilities considers workers to have different innate abilities in CS and other college occupations, which affects their occupational choices.

increases by half of what it does in the baseline, as it loses the most skilled CS while those who switch into CS, but stay in India are relatively less productive.

7.6 Additional Model Specifications

In Appendix F we evaluate other specifications. First, in Section F.4 we show that in a world with lower trade costs, welfare gains from immigration are lower for both countries. Trade mitigates some of the negative impacts of migration restrictions. Incorporating remittances, or not having age-dependent switching cost ($\chi_1 = 0$) does not change our quantitative results.

Second, in Section F.5 we compute welfare changes for different cap changes. The changes in welfare are roughly proportional to the magnitude of the cap change. Finally, we look at how welfare changes depend on what period the immigration restrictions start, and whether or not the Indian IT sector is already developed.

8 Discussion

India experienced a dramatic expansion in IT employment and a structural transformation in production over the 1990s and 2000s. Many factors contributed to this boom but our work suggests that, surprisingly, policies from halfway around the world played a critical role. We study how US immigration policy, combined with the US tech boom, enabled the IT boom in India. The prospect of high wages in the US incentivized students and workers in India to choose CS degrees and occupations. Those returning from the US after the expiration of their H-1Bs also contributed to the growing Indian workforce. These movements increased overall IT productivity in India and shifted the production of IT goods away from the US.

We explicitly test the explanatory power of certain conditions under which US policy stimulated growth in Indian IT. We do this by specifically focusing on four features over this period that created important incentives and constraints for Indian students and workers. First, technological innovations and changing consumer preferences generated strong demand for IT workers in the US. Second, the wage differential between the US and India was large, especially for IT workers. Third, US immigration policy, as embodied by the H-1B program, strongly favored skilled migrants. Finally, H-1B visas only last three to six years, obligating many to return to India with accumulated human capital and technical knowhow. Together, these features helped spread the boom across the world from the US to India.

The average worker in each country is better off because of migration. Yet, there are significant distributional consequences, where workers who are close substitutes are adversely affected while others benefit. These distributional effects have been at the forefront of political and academic discussion (Borjas, 1999; Peri and Sparber, 2011). We find that the overall gains outweigh the losses as the combined incomes of the US and India rise under the H-1B program by 0.14%. This net gain is consistent with a literature reviewed in Clemens (2011).[24] The welfare gains are approximately \$6.6 billion in total, a large fraction of which accrues to the migrants. US natives were better off by \$292 million in 2010 because of the H-1B program.

The gains are mostly driven by the development of the Indian IT sector. In a world with North-South trade, developing countries may specialize in less productive sectors, hindering economic growth (Matsuyama, 1992). Contrarily, we find that US immigration policy, coupled with the US tech boom, helped develop the Indian IT sector, boosting IT exports and raising average incomes. The prospect of migrating to the US was a considerable driver of this phenomenon and led to a 'brain gain' that outweighed the negative impacts of 'brain drain' (Dinkleman and Mariotti, 2016; Stark, 2004; Stark et al., 1997).

One striking result is that as production shifts (or is outsourced) to India, US IT market share actually falls. A driving feature of this is that increases in the Indian CS workforce increases the relative productivity of Indian IT. Such dynamics are discussed a rich literature on trade. In Krugman (1979) and Vernon (1966), richer countries initially have a monopoly over new products given their technological superiority and rate of innovation. Developing economies catch up with technological diffusion and over time export these very products to the developed world. As the rate of technological diffusion increases, living standards may actually fall in richer countries. With quality differentiation in products, technical progress in the South brings a decline in the North's wages, harming workers as production moves abroad (Acemoglu et al., 2015; Flam and Helpman, 1987). Therefore, as Samuelson (2004) notes, technical progress in the South erodes the US comparative advantage and lowers US incomes.

The labor literature also emphasizes these channels. For instance, Johnson and Stafford (1993) show how the effect of foreign competition from abroad lowers aggregate incomes in the US. In fact, Freeman (2006a) argues that the growth in high-tech labor abroad adversely affects US industry and

[24] Specifically, see Iregui (2005); Klein and Ventura (2007); Moses and Letnes (2004); van der Mensbrugghe and Roland-Holst (2009); Walmsley and Winters (2005).

workers, and immigration can help maintain the US's lead by attracting overseas talent. However, this literature does not account for 'brain gain' or return migration, which we show have important consequences. Davis and Weinstein (2002) show how in a Ricardian framework, such as ours, a country that experiences immigration due to technological superiority loses from such migration through a deterioration in the terms of trade (for us, the US's terms of trade deteriorate as immigration lowers the IT price). Mobility tends to equalize wages across countries and, therefore, hurt workers in the country with superior technology.

In this way, our results confirm canonical predictions in the literature. Even though migration increases the welfare of the average US and Indian worker, these averages hide significant distributional changes in each country. Academic discourse that ignores endogenous skill acquisition (in response to migration opportunities) in sending countries, trade, innovation, dynamic labor supply decisions, and price changes will miss important aspects of this discussion.

References

Abarcar, P. and C. Theoharides (2020). Medical Worker Migration and Origin-Country Human Capital: Evidence from US Visa Policy. *Working Paper*.

Acemoglu, D. (1998). Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality. *Quarterly Journal of Economics 113*(4), 1055–1089.

Acemoglu, D., G. Gancia, and F. Zilibotti (2015, July). Offshoring and Directed Technical Change. *American Economic Journal: Macroeconomics 7*(3).

Adao, R., M. Kolesar, and E. Morales (2019). Shift-Share Designs: Theory and Inference. *The Quarterly Journal of Economics 134*(4), 1949–2010.

Agrawal, A., D. Kapur, J. McHale, and A. Oettl (2011). Brain Drain or Brain Bank? the Impact of Skilled Emigration on Poor-country Innovation. *Journal of Urban Economics 69*(1), 43–55.

Ahlfeldt, G. M., S. K. Redding, D. M. Sturm, and N. Wolf (2015). The Economics of Density: Evidence from the Berlin Wall. *Econometrica 83*(6), 2127–2189.

Allen, T., C. Arkolakis, and Y. Takahashi (2020). Universal Gravity. *Journal of Political Economy 128*(2), 393–433.

Alvarez, F. E., F. J. Buera, and R. E. Lucas (2013, July). Idea Flows, Economic Growth, and Trade. *NBER Working Paper Series, No. 19667 7*(3).

Arora, A. and S. Athreye (2002). The Software Industry and India's Economic Development. *Information Economics and Policy 14*(2), 253–273. Social Science Research Network.

Arora, A., A. Gambardella, and S. Torrisi (2001). In the Footsteps of the Silicon Valley? Indian and Irish Software in the International Division of Labour. *Building High-Tech Clusters*, 78–120. (eds. Timothy Bresnehan and Alfonso Gambardella).

Athreye, S. S. (2005). The Indian Software Industry and its Evolving Service Capability. *Industrial and Corporate Change Advance Access 14*(3), 393–418.

Azam, M., A. Chin, and N. Prakash (2013). The Returns to English-Language Skills in India. *Economic Development and Cultural Change 61*(2), 335–367.

Bahar, D. and H. Rapoport (2018). Migration, knowledge diffusion and the comparative advantage of nations. *The Economic Journal 128*(612), F273–F305.

Banerjee, A. and E. Duflo (2000). Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry. *The Quarterly Journal of Economics 115*(3), 989–1017.

Barrett, A. and P. O'Connell (2001). Is There a Wage Premium for Returning Irish Migrants? *The Economic and Social Review 32*(1), 1–21.

Beine, M., F. Docquier, and H. Rapoport (2001, February). Brain Drain and Economic Growth: Theory and Evidence. *Journal of Development Economics 64*(1), 275–289.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003, September). Plants and Productivity in International Trade. *The American Economic Review 93*(4), 1268–1290.

Bhatnagar, S. (2006). India's Software Industry. *Technology, Adaptation and Exports: How Some Developing Countries Got it Right. World Bank.*, 95–124. (Eds. Vandana Chandra), World Bank, Washington DC.

BLS (1996). Abundant career opportunities projected in information technology on the internet. *Bureau of Labor Statistics, U.S. Department of Labor, Employment Projections*.

Borjas, G. J. (1999). The Economic Analysis of Immigration. *Handbook of Labor Economics 3*(28). (eds Orley Ashenfelter and David Card), North-Holland.

Borusyak, K., P. Hull, and X. Jaravel (2020). Quasi-Experimental Shift-Share Research Designs. *Working Paper*.

Bound, J., B. Braga, J. Golden, and G. Khanna (2015). Recruitment of Foreigners in the Market for Computer Scientists in the US. *Journal of Labor Economics 33*(S1), 187–223.

Bound, J., M. Demirci, G. Khanna, and S. Turner (2014). Finishing Degrees and Finding Jobs: U.S. Higher Education and the Flow of Foreign IT Workers. *Innovation Policy and the Economy 15*(1), 27–72.

Bound, J., G. Khanna, and N. Morales (2016). Understanding the Economic Impact of the H-1B Program on the US. *High-Skilled Migration to the United States and Its Economic Consequences*. (eds. G Hanson, W Kerr, and S Turner), University of Chicago Press.

Burstein, A., G. Hanson, L. Tian, and J. Vogel (2020, May). Tradability and the labor-market impact of immigration: Theory and evidence from the u.s. *Econometrica 88*(3), 1071–1112.

Burstein, A., E. Morales, and J. Vogel (2019). Changes in Between-Group Inequality: Computers, Occupations and International Trade. *American Economi Journal: Macroeconomis*.

Byrne, D. M., S. D. Oliner, and D. E. Sichel (2013). Is the Information Technology Revolution Over? *International Productivity Monitor. 25*, 20–36.

Caliendo, L., L. Opromolla, F. Parro, and A. Sforza (2020). Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement. *Working Paper*.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics 19*(1), 22–64.

Card, D. and T. Lemieux (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *Quarterly Journal of Economics 116*(2), 705–746.

Clemens, M. A. (2011, Summer). Economics and Emigration: Trillion-Dollar Bills on the Sidewalk. *Journal of Economic Perspectives 25*(3), 83–106.

Clemens, M. A. (2013). Why Do Programmers Earn More in Houston than Hyderabad? Evidence from Randomized Processing of U.S. Visas. *American Economic Review Papers & Proceedings 103*(3), 198–202.

Colas, M. (2019). Dynamic Responses to Immigration. *Working Paper*.

Computerworld (2015). With H-1B visa, Diversity Doesn't Apply: In Computer Occupations, India Dominates H-1B Visa Use. *By Patrick Thibodeau and Sharon Machlis*.

Dasgupta, K. (2012, July). Learning and Knowledge Diffusion in a Global Economy. *Journal of International Economics 87*(2), 323–336.

Dataquest (2003). The Made-in-India Brigade. *Cyber Media (India) Limited, Gurgaon 20*(15).

Davis, D. R. and D. E. Weinstein (2002, May). Technological Superiority and the Losses from Migration. *NBER WP No. 8971*.

Desai, A. (2003). The dynamics of the Indian information technology industry. *DRC Working Paper No. 20, Centre for New and Emerging Markets*. London Business School.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018, June). The Geography of Development. *Journal of Political Economy 126*(3), 903–983.

di Giovanni, J., A. Levchenko, and F. Ortega (2015, February). A Global View of Cross-Border Migration. *Journal of the European Economic Association 13*(1), 168–202.

Dinkleman, T. and M. Mariotti (2016). The Long Run Effect of Labor Migration on Human Capital Formation in Communities of Origin. *American Economi Journal: Applied Economics 8*(4), 1–36.

Dix-Carneiro, R. (2014). Trade liberalization and labor market dynamics. *Econometrica 82*(3).

Doran, K. B., A. Gelber, and A. Isen (2017). The Effects of High-Skill Immigration Policy on Firms: Evidence from Visa Lotteries. *NBER*. Working Paper 20668.

Easterly, W. and Y. Nyarko (2009). Is the Brain Drain Good for Africa? *Skilled Immigration Today: Prospects, Problems, and Policies*, 1741–1779. edited by Bhagwati et al., Oxford University Press.

Eaton, J. and S. Kortum (2002). Technology, Geography and Trade. *Econometrica 70*, 1741–79.

Flam, H. and E. Helpman (1987, December). Vertical Product Differentiation and North-South Trade. *The American Economic Review 77*(5), 810–822.

Freeman, R. (2006a). Does Globalization of the Scientific/Engineering Workforce Threaten U.S. Economic Leadership? *Innovation Policy and the Economy 6*.

Freeman, R. (2006b). People Flows in Globalization. *Journal of Economic Perspectives 20*(2), 145–170.

Goldin, C. and L. F. Katz (2007). Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing. *Brookings Papers on Economic Activity* (2), 135–65.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Hazans, M. (2008). Effect of Foreign Experience on Return Migrants' Earnings: Causality and Heterogeneity. *Working Paper*.

Heeks, R. (1995). From Regulation to Promotion: The State's Changing but Continuing Role in Software Production and Export. *Development Policy and Practice*. Development Policy and Practice Research Group, The Open University, UK.

Hunt, J. and M. Gauthier-Loiselle (2010). How Much Does Immigration Boost Innovation? *American Economic Journal: Macroeconomics 2*(2), 31–56.

Indian Express (2020). 20 years on, where are the Board toppers? Over half are abroad, most in science and technology. *Indian Express Group, Mumbai*. Ritika Chopra.

Iregui, A. M. (2005). Efficiancy Gains from the Elimination of Global Restrictions on Labour Mobility. *Poverty, International Migration and Asylum*, 211–258. (ed. Borjas and Crisp).

Jaeger, D. A., J. Ruist, and J. Stuhler (2018). Shift-share instruments and the impact of immigration. *National Bureau of Economic Research, WP 24285*.

Jagnani, M. and G. Khanna (2020). The Effects of Elite Public Colleges on Primary and Secondary Schooling Markets in India. *Journal of Development Economics*. Forthcoming.

Johnson, G. E. and F. P. Stafford (1993, May). International Competition and Real Wages. *The American Economic Review 83*(2), 127–130.

Jorgenson, D. W., M. S. Ho, and J. D. Samuels (2016). Educational Attainment and the Revival of US Economic Growth. *NBER*. Working Paper No. 22453.

Katz, L. F. and K. M. Murphy (1992, February). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics 107*(1), 35–78.

Kerr, S. P., W. R. Kerr, and W. F. Lincoln (2015). Skilled Immigration and the Employment Structures of U.S. Firms. *Journal of Labor Economics 33*(S1), S147–S186.

Kerr, W. (2013a, August). U.S. High-Skilled Immigration, Innovation, and Entrepreneurship: Empirical Approaches and Evidence. *Harvard Business School*. Working Paper, No. 14-017.

Kerr, W. and W. Lincoln (2010). The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention. *Journal of Labor Economics 28*(3), 473–508.

Kerr, W. R. (2013b). Heterogeneous Technology Diffusion and Ricardian Trade Patterns. *NBER*. Working Paper Series, No. 19657.

Khanna, G. and M. Lee (2018). High-skill immigration, innovation, and creative destruction. *National Bureau of Economic Research, Working Paper 24824*.

Klein, P. and G. Ventura (2007). TFP Differences and the Aggregate Effects of Labor Mobility in the Long Run. *The B.E. Journal of Macroeconomics 7*(1).

Krugman, P. (1979). A Model of Innovation, Technology Transfer, and the World Distribution of Income. *Journal of Political Economy 87*(2), 253–266.

Kumar, N. (2001). Indian Software Industry Development: International and national perspective. *Economic and Political Weekly*. Sameeksha Trust Publication, Mumbai.

Kumar, N. (2006). Emerging Multinationals : Trends, Patterns and Determinants of Outward Investment by Indian Enterprises. *East Asian Bureau of Economic Research*. Microeconomics Working Papers 22108.

Lagakos, D., A. M. Mobarak, and M. E. Waugh (2018). The Welfare Effects of Encouraging Rural-urban Migration. *NBER Working Paper No. 24193*. January.

Lagakos, D. and M. E. Waugh (2013). Selection, agriculture and cross-country productivity differences. *American Economic Review 103*(2), 948–80.

Lee, H. (2016). Quantitative Impact of Reducing Barriers to Skilled Labor Immigration: The Case of the US H-1B Visa. *Working Paper*.

Levchenko, A. and J. Zhang (2016, April). The Evolution of Comparative Advantage: Measurement and Welfare Implications. *Journal of Monetary Economics 78*, 96–111.

Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labor. *Manchester School*.

Llull, J. (2018). Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model. *The Review of Economic Studies 85*(3), 1852–1896.

Matsuyama, K. (1992). Agricultural Productivity, Comparative Advantage, and Economic Growth. *Journal of Economic Theory 58*, 317–334.

McFadden, D. L. (1984). Econometric analysis of qualitative response models. *Handbook of Econometrics 2*, 1395–1497.

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

Morales, N. (2019). High-skill migration, multinational companies and the location of economic activity. *Working Paper*.

Moses, J. W. and B. Letnes (2004). The Economic Costs to International Labor Restrictions: Revisiting the Empirical Discussion. *World Development 32*(10), 1609–1626.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1989). Industrialization and the big push. *Journal of Political Economy 97*(5).

NASSCOM (2003). The IT Industry in India - Strategic Review. *New Delhi*.

NASSCOM (2012). The IT Industry in India - Strategic Review. *New Delhi*.

Ottaviano, G. I. P. and G. Peri (2012, February). Rethinking the effect of immigration on wages. *Journal of the European Economic Association 10*(1), 152–197.

Peri, G., K. Shih, and C. Sparber (2015). STEM Workers, H-1B Visas, and Productivity in US Cities. *Journal of Labor Economics 33*(S1), S225 – S255.

Peri, G. and C. Sparber (2011). Highly-Educated Immigrants and Native Occupational Choice. *Industrial Relations 50*(3), 385–411.

Perla, J., C. Tonetti, and M. E. Waugh (2015). Equilibrium Technology Diffusion, Trade, and Growth.

NBER. Working Paper 20881.

Reinhold, S. and K. Thom (2013). Migration Experience and Earnings in the Mexican Labor Market. *Journal of Human Resources 48*(3).

Rosenzweig, M. (2006). Global Wage Differences and International Student Flows. *Brookings Trade Forum*, 57–96. Global Labor Markets? (eds. Mark R. Rosenzweig, Douglas A. Irwin and Jaffrey G. Williamson).

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica 55*(5), 999–1033.

Rybczynski, T. (1955). Factor Endowment and Relative Commodity Prices. *Economica 22*(88), 336–341.

Ryoo, J. and S. Rosen (2004, February). The Engineering Labor Market. *Journal of Political Economy 112*(S1), S110–S140.

Samuelson, P. A. (2004). Where Ricardo and Mill Rebut and Confirm Arguments of Mainstream Economists Supporting Globalization. *Journal of Economic Perspectives 18*(3), 135–146.

Saxenian, A. (1999). *Silicon Valley's New Immigrant Entrepreneurs*. San Francisco CA: Public Policy Institute of California.

Shrestha, S. (2016). No Man Left Behind: Effects of Emigration Prospects on Educational and Labor Outcomes of Non-migrants. *Economic Journal 127*, 495–521.

Somale, M. (2014). Comparative Advantage in Innovation and Production. *Working Paper*.

Stark, O. (2004). Rethinking the Brain Drain. *World Development 32*(1), 15–22.

Stark, O., C. Helmstein, and A. Prskawetz (1997). A Brain Gain with a Brain Drain. *Economics Letters 55*(2), 227–234.

Subramanian, C. (1992). India and the Computer: A Study of Planned Development. *Oxford University Press: New Delhi.*.

Traiberman, S. (2019). Occupations and Import Competition: Evidence from Denmark. *American Economic Review 109*(12), 4260–4301.

U.S. Department of Commerce (2000). Digital economy 2000. Technical report.

USCIS (2014). Characteristics of H-1B Speciality Occupation Workers. *US DHS*.

USINS (2000). Characteristics of Specialty Occupation Workers (H1B). *U.S. Immigration and Naturalization Service*. Washington, D.C.

van der Mensbrugghe, D. and D. Roland-Holst (2009). Global Economic Prospects for Increasing Developing-Country Migration into Developed Countries. *United National Development Programme Human Development Research Paper 50*. UNDP.

Ventura, J. (1997, February). Growth and Interdependence. *The Quarterly Journal of Economics 112*(1), 57–84.

Vernon, R. (1966, May). International Investment and International Trade in the Product Cycle. *Quarterly Journal of Economics 80*, 190–207.

Walmsley, T. L. and A. L. Winters (2005). Relaxing the Restrictions on the Temporary Movement of Natural Persons: A Simulation Analysis. *Journal of Economic Integration 20*(4), 688–726.

Appendix

Table of Contents

A Details of the Data Used

A.1 Data for Descriptive Figures

Figures 1a to 4b are constructed using a variety of sources from India and the US. Figures 1a, and 1b use data from the March CPS, obtained from the IPUMS and NBER websites. The sample consists of employed persons with at least a BA degree. A person is defined as 'foreign' if he/she was born outside the United States and immigrated after the age of 18.

Figure 2a on the H-1B cap series is calculated by the authors from USCIS reports. Figure 3a, on the expected relative wages in India is calculated using the March CPS for the US wage component and the National Sample Survey (NSS) for the wages in India. More details in the NSS are described in Section A.3. To compute Figure 3b, we use various sources on education in India. We combine data from the Ministry of Human Resources and Development, which records number of degrees and universities by type of degree (for example, engineering degrees) with reports from the All India Council for Technical Education (AICTE) and the National Association of Software and Service Companies (NASSCOM) to also look at the growth in Masters for Computer Application (MCA) degrees.

To get total IT output (and export numbers which we corroborate with our exports data) for Figure 4a, we use data from the Electronic and Information Technology Annual Reports, and the Indian Department of Electronics reports. Much of this data has been collated and standardized by the Center for Development Informatics at the University of Manchester, UK. The remaining data for summary tabs and graphs are from National Association of Software and Service Companies (NASSCOM). Finally, for the exports data for Figure 4b, we use the OECD Trade in Value Added statistics, 2016 Edition.

Finally, Figure 2b uses data from the ACS from the US and the NSS from India, which is described in detail in section A.3. For the US, we calculate the number of workers with college degrees who migrated to the US after the age of 18. We use the Census of 1990, and the ACS for 2000, 2006, 2008, 2010, and 2012. For the NSS, we use the samples of 1994, 2000, 2006, 2008, 2010, and 2012. We present the ratio for the IT sector and for all other economic sectors. The IT sector is defined as NIC04 code 72 for India and Census ind1990 code 732 for the US.

Table A1: Summary statistics for key variables

	Mean	**Median**	**Std. Dev**
Levels			
College employment - Total	20153	4721	41547
College employment - Young	13114	2885	28492
N Indians in US in past 5 years	3227	747	10234
Logs			
College employment - Total	8.25	8.46	2.20
College employment - Young	7.77	7.97	2.20
N Indians in US in past 5 years	6.63	6.62	1.69

Notes: Table presents summary statistics for main variables in equation 5, in levels and logs. Mean, median and standard deviation are calculated across region-occupation-years triplets. Sample is restricted to occupations that are high-skill intensive. Final sample includes 38 occupations, 30 Indian states and 5 periods. We only include occupation-region pairs where workers are found in at least one year. College employment in India comes from the National Sample Surveys (1994, 2000, 2005, 2010 and 2012). "College employment - Young", refers to college graduates who are between 25 and 40. "N Indians in US in past 5 years" varies at the occupation-time level and is defined as the number of Indian college graduates, who migrated to the US in the past 5 years to work in a given occupation. Data comes from the US Census (1990, 2000), and the American Community Survey (2005, 2010, 2012).

A.2 Data for Empirical Results on College Major Choice

To construct the number of exam-takers every year, we use examination reports published by the University Grants Commission (UGC) in India who keeps records of the number of students who take the annual examinations for every degree and university in India after the completion of each academic year. The annual examination reports for the years 2006-2010 can be downloaded directly from UGCs website. We supplement these reports with digitized reports from years 2001 to 2005, which were obtained from libraries in India. We clean the data in multiple ways. First, we harmonize the names of the universities to make them consistent across years. Second, we collapse the variety of degrees available into ten categories: "Commerce", "Architecture", "Medicine", "Education", "Law", "B.Sc/M.Sc", "Business", "Engineering/Tech", "B.A / M.A", and "Other". We then use as our preferred measure of student enrollment the "Total Appeared" by university, degree, and year.

To construct the migration exposure, we use information on the universe of LinkedIn profiles for Indian migrants in the US obtained from Revelio Labs, a corporate data provider that collects online curriculum vitae data for over 100 million users, and predominantly includes publicly available data from LinkedIn. The dataset includes: 1) reported education history, including university names, field of study, and degree type; 2) Full employment history, including firm name, start and end date, job title, and position changes; and 3) User-level data, including current location, first and last name, imputed gender and imputed race among others. While the data is incredibly rich, it does not include precise identifiers on whether individuals are migrants or not. Because of that, we infer an individual is a migrant when we observe them obtaining some university degree in one country, and then working or reporting their current location in another country. Hence, if we observe a user with a degree from an Indian university but reporting that they are currently in the US, we label them as an Indian migrant in the US.

While the users in LinkedIn are a selected sample of the universe of college graduates in the US, we show that its coverage is significant when compared to the population counts in the ACS. As shown in Table A2, LinkedIn users account for 54.3% of Indian migrants, and 62% of non-Indian migrants. The sample has particularly good coverage of those with engineering, computer science, business and economic degrees, where the coverage is above 60%. For other fields such as law and medicine, the coverage is lower given that many in those professions do not use LinkedIn for business networking and job search. As expected, young workers are more likely to show up in the dataset than older workers. Finally, the total share of migrants is somewhat lower than what we predict in the ACS, likely because we cannot identify migrants that did all of their studies in the US.

Table A2: College graduates in the ACS vs LinkedIn

	ACS	LinkedIn	Coverage		ACS	LinkedIn
Number of college workers				**Shares**		
Indian immigrants	1,223,542	664,553	54.3%	Female	52.2%	49.3%
Non-Indian immigrants	48,822,653	30,273,428	62.0%	Migrants	10.7%	7.3%
				Graduate degree	35.9%	31.6%
Numbers by Major				Age 22-30	23.2%	43.4%
Engineering and CS	6,726,992	4,077,538	60.6%	Age 31-40	29.7%	29.4%
Business and Economics	13,028,962	8,519,981	65.4%	Age 41-50	25.8%	16.5%
Other	14,689,959	3,168,242	21.6%	Age 51-60	21.2%	10.7%

Notes: We compare the total counts between our data from LinkedIn profiles and the ACS in 2019. An Indian immigrant is defined as someone whose reported current location is the United States but at least one university degree is from an institution located in India. For the ACS we define as an immigrant someone who migrated after the age of 22 and has a college degree.

A.3 Data for Empirical Results on Occupational Choice

Indian micro data comes from the National Sample Survey (NSS). We use the Employment / Unemployment surveys from rounds 50 through 66, which cover 1987 through 2012 with gaps in between. NSS is a nationally representative survey and the largest household survey in the country, asks questions on weekly activities for up to five different occupations per person, and weekly earnings for each individual. Computer scientists are defined as "213 Computing Professionals" and "312 Computing Associate Professionals" based on the National Classification of Occupations (NCO) of 2004. We use the earnings data for the primary occupation only. NCO 2004 codes follow the same structure as ISCO-88 3-digit occupations, which makes it useful to construct crosswalks with US and UK data between occupations. The IT sector is restricted to be "Computer and Related Activities', code 72 in the National Industrial Classification of 2004 (NIC).

For the regressions presented in Section 3 we focus on the NSS samples of 1994, 2000, 2005, 2010 and 2012. In certain robustness specifications we use 1987 to calculate baseline shares. These rounds are often split over two years. For instance, the 2005 round was based on data collected between the end of 2005 and the first half of 2006. The NCO broadly groups occupation by skill. We keep 'college-graduate occupations' which include all NCO 2004 codes from 211 to 422. This includes professionals, technicians, associate professionals and clerical support workers. As the H-1B focuses on high-skilled occupations, we exclude from the analysis service and sales, agricultural, craft and trades, plant and machine operators, elementary occupations and armed forces from the analysis. We also exclude managers since the mapping with occupational classifications from other countries is less clear. Finally, we exclude physicians from the analysis as they migrate to the US predominantly through J visas, and are not subject to the H-1B cap. We construct crosswalks to get all waves of the NSS in consistent NCO 2004 codes and NIC 2004 codes.

We end up with 38 occupations, 30 Indian states and 5 periods for the analysis. However, not every occupation-state has observations for every year. To construct the Bartik local demand control in equation 3, we consider 1-digit industries. We calculate the share of employment of each industry in the total employment of a given occupation-region pair in 1994 and interact it with the national level growth of industry employment excluding (leave-one-out) the individual occupation-region-industry count from the aggregate.

The US micro data used in Section 3 comes from the American Community Survey (ACS). We use the Census for 1990, 2000 and all ACS samples between 2005 to 2012. We use the 1990 Census Bureau occupational classification scheme which provides consistent occupation codes for our period of study. As we are looking at a very specific group of workers (Indian college graduates in US by occupation who migrated in the past 5 years), we pool years 2005-2006, 2009-2010 and 2011-2012, with the requisite weights to match the NSS years. We use a crosswalk to map the ACS occupational classification to ISCO-88 3-digit codes which allows us to connect the 38 occupations in India with occupation codes in the US.

A.4 Data for Calibration

The time-varying parameters described in Section 5.2 and the targeted moments in Section5.3 are computed for 4 years: 1995, 2000, 2005 and 2010. For moments in India we use the NSS for years 1994, 2000, 2005 and 2010. For the US, since the ACS does not have information for 1995, we use the IPUMS CPS for 1995, 2000, 2005 and 2010. This harmonized dataset is highly compatible with the ACS and we can use the same classification of occupations and industries. For information on imports, exports and RoW consumption of IT from the US and India we use the OECD Trade in Value Added statistics, 2016 Edition. The dataset uses World input-output tables for the period 1995 to 2011. We use gross exports, gross imports and total GDP data for the "C72: Computer and Related Activities" industry in addition to aggregate numbers by country across industries. We use trade data for 1995, 2000, 2005, and 2010.

B Empirical Design, Alternative Specifications, Instruments & Falsification Tests

B.1 Alternative Specifications for Enrollment Decisions in India

Table B1: The Effect of the H-1B cap on Enrollment in Majors: Alternative Leads

	Dependent variable: Log(Exam Eligible$_{s,f,t+1}$)			
Migration Intensity*Log(H1B Stock)	0.125***	0.094*	0.124*	0.085*
	(0.045)	(0.050)	(0.065)	(0.049)
School-Field FE	Y	Y	Y	Y
Time FE	Y			
Field-Time FE		Y		
State-Time FE		Y		
Field-State-Time FE			Y	Y
School-Time FE				Y
N	9513	9509	8772	8654
R-sq	0.900	0.908	0.917	0.947

Dependent variable is the number of students eligible to appear for an examination in school s, located in state r, studying in field f, and year $t+1$. Migration exposure is defined as the number of pre-2000 graduates from school s and field f working in the US, divided by the total pre-2000 enrollment at school s and field f. The H-1B stock is the cap-predicted number of H-1B workers in the US. Data for the migration exposure are from the universe of LinkedIn users that studied in India, and work in the US. Data on examinations are from annual Indian Government reports. Data details are in Appendix A.2. Standard errors clustered at the school-field level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table B2: The Effect of the H-1B Indian Migration on Enrollment in Majors

	Dependent variable: Log(Exam Eligible$_{s,f,t+2}$)			
Migration Intensity*Log(H1B Indians)	0.088***	0.067**	0.085*	0.058*
	(0.031)	(0.033)	(0.044)	(0.032)
School-Field FE	Y	Y	Y	Y
Time FE	Y			
Field-Time FE		Y		
State-Time FE		Y		
Field-State-Time FE			Y	Y
School-Time FE				Y
N	9513	9509	8773	8655
R-sq	0.900	0.908	0.917	0.947

Dependent variable is the number of students eligible to appear for an examination in school s, located in state r, studying in field f, and year $t+2$. Migration exposure is defined as the number of pre-2000 graduates from school s and field f working in the US, divided by the total pre-2000 enrollment at school s and field f. The H-1B Indians is the number of H-1B workers from India in the US. Data for the migration exposure are from the universe of LinkedIn users that studied in India, and work in the US. Data on examinations are from annual Indian Government reports. Data details are in Appendix A.2. Standard errors clustered at the school-field level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table B3: The Effect of the H-1B non-Indian Migration on Enrollment in Majors

	Dependent variable: Log(Exam Eligible$_{s,f,t+2}$)			
Migration Intensity*Log(H1B Non-Indians)	0.110***	0.084**	0.109*	0.073*
	(0.039)	(0.043)	(0.055)	(0.042)
School-Field FE	Y	Y	Y	Y
Time FE	Y			
Field-Time FE		Y		
State-Time FE		Y		
Field-State-Time FE			Y	Y
School-Time FE				Y
N	9513	9509	8773	8655
R-sq	0.900	0.908	0.917	0.947

Dependent variable is the number of students eligible to appear for an examination in school s, located in state r, studying in field f, and year $t+2$. Migration exposure is defined as the number of pre-2000 graduates from school s and field f working in the US, divided by the total pre-2000 enrollment at school s and field f. The H-1B Non-Indians is the number of H-1B workers from other countries (not India) in the US. Data for the migration exposure are from the universe of LinkedIn users that studied in India, and work in the US. Data on examinations are from annual Indian Government reports. Data details are in Appendix A.2. Standard errors clustered at the school-field level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table B4: State Level: The Effect of the H-1B Cap on Enrollment in Majors

	Dependent Variable: Log(Enrollment)$_{rf,t+2}$		
Migration Intensity*Log(H1B Shifter)	11.715***	11.001***	5.736*
	(4.084)	(3.813)	(3.088)
Shifter	H1B Non-Indians	H1B Indians	H1B cap
State-Field FE	Y	Y	Y
Field-Time FE	Y	Y	Y
State-Time FE	Y	Y	Y
N	670	670	950
R-sq	0.966	0.966	0.968

Dependent variable is the number of students eligible to appear for an examination in state r, studying in field f, and year $t+2$. Migration exposure is defined as the number of pre-2000 graduates from school s and field f working in the US, divided by the total pre-2000 enrollment at school s and field f. The H-1B shifter is different across columns: the number of non-Indian H1B migrants, the number of H1B migrants from India, and the simple cap. Data for the migration exposure are from the universe of LinkedIn users that studied in India, and work in the US. Data on examinations are from annual Indian Government reports. Data details are in Appendix A.2. Standard errors clustered at the school-field level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

B.2 Components of Local Occupational Choice

Equation 2 in Section 3, describes occupation choice in India. While we are most interested in the demand from the US component, the alternative components are both interesting in and of themselves, but may also be confounding our estimates of γ_2. We elaborate on how we account for these components in the text, with the help of instrumental variables and additional controls. But to fix ideas, we describe here in detail the role played by the components.

As a first step, we parameterize the component of supply and preferences as in equation B1:

$$\text{Supply and Preferences}_{ort} = \delta^S_{or} + \delta^S_{rt} + \varepsilon^S_{ort} \tag{B1}$$

These preferences may be for specific occupations, in certain regions, and may vary across occupations and regions δ^S_{or}. For instance, certain regions may be more likely to produce certain types of workers given their institutions and type of capital. Similarly, certain regions may see differential growth in employment related preferences δ^S_{rt} for high-skill work, as they shift out of low-skill work. We do not observe ε_{ort}, and as such, in the absence of instrumental variables that isolate demand-shocks, may bias estimates of γ_2, given the likely correlation with Migration prob_{ort}. ε^S_{ort} also includes any unobserved components that vary at the occupation-time level.

We describe the component of labor demand from India as in equation B2:

$$\text{Demand India}_{ort} = Local_{ort} + \delta^D_{or} + \delta^D_{rt} + \varepsilon^D_{ort} \tag{B2}$$

This includes time varying local demand shocks that vary by occupation $Local_{ort}$. It also accounts for the propensity to hire workers from certain regions, to hire workers of certain occupations, and the propensities for certain regions to hire workers in certain occupations δ^D_{or}. Similarly, we account for local economic growth (again, not just in aggregate but also that vary across regions), that drive the increase in demand for overall employment in those regions δ^D_{rt}. The error term ε^D_{ort}, includes unobserved demand shocks that affect Indian workers and can be correlated with the migration probability for different occupations and regions.

Together, equations 2-4 and B1-B2, allow us to reformulate our primary relationship of interest:[25]

$$\begin{aligned} Ln(\text{N}_{ort}) = \gamma_0 + \gamma_1 \text{Local}_{ort} + \gamma_2 Ln(\text{Indians in US}_{ort}) + \underbrace{\left(\gamma_1 \delta^D_{or} + \gamma_3 \delta^S_{or}\right)}_{\delta_{or}} \\ + \underbrace{\left(\gamma_1 \delta^D_{rt} + \gamma_3 \delta^S_{rt} - \gamma_2 Ln(Total\ US\ Mig\ App)_t\right)}_{\delta_{rt}} + \underbrace{\left(\gamma_1 \varepsilon^D_{ort} + \gamma_3 \varepsilon^S_{ort}\right)}_{\varepsilon_{ort}} \end{aligned} \tag{B3}$$

This produces our final estimation equation 5. Yet, given the possible correlations between the error ε_{ort} and our primary variable of interest $Ln(\text{Indians in US}_{ort})$, we still require instrumental variables to estimate γ_2, as in Section 3.2.1.

B.3 Construction of main instrument and variables

To compute our endogenous main variable "Log number of Indian college graduates in US, who migrated in the past five years by occupation and year" we use data from the US. The US data provides information on the occupation of migrants but does not include origin region in India. Hence, in our baseline specification, our dependent variable varies at the occupation-year level. In subsequent specifications discussed in Appendix 3.2.4, we use Indian data from the prior five years to reweight this measure by the number of workers from region r that work in occupation o, creating a variable at the occupation-region-time level.

Instrument: For the baseline instrument, we use US data. We first compute the initial share of each

[25] Note that the additional advantage of the fixed effects are that it controls for other components that vary across occupation-region (such as the baseline propensity of regions to send migrants of certain occupations abroad) or across regions over time (say if there is overall increase in emigration from certain regions).

occupation among total Indian college graduate employment in the US in 1990. As shown in equation B4, "N Indians in US" is the stock of Indian college graduates in the US in 1990.

$$\text{H-1B Instrument}_{ot} = \underbrace{\frac{(\text{N Indians in US})_{o,1990}}{(\text{N Indians in US})_{1990}}}_{\text{Initial share}} \times \underbrace{(\text{Policy-predicted H-1B stock})_t}_{\text{time shifter}}, \tag{B4}$$

Our main time shifter is the cumulative H-1B stock constructed using the official H-1B cap between 1995 and 2012. In Appendix B.5, we describe other time shifters such as the total number of non-Indian college graduate immigrants in the US and the total number of non-Indian college graduate immigrants in the US by occupation. We also show robustness to an instrument that creates the policy-predicted H-1B stock after accounting for return migration in Appendix B.4.

As shown in Figure 2a, between the period of 1994 to 2012 we can exploit four sharp changes to the H-1B cap. It went from 65,000 to 115,000 in 1999, to 195,000 in 2001 and decreased to 65,000 in 2004. In 2006, an additional 20,000 visas were added to the cap for those who have graduate degrees granted by a US institution. The cumulative H-1B stock is calculated as follows: we assume the H-1B stock in 1992 is of 65,000 (same as the cap). For 1993, we add 65,000 to the stock in 1992. In 1999, where the cap increases from 65,000 to 115,000, we add 115,000 to the stock in 1998. We calculate the cumulative stock every year until 2012 to construct our main time-shifter. Since the regression controls for time fixed effects, changes in the H-1B cap are the main source of identifying variation captured by the instrument.

For the region-reweighted analysis in Section 3.2.4, and particularly equation 9, we need to measure the stock of workers from the prior 5-year period. We use the NSS in 1987 to compute the shares in 1994, the 1994 to compute 2000, 2000 to compute 2005, 2005 to compute 2010 and 2008 to compute 2012. For the instrument in equation 9, we use data from the Indian Census in 1991, to compute the baseline distribution of Indian college graduates across regions and occupations in India.

B.4 Alternative Specifications

In Table B5, we describe a few alternative specifications. First, we incorporate into our H-1B cap instrument the length of the visas and the probability of returning. Visas are granted for a period of three years, renewable for three further years. If a company sponsors a green card, Indian nationals may continue to renew their H-1Bs till the green card is granted. Yet, the OECD estimates that 23.5% of high-skill immigrants in the US return to their home countries after a six-year period (Lee, 2016). To construct this instrument, we subtract $0.235 \times \text{H-1B cap}_{t-6}$ from the policy predicted H-1B stock in equation 6. In this case, we construct the H-1B stock as follows: like before we assume the H-1B stock in 1992 is of 65,000. For 1993, we assume an additional 65,000 join the stock, while 23.5% of the cap 6 years ago return (following our return rate calibration in Section 5.2). We continue updating the stock every year with the current value of the cap and subtracting an estimate of those who return until we reach 2012. Hence, the time-varying component of the instrument will also account for those high-skill immigrants that return to their home countries. In Panel A of Table B5, we compare our original instrument with a version that includes the return rates, and find them both to be similar.

Next, we include occupation-year observations for which there were no Indians that migrated to the US in the preceding five years. In Panel B of Table B5, we take the inverse hyperbolic sine of our primary explanatory variable, and show the results to be robust across both versions of the H-1B instrument, and age groups.

In Panel C of Table B5, we consider the role played by the local demand control in our estimation, and find that in removing the control, our effects are similar to before. Our measure of local demand shocks is itself meaningful across all tables, and as expected, has a direct effect on occupational choice. But it does not seem to affect the relationship between the probability of migration and occupational choice. This control not only accounts for local industrial growth but also other nationwide shocks. For instance, if there was a nationwide demand shock in the IT industry, the IT component of the shift-share will increase and those region-occupation pairs more intensive in IT will show a larger overall labor demand shock than less intensive region-occupation pairs. This way, the control also accounts for local demand changes for each occupation over time.

Table B5: Alternative Specifications

Panel A: With return migrants	**Baseline IV**		**With return migrants**	
	Log Employment All	Log Employment Young	Log Employment All	Log Employment Young
Log Indians in US	0.740***	1.065***	0.704***	1.038***
migrated in past 5 years	(0.262)	(0.311)	(0.258)	(0.309)
Demand Control	0.724***	0.833***	0.719***	0.830***
	(0.225)	(0.259)	(0.224)	(0.258)
N	3,114	3,114	3,114	3,114
1st stage F-stat	67.64	67.64	67.97	67.97

Panel B: With zero shares inverse hyperbolic sine	**Baseline IV**		**With return migrants**	
	Log Employment All	Log Employment Young	Log Employment All	Log Employment Young
Log Indians in US	0.543**	0.822***	0.503**	0.784***
migrated in past 5 years	(0.227)	(0.279)	(0.218)	(0.270)
Demand Control	0.950***	1.214***	0.932***	1.196***
	(0.255)	(0.301)	(0.250)	(0.296)
N	3,224	3,224	3,224	3,224
1st stage F-stat	39.74	39.74	39.81	39.81

Panel C: No demand control	**Baseline IV**		**With return migrants**	
	Log Employment All	Log Employment Young	Log Employment All	Log Employment Young
Log Indians in US	0.875***	1.178***	0.835***	1.150***
migrated in past 5 years	(0.261)	(0.309)	(0.258)	(0.308)
N	3,380	3,380	3,380	3,380
1st stage F-stat	68.44	68.44	68.62	68.62

Panel D: Adding trade control	**Baseline IV**		**With return migrants**	
	Log Employment All	Log Employment Young	Log Employment All	Log Employment Young
Log Indians in US	0.732***	1.059***	0.698***	1.034***
migrated in past 5 years	(0.260)	(0.309)	(0.257)	(0.308)
Demand Control	0.638***	0.773***	0.634***	0.770***
	(0.224)	(0.261)	(0.223)	(0.260)
US Imports Control	0.0348***	0.0262*	0.0346***	0.0260*
	(0.0125)	(0.0151)	(0.0125)	(0.0150)
N	3,114	3,114	3,114	3,114
1st stage F-stat	67.44	67.44	46.24	46.24

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$ All regressions include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Columns 1 and 3 show 2SLS regressions for the dependent variable of log employment of all college graduates in India while columns 2 and 4 count only college graduates who are between 25 and 40. Main explanatory variable is the log number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o. Demand Control is defined in equation 3. Sample is restricted to occupations that are high-skill intensive. Final sample includes 38 occupations, 30 Indian states and 5 periods. We only include occupation-region pairs where workers are found in at least one year. The instrument exploits variation in the H-1B cap as a time shifter. Data for India, from the National Sample Surveys (1994, 2000, 2005, 2010 and 2012), and data for the US from the US Census (1990, 2000), and the American Community Survey (2005, 2010, 2012). The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3. Panel A introduces an alternative formulation of the instrument, where in columns 3 and 4 the instrument includes the fixed probability of return migration. Panel B uses the inverse hyperbolic sine of the explanatory variable to include observations with 0 value. Panel C shows the results without controlling for the local demand controls. Panel D includes as an additional control, the demand shifter using US imports by industry from India as in equation B5.

If there is some component of the shock that is correlated with the instrument and not fully captured by the demand control, we would expect that if we remove the demand control, it would meaningfully change the estimates, as these unobserved shocks would be correlated with both the demand control and instrument. It is reassuring then that when we remove the demand control, the point estimates of our main variable do not change significantly, suggesting that it is unlikely that an unobserved demand shock correlated with the demand

Table B6: Alternative Instrumental Variables

Panel A: Inverse Hyperbolic Sine	Non Indian Migrants IV Log Employment All	Non Indian Migrants IV Log Employment Young	Non Indian Migrants by Occupation Log Employment All	Non Indian Migrants by Occupation Log Employment Young
Log Indians in US	0.562***	0.701***	0.432**	0.719***
migrated in past 5 years	(0.116)	(0.136)	(0.207)	(0.260)
Demand Control	0.959***	1.157***	0.899***	1.166***
	(0.248)	(0.280)	(0.244)	(0.290)
N	3,224	3,224	3,224	3,224
1st stage F-stat	138.5	138.5	38.83	38.83

Panel B: No demand control	Non Indian Migrants IV Log Employment All	Non Indian Migrants IV Log Employment Young	Non Indian Migrants by Occupation Log Employment All	Non Indian Migrants by Occupation Log Employment Young
Log Indians in US	0.635***	0.763***	0.685***	1.006***
migrated in past 5 years	(0.117)	(0.137)	(0.238)	(0.290)
N				
1st stage F-stat	180.6	180.6	67.18	67.18

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$ All regressions include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Columns 1 and 3 show 2SLS regressions for the dependent variable of log employment of all college graduates in India while columns 2 and 4 count only college graduates who are between 25 and 40. Main explanatory variable is the log number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o. Demand Control is defined in equation 3. Sample is restricted to occupations that are high-skill intensive. Final sample includes 38 occupations, 30 Indian states and 5 periods. We only include occupation-region pairs where workers are found in at least one year. The instrument exploits variation in the H-1B cap as a time shifter. Data for India, from the National Sample Surveys (1994, 2000, 2005, 2010 and 2012), and data for the US from the US Census (1990, 2000), and the American Community Survey (2005, 2010, 2012). Details of the data construction can be found in Appendix A.3. Panel A uses the inverse hyperbolic sine of the explanatory variable. Panel B shows the results without controlling for the local demand controls. In columns 1 and 2, as in equation 7, we capture the demand for migrants by looking at the flow of college-educated workers from all other countries to the US. In columns 3 and 4, we modify the instrument to be the number of migrants from other countries *by occupation*.

control is also correlated with out instrument.

B.5 Alternative Instrumental Variables & Correlated Demand Shocks

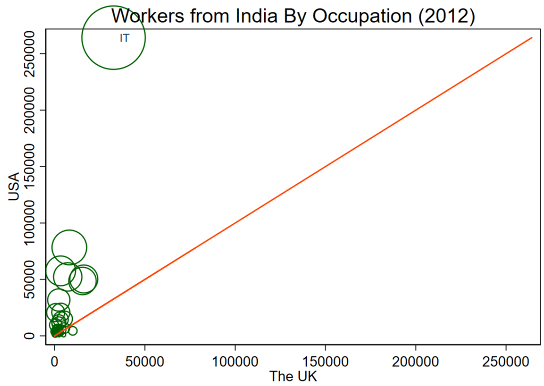
Table B6 describes results for the alternative instruments that leverage variation in the overall demand for migrants from US firms. In panel A we use the inverse hyperbolic sine of the independent variable, and panel B we show the results without the local demand control. Across specifications, instruments and age cohorts, our estimates are once again similar to our baseline specification.

However, with an instrument that does not rely on the H-1B policy changes, we may be concerned that the demand for Indian IT workers comes from correlated productivity shocks (say, in IT) in other countries. To that end, we obtain microdata for top three destinations for Indian migrants. The scale of migration to the US, however, is a lot larger to other destinations – in 2012 the stock of college-educated Indian migrants in the US was 4 times that of the UK and 4.8 times that of Canada.[26]

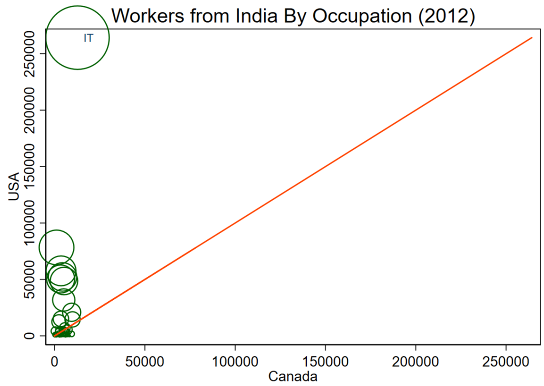
Importantly, for our purposes, the specific demand for IT workers from India is swamped by the US, not just because the US is larger, gets more migrants, but also because the IT boom was concentrated in the US. Figure B1 describes the demand for Indian workers by occupations in the top destinations. In Figures B1a and B1b, it is clear that even in 2012, when other countries have had time to develop their IT sectors, the US (on the vertical axes) swamps overall demand when compared to the UK or Canada (on the horizontal axes). Of course, in 2000, when only the US IT sector was developed, the scale of the difference in the demand for IT from the US was many magnitudes larger than the UK (Figure B1c). Indeed, the largest occupation for college graduate Indian migrants in the UK, in our data, was that of physicians, and not IT workers. Furthermore,

[26]Since we have many waves of UK data, we can also check the relative stocks in previous years: there were 3.5 times more Indians in the US than in the UK in 2000.

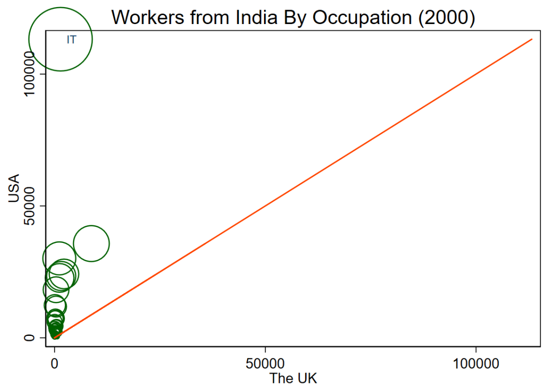
Figure B1: Demand from the UK and Canada relative to the US



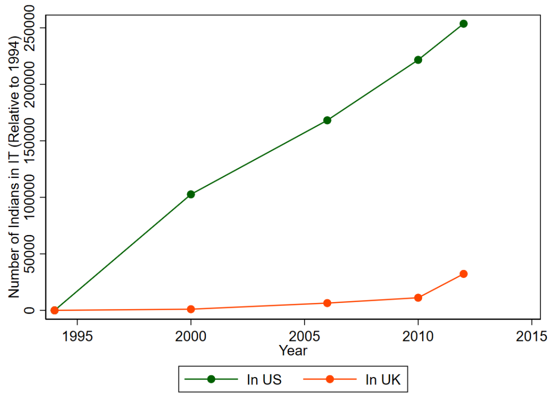
(a) Indians in the US and UK (2012)



(b) Indians in the US and Canada (2011-12)



(c) Indians in the US and UK (2000)



(d) Indians in the US and UK (1994-2012)

Notes: Sample of college educated. All figures with US data use the American Community Survey, Figures B1a, B1c and B1d use the UK Labor Force Survey, and Figure B1b is based on 2011 National Household Survey. Each bubble represents an occupation weighted by the size of the number of migrants from India in the US. Figure B1d plots the change in the number of workers over time relative to 1994. For instance, the last observation is the number of workers in 2012 minus the number of workers in 1994.

despite starting with a much larger scale at baseline, the US's growth in demand for IT workers from India was also a lot larger than the UK's (Figure B1d).

All of this suggests that the US was somewhat unique in its substantial demand for IT workers from India, indicating that the role played by the demand for Indian tech workers from other countries was relatively small.

As a final check, we want to make sure that our results are robust to controlling for potentially correlated global shocks between the US and India. If migration from India to the US is correlated with production relationships between the two countries, our estimates of brain gain could be biased. While our main specifications already include the $Local_{ort}$ control, we consider an additional control derived from product-level flows between the US and India. We construct a demand control that leverages the US imports from India across industries as in equation B5:

$$\text{US Trade Control}_{ort} = Ln\left(\sum_{ind} \frac{N^{ind}_{or,1994}}{N_{or,1994}} \times \text{Total value of US imports from India}_{ind,t}\right) \tag{B5}$$

To construct the demand control, we use data from the World Input Output Database (WIOD) to compute the total value of US imports from India over time for 35 industries categories that cover all economic activity.

Similar to the local demand control from equation 3, we re-weight each industry trade shock with the share of Indian workers in each occupation-region pair at baseline period 1994. The goal of this measure is to capture other demand channels through which occupation-region pairs in India respond to demand shocks in the US. As shown in Table B5, panel D, the US trade control is positive and significant (as expected), but does not alter the estimates for brain gain.

B.6 Leveraging Occupation-specific Wages in the US

Over this period, changes in (policy-driven) demand from abroad are reflected in the changing probability of migration, as in equation 5. Yet, one may expect occupation-specific demand-shocks in the US may change occupation specific wages as well, and influence the occupational-choice of workers in India. For instance, we could define the demand from the US to be of the form: Migration $\text{prob}_{ort} \times \text{wage}^{US}_{ot}$, then (after taking logs) our new equation 5 will require us to control for the wage in the US by occupation, wage^{US}_{ot}.

We can measure the occupation-specific wage in a few different ways. Using the ACS data, we can define wage^{US}_{ot} to be the average earnings of Indians in the US by occupation. Yet, such a variable may not necessarily only capture demand shocks, but also the consequences of more Indians in the US. We therefore propose an alternative measure of wages, the earnings of non-Indian migrants to the US.

Table B7 Panel A shows the results of this exercise where we include the log wages for non-Indians by occupation. The elasticity with respect to wages of other migrants is not meaningfully different from zero. This perhaps reflects the fact that the primary driver of occupational choice in India is the change in the probability of migrating. For low baseline probabilities of migrating, but high wage differentials between the US and India, we would expect this to be the case: an increase in the probability of migrating increases expected wages a lot more than an increase in the wage. Importantly, the elasticity with respect to the probability of migrating does not seem to be affected by this reformulation.

If changes in probabilities are the primary driver, perhaps we may expect that the impacts of the changes in these probabilities matter more for occupations where Indians in the US earn higher wages. To test this, we reformulate our primary endogenous variable to be the $Ln(\text{Wage Bill of Indians in the US})_{ot}$, the wage bill of Indians that migrated in the past five years. This suggests that when there are changes in the probability of migration, perhaps the occupations that respond more in India, are those that earn higher wages in the US.

Panels B and C of Table B7 shows these results across our various instruments, for this new endogenous variable of interest. While we show these results for our main instruments, we also show robustness to alternative formulations of our H-1B instrument, where the baseline shares are reweighted to include the wage bill share (rather than just the employment share) of Indians in the US. This is in Panel C. Across specifications, once again, our elasticities are similar to before.

B.7 Specification Sensitivity, Pre-trends and Baseline Correlations

We conduct other important sensitivity checks. We show that our results across specifications are consistent when using levels instead of logs for all our variables (dependent, independent, instrumental and controls).[27] In Table B10 we show that across instruments (Panel A and B), regional reweighting (Panel C) and occupation-by-region specific trends (Panel D), our results again show that demand from the US induces occupation choice in India.

We should note that when taking levels, the interpretation of our coefficients changes. While we can calculate the elasticity with respect to the change in the number of Indian migrants in the US, without the log transformation, this is not the same as the elasticity with respect to the change in the probability of migration.

Next, in Table B11, we test for pre-trends, where we change the outcome variable to be five years preceding the independent variable (and instrument). The lack of a relationship indicates that future changes in the probability of migrating to the US do not affect current occupation choices in India. This is further indicative

[27] To be conservative, we only include occupation-region pairs where workers are found in at least one year. Not doing so (including more zeros) only seems to increase precision without contributing to identification.

Table B7: Including US wages in the demand from abroad

Panel A: With wage of other migrants	**Baseline IV** Log Employment All	Log Employment Young	**With return migrants** Log Employment All	Log Employment Young
Log Indians in US migrated in past 5 years	0.717***	1.062***	0.681***	1.035***
	(0.263)	(0.313)	(0.260)	(0.311)
Demand Control	0.727***	0.834***	0.723***	0.830***
	(0.224)	(0.259)	(0.223)	(0.259)
Log(Wage of other migrants)	0.147	0.0178	0.151	0.0212
	(0.250)	(0.294)	(0.248)	(0.292)
N	3,114	3,114	3,114	3,114
1st stage F-stat	68.99	68.99	69.18	69.18

Panel B: Endogenous wage bill H-1B Policy IVs	**Baseline IV** Log Employment All	Log Employment Young	**With return migrants** Log Employment All	Log Employment Young
Log Wage bill of Indians in US migrated in past 5 years	0.553***	0.796***	0.532***	0.784***
	(0.196)	(0.234)	(0.196)	(0.235)
Demand Control	0.611***	0.670**	0.611***	0.670***
	(0.226)	(0.260)	(0.225)	(0.259)
N	3,114	3,114	3,114	3,114
1st stage F-stat	66.71	66.71	66.11	66.11

Panel C: Endogenous wage bill Alternative IVs	**Baseline IV using wage-bill** Log Employment All	Log Employment Young	**Non Indian Migrants by Occupation** Log Employment All	Log Employment Young
Log Wage bill of Indians in US migrated in past 5 years	0.428**	0.729***	0.528***	0.646***
	(0.196)	(0.237)	(0.101)	(0.117)
Demand Control	0.614***	0.672***	0.612***	0.675***
	(0.221)	(0.256)	(0.224)	(0.252)
N	3,114	3,114	3,114	3,114
1st stage F-stat	72.58	72.58	202.8	202.8

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$ All regressions include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Columns 1 and 3 show 2SLS regressions for the dependent variable of log employment of all college graduates in India while columns 2 and 4 count only college graduates who are between 25 and 40. Main explanatory variable is the log number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o. Demand Control is defined in equation 3. Sample is restricted to occupations that are high-skill intensive. Final sample includes 38 occupations, 30 Indian states and 5 periods. We only include occupation-region pairs where workers are found in at least one year. The instrument exploits variation in the H-1B cap as a time shifter. Data for India, from the National Sample Surveys (1994, 2000, 2005, 2010 and 2012), and data for the US from the US Census (1990, 2000), and the American Community Survey (2005, 2010, 2012). The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3. Panel A includes the Log(Wage of other migrants)$_{ot}$ in the US, as an explanatory variable. Panel B reformulates the explanatory variable of interest to include the baseline wage bill of Indians in the US, by occupation, when measuring baseline shares. Panel C includes additional instrumental variables, including a new instrument where we also include the baseline wage-bill when defining the baseline share of the instrument.

of the fact that changes to H-1B policy were unexpected and non-monotonic.

Indeed, in Table B12 we go one step further and show that the instrument does not predict baseline levels of education, and nor is it associated with other local labor market characteristics that have little to do with occupational choice.

In Figure B2 we show an additional robustness exercise where when we drop each occupation one-at-a-time (Figure B2a), and each region one-at-a-time (Figure B2b), our results are similar to before. This indicates that even though there may be few major occupations that determine migration patterns, none are the sole driver of this elasticity.

B.8 Tests of Shift-share Style Instruments

Our setup is not precisely a conventional shift-share, but we can still draw lessons from the shift-share literature. Our 'shifters' vary over time, and resembles most closely the description of the panel-data shift-share setups

Table B8: Reweighted by Regional Shares

	Non Indian migrants by Occupation		**H-1B cap IV with share returning**	
	Log Employment All	Log Employment Young	Log Employment All	Log Employment Young
Log Indians in US	0.794**	1.108**	0.870**	1.188**
migrated in past 5 years	(0.381)	(0.457)	(0.435)	(0.509)
Demand Control	0.728***	0.824**	0.726**	0.823**
	(0.277)	(0.346)	(0.289)	(0.360)
N	3,087	3,087	3,087	3,087
1st stage F-stat	9.710	9.710	9.701	9.701

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$ All regressions include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Columns 1 and 3 show 2SLS regressions for the dependent variable of log employment of all college graduates in India while columns 2 and 4 count only college graduates who are between 25 and 40. Main explanatory variable is the log number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o. Demand Control is defined in equation 3. Sample is restricted to occupations that are high-skill intensive. Final sample includes 38 occupations, 30 Indian states and 5 periods. We only include occupation-region pairs where workers are found in at least one year. The instrument exploits variation in the H-1B cap as a time shifter. Data for India, from the National Sample Surveys (1994, 2000, 2005, 2010 and 2012), and data for the US from the US Census (1990, 2000), and the American Community Survey (2005, 2010, 2012). The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3. In this table, we reformulate the explanatory variable and the instrumental variable to be reweighted by the baseline region-share of Indians in a specific occupation, as in equations 9 and 8. Columns 1 and 2 show the results where the time-shifter in the instrument is the number of non-Indian migrants to the US by occupation. Columns 3 and 4 show the results with the primary policy-induced H-1B instrumental variables accounting for return migrants.

described by Borusyak et al. (2020). Yet, we perform tests from all three sets of shift-share papers (Adao et al., 2019; Borusyak et al., 2020; Goldsmith-Pinkham et al., 2020) as we believe them to be informative of the underlying variation.

There are a few suggestions these papers recommend. First, researchers are encouraged to control for the baseline shares (if they do not add up to one). Since we include occupation-by-region fixed effects, we already control for these baseline shares in a non-parametric fashion. Related to this is what Borusyak et al. (2020) point out, that in this setup, these fixed effects purge time-invariant unobservables and isolate the time-varying component of the shock.

Second, we are careful to not have to construct our shocks from the same database as our shares as in the traditional migration shift-shares (Card, 2001). Indeed, we are considerably better placed in that our endogenous variables and instruments are based on US data and policy, whereas our outcomes are based on data from halfway across the world in India. This relaxes identification concerns. For instance, unlike previous work that leverages the cap, we do not require the cap to be otherwise uncorrelated with other US outcomes, but simply to not be correlated with other drivers of occupation choice in India.

Furthermore, in the traditional framework, the researcher does not observe the reason behind the time-shifters (say push or pull factors), but rather assumes an underlying set of shifters that are driving the aggregate trends (in migration to the US). As Jaeger et al. (2018) point out, since these shocks are estimated in-sample, they may potentially introduce mechanical biases. In our framework, the shifters (H-1B administrative caps, or US ACS other-country migrants) and outcomes (India's NSS) are from different sources.

Third, as Goldsmith-Pinkham et al. (2020) suggest, it may be useful to understand what variables at baseline are correlated with the shares, as we may not want trends in other characteristics driving our results. While this may be more important when arguing for share-orthogonality, as Borusyak et al. (2020) point out, these are informative for shock-orthogonality (when data are *iid*). As such, we still find it informative that in Table B13, a large set of observable characteristics (demographics, work structure, education characteristics) only explain about 3.8% of the variation in the baseline shares. Including the share of college graduates and a whole set of region fixed effects, not surprisingly raises this slightly to about 10.5%. Yet, most of the variation in the baseline shares are still unexplained, and there are no meaningful correlations seen in Figure B4.

Fourth, most papers suggest a pre-trend test (albeit each for slightly different reasons), which we take seriously. In Table B11 we conduct these tests for our main specifications with the various specifications, and show a lack of pre-trends.

Table B9: Occupation-by-Time Level Regressions

Log-log specification	Main Specification		With Occupation trends	
	Log Employment All	Log Employment Young	Log Employment All	Log Employment Young
Log Indians in US migrated in past 5 years	0.882*** (0.293)	0.974*** (0.344)	0.900*** (0.306)	0.995*** (0.359)
N	187	187	177	177
1st stage F-stat	16.50	16.50	15.39	15.39

Levels (H1B cap IV)	Main Specification		With Occupation trends	
	Employment All	Employment Young	Employment All	Employment Young
Indians in US migrated in past 5 years	31.46*** (7.784)	23.23*** (5.421)	31.55*** (7.834)	23.30*** (5.460)
N	190	190	190	190
1st stage F-stat	30.12	30.12	29.82	29.82

Levels (Non Indian Migrants to US IV)	Main Specification		With Occupation trends	
	Employment All	Employment Young	Employment All	Employment Young
Indians in US migrated in past 5 years	15.86*** (3.619)	11.93*** (2.519)	15.78*** (3.625)	11.88*** (2.525)
N	190	190	190	190
1st stage F-stat	161.6	161.6	161.6	161.6

Levels per worker (H1B cap IV)	Main Specification		With Occupation trends	
	Employment per US worker	Young per US worker	Employment per US worker	Young per US worker
Indians in US per US workforce migrated in past 5 years	33.60*** (8.925)	25.86*** (6.705)	33.74*** (9.023)	25.98*** (6.786)
N	190	190	190	190
1st stage F-stat	21.03	21.03	20.71	20.71

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$ Regressions include time fixed effects, and occupation fixed effects (columns 1 and 2) or occupation-specific trends (columns 3 and 4). Heteroskedastic robust standard errors. Columns 1 and 3 show 2SLS regressions for the dependent variable of employment of all college graduates in India while columns 2 and 4 count only college graduates who are between 25 and 40. Main explanatory variable is the number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o (log of this variable in the top panel). All specifications include a demand control that is in the levels formulation of equation 3. Sample is restricted to occupations that are high-skill intensive. Sample includes 38 occupations and 5 years. Data for India, from the National Sample Surveys (1994, 2000, 2005, 2010 and 2012), and data for the US from the US Census (1990, 2000), and the American Community Survey (2005, 2010, 2012). The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3. In this table, the data are only at the occupation-time level. The top panel is a log-log specification. The bottom few panels are the levels specification across instruments.

Fifth, the papers suggest doing balance tests, where we look at baseline outcomes, but also contemporaneous outcomes that we should not expect to change. For instance, the fraction unemployed, or self-employed are likely to be driven by local demand shocks rather than supply-side changes. So the lack of a change in such outcomes (and in baseline education levels), seen in Table B12 is also supportive of our analysis.

Sixth, as Borusyak et al. (2020) suggest, we should control for confounders. This has been one of the motivations behind our local-demand shocks controls, where we control for within-India demand shocks that affect our occupation-region shares over time. As such, occupations-by-regions that were differentially affected by industrial growth in India are accounted for. It turns out that these controls do not change the magnitudes nor significance of our estimated parameters.

Seventh, despite our setup being different from the Adao et al. (2019) framework, the intuition from the three papers when doing inference can be examined further. For instance, we can test for correlation in the residuals of the changes in outcomes over time as a function of the baseline shares. Figure B3a suggest that the changes (in residual employment of college graduates) over time are not correlated with the baseline shares. Yet, we should note that since we are not aggregating multiple shocks, there is no standard error correction needed.

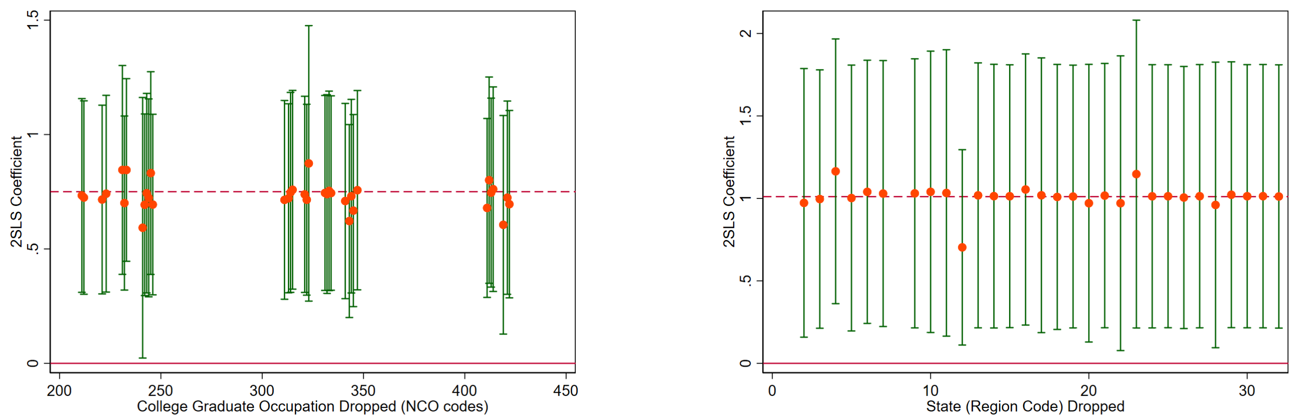
Eighth, we perform a few other tests, inspired by the intuition underlying the Goldsmith-Pinkham et al.

(2020) tests. We are interested in the occupation-choice elasticity between CS and other occupations. Hence, we do not want trends in some other major occupations to be driving our elasticities. In Figure B2a we drop, one-at-a-time, every other occupation and see that our elasticity measure stays constant. We also drop each region, one-at-a-time to show robustness in Figure B2b.

Figure B2: Sensitivity to Occupations and States

(a) Dropping Occupations One at a Time

(b) Dropping Each State One at a Time

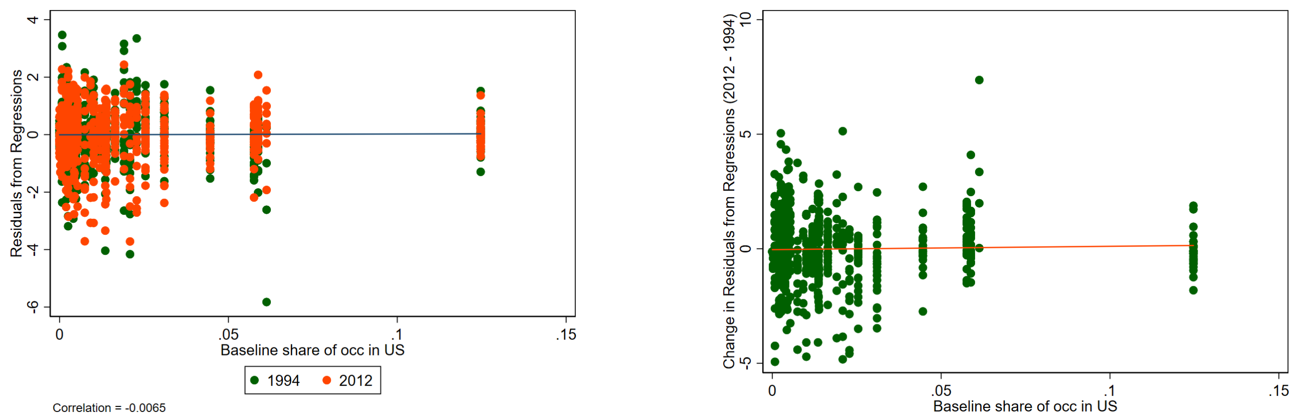


Note: We show robustness of our primary specification dropping one occupation at a time (left panel), and one region at a time (right panel). All regressions include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Main explanatory variable is the log the number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation *o*. Sample is restricted to occupations that are high-skill intensive. We only include occupation-region pairs where workers are found in at least one year. Data for India, from the National Sample Surveys, and data for the US from the US Census, and the American Community Survey. The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3

Figure B3: Correlations Between Growth and Baseline Shares

(a) Employment vs Baseline Share Weight

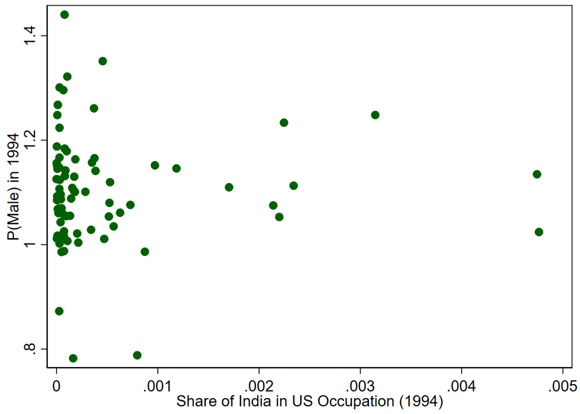
(b) Change in Employment vs Baseline Share Weight



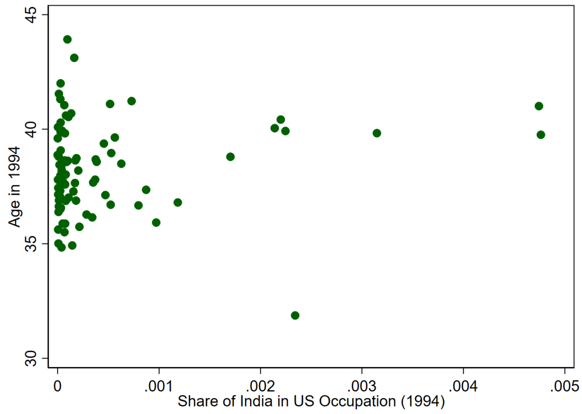
Note: We show the correlation between the baseline occupation-region shares (horizontal axis), and the residual of a 2SLS regression where the primary outcome is the log of employment in an occupation-region, and the explanatory variable is number of Indian migrants in the US (instrumented with the H-1B IV). Data for India, from the National Sample Surveys, and data for the US from the US Census, and the American Community Survey. The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3

Figure B4: Correlations Between Other Outcomes and Baseline Shares

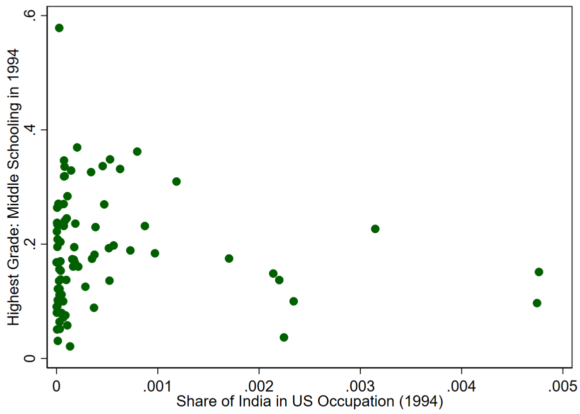
(a) Fraction Male



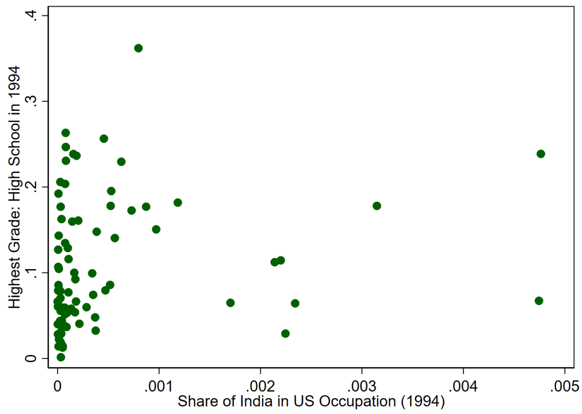
(b) Age Distribution



(c) Highest Grade: Secondary School



(d) Highest Grade: High School



Note: We show the correlation between the baseline shares of the occupation in the US (horizontal axis), and various outcomes in India. Data for India, from the National Sample Surveys, and data for the US from the US Census, and the American Community Survey. The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3

Table B10: Regression in Levels

Panel A: Main Specification	**Baseline IV**		**With return migrants**	
	Employment All	Employment Young	Employment All	Employment Young
Indians in US	2.085***	1.537***	2.004***	1.484***
migrated in past 5 years	(0.392)	(0.290)	(0.376)	(0.279)
N	3,224	3,224	3,224	3,224
Elasticity	0.332	0.375	0.319	0.363
SE	(0.062)	(0.071)	(0.060)	(0.068)
1st stage F-stat	61.04	61.04	62.03	62.03

Panel B: Alternative IVs	**Non Indian Migrants**		**Non Indian Migrants by Occupation**	
	Employment All	Employment Young	Employment All	Employment Young
Indians in US	1.360***	1.019***	1.844***	1.376***
migrated in past 5 years	(0.303)	(0.209)	(0.353)	(0.261)
N	3,224	3,224	3,224	3,224
Elasticity	0.216	0.249	0.293	0.336
SE	(0.048)	(0.051)	(0.056)	(0.064)
1st stage F-stat	51.98	58.23	51.98	58.23

Panel C: With regional reweighting	**H-1B cap IV**		**Non Indian Migrants to US**	
	Employment All	Employment Young	Employment All	Employment Young
Indians in US	48.57***	36.89***	31.42***	23.04***
migrated in past 5 years	(13.63)	(10.59)	(9.825)	(6.727)
N	3,087	3,087	3,087	3,087
Elasticity	0.452	0.528	0.293	0.33
SE	(0.127)	(0.152)	(0.092)	(0.096)
1st stage F-stat	11.16	11.16	14.96	14.96

Panel D: with occupation- -by-region specific trends	**H-1B cap IV**		**Non Indian Migrants to US**	
	Employment All	Employment Young	Employment All	Employment Young
Indians in US	1.857***	2.596***	3.371***	1.420***
migrated in past 5 years	(0.668)	(0.905)	(1.175)	(0.450)
N	3,224	3,224	3,224	3,224
Elasticity	0.295	0.634	0.536	0.347
SE	(0.106)	(0.221)	(0.187)	(0.110)
1st stage F-stat	32.38	32.38	6.729	6.729

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$ Levels form of the primary (log-log) specifications. Panel A-C include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Panel D include region-time fixed effects, and trends for each occupation-region pair. Columns 1 and 3 show 2SLS regressions for the dependent variable of *level of* employment of all college graduates in India while columns 2 and 4 count only college graduates who are between 25 and 40. Main explanatory variable is the *level of* the number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o. Demand Control is the level formulation of equation 3. Sample is restricted to occupations that are high-skill intensive. Final sample includes 38 occupations, 30 Indian states and 5 periods. We only include occupation-region pairs where workers are found in at least one year. Data for India, from the National Sample Surveys (1994, 2000, 2005, 2010 and 2012), and data for the US from the US Census (1990, 2000), and the American Community Survey (2005, 2010, 2012). The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3. Panel A shows the results with the primary policy-induced H-1B instrumental variables. Panel B shows the results where the time-shifter in the instrument is the number of non-Indian migrants to the US. Panel C shows the formulation where we reweight the explanatory variable and instrument by the baseline occupation-specific region-share in India. Panel D controls for occupation-specific trends.

Table B11: Falsification Tests of Pre-trends

	Baseline		Without CS Occupations		No Demand Controls	
	Log Emp All	Log Emp Young	Log Emp All	Log Emp Young	Log Emp All	Log Emp Young
Log Indians in US (t+1)	-0.149	0.468	-0.291	0.464	-0.316	0.258
migrated in past 5 years	(0.322)	(0.434)	(0.388)	(0.536)	(0.330)	(0.427)
Demand Control	1.093***	1.379***	1.185***	1.387***		
	(0.360)	(0.415)	(0.357)	(0.407)		
Observations	1,833	1,833	1,811	1,811	1,833	1,833
F stat	37.72	37.72	31.45	31.45	41.92	41.92

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$. All regressions include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Columns 1, 3 and 5 show 2SLS regressions for the dependent variable of log employment of all college graduates in India while columns 2, 4 and 6 count only college graduates who are between 25 and 40. Main explanatory variable is the log the number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o, but artificially moved forward by one period (5 years). Demand Control is as in equation 3. Sample is restricted to occupations that are high-skill intensive. We only include occupation-region pairs where workers are found in at least one year. Data for India, from the National Sample Surveys, and data for the US from the US Census, and the American Community Survey. The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3. Columns 1 and 2 show the baseline specification. Columns 3 and 4 exclude computer science (CS) occupations. Column 5 and 6 exclude demand controls. Year for explanatory variables include 2000, 2005 and 2010, whereas years for dependent variables include 1994, 2000 and 2005. As such, the explanatory variables are moved forward by one period. Periods for the control variables are kept concurrent with the outcome variables.

Table B12: Correlations with Baseline Shares and Labor Market Changes

	Education Levels at Baseline			Labor Characteristics		
	Primary	Secondary	Upper Secondary	Unemployed	Family Enterprise	Domestic Work
Log Indians in US	-0.0157	0.0535	0.00366	-0.00760	-0.00943	-0.0113
migrated in past 5 years	(0.0273)	(0.0465)	(0.0628)	(0.00463)	(0.00991)	(0.0161)
Observations	1,154	1,154	1,154	3,114	3,114	3,114
F stat	33.43	33.43	33.43	67.64	67.64	67.64

Notes: $*** = p < 0.01, ** = p < 0.05, * = p < 0.1$. Correlations with education (first three columns) and other labor market outcomes (last three columns). All regressions include occupation-region and region-time fixed effects. SE clustered at the occupation-region level. Main explanatory variable is the log the number of Indian college graduates, who migrated to the US in the past 5 years to work in occupation o. Sample is restricted to occupations that are high-skill intensive. We only include occupation-region pairs where workers are found in at least one year. Data for India, from the National Sample Surveys, and data for the US from the US Census, and the American Community Survey. The H-1B cap changes are as recorded by the Department of Homeland Security. Details of the data construction can be found in Appendix A.3. In columns 1-3 we only examine outcomes in the first two periods (1994 and 2000) as we are studying education attainment at baseline. In columns 4-6 we use all periods.

Table B13: Correlations with Baseline Shares in the Shift-share

	Share of Indians in US (1994) x 100			
P(Male) in 1994	-0.412	-0.45	-1.12	-1.212
	(0.209)	(0.214)	(0.311)	(0.292)
Age in 1994	0.0328	0.0366	0.00716	-0.00549
	(0.00984)	(0.0102)	(0.0169)	(0.0153)
Highest Grade: Primary Schooling in 1994	-1.852	-2.108	-2.183	0.128
	(0.251)	(0.265)	(0.499)	(0.532)
Highest Grade: Middle Schooling in 1994	-0.521	-0.616	-1.694	-0.180
	(0.220)	(0.227)	(0.354)	(0.369)
Highest Grade: High School in 1994	1.139	1.203	-0.369	0.725
	(0.307)	(0.314)	(0.454)	(0.452)
Unemployed and Looking for Work 1994	-12.04	-9.097	-142.8	-87.91
	(26.98)	(27.20)	(193.4)	(183.4)
Self Employed 1994	-0.229	-0.185	-0.936	-0.408
	(0.177)	(0.181)	(0.364)	(0.347)
Family Worker 1994	-0.870	-0.830	3.661	6.731
	(0.633)	(0.646)	(2.032)	(1.937)
Domestic Worker 1994	0.279	-0.846	-3.546	-3.317
	(6.843)	(6.947)	(7.956)	(7.380)
College Graduates 1994				2.323
				(0.306)
Constant	0.642	0.580	3.511	2.106
	(0.472)	(0.490)	(0.809)	(0.772)
Region Fixed Effects	No	Yes	Yes	Yes
Occupations	All	All	Skilled	Skilled
Observations	2,432	2,432	954	954
Adjusted R-squared	**0.0386**	**0.0331**	**0.0278**	**0.105**

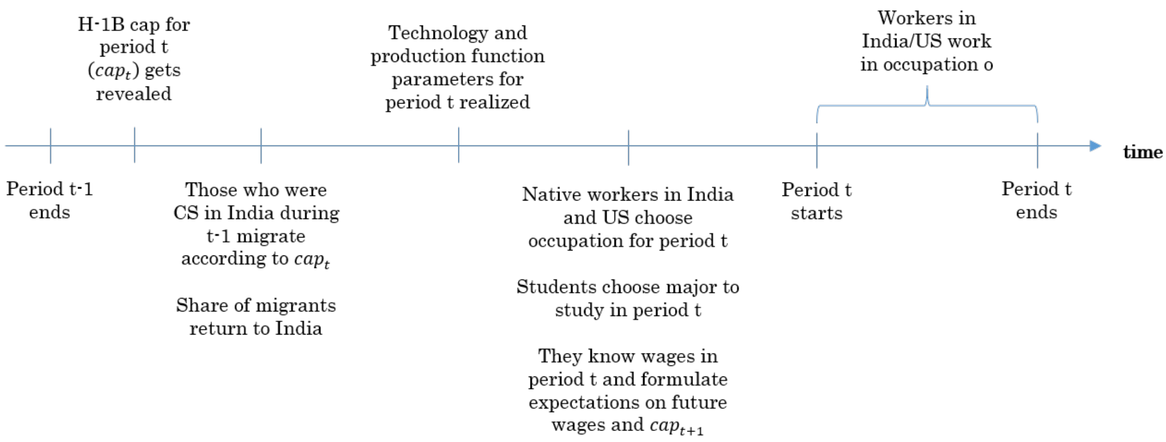
Notes: Correlations between the baseline share and other outcomes. Sample is the last two columns is restricted to occupations that are high-skill intensive. Last three columns include region fixed effects. We only include occupation-region pairs where workers are found in at least one year. Data for India, from the National Sample Surveys, and data for the US from the US Census, and the American Community Survey. Details of the data construction can be found in Appendix A.3.

C Model Timeline, CES Derivations and Equilibrium

C.1 Model Timeline

Figure C1 summarizes the timing of the decisions each period, for college graduates in the model.

Figure C1: Decision timeline for college graduates in the model.



C.2 Combining Occupations into Groups

In the empirical section we work with multiple occupations in order to leverage sufficient variation on the migration probabilities. To keep the model tractable, we focus on computer science and non-CS occupations. Occupations in the production function are aggregated in a CES manner, so it is possible to interpret the group of "other occupations" as a CES composite with the same elasticity of substitution between Other and CS. We reconcile this in the equations below:

$$\underbrace{\left[\sum_{o} \gamma_o x_o^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}}}_{\text{Regression-based production function}} = \left[\gamma_{cs} x_{cs}^{\frac{\lambda-1}{\lambda}} + \sum_{o' \neq cs} \gamma_{o'} x_{o'}^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}} = \left[\gamma_{cs} x_{cs}^{\frac{\lambda-1}{\lambda}} + \left(\sum_{o' \neq cs} \gamma_{o'} x_{o'}^{\frac{\lambda-1}{\lambda}}\right)^{\frac{\lambda}{\lambda-1}\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}}$$
$$= \left[\gamma_{cs} x_{cs}^{\frac{\lambda-1}{\lambda}} + \bar{\gamma} G^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}}, \quad \text{where } G \equiv \left[\sum_{o' \neq cs} \frac{\gamma_{o'}}{\bar{\gamma}} x_{o'}^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}} \tag{C1}$$

and x_o is the number of workers in each occupation. G is interpreted as a composite of occupations. Given the CES assumption the elasticity of substitution between CS and any other occupation $o' \neq CS$ is the same as the elasticity of substitution between the composite and CS. If we assume all college occupations other than CS are identical for workers and have an equal productivity in the production function ($\gamma_{o'} = \gamma$), the wage for each occupation other than CS would be the same and it is possible to re-write the G aggregate as in equation C2. If the total number of non-CS occupations is N and the total number of workers in non-CS occupations is $\bar{G} = \sum_{o'} x_{o'}$:

$$G = \left[\sum_{o' \neq cs} \frac{\gamma_{o'}}{\bar{\gamma}} x_{o'}^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}} = \left[N^{\frac{1}{\lambda}} \frac{\gamma}{\bar{\gamma}} (N x_o)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}} = \underbrace{N^{\frac{1}{\lambda}} \frac{\gamma}{\bar{\gamma}}}_{\delta_g} \bar{G} \tag{C2}$$

Replacing the result in equation C2 on the CES aggregate of college graduates gives us the function used in the model:

$$\left[\gamma_{cs} x_{cs}^{\frac{\lambda-1}{\lambda}} + \delta_g \bar{\gamma} \bar{G}^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}} \equiv \underbrace{\left[\delta\left(cs_{v,y}\right)^{\frac{\lambda-1}{\lambda}} + (1-\delta)\left(g_{v,y}\right)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}}}_{\text{Model-based function}} \tag{C3}$$

C.3 Equilibrium

Equilibrium in each period can be defined as a set of prices and wages ($P_{t,c,k}$, $P_{t,y,k}$, $w^{\ell}_{t,k}$, $w^{g}_{t,k}$, $w^{cs}_{t,k}$), quantities of output and labor ($C_{t,y,k}$, $Y_{t,k}$, $L_{t,k}$, $G_{t,k}$, $CS_{t,k}$), and the level of technology ($T^{s}_{t,k}$) such that: (1) consumers in the US, India, and the rest of the world, maximize utility by choosing $Y_{t,k}$ taking prices as given, (2) college graduates in the US and India choose their field of major and occupations, taking wages as given, and forming expectations, (3) firms in both the IT and the final goods sector maximize profits taking wages and prices as given, (4) trade between the three regions is balanced, and (5) output and labor markets clear.

Given the Frechet distribution assumption, we can aggregate across varieties and write the probability of country k buying goods of sector s from country b as in equation C4:

$$\pi^{s}_{t,k,b} = \frac{T^{s}_{t,b}(d^{s}_{t,k,b}\xi^{s}_{t,b})^{-\theta}}{\sum_b T^{s}_{t,b}(d^{s}_{t,k,b}\xi^{s}_{t,b})^{-\theta}}\,, \tag{C4}$$

where $\xi^{s}_{t,b}$ is the unit cost of production in country b, $d^{s}_{t,k,b}$ the iceberg trade cost of trading goods from b to k and $T_{t,b}$ is country b technology level. The price index in country k, sector s is a combination of production costs and technologies of different suppliers weighted by trade costs between each supplier and country k:

$$P^{s}_{t,k} = \bar{\gamma}\left(\sum_{k'} T^{s}_{t,k'}(d^{s}_{t,k,k'}\xi^{s}_{t,k'})^{-\theta}\right)^{-\frac{1}{\theta}}\,, \tag{C5}$$

where $\bar{\gamma} = \left(\Gamma\left(\frac{1-\iota}{\theta}+1\right)\right)^{\frac{1}{1-\iota}}$, and Γ is the Gamma function. As we assume trade balance, total income from sector Y in country k has to equal the sales to each of the markets. Similarly for the IT sector, we have that total income earned has to equal the the total sales of intermediate IT goods sold to each country.

Labor markets clear as long as total demand for each occupation in country k equals the total supply of labor for that occupation. Non-college workers' supply is fixed at $\bar{L}_k$ in both countries. Native college graduates in both countries face the decision of whether to work as CS or in non-CS college occupations. This decision has an inter-temporal dimension that requires us to define the dynamic equilibrium in the labor market for college graduates. Using the properties of the Type I Extreme Value, we derive the share of workers with occupation $o_{t-1} = o$ who choose occupation o' in period t as in equation C6:

$$\pi^{o,o'}_{t,k,a} = \frac{exp(\frac{1}{\sigma_k}\bar{V}^{o,o'}_{t,k,a})}{exp(\frac{1}{\sigma_k}\bar{V}^{o,o'}_{t,k,a}) + exp(\frac{1}{\sigma_k}\bar{V}^{o,o}_{t,k,a})}\,, \tag{C6}$$

where $\bar{V}^{o,o'}_{t,k,a} = w^{o'}_{t,k} + \chi_k(a)\times \mathbb{1}(o\neq o') + \zeta_k \times \mathbb{1}(o'=g) + \beta\mathbb{E}_t[V^{o'}_{t+1,a+1,k}]$. As in Bound et al. (2015), this equilibrium is characterized by the system of equations (10-15 and C4-C6) and a labor demand shifter Ω_t through which we characterize the expectations of workers with respect to future career prospects. A unique equilibrium is pinned down in each period by an aggregate labor demand curve in each country for native computer scientists relative to other college graduates. The labor demand shifter Ω_t represents the change over time in the production function parameters in both countries ($\delta_{k,t}$, $\alpha_{t,k,y}$, $T^{s}_{k,t}$, $d^{s}_{k,b,t}$, $\gamma_{k,t}$, and cap_t) that shift the relative labor demand curve of CS relative to other college occupations. Intuitively, the change in the production function parameters reflects the skill-biased technological change toward CS and IT that can be interpreted as a innovation shocks (as seen by the IT boom) to the labor market that push workers to switch to CS occupations. Individuals have rational expectations on the evolution of these demand shocks and make decisions based on that, and the expectations they have over their preference shocks and migration probabilities.

The equilibrium in the labor market is a mapping from the exogenous demand shifter, Ω_{t-1} and state variables: $s = \{\mathbf{CS^a_{t-1,k}}, \mathbf{G^a_{t-1,k}}\}$ for all ages a to the values of $CS_{t,k}$, $G_{t,k}$, $w_{t^{cs},k}$, $w^g_{t,k}$, and $\mathbf{V_t}$, the vector of career prospects at different occupations for different ages, that satisfies the system of equations 10-15, C4-C6 as well as each period's relative demand curve.

A note on uniqueness: While solving for equilibrium with multiple different initial values, our algorithm converges to the same unique solution. Issues of multiple equilibria arise, for instance, in models of economic geography, where agglomeration forces in a city attract workers to that city away from other cities (Allen et al., 2020). In such instances, authors suggest they study a counterfactual equilibrium that lies near the observable steady state real world (Ahlfeldt et al., 2015). Yet, in our context, even if β (the spillover parameter) were high, it increases demand for both CS *and* non-CS graduates as the spillover raises the productivity of *all* workers. So our analogous occupation-choice (instead of city-choice) is unlikely to generate multiple equilibria. In addition, since the total supply of workers is fixed, and the number of migrants is capped (by policy), this imposes a limit on how much companies can produce, as wages will start increasing until unit costs are sufficiently high that it will be optimal to not increase production.

D Estimation of Other Model Parameters

D.1 Estimating τ

For the elasticity between college and non-college graduates, we replicate Card and Lemieux (2001) using the India NSS data and estimate an elasticity of complementarity of 0.55 (see Table D1). This corresponds to $\tau = 1.8$, and is statistically indistinguishable from 1.7, which is the value used in the calibration for both countries.

Table D1: Estimating τ: Doing Card and Lemieux (2001) in India

	Log(Col Wage / HS wage)
Log(Col L / HS L)	-0.553***
	(0.140)
Log(Col L / HS L) – Log(Col by age / HS by age)	0.322**
	(0.137)
Observations	60
R-squared	0.857
Fixed effects	Cohort, Year
Elas of Sub ($\hat{\tau}$)	1.8
Prob > χ^2	0.000

We estimate τ in India using the National Sample Survey. We follow Card and Lemieux (2001) and divide the working-age population into 10 equally spaced age groups, and by whether or not they are college graduates. 'Col' represents having a college degree, whereas 'HS' is only a high school graduate. L is the number of workers in a college-age bin. The elasticity of substitution is the inverse of the estimated coefficient on $Log(Col\ L/HS\ L)$. This elasticity is precisely measured as indicated by the χ^2 test. It is statistically different from 0 but not 1.7.

D.2 Endogenous Technology: Patenting Response to CS

In Section 4.2.4, equation 23, we mentioned that the level of technology of the IT sector depended on the number of CS working in IT. Parameter β is the elasticity of the technology level with respect to the number of CS working for the IT sector in country k, time t. One advantage of the procedure we use, is that we can estimate the $T^s_{t,k}$ in equilibrium so our estimate will already capture the baseline level of technology plus any endogenous component that affects the overall level of technology.

To estimate the elasticity of technology with respect to the number of CS workers, we use a shift-share instrumental variables strategy that interacts the industry-specific dependence on immigrant CS workers in a pre-period with the total number of immigrant CS in the US each year. Our instrument leverages variation in US immigration policy (like changes to the H-1B cap), sending country shocks, and the fact that immigrants are more likely to be CS.

Our aim is to estimate the parameter β that is shown in equation 23. We first combine our data on the number of CS workers by industry with data on patenting from the US Patent and Trademark Office (PTO) to proxy for the technology level. We use firm-level measures of patents granted from US PTO, match the firms to Compustat data, and then use the Compustat industry identifiers to compute industry-level measures of patenting. However, a simple OLS regression of patenting on the number of CS workers would be biased, as when industries increase investments in R&D they may concurrently increase hiring of CS workers.

To isolate variation in the size of the CS workforce by industry that is not driven by confounding factors, we use the fact that immigrants are concentrated in CS occupations, and the H-1B cap fluctuations affect the size of immigrant flows. In the vein of a modified shift-share instrument, we use the baseline dependence of an industry on immigrant computer scientists interacted with the differential growth in immigrant CS across industries, as an instrument for the CS workforce by industry.[28] Equation D1 captures the first stage of our strategy:

$$CS_{j,t} = \delta_j + \delta_t + \gamma \left(\frac{Imm\ CS_{j,0}}{Emp_{j,0}} \right) Imm\ CS_t + \varepsilon_{j,t}\ , \tag{D1}$$

where $CS_{j,t}$ is the number of computer scientists in industry j and year t, and $\left(\frac{Imm\ CS_{j,0}}{Emp_{j,0}} \right)$ is the baseline (in 1994) share of the workforce in industry j that is an immigrant computer scientist. $Imm\ CS_t$ is the number of immigrant CS workers in the US over time. The interaction between these two terms is the excluded instrument, conditional on industry δ_j and year δ_t fixed effects. Importantly, our instrument leverages variation in US immigration policies (say, changes to the H-1B cap), and the fact that immigrants are more likely to be CS. In our second stage, we study patenting activity:[29]

$$Log(Patents)_{j,t} = \delta_j + \delta_t + \beta_{tech}\widehat{CS_{j,t}} + \varepsilon_{j,t}\ , \tag{D2}$$

The results of this exercise are shown in Table D2. Our first stage is strong, and our 2SLS analysis produces an elasticity that lies between 0.226 and 0.24. We conduct a variety of robustness checks, where we vary the controls in the regression, look at only the flow of new patent filings, and exclude the truncated patent data.[30] Importantly, our estimated elasticity is very close to similar findings in the literature. Peri et al. (2015) estimate that a 1% increase in total US STEM workforce would increase average TFP by 0.27%, whereas Kerr and Lincoln (2010) find patenting elasticities that lie between 0.1 and 0.4. In our earlier work, Bound, Khanna, and Morales (2016), we use an elasticity of 0.23 that we measure by studying how the price of IT goods change with changes in the CS workforce. In recent work, Khanna and Lee (2018) find an elasticity of 0.2 when using measures of innovation derived from the Schumpeterian growth literature. In Section 7.2, we discuss how our results change for different values of β, including the case where there is no spillover in India.

D.3 Calibrating Time-Varying Parameters Parameters

The Cobb Douglas parameters γ_k represent the share of income from the final goods sector spent on varieties of the IT sector. We determine the parameters for $k = \{us, in\}$ from the share of IT output to total output in each country, using data from the OECD, and get values: 0.7%-1.6% for US and 0.2%-1.5% for India. By solving for these parameters every year, we capture changes in demand for IT varieties as an input into the final good production, which is increasing for both countries throughout the period. We do the same exercise to determine the IT share in the rest of the world (RoW).

The share parameter of non graduates in the production function, $\alpha_{t,k}$, is determined in both India and the US such that it matches the observed share of expenditures from the final goods sector in non graduates. Specifically, from the US March CPS and the Indian NSS data, we calculate the share of expenditures on non-graduates $\vartheta_{t,k}$ and the number of graduates and non graduates in the final goods industry $\bar{L}_{t,k}$, $\bar{H}_{t,k}$, and using equation D3 we estimate $\alpha_{t,k}$:

$$\vartheta_{t,k} = \frac{\alpha_{t,k}(\bar{L}_{t,k})^{\frac{\tau-1}{\tau}}}{\alpha_{t,k}(\bar{L}_{t,k})^{\frac{\tau-1}{\tau}} + (1-\alpha_{t,k})(\bar{H}_{t,k})^{\frac{\tau-1}{\tau}}} \tag{D3}$$

Importantly, $\alpha_{t,k}$ decreases over time in both countries, capturing how skill-biased technological change shifts production to college-graduate occupations over time.

[28] The Card (2001) method derives an instrument for immigrants by region. Instead, here we create an instrument for CS workers by industry. Instead of exploiting regional migrant-networks, our underlying variation is driven by changes to H-1B caps and the fact that immigrants are likely to be CS.

[29] We use NBER patent data project which matches patent assignees to the North American Compustat database between 1976-2006. This also includes utility patents. We then match these data to the Compustat database, which includes industry (NAICS) codes. We create a crosswalk between NAICS and the 1990 Census codes available on IPUMS, using the 2000 Census, which helps us measure the number of patents filed by industry in each year till 2006. We exclude 2006 as we believe the data are artificially censored in the last year, as evidenced by the small number of patents.

[30] As there is a lag between patents filed and granted, we exclude the last year as a robustness check.

Table D2: Migration, Computer Scientists and Patenting by Industry

	CS workers	Log(Patents)	Log(Patents)	Log(Patents)	Log(Patents)	Log(New Patents)
Shift Share	1.258***					
	(0.309)					
CS workers		5.45e-06**	5.79e-06**	3.99e-06**	5.53e-06**	5.90e-06*
		(2.59e-06)	(2.56e-06)	(1.89e-06)	(2.27e-06)	(3.10e-06)
Observations	275	275	275	250	250	250
R-squared	0.877	0.967	0.967	0.971	0.968	0.329
Number of Industries	25	25	25	25	25	25
Additional Controls	No. of Firms	No. of Firms	None	No. of Firms	None	No. of Firms
F stat		16.67	16.31	12.54	11.11	12.54
Elasticity		0.226	0.240	0.162	0.225	0.240
SE		(0.107)	(0.126)	(0.076)	(0.092)	(0.126)

Notes: Two-staged least squares regressions of Log(Patents) on the number of computer science workers by industry. Years 1994 to 2005. Controls include year and industry fixed effects, size of total industry workforce, and when mentioned, the number of firms. Sample restricted to (the 25 top) industries that have at least a total of 50 patents over the entire period. In Columns 4-5 we exclude 2005 for robustness (as the data records all patents granted by 2006, there may be truncation based on patents applied for in 2005 but not granted by the end of 2006.) The last specification has the natural log of new patents (i.e. $patents_{t,i} - patents_{t+1,i}$) as the dependent variable. Standard errors clustered by industry.

The distributional parameter between CS and non-CS college graduates $\delta_{t,k}$ is calibrated so that it matches the within-country relative wages between CS and non-CS college graduates observed in the data. Empirically, $\delta_{t,k}$ increases over time, capturing how shifts in skill-biased technology increase the labor share of CS workers. The additional distributional parameter in the IT sector $\Delta_{t,k}$ captures the extra intensity of CS in IT. We calibrate $\Delta_{t,k}$ to be proportional to $\delta_{t,k}$ every period such that it matches the within-country relative share of CS between the IT and non-IT sector in 1995.

To estimate the productivity levels ($T^s_{t,k}$) and bilateral trade costs ($d^s_{t,k,b}$) for each country-sector pair we use trade data such that we match the observed trade flows every year. We follow standard estimation procedures in the trade literature – specifically, the approach of Eaton and Kortum (2002) and Levchenko and Zhang (2016), by using the gravity equations of the model to estimate trade costs and technology parameters. As a first step we use equation C4 and take the ratio between the probability of country k buying from country b and the probability of country k buying from itself, which yields the gravity equation D4:

$$\frac{\pi^s_{t,k,b}}{\pi^s_{t,k,k}} = \frac{EX^s_{t,k,b}}{EX^s_{t,k,k}} = \frac{T^s_{t,b}(d^s_{t,k,b}\xi^s_{t,b})^{-\theta}}{T^s_{t,k}(d^s_{t,k,k}\xi^s_{t,k})^{-\theta}}\,, \tag{D4}$$

where $EX_{t,k,b}$ is the value of expenditures that country k has on products from country b in sector s at time t. Using data on bilateral trade flows and domestic consumption by sector and year for the US, India, and a series of 57 countries, we use equation D4 to estimate trade costs and a term that combines the technology level and the unit cost of production $T^s_{t,k}(\xi^s_{t,k})^{-\theta}$.

First, we parametrize the trade costs as in equation D5. Following Levchenko and Zhang (2016) we define log of trade costs as a function of distance ($dist_{k,b}$), an indicator on whether the two countries share a border $border_{k,b}$, an indicator on whether the two countries belong to a currency union $CU_{t,k,b}$ and an indicator for participating in a regional trade agreement $RTA_{t,k,b}$. We also allow the trade costs to be affected by an exporter fixed effect $exp_{t,k}$ and an error term $v_{t,k,b}$.

$$log(d^s_{t,k,b}) = dist_{k,b} + border_{k,b} + CU_{t,k,b} + RTA_{t,k,b} + exp^s_{t,k} + v^s_{t,k,b} \tag{D5}$$

To estimate trade costs and technology, we take logs of equation D4 and get equation D6 which can be estimated by OLS, and will allow us to back out the trade costs and a term that combines the technology level and the unit cost of production $T^s_{t,k}(\xi^s_{t,k})^{-\theta}$.

$$log\left(\frac{EX^s_{t,k,b}}{EX^s_{t,b,b}}\right) = \underbrace{log\left((T^s_{t,k}(\xi^s_{t,k})^{-\theta}\right) - \theta exp_{t,k}}_{\text{Exporter fixed effect}} - \underbrace{log\left((T^s_{t,b}(\xi^s_{t,b})\right)^{-\theta}}_{\text{Importer fixed effect}} \tag{D6}$$
$$-\theta(dist_{k,b} + border_{k,b} + CU_{t,k,b} + RTA_{t,k,b} + \nu_{t,k,b})$$

The distance variable $dist_{k,b}$ is a group of six indicator variables that take the value of 1 if the distance between k and b falls within each of the following intervals measured in miles: [0, 350], [350, 750], [750, 1500], [1500, 3000], [3000, 6000], [6000, max), and 0 otherwise. We also assume trade costs are equal to 1 if the country is buying from itself, so trade costs only arise due to international trade.

We estimate equation D6 separately by sectors Y and C. Distance, currency union, and regional trade agreement data comes from the CEPII gravity database and data on industry-specific trade flows and GDP come from the OECD. We run the regression separately for 1995, 2000, 2005, and 2010. From the estimates of equation D6 we can back out the trade costs conditional on our preferred value of θ. The fixed effects represent a convolution of the relative technology levels and unit costs of a country. To estimate them, we normalize the fixed effect for the US, and interpret the estimate for each country as the relative technology levels and unit costs between a country k and the US, and derive estimate $\hat{\Xi}^s_{t,k,us} = \left(\frac{T^s_{t,k}\xi^s_{t,k}}{T^s_{t,us}\xi^s_{t,us}}\right)^{-\theta}$.

When computing our GE model, we feed the estimated expressions for $\hat{\Xi}^s_{t,k,b}$ together with our preferred value of θ and total labor quantities for each occupation. This allows the model to endogenously calculate wages based on the labor market clearing conditions and pin down the unit costs, which in turn allows us to back out the level of technology relative to the the US. We assume the technology level for the US is 1 for both sectors. For RoW, we get estimates for 57 countries. For each country in the RoW, we calculate the country weights based on the country-sector GDP. We then use these weights to compute a weighted average of trade costs and $\hat{\Xi}^s_{t,k}$.

For relative technology in the non-IT sector between India and the US, we choose $\left(\frac{T^y_{t,in}}{T^y_{t,us}}\right)$ such that we match relative wages of non graduates between India and the US. Finally, we determine the number of college and non-college workers in each country using the US March CPS and the Indian NSS data. The total number of Indian CS who migrate to the US is calculated using administrative H-1B data, March CPS, and the American Community Survey, which provides information on birthplace and occupation. We calculate the observed migration cap every period and match the net change in the number of Indian CS in the US.

D.4 Identifying Dynamic Labor Supply Parameters

The Simulated Method of Moments (SMM) approach implies that all parameters are jointly estimated to match a series of targeted moments. For each parameter to be identified separately, we need each of the targeted moments to respond differently to changes in each parameter. In Section 5.3, we provide intuition on what moments in the data help us identify each labor supply parameter. In this section, we provide further evidence that confirms our intuition and explicitly relates each parameter to specific targeted moments.

We follow Lagakos et al. (2018) and calculate the elasticity of each targeted moment to a 1% change in each parameter. Once we estimate the supply side parameters, we exogenously increase each parameter estimate by 1% and compute the targeted moments predicted by the model. We then compute the change in the targeted moment with respect to the 1% increase in each of the parameters as shown in Table D3. As an example, in the first column, a 1% increase in the mean taste for non-CS in the US ζ_{us} causes a 0.59% decrease in the CS share in 1995, a 2.62% decrease in the CS share in 2010, and a 2.32% decrease in the transition rate across occupations in the US.

While it is true that changes in some parameters seem to have an effect on many moments, different parameters tend to affect disproportionately specific moments, in a manner consistent with our intuition. For example, ζ_{us} and σ_{us} tend to affect moments in the US and India since they determine the occupation choices

Table D3: Elasticity of targeted moments with respect to parameters (%)

	ζ_{us}	σ_{us}	$\chi_{0,us}$	$\chi_{1,us}$	$\bar{F}_{us}$	ζ_{in}	σ_{in}	$\chi_{0,in}$	$\chi_{1,in}$	$\bar{F}_{in}$	κ
Share CS 1995 - US	-0.59	0.37	-0.10	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Share CS 2010 - US	-2.62	1.30	-0.10	-0.06	-0.03	0.00	0.00	0.00	0.00	0.00	0.00
Ratio CS share [25-30]/[31-60] - US	-1.91	1.63	-0.15	-0.34	-0.04	0.00	0.00	0.00	0.00	0.00	0.00
Transition rate - US	-2.32	2.37	-1.01	-0.30	-0.09	0.00	0.00	0.00	0.00	0.00	0.00
Ratio CS share [45-60]/[31-60] - US	2.24	-1.05	0.10	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00
Share CS 1995 - India	-0.06	0.01	0.01	0.00	0.00	-1.42	3.57	-2.05	-0.24	0.01	0.22
Share CS 2010 - India	0.76	-0.37	0.03	0.02	0.01	-2.05	0.20	0.31	-0.04	0.00	1.41
Ratio CS share [25-30]/[31-60] - India	-1.11	0.55	-0.01	-0.02	-0.01	0.86	1.37	-1.84	-0.05	-0.03	-0.30
Transition rate - India	-0.19	0.14	-0.05	-0.02	0.00	0.39	2.54	-3.20	-0.05	-0.04	0.33
Ratio CS share [45-60]/[31-60] - India	-1.02	0.48	-0.02	-0.02	-0.01	0.82	3.92	-2.99	-0.27	0.00	-1.33
Supply response to migration	0.09	-0.08	0.02	0.00	0.00	0.60	-0.64	0.06	0.03	0.00	0.22

We calculate percent changes in each targeted moment when each supply-side parameter changes by 1%. ζ_{us} and ζ_{in} are the mean taste parameters for non-CS occupations, σ_{us} and σ_{in} are the taste dispersion parameters, $\chi_{0,us}$ and $\chi_{0,in}$ are the baseline occupation switching costs, , $\chi_{1,us}$ and $\chi_{1,in}$ are the age-specific occupation switching costs, $\bar{F}_{us}$ and $\bar{F}_{in}$ are the CS education fixed costs, and κ is the utility weight on US wages.

for US workers over time. However, the CS shares in the US are the moments that predominantly respond to changes in those parameters. Changes in the occupation costs $\chi_{0,us}$ and $\chi_{0,in}$ lower the transition rates in each country respectively. The age specific switching costs $\chi_{1,us}$ increases the "Ratio CS share [45-60]/[31-60]" for the US by preventing older workers to switch to other occupations while in India, parameter $\chi_{1,in}$ decreases such ratio by preventing older workers to switch into CS. Higher education costs $\bar{F}_{us}$ and $\bar{F}_{in}$ decrease the CS share of young workers between 25-30 relative to older ones in each country. Finally, a 1% increase in κ, increases the supply response to immigration by 0.22%.

E Additional Tables for Model Counterfactuals

Table E1: Compensating Variation in 2010 By Worker

	Total Welfare (USD mn)		**Welfare per Migrant (USD)**	
	US	**India**	**US**	**India**
Always CS	-1268	-348	-9661	-2651
Switchers from CS to non-CS	-38	-121	-290	-923
Always non CS	-675	768	-5147	5852
Non college	2273	41	17324	316
Return CS				
Total Welfare non-migrants	292	310	2227	2364
Welfare of Migrants	6012		45809	
Total	6614		50400	

Compensating Variation in USD (total and per migrant) defined as the amount in USD that must be provided to agents in a world with restricted migration to provide them with the same welfare as in a world with H-1B migration. In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995.

Table E2: Effect of Migration Over Time

	1995	2000	2005	2010
Wages				
US CS	0.02%	-0.10%	-0.71%	-0.97%
US non CS college grad	0.00%	-0.02%	-0.03%	-0.03%
US non college grad	0.00%	0.03%	0.07%	0.11%
India CS	1.5%	-9.4%	-9.5%	-11.5%
India non CS college grad	-2.5%	7.2%	7.7%	2.2%
India non college grad	-0.7%	0.9%	1.3%	1.3%
Occupational Choice				
US CS (native plus immigrant)	0.6%	0.9%	1.4%	1.6%
US CS native	-0.5%	-2.2%	-4.2%	-5.9%
India CS	-13.9%	60.4%	66.0%	56.6%
US non CS college grad	0.01%	0.08%	0.16%	0.24%
India non CS college grad	0.02%	-0.67%	-0.96%	-1.04%
IT production				
US IT output	0.9%	1.0%	0.6%	0.1%
India IT output	-9.5%	24.2%	28.7%	35.6%
World IT output	-3.6%	15.3%	21.6%	9.4%
US IT price	-0.2%	-0.3%	-1.0%	-1.3%
India IT price	1.2%	-2.7%	-3.3%	-14.4%
Welfare				
Welfare of US natives	0.003%	0.004%	0.005%	0.007%
Welfare of migrants	25.4%	45.4%	50.0%	41.4%
Welfare in India	-0.26%	1.19%	1.09%	0.13%
Combined welfare	-0.02%	0.24%	0.31%	0.14%

Percent difference on main outcomes when we go from a scenario with restricted migration (counterfactual) to a scenario with full migration (real). In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995.

Table E3: Effect of Migration: Robustness to Parameters

	Baseline	$\tau = 3$	$\theta = 4$	$\lambda = 2.5$	$\varepsilon = 10$
Occupational choice					
US CS (native plus immigrant)	1.6%	1.3%	0.6%	3.5%	1.5%
US CS native	-5.9%	-6.0%	-6.2%	-3.9%	-5.9%
India CS	56.6%	55.8%	50.1%	62.2%	56.3%
Welfare					
Welfare of US natives	0.007%	0.021%	-0.002%	0.021%	0.006%
Welfare in India	0.13%	0.15%	0.09%	0.15%	0.14%
Combined welfare	0.13%	0.11%	0.13%	0.11%	0.14%

Percent difference on main outcomes when we go from a scenario with restricted migration (counterfactual) to a scenario with full migration (real). In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995. We examine the effect of changing elasticities that we determine from the literature. This includes τ (the elasticity of substitution between non-college and college graduates), λ (the elasticity of substitution between CS and non-CS graduates), θ (the dispersion parameter of the Frechet distribution), and ε (the elasticity of substitution between regular CS and return migrant CS in India).

Table E4: Brain Drain vs Brain Gain: Welfare

	Baseline	No occupational choice in India	No migration but reallocation	No return migration
US Welfare				
Always CS	-0.97%	-0.80%	-0.12%	-1.34%
Always non CS	-0.03%	-0.03%	0.00%	-0.03%
Non college	0.11%	0.11%	0.01%	0.15%
India welfare				
Always CS	-12.6%	1.50%	-14.3%	-8.81%
Always non CS	0.93%	-0.17%	0.97%	0.75%
Non college	0.03%	-0.03%	0.03%	-0.01%
Total Welfare				
Welfare US natives	0.007%	0.009%	-0.002%	0.012%
Welfare in India	0.13%	-0.05%	0.15%	0.11%
Combined (with migrants)	0.14%	0.10%	0.04%	0.17%

Percent difference on main outcomes when we go from a scenario with restricted migration (counterfactual) to a scenario with full migration (real). In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995. The 'no occupational choice in India' scenario restricts occupational choice and inhibits 'brain gain' in response to migration opportunities but allows for emigration ('brain gain'). 'No migration but reallocation' restricts migration (no 'brain drain') but allows for responses to migration opportunities ('brain gain'). 'No return migration' assumes return rate $\rho = 0$.

F Alternative model specifications

F.1 Endogenous Technology - Details

Table F1: Varying the Elasticity of Endogenous Technological Spillovers

	Baseline	No spillovers	US spillover=0.23 No spillover India	Spillover = 0.1 in both countries
IT production				
US IT output	0.11%	0.22%	0.28%	0.18%
India IT output	35.55%	21.81%	22.38%	27.56%
Welfare				
Welfare of US natives	0.007%	0.003%	0.006%	0.005%
Welfare in India	0.130%	-0.050%	-0.052%	0.026%
Total (with migrants)	0.142%	0.085%	0.087%	0.111%

Percent difference on main outcomes when we go from a scenario with restricted migration (counterfactual) to a scenario with full migration (real). In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995. In the baseline specification, the spillover elasticity is 0.23. In the scenario where there are no spillovers, the elasticity is 0 for both countries. In the scenario where the US has baseline spillovers and India has no spillovers, the US elasticity is 0.23. We further include a robustness exercise where the spillover elasticity is 0.1 in both countries.

F.2 Imperfect Substitution between Immigrants & Natives - Details

To evaluate how the quantitative results change when we allow for imperfect substitution, we change the US production function as in equation F1:

$$CS_{j,s} = \left[\left(CS^n_{j,s} \right)^{\frac{\upsilon-1}{\upsilon}} + \left(CS^f_{j,s} \right)^{\frac{\upsilon-1}{\upsilon}} \right]^{\frac{\upsilon}{\upsilon-1}}, \tag{F1}$$

where υ is the elasticity of substitution between natives and migrants, which we set to 10 following Ottaviano and Peri (2012).[31] $CS^n_{j,s}$ is the number of native CS workers in the US in firm j sector s, and $CS^f_{j,s}$ is the number of immigrants from India working as CS in firm j, sector s in the US. Under this model specification immigrants and native CS get different wages when working in the US. Results are in Table F2. Combined welfare from migration is higher in a world with imperfect substitution, as when immigrants work with (and complement) natives, overall production increases.

F.3 Endogenous College - Details

In our baseline specification, the decision to go to college is made outside of the model and treated as exogenous. To relax that assumption, we allow individuals to choose at the age of 20, whether they want to join the labor market immediately and work as non-college graduates for the rest of their careers, or go to college, choose either CS or another college major, and join the labor market next period as in the baseline model. We present the college decision in equation F2:

$$\max\{\beta \mathbb{E}_t V^{coll}_{t+1} + \bar{F}_{coll} + \sigma_c \eta^{coll}_{i,t}, w_{ncoll,t} + \beta \mathbb{E}_t V^{ncoll}_{t+1} + \sigma_c \eta^{ncoll}_{i,t}\} \tag{F2}$$

$\eta^{coll}_{i,t}$ and $\eta^{ncoll}_{i,t}$ are *iid* preference draws from a Type I Extreme Value distribution for college and non-college respectively. This specification introduces two additional parameters for each country. $\bar{F}_{coll}$ is a fixed cost of going to college that is necessary to capture the large share of workers not going to college despite the higher wages. σ_c is the preference parameter that controls how sensitive individuals are to the preference shocks. If the value of not going to college is larger, individuals join the labor force immediately and work as non-college graduates until the end of their careers earning wage $w_{ncoll,t}$. If they choose to go to college, they follow equation 10, and choose their college major. Then they follow the baseline model and join the labor force next period either as CS or other college graduate occupation. If they choose to go to college their value functions and decisions are identical to the baseline model.

To estimate the additional four parameters in the SMM routine we add 4 targeted moments to identify them. We use the shares of college graduates among all workers for each country in 1995 and 2010. A high $\bar{F}_{coll}$ implies the shares for both years will be lower on average. A low σ_c implies workers are not as sensitive to preferences and respond more to changes in the wage. The change in the shares between 1995 and 2010 conditional on wages help identify this parameter.

F.4 The Role of Trade, Remittances and the Age-specific switching cost

We also look at the relevance of international trade and remittances. Column 2 of Table F3 shows the impacts of migration for a model where bilateral trade costs between the three regions are 10% lower than in the baseline. We show that welfare for both countries increases by less when there is more trade, as migration becomes less relevant for welfare. This is because when migration gets restricted, consumption cost increases and workers can soften the impact on welfare by buying more products from abroad. The US decreases its IT production when there is more trade, going from an increase of 0.11% in the baseline to a decrease of 0.07% in the more open scenario.

As an additional channel we test how our conclusions would change if we incorporate remittances into the

[31] When looking at educated workers, Ottaviano and Peri (2012) find an elasticity of substitution of 12.6. Burstein et al. (2020) estimate a within-occupation elasticity of 5.6 and an overall elasticity of 10.1.

Table F2: Imperfect substitution between native and migrants, and endogenous college

	Baseline	Imperfect substitution	Endogenous college
IT production			
US IT output	0.11%	0.93%	0.20%
India IT output	35.55%	33.27%	31.57%
Occupational choice			
US CS (native plus immigrant)	1.65%	1.79%	1.69%
India CS	56.64%	52.29%	48.13%
US college non-CS	0.24%	0.24%	0.17%
India college non-CS	-1.04%	-0.91%	1.81%
US high-school	-	-	0.03%
India high-school	-	-	-0.18%
Welfare			
Welfare US natives	0.01%	0.04%	0.01%
Welfare in India	0.13%	0.12%	0.14%
Combined (with migrants)	0.14%	0.18%	0.14%

Percent difference on main outcomes when we go from a scenario with restricted migration (counterfactual) to a scenario with full migration (real). In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995. The 'Imperfect substitution' model allows for imperfect substitution between native and migrant workers. For 'Endogenous college' individuals choose whether to go to college before joining the labor market.

model. To do this, we assume that Indians who migrate to the US remit a fixed share of their income to India, and we calibrate the fixed share to 3.83% of labor income, to match the share of Indian CS wage bill remitted.[32] Such income, is distributed evenly among those residing in India. Given the small share of income remitted, results for the remittances model are very close to those of the baseline, as shown in Table F3, column 3.

Finally, we also evaluate the role of our age-specific switching cost χ_1, by running an alternative specification where $\chi_1 = 0$. As shown in Table F3, the US gains slightly less when not having an age-specific switching cost. This is because older CS workers in the US now switch more to other occupations while with the age-specific switching cost they were more likely to stay in CS and provide more of the spillover through US IT.

F.5 Different Cap Sizes and Starting Restriction-Periods

Our main counterfactual reduces the cap for Indian CS by 50% every year since 1995. We explore how our results would change when trying alternative counterfactuals, such as varying the size of the cap, and beginning the migration restriction in later years. Figure F1 shows how overall welfare for US and Indian natives (excluding migrants) would change for alternative changes in the migration cap. On the horizontal axis we plot the reduction in the cap (i.e. -25% is a world where immigration is restricted to 75% of the current cap). On the vertical axis we plot the welfare change of going from a scenario with an alternative cap (counterfactual) to a scenario with the observed cap (real). In general, the results are monotonic – more migration leads to higher welfare for US natives and Indian non-migrants. Yet, there are some interesting non-linearities when studying the welfare of US natives.[33]

[32]The actual value remitted by Indian CS workers is not available. To compute this, we use the OECD database on Global Remittances to find the total value of remittances from the US to India in 2010. We then compute the total value of remittances relative to total labor income of Indians in the US, which yields 3.83%.

[33]Here we assume that as we raise the cap, the cap will still bind (i.e. supply from abroad is infinitely elastic).

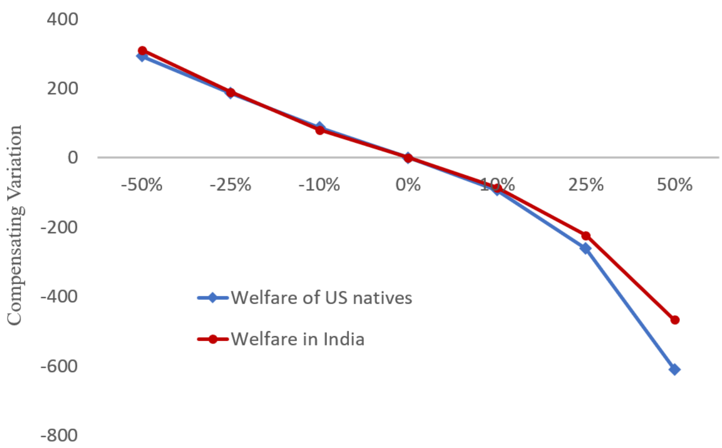
Table F3: Increased trade, remittances and age-specific switching cost

	Baseline	**10% lower trade costs**	**Remittances**	**No age-specific switching cost**
IT production				
US IT output	0.11%	-0.07%	0.11%	-0.06%
India IT output	35.55%	20.28%	35.61%	35.45%
Welfare				
Welfare of US natives	0.007%	0.005%	0.006%	0.005%
Welfare in India	0.130%	0.123%	0.132%	0.132%
Combined (with migrants)	0.142%	0.136%	0.142%	0.141%

Percent difference on main outcomes when we go from a scenario with restricted migration (counterfactual) to a scenario with full migration (real). In the counterfactual, the H-1B cap is reduced by 50% in every period, starting in 1995. The '10% lower trade costs' model has 10% lower bilateral trade costs between the three regions in both industries. 'Remittances' incorporates the feature that Indians who migrate to the US remit a share of their income to India.

In our baseline results we start our counterfactual experiment from 1995 onward. However, from a policy point of view, an important question may be what are the impacts of migration if we started the migration restriction in later years, once the Indian IT sector was already developed. As shown in Figure F2, a cap that starts later has similar effects on India since workers make expectations on future cap increases that are consistent with reality and start getting into CS before the cap comes into effect. For the US, if the cap restrictions start in later periods, workers initially gain as US workers start switching back to CS while the number of migrants is not yet restricted, generating additional spillover effects.

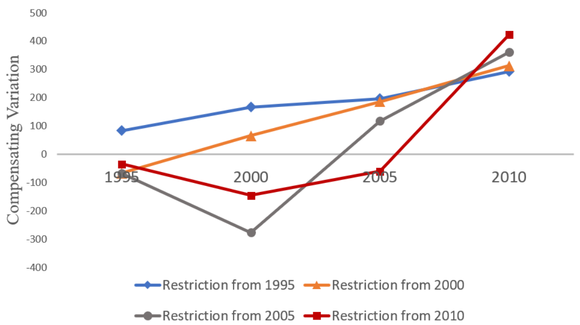
Figure F1: Welfare Loss by Different Cap Sizes (USD mn)



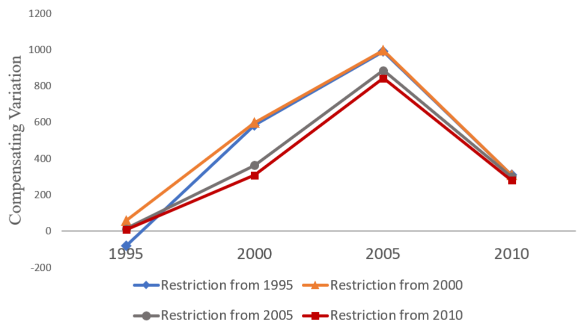
Percent difference on main outcomes when we go from a scenario with an alternative cap (counterfactual) to a scenario with the observed cap (real). We solve the model for different counterfactual scenarios to elicit the welfare loss for each intensity of migration restriction. For counterfactuals we vary the cap from a 50% reduction in the current H-1B cap to a 50% increase. Compensating variation is defined as the amount of USD that must be provided to agents in a world with restricted migration to provide them with the same welfare as in a world with H-1B migration.

Figure F2: Expectations and Starting Restrictions in Different Periods

(a) Welfare Change: US natives (USDmn)



(b) Welfare Change: India non-migrants (USDmn)



Graphs show the compensating variation of restricting H-1B migration by 50%, where we vary the first period of when the cap is lowered. All agents correctly expect the cap to be lowered in the corresponding future period. Compensating variation is defined as the amount of USD that must be provided to agents in a world with restricted migration to provide them with the same welfare as in a world with H-1B migration.

G Out-of-Sample Tests of Model Fit

As we see in Figure G1, cross-country differences in wages for the different types of workers, and within country wage premiums closely match the data. While we never explicitly match the relative wages for CS and non-CS college graduates between countries, we see in Figures G1a and G1b that the model does fairly well in predicting the trends and level differences between the wages in both countries. In Figure G1a, the CS 'place premium' in our model closely matches the data, and quasi-experimental results that show a sixfold increase in wages for H-1B lottery winners (Clemens, 2013). Similarly, in Figure G1b we show the non-CS college graduate wage premium between the US and India is in line with the data.

We also study how IT output and prices, and the location of IT production evolves over time. Figure G1e and G1f display a close match between the model and data for relative US IT prices and output growth. We also predict the relative shares if CS employment in IT between the US and India, even though we do not match these data when solving the model (Figure G1c). Together, the success of the out-of-sample matching gives us additional confidence in our modeling exercise to perform counterfactual tests.

H Why India?

While not the direct objective of our analysis, we find it both worthwhile and interesting to document why this spread of the tech boom was limited to certain countries like India in the period of our study. The other countries (Israel, Ireland and Germany) that became important IT exporters, did not have trajectories that were linked to migration to the US.[34] For instance, a lot of Ireland's IT exports were simply US based firms located there, initially for tax purposes (Athreye, 2005).

What then was special about India's relationship with the US in driving such a spread of the tech boom? Our paper quantitatively confirms qualitative work, discussed below, on the role played by migration of technology workers from India to the US. Such qualitative work by other researchers was largely conducted during the IT boom, and allows us to unpack how India alone could take such great advantage of the H-1B program. Indeed, by the end of our period 86% of computer science H-1Bs were awarded to Indians.

The qualitative literature relies on historiography, and on interviews and small-sample detailed surveys of industry leaders in Silicon Valley and India. While we unpack this rich evidence below, that complements our quantitative analysis, a few salient features of the drivers can be summarized:

First, India had made major investments in top Engineering schools in the 1950s that grew to have worldwide reputation. In later years, enrollment was driven by the prospect of migrating abroad as the domestic market was small. For instance, consistent with our hypothesis, Bhatnagar (2006) notes that "*growth (in training and degrees) was also driven by larger salaries in the IT industry abroad.*"

Related to strong training, was the fact that the urban population was widely comfortable (and trained) in the English language (Azam et al., 2013), giving India an advantage of other Asian giants, like China, in being a source of migrant workers. As such, most multinationals recruited directly from these Engineering colleges (Desai, 2003).

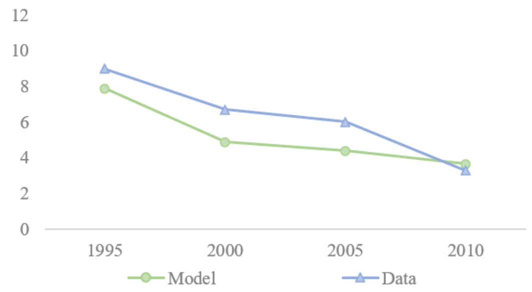
Second, India sent a lot of these top engineers during the earlier hardware-boom of the 1970s and 1980s. This diaspora helped establish strong connections and a reputation for well-trained workers (Arora et al., 2001; Saxenian, 1999). Bhatnagar (2006) notes that Indian professionals in Silicon Valley "*built personal networks and valuable reputations and used their growing influence within US companies to help Indian companies get a foot in the door*" in the expanding IT sector. Consistent with our shift-share approach, when migration caps were raised, it is then in these occupations with strong connections that saw large emigration, as these tech workers "*migrate to better paid jobs in other countries*" (Kumar, 2006).

Finally, (as we model) wages were lower in source countries like India (relative to countries like Germany, Israel or Ireland), and as such there was a ready supply of trained, English-speaking workers willing to work

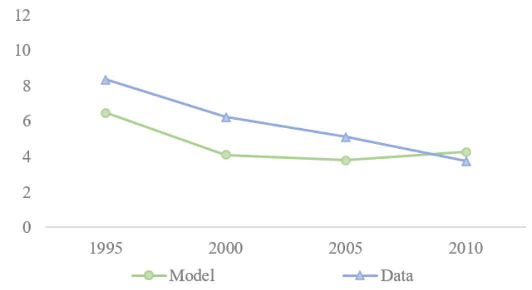
[34] Authors' calculations using the OECD Trade in Value Added (TiVA) statistics suggest that IT (code C72ITS) export growth was largest in these countries.

Figure G1: Model Fit on Wages, Production and Prices in India and the US

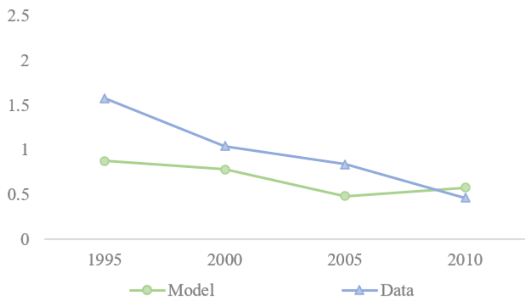
(a) Relative Wage for Computer Scientists: US to India



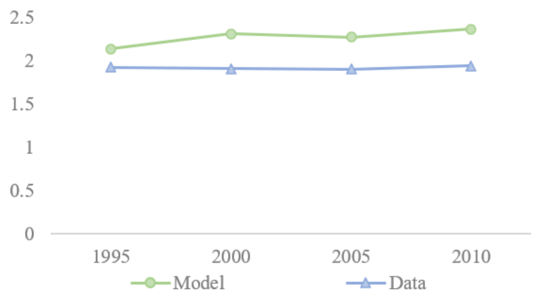
(b) Relative Wage for non-CS College Graduates: US to India



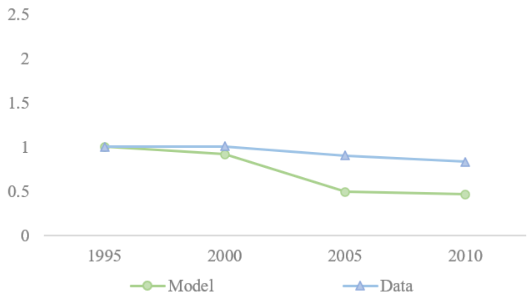
(c) Relative IT share in CS (US to India)



(d) College Wage Premium: US



(e) Price Index IT US



(f) IT Output US

Figures plot the simulated model output and the actual data for the endogenous variables of interest. Top panels show the relative wages for CS and non-CS college graduates between the US and India. Middle panels show the relative fraction between US and India of all CS workers who work in IT and the the college wage premium in the US. The college wage premium is defined as the weighted average of college graduate wages to non-college graduate wages. Bottom panels show IT prices and output in the US, where the indices are normalized to be 1 in 1995. For data sources please refer to Data Appendix A

at competitive wages (Heeks, 1995; Subramanian, 1992). Related to this, there was a large population to draw from, and so firms thought it meaningful to invest in the fixed costs of setting up recruitment systems from

India.

While we also discuss other reasons mentioned by qualitative literature, we find those to be more likely to be read as anecdotes. For instance, Kumar (2001) notes another advantage for the Indian industry – the 12-hour time lag between India and the US virtually doubled the working time per day and cut the software development life-cycle by half.

H.1 A Brief History of Indian IT

The Indian software industry was inextricably tied to the US. It was born in the 1970s when the Tata Consultancy Services (TCS) opened shop and started sending Indian engineers abroad to the US to do software programming, referred to as 'bodyshopping.' Some companies were helped at the start by Indian Government policies that permitted duty-free imports of computer systems if importers would promise to export software and services (usually to the US) worth twice the value of imports within a specified time period. In 1978, however, IBM had to exit the market in violation of Foreign Exchange regulations, serving a blow to firms that required their hardware.[35] There was little seen of the US hardware boom in India as the industry was not mature and did not acquire the technical knowhow of firms like IBM, that had virtually a monopoly in mainframe computers (Bhatnagar, 2006).[36]

As the personal computer became more popular in the US, the demand for software programming services grew rapidly, especially for low-cost workers from India.[37] However, the lack of a domestic market and of the advanced technological capability required to write software meant that 90% of software revenues came from US on-site work (NASSCOM, 2003). These counted towards exports, and exports grew from about \$50 million in the late 1980s to about \$200 million by 1993, at a rate of about 30% a year.

In the early 1990s, with the start of the H-1B program, there was a spurt in the entry of multinational firms and demand for software services that were outsourced from the US. Yet, on-site work dominated because otherwise software had to be transported on tapes which faced heavy import duties. Kumar (2001) notes another significant advantage for the Indian industry in this period – the 12-hour time lag between India and the US virtually doubled the working time per day and cut the development life-cycle by half. In 1992, satellite links were set up in Software Technology Parks (STP) to facilitate the smooth workflow process.

The Y2K threat was not to the detriment of the Indian industry (Desai, 2003). In fact, "*Y2K projects were an important source of revenue for Indian firms*" (Arora et al., 2001), and this helped build reputation with their US counterparts. One commentator notes that the industry "*grew on the strength of Y2K and never looked back*" (Dataquest, 2003).

By 2001 exports had reached about \$6 billion, growing at about 50-60% annually from the mid-1990s. By this time, only five of the top twenty exporters were subsidiaries of foreign firms, indicating that software exports were largely products of Indian firms. This was in stark contrast to the Irish software industry, where the bulk of the exports were US firms based in Ireland (Athreye, 2005).

Even at the end of our period, though, in 2011 90% of the population was still not connected (World Bank, World Development Indicators). As such, most of the growth was driven by exports to the US (Figure 4a).

H.2 The Indian IT Firm and its Relationship with the US

Unlike most Indian industries that focus on the large domestic market, the Indian IT firm was significantly export oriented; catering to a consumer base abroad that has the purchasing power for its products (Figure 4a). It is clear that most of the early-growth was export-led growth since by the turn of the century, software exports accounted for 26% of all exports, whereas in 1995 it was only 2% of all exports. Moreover, till about the end of the 1990s, most of these exports involved the physical presence of Indian workers at an overseas

[35] They were back as a joint venture with Tata Information Systems at the start of the H-1B program in 1992.

[36] At this time personal computers were absent from the Indian market altogether.

[37] In 1989, an Indian computer professional earned about \$5486, whereas Microsoft was offering \$40,000 plus relocation benefits and a green card for software engineers (Subramanian, 1992).

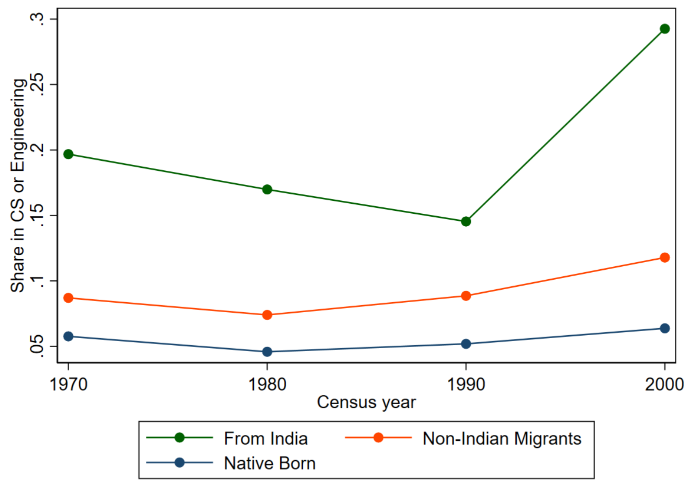


Figure H1: Share of US-based Workers in CS or Engineering (US Census 1970-2000)

work-site. Over time, however, Indian IT firms moved from providing low-cost programming abroad to more comprehensive software development services for their overseas clients that was directly exported from India. Bhatnagar (2006) describes how, in 1995, 66% of all Indian IT exports involved an Indian worker on a foreign work-site, but this number fell to 29% by 2005, after the IT boom had spread to India.

The low-wage advantage is one of the earliest explanations advanced to describe the growth in Indian IT (Heeks, 1995). Arora et al. (2001) note that by the turn of the century, India had the largest number of people working in the industry and the highest revenue growth, but also the lowest revenue per employee. However, this low productivity could be a reflection of the types of services they provided – maintenance of data/legacy systems and other low-value services (Athreye, 2005).

As corroborated by Figure H1, India always sent a disproportionate fraction of Engineers to the US even during the earlier hardware boom (in the 1970s and 1980s). Yet, since immigration was less widespread, the numbers were lower, even as the share of engineers as a fraction of the migrants was large. These engineers helped build connections and spread the reputation of Indian engineering graduates.

As such, a large part of the success of Indian firms is attributed to high-skilled Indian immigrant networks in the US. Bhatnagar (2006) notes that Indian professionals in Silicon Valley "*built personal networks and valuable reputations and used their growing influence within US companies to help Indian companies get a foot in the door*" of the expanding IT work.

We should note that our shift-share strategy implicitly leverages these networks and connections. As Indians established connections in certain engineering-related occupations, the subsequent growth (under more migration as policy caps are raised) would be in these occupations as well (Figure H1).

This reputation was largely built in the on-site consulting phase of the early 1990s. As Banerjee and Duflo (2000) note, reputation is essential in an industry like this because a lot of contracts are for customized software and can lead to hold-ups which a court of law may find difficult to arbitrate over.

Saxenian (1999) shows that in 1998, persons with Indian names headed 774 high-technology Silicon Valley firms overlooking sales of about $3.6 billion, which is roughly the same amount as the entire Indian IT sales bill in 1998. By the turn of the century, Indians headed 972 Silicon Valley firms, accounting for nearly 26,000 jobs. Bhatnagar (2006) claims that Indians headed about 3% of tech companies started between 1980 and 1985, but by 1995 they headed about 10% of them. At around the same time, NASSCOM estimated that about 200,000 Indian software professionals were working on H-1B visas.

The US has historically been the largest exporter of software products, and continued to produce the largest number of patents in the industry. US multinationals entered the Indian market by setting up liaison offices and subsidiaries. While they initially intended to sell to the Indian market, they rapidly shifted to using

India as a place for software development (Arora et al., 2001). By 1997, the US accounted for about 58% of the all export revenues whereas all the European countries combined accounted for only 21%. By the mid-2000s, however, India overtakes the US as the major exporter of IT products (Figure 4b).

Indian firms could use the H-1B program as a method to set up a base in the US with a ready supply of workers from India. Even as late as 2013, Indian firms were the largest sponsors of H-1B visas to the US. Even non-Indian firms are big employers of H-1Bs, some of which have Indians as their largest employment-base. Table H1 shows that 10 out of the top 11 H-1B firms have Indians as their main employment base. Indian citizens, are therefore, the largest beneficiaries of the H-1B visa program, with about 70% of all H-1Bs in 2014 being awarded to Indians (USCIS, 2014).

Table H1: Number of H-1Bs by Firm (Approved)

Rank	Company	Headquarters	Employment Base	2012	2013
1	Infosys	India	India	5600	6298
2	Tata Consultancy Services	India	India	7469	6258
3	Cognizant	USA	India	9281	5186
4	Accenture Inc	Bahamas	India	4037	3346
5	Wipro	India	India	4304	2644
6	HCL Technologies Ltd	India	India	2070	1766
7	IBM(India, Private Ltd.)	USA	India	1846	1624
8	Mahindra Satyam	India	India	1963	1589
9	Larsen & Toubro Infotech	India	India	1832	1580
10	Deloitte	USA	US	1668	1491
11	IGATE(Patni)	USA and India	India	1260	1157

Source: Author's calculations using USCIS reports (2012-13). Rank is based on 2013 number of visas approved. The last two columns indicate the number of H-1B visas that were approved for each year. To measure the employment base, we use the number of workers in each location by firm. For instance, the largest H1B benefactor at the end of our period: Infosys received 5,600 H-1B visa approvals in 2012. In 2012, the company has 156 thousand employees worldwide, out of which 141 thousand are in India, making India its main employment base (see https://www.infosys.com/investors/reports-filings/quarterly-results/2012-2013/Q3/Documents/fact-sheet.pdf).

These workers come in search of higher wages, as working in the US provides substantial productivity premiums (Clemens, 2013). However, there is a cap on H-1B visas, and the visas are only for 3-6 years and, a large number of potential engineers must seek work elsewhere. Workers who are unable to obtain an H-1B visa because of the cap, or return to India after their visa expires are a ready supply of labor for firms in India. This leads to a large skilled workforce in India, and enables the Indian IT sector to expand, tapping in to this growing educated workforce. The growth in Indian firms are therefore strongly linked to larger exports to the US, and in the ready supply of labor.

H.3 Indian Students and College Choice

The boom in the US also affected the education sector in India. Consistent with the hypothesis posited in our analysis, Bhatnagar (2006) notes that "*Since engineers were willing to work as programmers in a domestic environment with few job opportunities, growth (in training and degrees) was also driven by larger salaries in the IT industry abroad.*" To meet the rising demand for workers, engineering schools introduced more computer science oriented degrees, and companies started their own training divisions in the 1980s, building technical skills for the industry (Figure 3b).

In later years, as demand for such degrees expanded, India was somewhat uniquely placed in rapidly expanding the supply of Engineering colleges to meet rising demand, as we describe in other work (Jagnani and Khanna, 2020).

While the abundant stock of programmers had induced recruiters to come to Indian in the early 1990s, this was sustained till the end of the decade by a steady increase in freshly trained programmers (Desai, 2003). In 2002, these software engineers were young with a median age of 26.5 years, and 58% of them had less than three years of experience (NASSCOM, 2003). In India, most programmers and the chief executives in IT companies are predominantly trained as engineers (Desai, 2003). Science graduates and those with master's degrees in computer applications make up the rest. A NASSCOM-Hewitt survey found that 88% of firms (primarily multinationals) visited engineering colleges to recruit, and 47% recruited *only* there.

Many of these recruiters themselves had been trained at these very universities, and others knew of colleagues from such reputable schools like the Indian Institutes of Technology (IITs), that were established in the 1950s. Arguably, such a strong Engineering-school sector gave Indian workers an edge over other countries (Desai, 2003).

A survey by Arora and Athreye (2002) found that 80% of all software professionals employed had engineering degrees, and show that over time a number of engineering colleges have increased their emphasis on IT and even IT management. This, however, has meant that the number of PhDs in engineering disciplines has actually fallen from about 675 in 1987 to 375 in 1995. The industry was therefore attracting some of the brightest young graduates, with little academic bent, and with only industrial ambitions. The salaries were among the highest across industries, growing at a steady rate, and some firms even offer stock options. Despite this, the attrition was quite high, as they "*migrate to better paid jobs in other countries*" (Kumar, 2006).

The bulk of Indian workers get their degrees at Indian universities. India has historically been better at technical education like engineering and medicine (Arora et al., 2001). Furthermore, it has a linguistic advantage over East Asian countries, due to a vast majority of Indians being fluent in English. Over the last few decades, there has also been consistent growth in the number of new undergraduate engineering schools being opened to cater to the burgeoning demand (NASSCOM, 2012). This growth in engineering schools drives the growth in undergraduate enrollment in technical institutions (Figure 3b), and a similar trend can be seen for graduate and diploma-granting institutions.

A large number of foreign students also come to the US for higher education, plausibly also exploring this as a pathway to the US labor market. Many students stay on to obtain work visas (Bound et al., 2014). 20,000 H-1B visas are granted to students who obtained their master's (or higher) degrees from US institutions, and this may incentivize students to obtain their degrees in US universities so that they may avail of the higher wages that come with an H-1B visa (Rosenzweig, 2006).