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

Gaurav Khanna[†]
Nicolas Morales[‡]

Most recent version here

December 2018

Abstract

We study how US immigration policy coupled with the Internet boom affected not just the US economy, but also led to a tech boom in India. Indian students enrolled in engineering schools to gain employment in the rapidly growing US IT industry, via the H-1B visa program. Those who could not join the US workforce, due to the H-1B cap, remained in India, enabling the growth of an Indian IT sector. Those who returned with acquired human capital and technology after the expiration of their visas also contributed to the growing tech-workforce in India. The increase in IT productivity allowed India to eventually surpass the US in software exports. Our general equilibrium model captures firm-hiring across various occupations, innovation and technology diffusion, and dynamic worker decisions to choose occupations and fields of major in both countries. We identify key elasticities using an instrumental variables strategy, and show that we capture levels and trends of key variables in validation tests. In counterfactual exercises, we find that on average, workers in each country are better off because of high-skill migration. The H-1B program induced Indians to switch to computer science (CS) occupations, increasing the CS workforce and overall IT production in India by 15%. It also induced US workers to switch to non-CS occupations, reducing the US native CS workforce by 4.7%.

JEL: I25, J30, J61

Keywords: High-skill immigration, H-1B visas, India, computer scientists, IT sector

---

*We thank Ashish Arora, Dominick Bartelme, John Bound, Michael Clemens, Gordon Hanson, Andrei Levchenko, David McKenzie, Justin Sandefur, Jagadeesh Sivadasan, Sebastian Sotelo and seminar participants at Michigan, Purdue, Montevideo, Riverside, Raleigh (Society of Labor Economists), Western Ontario (Labour Day), Claremont Mckenna, Hawaii, Mt Holyoke, Center for Global Development, Clermont-Ferrand, Bonn (IZA), Yale McMillan Center, Delhi (ISI), Upjohn Institute, Bank of Italy, and the NBER Summer Institute (Cambridge, MA) for insightful comments. The authors are grateful to the Alfred P. Sloan Foundation and the NBER Fellowship on High-Skill Immigration for generous research support.

[†]University of California – San Diego, gakhanna@ucsd.edu, econgaurav.com
[‡]University of Michigan, moralesn@umich.edu, 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 the sending and receiving countries. For instance, native workers in the receiving country may benefit if they are complements to immigrants, or suffer if they are substitutes. The sending country may experience brain drain as human capital departs, or experience brain gain as the opportunity to migrate induces human capital accumulation. Immigration can attract global talent and lead to production growth in the receiving country, but migration-induced technological catch-up may contrarily shift production to the sending country. We resolve these ambiguities by modeling and estimating the long-term welfare consequences of high-skill migration under the H-1B program, and examine how the combination of the US IT boom and immigration policies led to a boom halfway across the world in India.

We model firm-production, trade and the forward-looking decisions of workers and students in both countries, and estimate key elasticities using instrumental variables. Innovation rapidly expanded the US IT sector in the early 1990s (Bound et al., 2015; Kerr, 2013a), and a few years later the IT sector in India quickly grew from 1.2% of GDP in 1998 to 7.5% in 2012 (NASSCOM, 2012). 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. Our estimated elasticities help match data from various sources and countries over the period 1994 to 2010, and we perform out-of-sample tests to show that our model accurately captures these trends.

We then conduct counterfactual exercises that change the number of immigrants allowed into the country. Given that 70% of H-1B visas went to Indian workers by 2014, our results indicate that the H-1B program and the tech boom had a powerful impact on IT sectors in both countries. By the early-2000s, many workers returned to India once their visas expired with newly acquired knowhow and connections. This additionally facilitated the US-led boom to spread to India, and by the mid-2000s India surpassed the US as the major exporter of software. Despite various distributional effects, we find that the average worker in each country is better off due to immigration.

In Section 2 we first use descriptive trends and background information to describe our narrative and ground our model. Starting in the early 1990s, innovation in the US IT sector led to a growth in IT firms, and computer science (CS) employment. An immigration policy that favored high-skill immigrants led to an increasing proportion of foreigners in the US computer-science workforce. The foreign fraction of CS workers grew considerably from 9% in 1994 to

24% in 2012; much faster than the foreign fraction of all workers in STEM occupations. By the mid-2000s more than half of all H-1B visas were awarded to Indians (USCIS, 2014). This fraction was higher among CS occupations: by 2014, 86% of all computer science H-1B visas were awarded to Indians (Computerworld, 2015), making India the largest contributor of foreign computer scientists. All the top firms that hired H-1Bs are in IT, and the top 9 had India as their primary employment base (USCIS, 2014).

Given the large wage differential between the US and India, and a non-trivial probability of migrating to the US, many more Indian students started enrolling in engineering schools. However, the number of available H-1B visas was capped, so a large number of Indian workers that would have preferred to work in the US, had to seek employment in India. Furthermore, since H-1Bs expire after 3 to 6 years, many of these workers returned to India, bringing with them their accumulated human capital, technological knowhow and connections, facilitating further technological diffusion (Kerr, 2008). This educated workforce in India enabled the Indian IT sector to grow rapidly, with new firms joining the race and older firms expanding, and over time, India became a major producer of software eroding the US dominance in IT exports. This boom missed many other countries but settled on India. India has not only historically had high quality engineering schools that train potentially lower-wage, English-speaking workers, but had also developed strong networks with the US sector during the earlier hardware boom (Arora et al., 2001; Bhatnagar, 2005).

1.1 The Model and Expected Impacts

We create a general equilibrium model that contains five crucial features, described in Section 4. First, we model how US firms hire both US and foreign workers, and Indian firms hire workers from India. Importantly, firms hire three different types of workers – computer scientists, non-CS college graduates and non college graduates. More skill-biased capital and better technology imply that wages are higher in the US and Indian CS workers wish to emigrate. 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.

Computer scientists, both domestic and foreign, are innovators and increase the overall productivity of firms in the IT sector via the generation of non-excludable ideas (Kerr, 2010). Under directed technological change, an increase in the size of the computer science workforce makes India more productive over time. In India, the return migrants are not perfect substitutes with those that never migrated, as they may return with acquired human capital.

Second, the IT sector produces a continuum of varieties, the productivities of which differ across countries. Restricting immigration, or rapid growth in the Indian IT sector can shift some production of these varieties from the US to India. Consumers benefit from lower prices, and the final goods sector of the economy uses software as an intermediate input in production; an expansion in the IT sector raises overall productivity in the rest of the economy.

Third, to capture the role of trade, we encapsulate the canonical Eaton and Kortum (2002) framework into our model. All goods are tradable with asymmetric trade costs and each country will have a comparative advantage in producing some varieties. Both countries are competing for the world market. The potential to trade allows India to grow, as workers switch to the high innovation IT sector. At the same time, in both countries, the wage impacts of immigration are muted by trade as resources are shifted across sectors (Ventura, 1997). Under the H-1B program, the price of IT falls worsening the US's terms of trade.

While the above three features capture the product and the labor demand aspects of the economy, the next few capture important labor supply decisions. The fourth feature is that students in both countries have heterogeneous preferences, and make dynamic decisions on choosing their college major given their expected future earnings in different occupations. Changes to expected earnings, driven by innovation shocks and immigration policies, have long-run effects on human capital accumulation and the labor supply elasticity.

Fifth, after graduation, workers (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 CS workers pay an additional cost of migration and earn higher wages in the US if they win the H-1B lottery. Importantly, as expected earnings change with immigration policy, workers switch occupations mitigating either the positive or negative wage impacts of immigration.

Our model includes many 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 as return migrants bring back acquired knowhow. 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.

Distributional impacts on different types of workers are similarly ambiguous. Computer science wage growth may be depressed by the rapid inflows of immigrants, but an increased CS workforce can lead to more innovation raising the demand and wages for all workers. The demand for non-CS workers may rise not just because of innovation, but also because they are complements to CS workers in the production process; however, depressed CS wages may encourage US born CS workers to switch to non-CS occupations, lowering non-CS wages as well. As the IT sector grows in both countries, lower IT prices hurt the US's terms of trade. However, consumers are better off because they have more efficient and affordable products, and sectors that use IT as an intermediate goods are more productive.

1.2 Identifying Key Parameters

To solve the model and resolve these theoretically ambiguous effects we estimate three important elasticities. We first show well-identified evidence supporting the underlying driving force behind our hypothesis: that labor demand shocks in the US affected human capital accumulation in India. In Section 3 we use an instrumental variables strategy to describe this response. We create a measure for the differential demand from the US over time, by interacting the baseline occupational share of migrants from India with the annual total number of migrants from all *other* countries. We estimate a robust response to these demand shocks in the US, and use our estimated parameter when solving our GE model.

Our second important parameter is the industry-level innovation change. We use an immigration shift-share instrument (Card, 2001), implicitly leveraging variation in US migration policy and migrant-supply shocks, to measure the response in innovation (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., 2015b).

The occupation switching elasticity and the college-major choice determine our third important parameter: the dynamic labor supply elasticity. This plays a crucial role in the distributional effects of immigration across different types of workers, and we estimate this elasticity using a Simulated Method of Moments (SMM). Given the dynamic nature of the 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 immigrant changes in other work will not pick up the entirety of the labor supply adjustments.

In Section 5 using data from 1994-2010 we empirically determine the model's parameters.[1] The production side help us determine the exogenous innovation shocks that shift the labor demand curve out every year, 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.

Given the complexity of the model it is important to do validation exercises. In Section 6 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.

1.3 Results and Implications for the Literature

In Section 7 we conduct counter-factual 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

[1] We stop our analysis in 2010 for two main reasons: First, after 2010, there was a shift in H-1B usage towards outsourcing firms. Second, after 2012 there was a rapid increase in Indian students enrolling in Engineering degrees at US universities.

only half the number of migrants over the 1994-2010 period. By shutting-down certain parts of our model we are able to ascertain how important each mechanism (trade, return migration, innovation, etc.) is in contributing to our results.

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. Those that returned after the expiration of their visas contributed to this growing CS workforce and enabled the increases in technological productivity in India. We show that the H-1B program led to an increase in the size of the non-migrant Indian CS workforce. However, the migration led US native CS workers to switch to non-CS occupations and is therefore associated with a fall in the US native CS workforce by as much as 4.7% in 2010.

An increase in the size of the Indian CS workforce also led to an increase in productivity in the Indian IT sector. Under the H-1B program, production shifts to India – US IT output is 0.5% lower, and Indian IT output 15% higher in 2010. The shift in production 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 0.12%, and the average worker in each country is better off in a world with skilled migration. The welfare gains from H-1B migration amount to about $53 thousand per migrant in 2010. These include a net welfare gain to US workers of about $927 per migrant, and to Indian non-migrants of $1582 per migrant.

We highlight important mechanisms and show the quantitative relevance of our modeling decisions. Not allowing for endogenous labor supply decisions will lead to large negative impacts on US born CS workers that cannot switch to other occupations, and ignoring the possibility of return migration will lead to significant overestimates of CS employment in India. Most importantly, not allowing Indian workers to respond to the possibility of migrating will lead to drastically different results. Modeling trade is also important: Indian IT is mostly exported and ignoring trade would underestimate the benefits of more CS working in India.

In Section 9 we discuss these results in relation to the larger literature on labor, trade, technological diffusion and migration. Our paper is innovative in a few ways. First, we incorporate migration and endogenous human capital accumulation into a model of trade and technological diffusion. In doing so we synthesize different insights from a broad literature, and add crucial features overlooked by previous work. On the one hand, North-South trade may hinder structural transformation as developing economies specialize in less productive sectors (Matsuyama, 1992). On the other hand, technological diffusion can help developing countries catch up with more developed ones (Krugman, 1979). Since migrants accumulate human capital and technical knowhow in the US and return with this knowledge to India, this speeds up technological diffusion and catch-up (Kerr, 2008). As Davis and Weinstein (2002) highlight, immigration-induced catch-up may also deteriorate the terms of trade for the country with superior technology – in this case, the US. As the price of IT falls, the US is hurt as it has a comparative advantage in exporting IT products. Furthermore, depending on the rate of technical change, offshoring will benefit workers in developing countries but may harm workers in developed economies (Acemoglu et al., 2015). Alternatively, Freeman (2006b) argues that immigration can help the US maintain its advantage by attracting global talent. Such analyses, however, miss the incentives to invest in human capital, and the corresponding growth in production for sending countries – features that play an important role in our analysis.

Second, our paper addresses crucial issues raised by the labor literature on the impacts of high-skill immigrants on the US economy. High-skill immigrants impart benefits to employers and consumers, complementary inputs used in production, and in general may be valuable innovators that improve technology (Foley and Kerr, 2013; Hunt and Gauthier-Loiselle, 2010; Kerr and Lincoln, 2010). However, they potentially impose costs on domestic workers who are close substitutes (Borjas, 1999; Doran et al., 2017). The magnitude of these costs may be substantially mitigated if US high skill workers have good alternatives to working in sectors most impacted by immigrants (Peri and Sparber, 2011). If high skill immigrants contribute to the generation of knowledge and productivity through patenting and innovation, then this serves to shift out the production possibility frontier in the US, and may slow the erosion of the US comparative advantage in high tech (Freeman, 2006b). In Bound, Khanna, and Morales (2016), we use a general equilibrium model of the US economy to look at the short-run effects on the welfare of workers in the receiving country, after deriving labor demand for each type of worker. Importantly, these papers abstract away from the role played by other countries. In our current paper we capture the long-run effects since we model the growth of the tech sector in India, 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.

Third, we provide novel evidence to understand major 'big push' sectoral transformations to high-skill production in emerging economies (Banerjee and Newman, 1993; Lagakos and Waugh, 2013; Lewis, 1954; Matsuyama, 1992; Murphy et al., 1989). India experienced a tech boom that substantially contributed to its rapid economic and export growth. The IT boom and immigration policy in the US, and the Indian growth-story are therefore closely linked, and studying this boom can help us understand how workforce skill transformations may come about in developing countries. Importantly, we show that US immigration policy can affect structural development half-way across the world, in India.

Last, we address the debate between brain-drain and brain-gain (Beine et al., 2001; Dinkleman and Mariotti, 2016; Shrestha, 2016; Stark, 2004; Stark et al., 1997). While many worry about the fact that a large number of educated Indians leave the country for the US, we show how better paid jobs abroad may incentivize students to choose certain majors and supply a highly-educated workforce to Indian firms. Migrants that return with newly acquired human capital and technical knowhow help develop the IT sector at home, contributing to the brain-gain.

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, the usage of the Internet for commercial purposes grew rapidly in the US (Leiner et al., 1997).[2] This led to an increase in demand for computer scientists (CS), and a rise in IT R&D. The growth of tech firms like Yahoo, Amazon and eBay helped sustain the boom in the IT sector till the end of the century. Such changes had a significant impact on the market for IT workers: The number of computer scientists or computer software developers increased by 161% between 1990 and 2000 (US Census), whereas the number of workers in other STEM occupations increased by 14%. As Figure 1a shows, CS employment, as a share of the college educated workforce and the STEM workforce, rose dramatically in the second half of the 1990s. By the turn of the century more than half of all STEM workers were CS.

CS occupations were the fastest growing occupations in the second half of the 1990s, and were expected to remain the fastest over the next decade (BLS, 1996). This growth was, in part, fueled by foreigners (Figure 1b). In 1994, foreigners were less represented among CS (about 9%) than in other STEM occupations, but given the dramatic growth in the second half of the 1990s, by 2010 foreigners account for more than 20% of the CS workforce.

Figure 1: High-Skill Immigration and the IT Boom

(a) Fraction of Computer Scientists in US Workforce

(b) Immigrants as Fraction of Workers by Occupation

Source: March Current Population Survey (CPS). STEM is science, technology, engineering, and mathematics. Immigrants defined as foreign born who migrated after the age of 18. STEM definition in data appendix C.

Freeman (2009) attributes some of these changes to the dramatic increase in college educated (science and engineering S&E) workers in India where the number of first degrees conferred in S&E rose from about 176 thousand in 1990 to 455 thousand in 2000. The second development was The Immigration Act of 1990 which established the H-1B visa program for temporary

[2] The decommissioning of the National Science Foundation Network in April of 1995 is considered crucial for introducing nationwide commercial traffic on the Internet.

workers in "specialty occupations," defined as requiring *theoretical and practical application of a body of highly specialized knowledge in a field of human endeavor.*

To hire a foreigner on an H-1B visa the firm must first file a Labor Condition Application (LCA), and pay them the greater of the actual compensation paid to others in the same job or the prevailing occupational compensation. USCIS may approve the petition for up to three years (extendable up to six years). Once the H-1B expires, employers can sponsor a green card but each country is eligible for only a specific fraction, and as such, wait times for Indian and Chinese nationals are many years. The U.S. General Accounting Office 2011 survey estimates the legal and administrative costs associated with each H-1B hire to be $2.3 - 7.5 thousand, suggesting that employers expect some cost or productivity advantage when hiring foreigners.

By the time the IT boom was starting in the mid-1990s, the 65,000-strong H-1B cap started binding and the allocation was filled on a first come, first served basis. According to the USINS (2000), the number of H-1B visas awarded to computer-related occupations in 1999 was about two-thirds of the visas, and U.S. Department of Commerce (2000) estimated that during the late 1990s, 28% of 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. Furthermore, 91% of IT workers with a Bachelor's degree and 76% of those with a Master's degrees were educated abroad (National Survey of College Graduates 2010). Given that such a large proportion obtain their degrees in other countries, the education sector abroad plays a major role in the US tech boom as well (Bound et al., 2014).

2.2 The Spread of the IT Boom to India

With large-scale economic reforms in the early 1990s, the IT industry in India was opened up. There was a spurt in the entry of multinational firms and demand for software services. In 1992, satellite links were set up in Software Technology Parks negating the need for some kinds of on-site work and this boosted the off-shoring of work to India.[3] One estimate suggests that by 1996, India had 16% of the globalized market in customized software, and more than 100 out of the Fortune 500s outsourced to them (Dataquest, 1996).

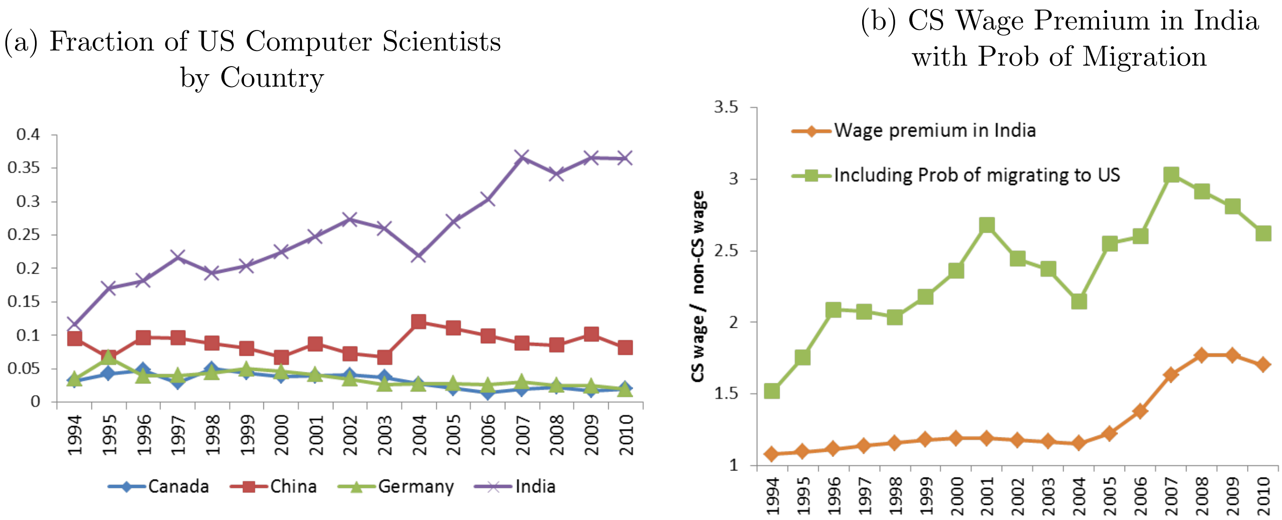
A large part of the success of Indian firms is attributed to high-skilled Indian immigrants in the US. Bhatnagar (2005) notes that Indian professionals in Silicon Valley "*built personal networks and valuable reputations and used their growing influence within US companies to help Indian companies get a foot in the door.*" This reputation was largely built in the early 1990s on-site consulting phase. By 2013, the top 9 H-1B firms had India as their primary employment base (USCIS, 2014).[4] Indian firms used the program as a way to set up base in the US with a

[3] Kumar (2001) notes another significant 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 development life-cycle by half.

[4] Indian computer scientists also became senior managers at tech firms (Saxenian, 1999). Indians headed

ready supply of workers from India. They had the advantage of extensive networks and an in-depth knowledge of the Indian labor market. Meanwhile, the US has a large market for client services and demand for software development, along with more capital, technology and industrial agglomeration, making it an attractive location for Indian firms.

Figure 2: Migration Prospects and Returns to Skill in India



Source: March Current Population Survey (CPS), National Sample Survey (NSS), and USCIS Statistics. In the left panel we restrict the sample to foreign-born workers in CS that immigrated after the age of 18. In the right panel we restrict the sample to college graduates, and compare the CS and non-CS wage in India (orange), and then the expected wage taking into account the probability of migrating and the wage differential in the US.

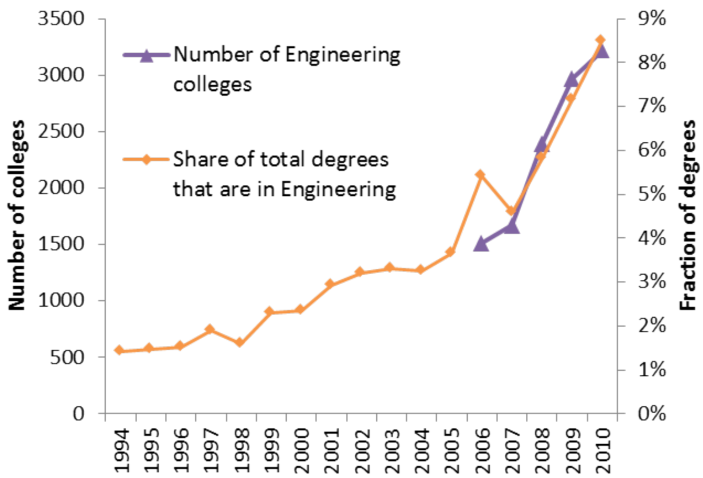
By 2014, more than 70% of all H-1B visas (USCIS, 2014), and 86% of all CS H-1B visas were awarded to Indians, whereas only 5% were awarded to Chinese nationals (Computerworld, 2015). These disproportionate flows are evident in the fraction of migrant CS workers from each country (Figure 2a). Given the 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 2b).

These higher returns to CS affected the education sector in India. Bhatnagar (2005) notes that "*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 CS oriented degrees (Figure 3), and companies started their own training, building technical skills for the industry. More than 80% of all software professionals employed have engineering degrees (Arora and Athreye, 2002), and over time a number of engineering colleges increased their emphasis on IT and IT management. The salaries were among the highest and fastest growing across industries, yet, many "*migrate to better paid jobs in other countries.*" (Kumar, 2006).[5]

about 3% of tech companies started between 1980-85, but by 1995 they headed about 10% of them. At the same time, NASSCOM estimated that about 200,000 Indian software professionals were working on H-1Bs.

[5]Indian students also come to the US for higher education purposes, plausibly exploring this as a pathway to the US labor market, and many stay on to obtain work visas (Bound et al., 2014). Nonetheless, before 2012,

Figure 3: Growth in Degrees and Colleges in India



Source: Ministry of Human Resources and Development and the All India Council for Technical Education

Even as Indians acquired skills that were valuable abroad, the H-1B visa was capped and many were unable to migrate to the US. Additionally, after six years a significant fraction of H-1B workers left the US as their employers were reluctant to sponsor a costly green card.[6] This meant that a highly-skilled workforce was not in the US in the longer term and many joined the Indian workforce. Indian firms tapped into this skilled workforce leading to an IT boom in India (Figure 4a). NASSCOM (2012) estimates that between 1996 and 2000, Indian IT generated about 60,000 jobs a year.

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 about 58% of the all export revenues. By the mid-2000s, however, India overtakes the US as the major exporter of IT products (Figure 4b). Indeed, Indian IT firms are significantly export oriented; catering to a consumer base abroad that has the purchasing power for its products (Figure 4a). Most of the early-growth was export-led growth: in 1995 software was only 2% of all exports, but by the turn of the century, it was 26%. By 2001 exports had reached about $6 billion, growing at about 50-60% annually from the mid-1990s.

Our hypothesis is that this spread in 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.

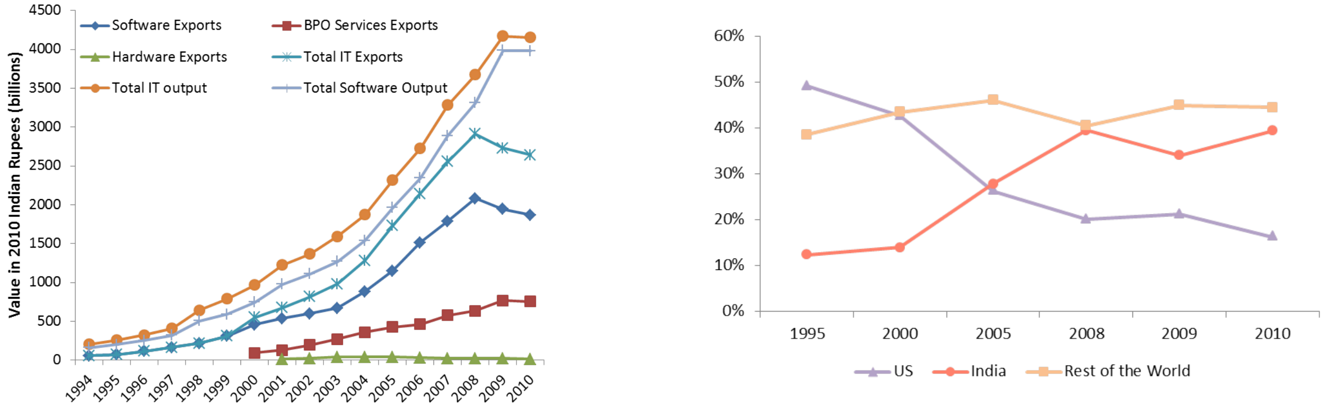
the bulk of Indian workers get their degrees at Indian universities. Based on our calculations using student visa data obtained via a Freedom of Information request to USCIS, we estimate that between 2004 and 2012, there were only 20,000 student visas granted to Indian STEM students (broadly defined). Since 2012 this number has rapidly increased. India has historically been better at technical education like engineering and medicine, and has the advantage of using the English-language over East Asian countries (Arora et al., 2001). 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 (Figure 3 and NASSCOM (2012)).

[6] Green cards have country quotas, and by the mid-2000s, Indians had to wait almost ten years.

Figure 4: The IT Boom Spreads to India

(a) Growth in the Indian IT sector

(b) IT Exports Over Time



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

3 US Demand and Occupational Choice in India

Even as we see the increase in Indians specializing in CS concurrent with the US tech boom, one may contend that such human capital investments among Indians was independent of demand shocks in the US. To test our hypothesis we first establish that labor demand shocks in the US did indeed lead to labor supply responses in India. We study how an increase in the migration of Indians to the US in particular occupations induced Indians into choosing such occupations. We restrict our analysis to skilled occupations, and estimate:

$$N_{o,r,t} = \delta_{o,r} + \delta_t + \beta Mig_{o,t} + \gamma Local_{o,r,t} + \epsilon_{o,t} \quad , \tag{1}$$

where $N_{o,r,t}$ is the Indian workforce in occupation o and time t in region r. $\delta_{o,r}$ is an occupation-region fixed effect and δ_t is a time fixed effect. $Mig_{o,t}$ is the number of college-educated Indian migrants in the US working in occupation o at time t. $Local_{o,r,t}$ is a control that captures local labor demand shocks for a given occupation, that may be correlated across countries. Here we construct a shift-share control using the initial industrial composition in a given region (share of occupation o in region r) and interacting it with national level growth by industry:

$$Local_{o,r,t} = \frac{N_{o,r,1991}}{\sum_o N_{o,r,1991}} \times \sum_{ind} \frac{N^{ind}_{r,1991}}{\sum_{ind} N_{r,1991}} \times \text{Total employment in } ind \text{ and } t \tag{2}$$

The increase in occupation-specific migration may be the result of either the demand from the US (say, driven by growth in certain US sectors) or an increase in supply from India (say, driven by unrelated investments in universities). To isolate the demand-from-abroad channel we use a research design that interacts two variables to create an instrument. First, is the share of Indian migrants within each occupation in the US at baseline (in 1990). This captures the propensity of Indians to be differentially represented in certain occupations at baseline, and varies only across occupations and not over time. The second, is the number of non-Indian college-graduate migrants in the US for each year. This captures the overall demand for migrant labor from the US. As this second measure excludes Indian migrants, we expect this demand to be devoid of preferences for Indians specifically.

$$Shift\ Share_{o,t} = \frac{N^{US}_{o,1990}}{\sum_o N^{US}_{o,1990}} \times \text{Number of non-Indian migrants in US at } t\ , \tag{3}$$

where $N^{US}_{o,1990}$ is the number of Indian college graduates who migrated to the US and work in occupation o in 1990. In Table 1, we see that the instrument has a strong first stage, and the labor-response elasticity is 0.451: for a 1% increase in migration from India in occupation o, there is about a 0.45% increase in employment in India for the same occupation o.

Table 1: US Demand Affects Supply in India

	OLS Emp in India	First Stage Emp Indians in US	2SLS Emp in India
Emp Indians in US	0.343***		0.504***
	(0.0948)		(0.104)
Shift share		3.949***	
		(0.352)	
Bartik control	0.0205**	-0.0111***	0.0227***
	(0.00838)	(0.00277)	(0.00852)
Observations	4,566	4,566	4,566
R-squared	0.140	0.280	0.132
F stat			125.9
Elasticity	0.307		0.451
SE	(0.085)		(0.093)

Notes: Tables show the OLS, first stage and 2SLS results for the impact of employment by occupation of Indians in the US on the employment by occupation of Indians in India. The instrument captures the demand for occupations in the US. 'Bartik control' controls for local labor demand demand shocks. Standard errors clustered by occupation-year. Robustness to alternative instruments and no controls can be found in Table A1.

In Table A1 we show the robustness of our exercise across different specifications. First we show that our result is unaffected by whether or not we control for local demand shocks. Next we use a log formulation for our outcomes. Last, we show robustness to two alternative versions of the instrument that capture labor demand shocks in the US, as in Equations 4 and 5:

$$Shift\ Share_{o,t} = \frac{N^{US}_{o,1990}}{\sum_o N^{US}_{o,1990}} \times \text{Number of US workers at } t \text{ in } o \tag{4}$$

$$Shift\ Share_{o,t} = \frac{N^{US}_{o,1990}}{\sum_o N^{US}_{o,1990}} \times \text{Number of non-Indian migrants in US at } t \text{ in } o\ , \tag{5}$$

The instrument in Equation 4 uses yearly variation on the number of US born college graduates *by occupation*, to capture shifts in US occupational demand that are independent of migration policy. Alternatively, the instrument in Equation 5 uses the yearly variation of non-Indian migrants who are college graduates and work in the US, by occupation, to capture the demand for migrants in certain occupations. Both measures are interacted with the share of Indian migrants within each occupation in the US at baseline (in 1990), and capture demand variation in the US that is independent from concurrent supply shocks in India.

Notice that this approach requires weaker assumptions than the traditional shift-share approaches (Goldsmith-Pinkham et al., 2018), where the temporal variation would be the total (aggregated over occupations) number of migrants from India to the US. Instead the specifications we show in Tables 1 and A1 are likely driven by the demand for workers in the US.

Recent work on shift-shares (Goldsmith-Pinkham et al., 2018; Jaeger et al., 2018) discuss specification tests which we conduct here. In Table A2 we examine how much of the variation in baseline shares are explained by baseline observables. We use demographics, employment status, and education outcomes and find that these only explain a small fraction of baseline shares (an R-squared of 0.038). As one would expect, the strongest explanatory power is from the share of college graduates in an occupation (but we restrict our specification to only college graduates). Importantly, in Table A3 we can check for pre-trends by looking at whether employment growth between 1987 and 1994 were correlated with the baseline occupational distribution of Indian migrants in US, and find no evidence of differential trends. Last, we find that our elasticity estimates are not driven by CS occupations alone (Table A4).

4 Model

Our model, consists of two main sections: 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 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 is assumed to have 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. 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 can either be college graduates or non-college graduates denoted by $Q_{t,k}$ and $H_{t,k}$ respectively, where k stands for a country subindex and t stands for time. College-attendance decisions are made outside the model such that the total number of college graduates is given. College graduates in both countries have the choice to become computer scientists ($L_{t,k}$) or work in some other skilled occupation ($G_{t,k}$).[7]

We allow for two type of decisions for college graduates. First, at the age of 20, 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, from age 25 to 65, workers choose every five years between working as a computer scientist and working in another occupation. We club 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 in each occupation.[8] Individuals are assumed to have perfect foresight regarding the evolution of future wages and know the distribution of taste shocks.[9] Individual i's decision problem, before joining the labor market, is summarized by Equation 6:

$$max\{\beta \mathbb{E}_t V^c_{t+1,25,k} + \bar{F}_k + \sigma_k \eta^c_{i,t,k},\ \beta \mathbb{E}_t V^o_{t+1,25,k} + \sigma_k \eta^o_{i,t,k}\} \tag{6}$$

Individuals compare the expected future payoffs of joining the labor force at age 25 with a CS degree $V^c_{25,k}$ with the future payoff of joining with some other degree $V^o_{25,k}$. Such choices are also affected by a fixed education cost for studying CS, $\bar{F}_k$, which can be positive or negative, and idiosyncratic taste shocks for studying each field: $\eta^c_{i,k}$ and $\eta^o_{i,k}$.[10] We assume that $\eta^c_{i,k}$ and $\eta^o_{i,k}$ are independently and identically distributed as a standard Type I Extreme Value distribution (Rust, 1987). The parameter σ_k controls the sensitivity of major and occupation choices to preference shocks. A smaller σ_k 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 beginning of each five-year period, individuals between ages 25 and 65 choose to work in CS or another occupation to maximize the expected present value of their lifetime utility. We denote occupational choice as $x = \{cs, other\}$ for the

[7] The labor supply side of the model is related to previous dynamic models of occupational choice such as Bound et al. (2015) who model the US labor market for computer scientists.

[8] To explore this in a more tangible manner, consider tabs from the American Community Survey (ACS) and National Survey of College Graduates (NSCG) to understand what occupations are CS workers more likely to switch into. Our calculations indicate that the majority of CS switchers move into managerial positions, including 'Managers and Administrators,' 'Management Analysts,' and 'Supervisors and Proprietors.'

[9] In Section 8.4 we introduce heterogeneity in abilities as well, and positive selection of migrants on ability.

[10] Getting a CS degree does not ensure a CS occupation, as shown in our model of occupation choice below.

individual's problem in Equation 7:

$$V^x_{t,a,k} = \max_x \{w^x_{t,k} + \chi_k \times \mathbb{1}(x_t \neq x_{t-1}) + \zeta_k \times \mathbb{1}(x_t = oth) + \beta \mathbb{E}_t[V^x_{t+1,a+1,k}] + \sigma_k \eta^x_{i,t,k}\} \quad , \tag{7}$$

where $V^x_{t,a,k}$ is the value of starting in occupation x 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 life stream of income. $\zeta_{1,k}$ is the taste attractiveness parameter for not working as a computer scientist, and χ_k, is the monetary cost of switching occupation. For simplicity we follow Artuc et al. (2010) and model the switching cost as a fixed cost. In the model, all workers retire at age 65 and their retirement benefits do not depend on their career choices. As a consequence, workers at age 60 face the same decision problem but, without consideration for the future. The current and future wages in both occupations $w^x_{t,k}$ are perfectly anticipated by the workers and taken as given.

Finally, we introduce the possibility of migration into the model. Indians can migrate to the US if they choose to work in CS. At the beginning of each period, Indian workers choose their occupation without knowing whether they will get to migrate. They make their decision based on the *expected wage*, by weighting the wage of CS in India $w^\ell_{t,in}$ and in the US $w^\ell_{t,us}$ with the probability of migrating $q^m_{t,us}$, as in Equation 8.

$$w^{cs,e}_{t,in} = q^m_{t,us}(w^\ell_{t,us} - \kappa) + (1 - q^m_{t,us})w^\ell_{t,in} \quad , \tag{8}$$

where κ is a fixed migration cost workers pay when migrating. The probability of getting a US job is determined by the number of Indian CS workers going to the US every year (determined by the H-1B cap, cap_t, set by US policy) and the number of CS in India that year ($L_{t,in}$):[11]

$$q^m_{t,us} = \frac{cap_t}{cap_t + L_{t,in}} \tag{9}$$

Once workers choose their occupation conditional on the expected wage, a share $q^m_{t,us}$ of all CS in India migrate to the US, and the rest remain in India. The value function for Indian college graduates is as in Equation 7, but using the expected wage $w^{cs,e}_{t,in}$ for CS. While we ignore the possibility of heterogeneous migration preferences, as long as there is a large wage premium in the US with respect to India, and the H-1B cap remains small relative to the total number of CS in India, it is a reasonable simplification to assume that there will be workers who want to migrate to the US. In our baseline specification we assume that if Indian CS migrate to the US, they stay there until the end of their careers. In Section 8.1 we explore the implications of

[11] In Section 8.4 we allow for positive selection into migration on ability. For now, we assume that migrating CS have similar skills to the CS that remain (we relax this in Section 8.4). Given the large number of H-1B applications, lottery winners are similar to losers. Furthermore, CS immigrants perform a wide variety of tasks across the CS wage distribution. In our NSS data CS workers have consistently lower variance in wages than similar occupation groups like doctors and professors, suggesting not a very wide distribution of skills.

allowing some migrants to return to India with newly acquired skills.

The uncertainty of migration when making the occupational decision is crucial. A decrease in the migration cap to the US lowers the expected wage for Indian graduates in CS, which in turn, lowers the total supply of Indian CS. 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, the long-run elasticity is higher.

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 $v \in [0, 1]$.[12]

$$Y = \left(\int_0^1 y_v^{\frac{\iota - 1}{\iota}} dv \right)^{\frac{\iota}{\iota - 1}} , \tag{10}$$

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

4.2.2 Final Goods Production

Each firm producing variety v in the final goods sector has a Cobb Douglas constant returns to scale technology over intermediate inputs from the IT sector $C_{v,y}$ and the labor aggregate x. Each variety can be produced with productivity level $z_{v,y}$:

$$y_v = z_{v,y} C_{v,y}^{\gamma} x_{v,y}^{1-\gamma} \tag{11}$$

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.[13] Following the framework

[12] 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 are time and country specific.

[13] 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.

introduced by Dornbusch et al. (1977) and Eaton and Kortum (2002), each country has a different level of efficiency in producing each variety, denoted by $z_{v,y}$. The final goods sector employs three types of labor denoted by subscript y. $x_{v,y}$ is a labor aggregate of non college graduates $h_{v,y}$, and an aggregate of college graduates $q_{v,y}$:

$$x_{v,y} = \left[\alpha(h_{v,y})^{\frac{\tau-1}{\tau}} + (1-\alpha)(q_{v,y})^{\frac{\tau-1}{\tau}}\right]^{\frac{\tau}{\tau-1}} \tag{12}$$

Using a nested CES format, the aggregate of college graduates $q_{v,y}$ can be represented by Equation 13, where $\ell_{v,y}$ is the number of CS hired in the final goods sector, and $g_{v,y}$ is the number of non CS 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.[14] Both sectors have the same elasticity of substitution between college and non college graduates τ, and between CS and non-CS college graduates λ.

$$q_{v,y} = \left[\delta \ell_{v,y}^{\frac{\lambda-1}{\lambda}} + (1-\delta) g_{v,y}^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}} \tag{13}$$

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 δ and sector-biased technical change α will shift over time with the innovation boom.

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.

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:

$$C_y = \left(\int_0^1 c_j^{\frac{\iota-1}{\iota}} dj\right)^{\frac{\iota}{\iota-1}} \tag{14}$$

IT firms have CES technology in the labor aggregate (Equation 15), where $\ell_{j,c}$ 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 parameter of the CES function. $\Delta > 0$ as IT is more intensive in CS than the final goods sector.

[14] In Section 8.3 we allow foreign CS to be imperfect substitutes to natives CS.

$$c_j = z_{j,c}\left[(\delta+\Delta)\ell_{j,c}^{\frac{\lambda-1}{\lambda}} + (1-\delta-\Delta)g_{j,c}^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}} \tag{15}$$

4.2.4 International Trade

We model the world economy as a set of three regions: 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 have the option of buying and selling products to a third region. While workers in RoW produce and consume both final goods and IT goods, we simplify the analysis by assuming they do not receive or send CS migrants.[15]

All three regions trade both goods 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 $z_{j,s,k}$ comes from the Frechet (Type II extreme value) distribution:

$$F_{s,k}(z) = e^{-T_{s,k}z^{-\theta}} \tag{16}$$

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

Consumers in each country will buy each variety from the lowest-price producer. If a consumer in k buys from b they pay an iceberg-trade cost, modeled as a share of the final good that gets lost when moving the good from k to b. Given each variety is produced in a perfectly competitive market, the price of a good produced in country k and sold in country b can be written as the marginal cost of production in each sector ξ_k^s times the iceberg trade cost $d_{k,b}$ divided by the variety-specific productivity in country k, $z_{s,k}$ as shown in Equation 17:

$$p_{j,k,b}^s = \frac{d_{b,k} \times \xi_k^s(w_{\ell,k}, w_{g,k}, w_{h,k})}{z_{j,s,k}}\ , \tag{17}$$

All else equal, a country becomes a more attractive provider of the good whenever one of three things happen: an increase in the technology (that allows for better draws of $z_{s,k}$), a decrease

[15]Since the labor market in RoW is not our focus, we avoid modeling their occupational distribution and assume output in both sectors is produced with a single type of worker that can freely move across sectors.

in trade costs $d_{b,k}$, or a decrease in labor costs.

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 country k, reflected by more imports in b from k. US-owned firms producing and exporting from India count towards Indian production and exports, and the same is true for Indian-owned firms in the US. A US firm that outsources to India may then import this good to the US.

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 mitigate the depression in CS wage growth due to immigration. Since the IT output is an intermediate input into the final goods sector, technological advances can increase the productivity of other sectors of the economy as well.

Innovation in India depends on the number of CS workers. The 'brain drain' of CS workers to the US is countered by the 'brain gain' of return migrants, and workers acquiring CS skills with the prospect of migrating. Workers that migrate from India to the US acquire human capital in the form of skills and technologies, and when they return they bring this knowledge with them. This spread of technology makes the Indian IT sector more productive, and over time the leading exporter of IT goods.[16] To capture this feature we model IT sector productivity in country k to be a function of the total number of CS IT workers in country k:

$$T_{c,k} = T(L_{c,k}) \;\; for \; k = \{us, in\} \tag{18}$$

4.3 Equilibrium

Equilibrium in each period can be defined as *a set of prices and wages ($P_{t,c,k}$, $P_{t,y,k}$, $w^{\ell}_{t,k}$, $w^{g}_{t,k}$, $w^{h}_{t,k}$), quantities of output and labor ($C_{t,y,k}$, $Y_{t,k}$, $L_{t,k}$, $G_{t,k}$, $H_{t,k}$), 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*

[16]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).

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 19:

$$\pi^s_{t,k,b} = \frac{T^s_{t,b}(d^s_{t,k,b}\xi^s_{t,b})^{-\theta}}{\sum_b T^s_{t,b}(d^s_{t,k,b}\xi^s_{t,b})^{-\theta}} \tag{19}$$

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:

$$P^s_{t,k} = \bar{\gamma}\left(\sum_{k'} T^s_{t,k'}(d^s_{t,k,k'}\xi^s_{t,k'})^{-\theta}\right)^{-\frac{1}{\theta}}, \tag{20}$$

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 $\bar{H}_k$ 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 which requires the definition of 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 $x_{t-1} = x$ that choose occupation x' in period t as in Equation 21:

$$\pi^{x,x'}_{t,k} = \frac{exp(\frac{1}{\sigma_k}\bar{V}^{x,x'}_{t,k})}{exp(\frac{1}{\sigma_k}\bar{V}^{x,x'}_{t,k}) + exp(\frac{1}{\sigma_k}\bar{V}^{x,x}_{t,k})}, \tag{21}$$

where $\bar{V}^{x,x'}_{t,k} = w^{x'}_{t,k} + \chi_k \times \mathbb{1}(x \neq x') + \zeta_k \times \mathbb{1}(x' = oth) + \beta\mathbb{E}_t[V^{x'}_{t+1,a+1,k}]$. As in Bound et al. (2015), this equilibrium is characterized by the system of Equations (6-9) 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 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 towards 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 perfect foresight on the evolution of these 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: $s = \{\mathbf{L^{20}_{t-1,k}}, \mathbf{L^{25}_{t-1,k}}, ..., \mathbf{L^{60}_{t-1,k}}, \mathbf{G^{25}_{t-1,k}}, ..., \mathbf{G^{60}_{t-1,k}}\}$ to the values of $L_{t,k}$, $G_{t,k}$, $w_{t^{\ell},k}$, $w^{g}_{t,k}$, and $\mathbf{V_t}$, the vector of career prospects at different occupations for different ages, that satisfies the system of Equations 6-9 as well as each period's relative demand curve.

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.[17] To make the exposition transparent, we separate this process into four building blocks of parameters: 1) Product market elasticities: τ, λ, ι and θ; 2) The elasticity of labor supply of college graduates in US and India that are pinned down by labor supply parameters: $\Theta = \{\sigma_k, \zeta_k, \chi_k, \bar{F}_k, \kappa\}$ for $k = \{us, in\}$; 3) 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, 4) the endogenous technology parameter that determines the CS innovation spillover.

The solution algorithm for the first three blocks consists of the following steps: First, we set the product market elasticities to be values estimated in the literature, or using other data (Block 1). Second, we guess the labor supply elasticity parameters $\hat{\Theta}_{guess}$ (Block 2) 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 3). 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 predicted data moments related to the labor supply. Our algorithm repeats the process, searching over the labor supply parameters until the distance between predicted and observed data moments is minimized (Block 2). In Section 5.1 - 5.3, we discuss how we determine the first three blocks of parameters, and in Section 5.4 we estimate the spillover parameter.

5.1 Product Market Elasticities

As a first step, we set the product market elasticities, and vary them in robustness checks. We set the elasticity of substitution between college and non college graduates, $\tau = 1.7$ based

[17] We stop in 2010 for two reasons: First, after 2010, there was a shift in H-1Bs to outsourcing firms (Park, 2015). Second, using administrative data, we see that despite being stable for many years, after 2012 there was a rapid increase in Indian students enrolling in Engineering degrees at US universities. While our model implicitly allows for these, we avoid going further than 2010 as such features are less explicitly modeled.

on an average of papers that estimate that parameter (Card and Lemieux, 2001; Goldin and Katz, 2007; Katz and Murphy, 1992), and explicitly estimate it using data from India.[18] For the elasticity of substitution between computer scientists and non-CS college graduates we set $\lambda = 2$ which is within the estimates of Ryoo and Rosen (2004).[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 $\iota = 4$. Finally, we use the trade elasticity $\theta = 8.28$, the value proposed by Eaton and Kortum (2002), and in robustness checks, we vary this parameter to $\theta = 4$ using the estimate of Simonovska and Waugh (2014).

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 C.

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 the 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 $\bar{H}_{t,k}$, $\bar{Q}_{t,k}$, and using Equation 22 we estimate $\alpha_{t,k}$:

$$\vartheta_{t,k} = \frac{\alpha_{t,k}(\bar{H}_{t,k})^{\frac{\tau-1}{\tau}}}{\alpha_{t,k}(\bar{H}_{t,k})^{\frac{\tau-1}{\tau}} + (1-\alpha_{t,k})(\bar{Q}_{t,k})^{\frac{\tau-1}{\tau}}} \tag{22}$$

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

The distributional parameter between CS and non-CS college graduates $\delta_{t,k}$ is calibrated so

[18] We replicate Card and Lemieux (2001) using the India data and estimate an elasticity of complementarity of 0.55 (see Table A6). 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.

[19] They estimate the elasticity of substitution between engineers and other graduates to be 1.2 - 2.2.

that it matches the within country relative wages between CS and non-CS college graduates observed in the data. This parameter 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 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 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 19 and take the ratio between the probability of country k buying from country b and the probability of country k buying form itself which yields the gravity Equation 23.

$$\frac{\pi^s_{t,k,b}}{\pi^s_{t,k,k}} = \frac{EX^s_{t,k,b}}{EX^s_{t,k,k}} = \frac{T^s_{t,b}(d^s_{t,k,b}\xi^s_{t,b})^{-\theta}}{T^s_{t,k}(d^s_{t,k,k}\xi^s_{t,k})^{-\theta}}\,, \tag{23}$$

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 23 to estimate trade costs and a term that combines the technology level and the unit cost of production $T^s_{t,k}(\xi^s_{t,k})^{-\theta}$. We use this term as a parameter in the model, which allows us to separately calculate the unit costs $\xi^s_{t,k}$ and the technology level $T^s_{t,k}$. In Appendix D.1 we provide more detail on how we estimate the parameters for trade costs and technology.

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 that 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.[20]

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

[20] We explore the possibility that the net change includes both Indians in India migrating to the US and Indians in the US returning to India in Section 8.1 where we allow for the possibility of return migration.

changes in technology and production function parameters as in Section 5.2. Such exogenous innovation shocks are captured by 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. If we define the product market parameters as in Sections 5.1 and 5.2 as $\hat{\Omega}$ we can calculate a vector of moments predicted by the model, $m(\hat{\Omega}, \hat{\Theta}_{guess})$, using parameters $\hat{\Omega}$ and labor supply guess $\hat{\Theta}_{guess}$. The algorithm will search over $\hat{\Theta}$ such that it minimizes the distance between predicted moments 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 predicted moments and W is the weighting matrix. $\hat{\Theta}$ is composed of 9 parameters, the taste dispersion parameters σ_k, the mean tastes for non-CS occupations ζ_k, the occupation switching costs χ_k, the education costs for CS $\bar{F}_k$ and the migration cost κ. To separately identify each parameter, we choose 9 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 moments and the moments observed in the data. Both for India and US we choose the following moments: the share of workers in CS in 1995 and 2010; the ratio between the CS share among those between 25-30 years old relative to the CS share among those 31 to 60 years old in 2010, and the net occupation switching rate between year 1995 and 2000.

The final (and somewhat, important) 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 10% increase in high-skill migrants working in a given occupation in the US would cause a 4.5% increase in the labor supply of that occupation in India (Table 1).

While the system uses all the data at the same time, there is strong intuition behind the identification of each parameter. The CS share in 1995 helps identify the mean taste for non-CS, since 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 the dispersion parameter σ_k. A higher σ_k means that individuals assign high weights to whatever idiosyncratic preference shocks they receive every period, making them less responsive to changes in the relative wage. How much the CS share changes conditional on wages will help identify the responsiveness of the CS share over time. 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 will mean that it is easier to join CS occupations at the initial period (paying the education cost) than later on by switching occupations. The net switching rate will primarily identify the switching costs χ_k, since higher switching costs

Table 2: Empirical vs. Simulated moments

	US		India	
	Data Moments	Simulated Moments	Data Moments	Simulated Moments
Supply response to migration	–	–	1.045 [0.88-1.2]	1.046
Transition rate	1.39% [1.1%-1.7%]	1.49%	0.16% [0.1%-0.2%]	0.18%
Ratio CS Share [25-30]/[31-60]	1.00 [0.83-1.17]	1.01	3.28 [1.77-4.8]	3.28
Share CS 1995	2.96% [2.7% -3.2%]	3.15%	0.10% [0.06%-0.14%]	0.28%
Share CS 2010	3.90% [3.7%-4.1%]	4.01%	1.99% [1.5%-2.5%]	1.96%

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 1. '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 3 are simultaneously estimated by matching all these moments across both countries. 95% Confidence intervals for empirical moments are in parenthesis.

are expected to decrease switching both in and out of CS. Finally, the employment elasticity with respect to migration will help identify the migration cost. When migration becomes more costly, the local CS labor supply in India will become less responsive to changes in the wage in the US, lowering the local response to migration.

We construct the nine 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.[21] Finally, for the employment elasticity with respect to migration, we use the estimated elasticity of 0.45 from Section 3, Table 1.

The Equation 24 minimization routine minimizes the distance between the predicted and empirical moments. The model is perfectly identified (9 data moments to identify 9 parameters). Table 2 shows the predicted moments match very closely to the empirical data moments:

The estimated labor supply parameters can be found in Table 3. 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. Labor market frictions in India are higher than in the US: India shows a more negative switching cost and a higher preference dispersion, indicating that occupational choices are less responsive to wage changes in India. Education cost of CS (relative to non-CS) is positive for both countries which can be interpreted as, conditional on individuals

[21]As an example, if one cohort has 10 individuals in CS and 20 in Other at $t-1$, and 15 individuals in CS and 15 in Other at t, then the net switchers are 5 and the net stayers are 25. The net switching rate with just this observation would be $5/25 = 0.2$

Table 3: Labor supply parameter estimates

	US	India
Occupation Switching Costs	-0.42	-2.44
Mean taste for non-CS Occupations	0.609	0.319
Preference dispersion Parameter	0.265	0.453
Education Cost for CS Degrees	0.638	0.020
Migration Cost from India to US		0.025

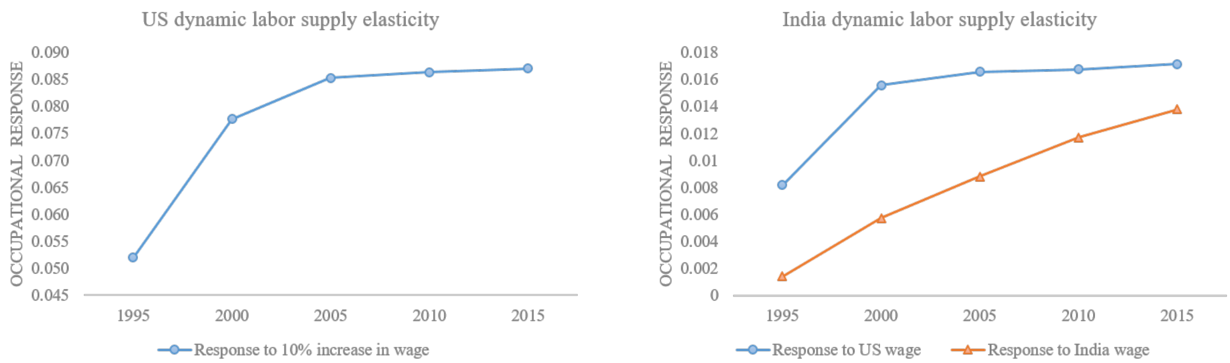
Estimated labor supply parameters based on the simulated method of moments exercise. These parameters jointly determine the short-run and long-run labor supply elasticities. 'Switching costs' are occupation switching costs between CS and non-CS occupations. 'Mean taste for non-CS' is the average preference for non-CS occupations. 'Preference dispersion' is the variance in tastes for 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).'Migration cost' is the cost of migration from India to the US in units of a thousand numeraire.

having a mean taste towards non-CS occupations, it is easier to enroll in CS than in other college majors. Finally, we estimate a small migration cost.

Together, the parameters estimated in Table 3 determine the dynamic labor supply elasticity in each country. We represent this elasticity in Figure 5, where we artificially raise the CS wage by 10% starting in 1995, and estimate the labor supply response over the years. In the short run, the labor supply curve is inelastic, as few people shift into CS occupations, but over the longer run, as more and more students major in CS degrees, and more workers switch into CS, the labor supply curve becomes more elastic. Estimating this dynamic labor supply elasticity is an important contribution of our work; previous reduced-form estimates in the immigration literature that find little effect on native occupation choice are essentially estimating a static elasticity, and as such, providing a limited picture of the labor supply response.

Figure 5: Dynamic occupational choice in response to wage increase

(a) Effect of 10% increase in CS wage (b) Effect of 10% increase in CS wage



Graphs show the dynamic labor supply response to a 10% increase in CS wages in 1995. The short and long run response depend on the labor supply parameters estimated in Table 3.

5.4 Endogenous Technology: Patenting Response to CS

In Section 4.2.4, Equation 17 we mentioned that the level of technology of the IT sector depended on the number of CS working in IT without specifying any functional form on this relationship. 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 effect that affects the overall level of technology. In our counterfactuals, we change the number of foreign CS that are in the US and India which presumably affects the level of technology, so we need an estimate on how technology changes when the number of CS changes.

To estimate the elasticity of technology with respect to number of CS workers, we use an instrumental variables strategy that exploits the industry specific dependence on immigrant CS workers. We first combine our data on immigrant CS workers by industry with data on patenting from the US Patent and Trademark Office (PTO).[22] 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 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-Card (2001) 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.[23] Equation 25 captures the first stage of our strategy.

$$CS_{j,t} = \alpha + \delta_j + \delta_t + \gamma \left(\frac{Imm\ CS_{j,0}}{Emp_{j,0}} \right) Imm\ CS_t + \epsilon_{j,t}\ , \tag{25}$$

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:

$$Log(Patents)_{j,t} = \mu + \delta_j + \delta_t + \beta_{tech} \widehat{CS_{j,t}} + \varepsilon_{j,t}\ , \tag{26}$$

[22]Details of the data can be found in the Appendix. 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.

[23]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.

Table 4: Migration, Computer Scientists and Patenting by Industry

	CS workers	Log(Patents)	Log(Patents)
Shift Share	1.258***		
	(0.309)		
CS workers		5.45e-06**	5.79e-06**
		(2.59e-06)	(2.56e-06)
Observations	275	275	275
R-squared	0.877	0.967	0.967
Number of Industries	25	25	25
Additional Controls	No. of Firms	No. of Firms	None
F stat		16.67	16.31
Elasticity		0.226	0.240
SE		(0.107)	(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 fifty patents over the entire period. Standard errors clustered by industry. Robustness to alternative specifications can be found in Table A5.

The results of this exercise are shown in Table 4. 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 found in Table A5, where we vary the controls in the regression, look at only the flow of new patent filings, and exclude the truncated patent data.[24]

Importantly, our estimated elasticity is very close to similar findings in the literature. Peri et al. (2015b) 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.[25] 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 our counterfactual simulations we estimate the change in $T^c_{t,k}$ to be consistent with the spillover parameter β_{tech} from Equation 26. Implicitly, we are assuming that the endogenous productivity growth happens through computer scientists working in the IT sector. This is consistent with the work of Jorgenson et al. (2016) and Byrne et al. (2013) who estimate that IT producing industries contributed more than 50% of the aggregate productivity growth in the US between 1995-2014. At the same time, Peri et al. (2015a) estimate that foreign STEM workers alone, contributed between 30% and 50% of the aggregate productivity growth between 1990-2010. As most foreign STEM workers come to work as computer scientists, it is reasonable to assume that overall, IT is a predominant driving force for productivity growth.

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

[25]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.

6 Endogenous Variables and Model Fit

We study the evolution of endogenous variables in our model over time, and evaluate how well it matches data. 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. Figures A1-A2 show that we match fairly well some of the key aspects that we are trying to capture.

As we see in Figure A1, 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 A1a and A1b that the model does fairly well in predicting the trends and level differences between the wages in both countries. In Figure A1a the CS 'place premium' in our model closely matches the data, and quasi-experimental results that show a 6-fold increase in wages for H-1B lottery winners (Clemens, 2013). Similarly, in Figure A1b we show the non-CS college graduate wage premium between the US and India is in line with the data. We do not plot the non-college graduate wage as we are explicitly matching that series in our exercise.

In Figure A2 we study how IT output and prices, and the location of IT production evolves over time. Figure A2a and A2b display a close match between the model and data for relative US IT prices and output growth. We also predict the US CS occupational share, and the relative shares between the US and India, even though we do not match these data when solving the model (Figure A2c - A2d). Together, the success of the out-of-sample matching gives us additional confidence in our modeling exercise to perform counterfactual tests.

7 Counterfactual Exercises

Policy makers often debate changing the H-1B cap. To evaluate the impact of the H-1B program on US and Indian economies we conduct a counterfactual exercise where, in the baseline specification, we reduce the H-1B cap by 50% every year starting in 1995. In Section 8.6 we vary the cap across different sizes. With a lower cap, some workers who may have been granted visas to the US are now forced to work in India. As our model focuses on migration of CS from India only, halving the cap, halves the number of CS migrants from India. Using our estimates parameters and the given set-up we can trace out what happens to all endogenous variables between 1995 and 2010. To discuss the results, we look at the changes of going from the counterfactual scenario with limited migration (50% of the observed 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 grow or facilitates the shift in production to India.

7.1 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 over 11,000 in 1995 and by over 113,000 in 2010. The increase in migration of CS drives down CS wages in the US leading to more than 81,000 US-born CS to switch to non-CS occupations. In India, college graduates respond to the migration incentives by getting into CS occupations and while some migrate many others stay in India, increasing the number of CS in India by 29,000 workers.

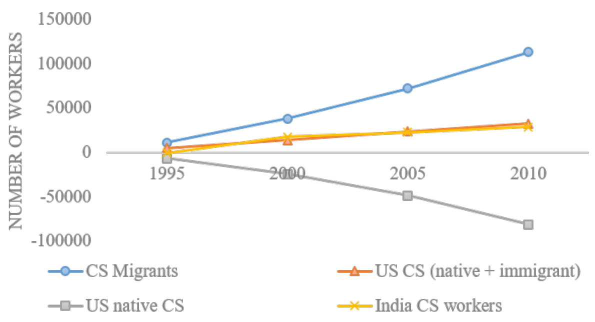
Both the US and India increase the total number of CS which is a complementary occupation to non-CS college graduates and high-school graduates. In addition, CS workers have a spillover effect that benefits all workers, making total net welfare gains positive for both countries as shown in Figure 6b. Welfare for US natives increases by $100 million by 2010 while Indian natives who stay in India gain $171 million when migration increases to the US. In Appendix Table A7 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 $927 per migrant, and of $1582 per migrant to Indian non-migrants. If we also include gains to migrants themselves, overall net welfare increases by $53 thousand per migrant.

Even though natives from both countries benefit from migration, there are distributional gains and loses. Since the total number of CS in both countries increases, those already working in CS lose from migration (despite the gains from innovation) as their wages decline with increased competition (Figures 6c-6d). 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 both countries benefit from migration (as prices are lower and they are complements in production), although more so in the US since the total number of college graduates increases, whereas it decreases in India (due to emigration).

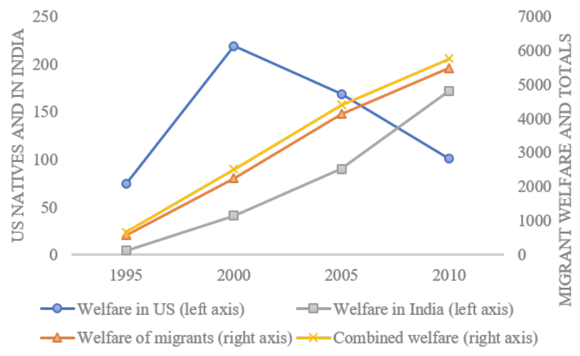
Accounting for the occupational response of Indian college graduates to changes in migration policy in the US is crucial to quantifying the gains and losses from migration. To see this, in Table 5 we compare the results between our baseline model (column 1), and alternative models that shut down the supply response in India. Column 2 shows the results for a model where

Figure 6: Effect of immigration on occupational choice and welfare

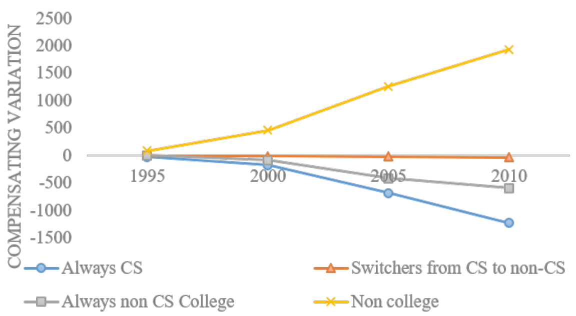
(a) Change in Number of Workers Due to Immigration



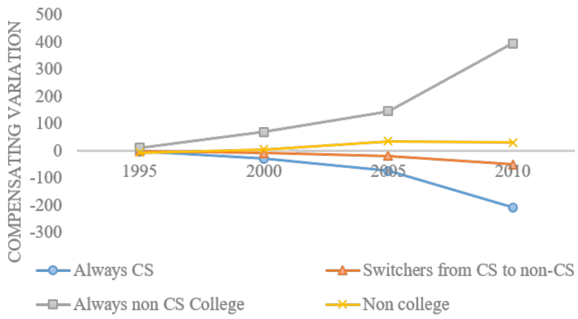
(b) Compensating Variation (USD mn)



(c) Welfare Change: US native Workers (USD mn)



(d) Welfare Change: Workers in India (USD mn)



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.

individuals in both 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, 4.72% of US college graduates in CS switch to other occupation when migration is increased, driven by a -0.89% reduction in the CS wage. In India, the total supply of CS increases (despite emigration) when migration to the US is allowed. This increases the size of the Indian IT sector by 15.07%, while the US IT sector decreases by 0.48% as natives switch away from CS, and there is increased competition from India for the world market.

If Indians were not allowed to respond to migration opportunities abroad (Column 3), India would only experience brain-drain as their CS leave the country, lowering IT output by 7.84%. As a consequence, the increase in total CS in the US would increase US IT output by 2.31% reversing the effect in the baseline model, as the US captures the market that was satisfied by

Table 5: Effect of Migration in 2010: Baseline vs No Occupational Choice

	Baseline	No occupational choice	
		In both countries	**In India only**
Wages			
US CS workers	-0.89%	-2.21%	-0.65%
India CS workers	-9.7%	3.6%	4.1%
Occupational Choice			
US CS (native & immigrant)	1.21%	6.0%	2.20%
US CS workers	-4.72%	–	-3.77%
India CS workers	31.2%	–	–
IT production			
US IT output	-0.48%	4.88%	2.31%
India IT output	15.07%	-8.8%	-7.84%
Welfare			
Welfare of US natives	0.002%	0.03%	0.01%
Welfare in India	0.08%	-0.05%	-0.05%

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. '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.

India. Shutting down the supply response in India also has significant consequences for welfare. The total welfare of US natives due to immigration is 5 times higher than in the baseline, while in India welfare *decreases* by 0.08% as the effects of brain-drain outweigh brain-gain. When restricting occupational choice in both countries (Column 2), total welfare in the US rises even more, as US CS cannot change occupations and the US captures all the spillover created by CS innovation. The results in Table 5 highlight the importance of the main mechanism introduced in this paper: taking into account supply responses in sending countries can drastically mute the welfare effects of increasing migration.

In Appendix Table A9 we study the sensitivity to varying the three elasticities we take from the literature, which correspond to parameters λ, τ and θ. Our results are similar to the baseline, even as some magnitudes vary with the elasticity values in the manner one would expect.

8 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 four main extensions in order to tease out the mechanisms underlying our base specification, and try alternative specifications. First, in Section 8.1, we separate out the mechanisms underlying brain drain and brain gain present in our model. We additionally incorporate the possibility of a fraction of migrants returning to India with newly acquired skills, as this is potentially an additional channel of brain gain. In Section 8.2 we explore the role played by endogenous technology by varying the technology elasticity, and in Section 8.3, we look into the possibility of migrant workers in the US being imperfect substitutes with native CS. Section 8.4 discusses the implications for a specification that has heterogeneity in abilities and positive selection into migration, while Section 8.5 discusses the role played by trade and remittances. Finally, Section 8.6 varies the size of the H-1B cap, and the starting period of immigration restrictions, to explore alternative counterfactuals.

8.1 Brain Drain, Brain Gain and Return Migration

The baseline results presented in Section 7.1 include a combination of brain drain and brain gain for India. On the one hand, some CS leave for the US creating brain drain. On the other hand, some college graduates choose CS with the prospects of migrating but end up staying, generating a spillover effect and hence, 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 making their occupational choices such that India experiences no brain gain. As shown in Table 6 column 2, if Indians do not respond to migration incentives, the CS workforce in India decreases by 12.08% 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 2.31%. Net welfare gains from migration are larger in the US and lower in India when compared to the baseline since India does not experience brain gain when migration is increased.

As a second exercise in Table 6 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 migration (i.e. allowing brain gain). The CS workforce in India then grows rapidly, and IT output expands by 20%. Such increases makes the US IT sector shrink by 2%. When there is no brain-drain total welfare in India increases by 0.14% while decreasing in the US by 0.01%.

Last, we explore the possibility of return migration as an additional channel of brain gain for India. The H-1B is a temporary visa that lasts between 3 to 6 years, after which workers need to be sponsored by their employers for a green card in order to stay in the US. While many do stay in the US beyond the H-1B, it is reasonable to consider some go back to India with new skills and connections acquired during their time in the US. To explore how this affects our results we assume a fixed share of those who migrate return a period later to work as computer scientists in India. In this version of the model, native and return migrant CS workers are not perfect substitutes as return migrants have a different set of skills given their work history abroad. This is captured in Equation 27:

Table 6: Brain Drain vs Brain Gain: Main Outcomes

	Baseline	No occupational choice in India	No migration but reallocation	With return migration
Occupational Choice				
US CS (native & immigrant)	1.21%	2.20%	-0.96%	3.52%
US CS native	-4.72%	-3.77%	-0.96%	-6.90%
India CS	31.23%	-12.08%	58.25%	7.91%
IT production				
US IT output	-0.48%	2.31%	-2.19%	1.90%
India IT output	15.07%	-7.84%	20.10%	1.29%
Total welfare				
US natives	0.002%	0.01%	-0.01%	0.02%
Welfare in India	0.08%	-0.05%	0.14%	0.02%

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 drain'). 'No migration but reallocation' restricts migration (no 'brain drain') but allows for responses to migration opportunities ('brain gain'). 'With return migration' introduces another aspect of 'brain gain' – that of returning migrants.

$$\ell_{j,s} = \left[n_{j,s}^{\frac{\epsilon-1}{\epsilon}} + r_{j,s}^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}} , \tag{27}$$

where $n_{j,s}$ is the number of CS workers in sector s that never went abroad and $r_{j,s}$ the number of return migrant CS workers employed by firm j in sector s. Here ϵ is the elasticity of substitution between the native CS workers and return migrants.

To determine the substitution elasticity between CS who never worked abroad and those who return from the US we follow the literature on return migration. We would expect this elasticity to be greater than the elasticity between CS and non-CS graduates. The scant literature on return migrants find that in other contexts, those who emigrated for labor reasons and return home earn a wage premium relative to those that never migrated (Barrett and O'Connell, 2001; Hazans, 2008; Reinhold and Thom, 2013). In our steady state year, we match the average premium of 15% across papers in this literature, which corresponds to a value of $\epsilon = 30$.

Indian data do not contain information that allows us to distinguish return migrants and non-migrants. We create the series of return migrants based on Indian CS that are working in the US since 1980. We use the 1980 and 1990 US censuses to calculate, on average, how much has the Indian CS population in the US increased every year. From the average yearly increase between 1980-1994 we assume that each year, ϱ fraction of migrants return to India and create the series as the cumulative of those that go back in 1980 till those that go back in 1994. Once we get our initial stock for 1995, every five year period we assume that the number of return migrants evolves according to Equation 28:

$$R_{t+1,in} = R_{t,in} + \varrho \; cap_t \; , \tag{28}$$

where ϱ is the return rate to India. According to Lee (2016), the OECD estimates that 23.5% of high-skill immigrants in the US return to their home countries after a 6-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 close to $\varrho = 23.5\%$, and use it as our baseline value.

Interestingly, as shown in Table 6 column 4, when return migration is allowed the growth in welfare, CS employment and IT output in India are lower than in the baseline. This is the result of two countervailing effects on India. On the one hand, return migrants bring acquired skills and technical knowhow and enlarge the size of the CS workforce in India. On the other hand, the likelihood of returning home to lower wages in India, lowers the returns from migrating in the first place. These lower long-term expected returns from migrating inhibit the accumulation of CS skills for Indian students and workers, and restrict the size of the CS workforce in India. The lower growth of Indian CS workforce increases the migration welfare gains in the US, and lowers the overall gains for India. Appendix Table A10 shows how welfare for different groups changes under each of these alternative scenarios.

8.2 Endogenous Technology

The endogenous technology elasticity that we estimate in Section 5.4 using an instrumental variables strategy has a value of 0.23, and is consistent with other values in the literature (Kerr and Lincoln, 2010; Khanna and Lee, 2018; Peri et al., 2015b). In Table 7 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 higher in the US than in India, in order to capture the fact that the US may do more R&D than India.

H-1B driven migration barely affects US welfare 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 have switched to CS with the prospect of migrating, but were unable to migrate due to the cap. There is, therefore, a sub-optimal oversupply of CS workers, and no gains from innovation, lowering welfare. For the same reason, when spillovers are lower (columns 2-4) than in 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.

Table 7: Varying the Elasticity of Endogenous Technological Spillovers

	Baseline	No spillovers	US spillover=0.23 India spillover=0.1	Spillover = 0.1 in both countries
IT production				
US IT output	-0.48%	0.82%	-0.03%	0.05%
India IT output	15.07%	4.02%	10.57%	7.89%
Welfare				
Welfare of US natives	0.002%	0.002%	0.004%	0.002%
Welfare in India	0.081%	-0.016%	0.026%	0.025%
Total (with migrants)	0.117%	0.094%	0.104%	0.101%

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 low spillovers, the US elasticity is 0.23 and India elasticity is 0.1. We further include a robustness exercise where the spillover elasticity is 0.1 in both countries.

8.3 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 as in Equation 29:

$$\ell_{j,s} = \left[n_{j,s}^{\frac{\upsilon-1}{\upsilon}} + f_{j,s}^{\frac{\upsilon-1}{\upsilon}} \right]^{\frac{\upsilon}{\upsilon-1}} , \tag{29}$$

where υ is the elasticity of substitution between natives and migrants which we set to 10 following the lower bound estimated by Ottaviano and Peri (2012).[26] $n_{j,s}$ is the number of native CS workers in the US in firm j sector s, and $f_{j,s}$ is the number of immigrants from India working as CS in firm j, sector s in the US.

As shown in Table 8 column 2, US welfare gains from increasing migration would be tenfold higher (due to worker complementarities) than the baseline of perfect substitutes. Both countries would increase IT production and US CS workers would be less likely to switch away from CS than in the baseline.

[26]When looking at educated workers Ottaviano and Peri (2012) find an elasticity of substitution of 12.6. Burstein et al. (2018) estimate a within occupation elasticity of 5.6 and an overall elasticity of 10.1.

8.4 Heterogeneity in Abilities

Next we explore how our results would change if we have heterogeneity in abilities and positive-selective 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 mitigate the gains in India. We assume college graduates in the US and India draw an ability for CS and non-CS occupations, $\phi_{\ell,i}$, $\phi_{g,i}$ from independent Normal distributions with mean 0 (a normalization) and variance $\sigma_{\phi,\ell,k}$ and $\sigma_{\phi,g,k}$. We can now write the wage that worker i receives for working in CS or Other college occupations as in Equation 30:

$$w_{i,t,x,k} = exp(\phi_{x,i}) \times w_{t,x,k} \quad , \tag{30}$$

where $exp(\phi_{x,i})$ is a parametrization of worker i's human capital in occupation x, and $w_{t,x,k}$ is the wage per effective unit paid for occupation x in country k at time t. This heterogeneity extension is a simplified version of the labor market presented in Dix-Caneiro (2014). To identify the additional parameters $\sigma_{\phi,\ell,k}$ 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 4 additional moments in our SMM procedure.

Table 8: Imperfect substitution between native and migrants, and heterogeneous abilities

	Baseline	Imperfect substitution	Heterogeneity in abilities
IT production			
US IT output	-0.48%	1.53%	1.01%
India IT output	15.07%	8.97%	11.72%
Occupational choice			
US CS (native & immigrant)	1.21%	2.08%	4.03%
US CS native	-4.72%	-4.24%	-6.10%
India CS	31.23%	26.66%	33.49%
Welfare			
US natives	0.002%	0.02%	0.01%
Welfare in India	0.08%	0.07%	0.10%
Combined (with migrants)	0.12%	0.13%	0.09%

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. 'Heterogeneity in abilities' models heterogeneous abilities and selection into occupations, and migration, based on ability draws.

The ability draws are constant throughout individual careers while idiosyncratic preference shocks are drawn every period. Hence, heterogeneity in abilities will act as a 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 31:

$$q^{m}_{t,us,i} = \frac{\bar{q}_t \times cap_t}{cap_t + L_{t,in}} \times F(\phi_{\ell,i}|\sigma_{\phi,\ell,in}) \tag{31}$$

where the baseline probability of migration from Equation 9, $\frac{cap_t}{cap_t+L_{t,in}}$ is multiplied by the Normal distribution function $F(\phi_{\ell,i}|\sigma_{\phi,\ell,in})$ evaluated at individual i's ability draw in CS, and reweighted by parameter $\bar{q}_t$. The sum of all the probabilities of migration among those that choose CS adds up to the total number allowed by the migration cap at time t.

As shown in column 3 of Table 8, under a scenario with heterogeneous abilities, the US IT sector would grow when we increase migration. This is partly driven by the most able CS workers being able to immigrate, which has a larger impact on IT in the US. Welfare for US natives is also higher than in the baseline scenario due to the US receiving more effective units of CS. India gains slightly more than in the baseline since those switching from other college occupations to CS are relatively less productive in non-CS occupations.

8.5 The Role of Trade and Remittances

Finally, we look at the relevance of international trade and remittances. Column 2 of Table 9 shows the impacts of migration for a model where bilateral trade costs between the three regions are 10% lower than in the baseline.

The increase in Indian IT production almost doubles in the case with lower trade costs when compared to the baseline. The US decreases its IT production even more, going from a decrease of 0.48% in the baseline to a decrease of 0.66% in the more open scenario. Welfare increase in the US is less when there is more trade, as IT growth in India shrinks the US IT sector. In sum, trade plays a crucial role in the shift in production from the US to India.

As an additional channel we test how our conclusions would change if we incorporate remittances into the model. To do this, we assume that Indians who migrate to the US remit a fixed share of their income to India and calibrate the fixed share to 3.83% of labor income, to match the share of Indian CS wage bill remitted.[27] Such income, is distributed evenly across those residing in India. Given the small share of income remitted, results for the model that includes remittances are very close to those of the baseline, as shown in Table 9, column 3.

[27] 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 US, which yields 3.83%.

Table 9: Increased trade and incorporating remittances

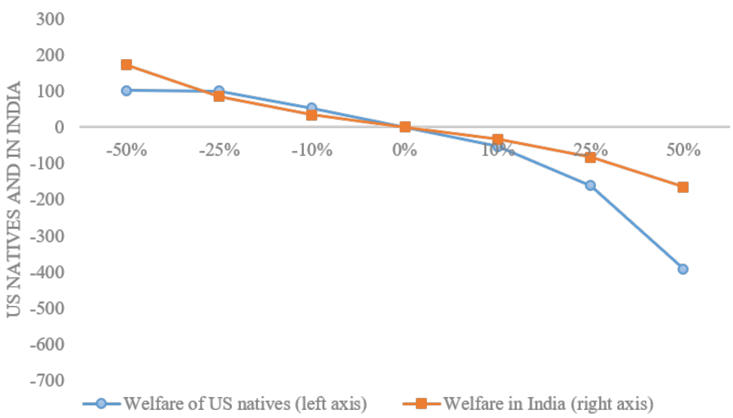
	Baseline	10% lower trade costs	Remittances
IT production			
US IT output	-0.48%	-0.66%	-0.47%
India IT output	15.07%	31.8%	15.1%
Welfare			
US natives	0.002%	0.001%	0.002%
Welfare in India	0.08%	0.10%	0.08%
Combined (with migrants)	0.12%	0.12%	0.12%

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.

8.6 Different Cap Sizes and Starting Periods

Our main counterfactual reduces the cap for Indian CS by 50% every year since 1995. In this section 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 7: Welfare Loss by Different Cap Sizes (USD mn)



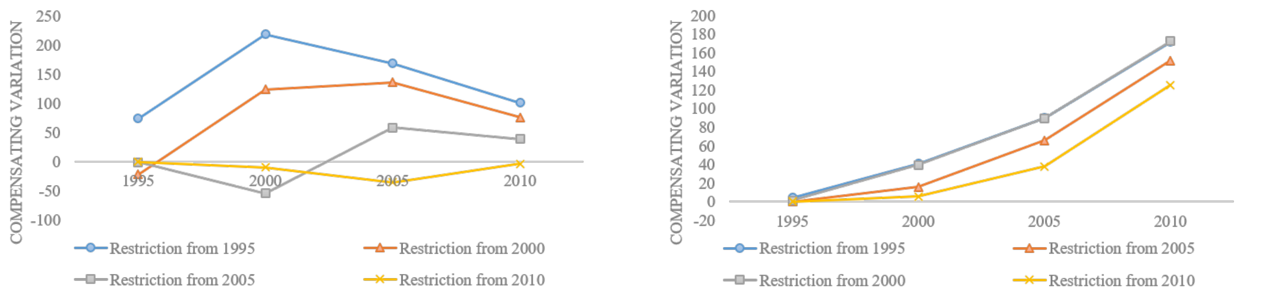
We solve the model for different counterfactual scenarios to elicit the welfare loss for each intensity of migration restriction. We vary the cap from a 50% reduction in the current H-1B cap to a 50% increase.

Figure 7 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 gain to not changing the cap. 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.[28]

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 8, a cap that starts later has a weaker effect on all outcomes and workers, since the cumulative growth of CS in India is lower.

Figure 8: 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.

9 Discussion

India experienced a dramatic expansion in IT employment, and 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. We capture the IT boom in the US and India with the help of a general equilibrium model. 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, and not unrelated, the wage differential between the US and India

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

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 3-6 years, obligating many to return to India with accumulated human capital and technical knowhow. Together, these features help spread the boom across the world from the US to India.

The average worker in each country is better off because of immigration. The program, however, has significant distributional consequences, where workers that are close substitutes are adversely affected while others benefit. These distributional effects have been at the forefront of academic discussion (Borjas, 1999; Peri and Sparber, 2011) and political debates. Importantly, certain countries may benefit more than others under such migration. In this paper 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 about 0.12%. This net gain is consistent with a long literature reviewed in Clemens (2011).[29] The welfare gains are approximately \$5.7 billion in total, the large fraction of which accrue to the migrants themselves. US natives are better off by about \$100 million in 2010 because of the H-1B program.

The gains, however, 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 somewhat striking result is that as production shifts (or is outsourced) to India, US IT output actually falls. A driving feature of this result is that an increase in the size of the Indian CS workforce increases the relative productivity of India's IT sector. Such reductions in US IT output have been discussed by a rich literature on the economics of trade and migration. Krugman (1979) and Vernon (1966) describe a North-South general equilibrium trade model where the North initially has a monopoly over new products given its technological superiority and rate of innovation. The South catches up due to technological diffusion and over time starts exporting to the North the very same products the North used to export. As the rate of technological diffusion increases, or the rate of innovation in the North declines, living standards will actually fall in the North. With quality differentiation in products, Flam and Helpman (1987) generate richer trade dynamics, but also show that technical progress in the South brings about a decline in the North's wage rate. Given certain rates of technical change, workers in the North may be harmed as production moves abroad (Acemoglu et al., 2015). Therefore, as Samuelson (2004) notes, such technical progress in the South erodes the US comparative

[29]Specifically, see Klein and Ventura (2007); Moses and Letnes (2004); van der Mensbrugghe and Roland-Holst (2009); Walmsley and Winters (2005). Some numbers are larger: Iregui (2005) shows that movement of skilled labor can increase world GDP by between 6-11%.

advantage and can permanently lower per capita incomes in the US.

The labor economics literature has also emphasized these channels. For instance, Johnson and Stafford (1993) show how the effect of foreign competition from abroad lowers aggregate real incomes in the US. In fact, Freeman (2006a) focuses on the global job market for high-tech workers and argues that the growth in such labor abroad adversely affects US industry and workers. In this analysis, immigration can help maintain the US's lead by attracting overseas talent. However, the analysis does not account for the effect of immigration on incentives to invest in sending countries and the role of return migration, which we show to have important consequences. Davis and Weinstein (2002) show how in a Ricardian trade 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 (in our case, the US's terms of trade deteriorate as immigration lowers the IT price). Mobility will tend to equalize wages across countries and, therefore, hurt workers in the country with superior technology.

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

References

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).

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, Rochester.

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).

Artuc, E., J. McLaren, and S. Chaudhuri (2010, June). Trade shocks and labor adjustment: A structural empirical approach. *American Economic Review*.

Banerjee, A. V. and A. F. Newman (1993). Occupational choice and the process of development. *Journal of Political Economy 101*(2).

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. (2005). India's Software Industry. *Working Paper*. Indian Institute of Management Ahmedabad.

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

Bollinger, C. and B. Hirsch (2007). How Well are Earnings Measured in the Current Population Survey? Bias from Nonresponse and Proxy respondents. In *North American Summer Meetings of the Econometric Society, Duke University.*

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

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. Gordon H. Hanson, William R. Kerr, and Sarah Turner), University of Chicago Press.

Burstein, A., G. Hanson, L. Tian, and J. Vogel (2018). Tradability and the labor-market impact of immigration: Theory and evidence from the u.s. *Working paper*.

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

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, May). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *The 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 US Visas. *American Economic Review, Papers & Proceedings 103*(3), 198–202.

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 (1996, 31 July). *Cyber Media (India) Limited, Gurgaon*, 43–44.

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

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-Caneiro, 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.

Dornbusch, R., S. Fischer, and P. Samuelson (1977, December). Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods. *American Economic Review 67*(5), 823–39.

Eaton, J. and S. Kortum (2002). Technology, Geography and Trade. *Econometrica 70*, 1741–1779.

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

Foley, F. and W. Kerr (2013). Ethnic Innovation and US Multinational Firm Activity. *Management Science 59*(7).

Freeman, R. (2006a). Does Globalization of the Scientific/Engineering Workforce Threaten U.S. Economic Leadership? *Innovation Policy and the Economy 6*. (eds. Adam B. Jaffe, Josh Lerner, and Scott Stern).

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

Freeman, R. (2009). What Does Global Expansion of Higher Education Mean for the US? *NBER*. Working Paper 14962.

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 (2018). Bartik instruments: What, when, why, and how. *National Bureau of Economic Research, WP 24408*.

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

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

Iregui, A. M. (2005). Efficiancy Gains from the Elimination of Global Restrictions on Labour Mobility. *Poverty, International Migration and Asylum*, 211–258. (ed. George J. Borjas and Jeff Crisp) New York: Palgrave Macmillan.

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

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, W. (2008, August). Ethnic Scientific Communities and International Technology Diffusion. *Review of Economics and Statistics 90*(3), 518–537.

Kerr, W. (2010). The Agglomeration of US Ethnic Inventors. *Agglomeration Economics*, 237–276. (ed. by Edward Glaeser), University of Chicago Press.

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. 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*.

Leiner, B. M., V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts, and S. S. Wolff (1997). The Past and Future History of the Internet. *Communications of the ACM 40*(2), 102–108.

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. *The Manchester School*. May.

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.

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 (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.

Park, H. (2015, November 10). How Outsourcing Companies Are Gaming the Visa System. *The New York Times*. https://www.nytimes.com/interactive/2015/11/06/us/outsourcing-companies-dominate-h1b-visas.html.

Peri, G., K. Shih, and C. Sparber (2015a). Foreign and Native Skilled Workers: What Can We Learn from H-1B Lotteries? *NBER*. Working Paper No. 21175.

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

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).

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

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.

Simonovska, I. and M. E. Waugh (2014). The Elasticity of trade: Estimates and Evidence. *Journal of International Economics 92*(1), 34–50.

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.

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

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

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.

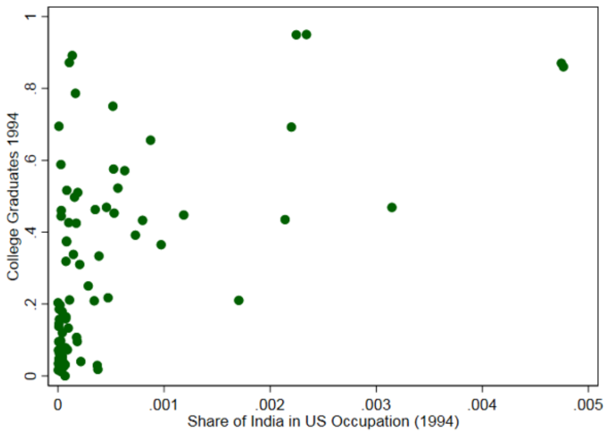
A Additional Statistics and Robustness Checks

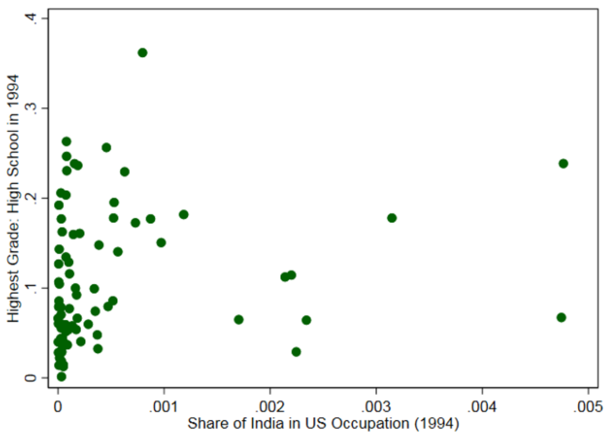
Table A1: US Demand Affects Supply in India: Alternative Specifications

	Emp in India	Emp in India	Emp in India	Log(Emp in India)
Emp Indians in US	0.491***	0.587***	0.469***	1.22e-05***
	(0.104)	(0.149)	(0.103)	(2.91e-06)
Bartik control		0.0239***	0.0223***	-2.12e-07
		(0.00860)	(0.00845)	(1.93e-07)
Time Variation	Non-Indian Migrants	US Natives by Occupation	Non-Indian Migrants by Occupation	Non-Indian Migrants
Observations	4,566	4,566	4,566	4,566
R-squared	0.127	0.120	0.135	0.208
F stat	123.6	82.01	92.00	125.9
Elasticity	0.440	0.525	0.42	0.152
SE	(0.093)	(0.13)	(0.09)	(0.036)

Notes: Tables show 2SLS regressions for the impact of employment by occupation of Indians in the US on the employment by occupation of Indians in India. Different instruments capture the demand for occupations in the US. 'Bartik control' controls for local labor demand demand shocks. Standard errors clustered by occupation-year. The main specification can be found in Table 1.

Table A2: Baseline Shift-Shares and Correlations





Shift-Shares and Education

	Shift Share Baseline	
P(Male)	-0.0157	-0.0147
	(0.00800)	(0.00732)
Age	0.00125	0.000378
	(0.000376)	(0.000346)
Primary Schooling	-0.0707	-0.00542
	(0.00958)	(0.00926)
Middle Schooling	-0.0199	0.000518
	(0.00839)	(0.00773)
High School	0.0435	0.0266
	(0.0117)	(0.0107)
Unemployed	-0.460	-0.149
	(1.031)	(0.943)
Self Employed	-0.00874	0.0122
	(0.00674)	(0.00624)
Family Worker	-0.0333	0.047
	(0.0242)	(0.0224)
Domestic Worker	0.0107	-0.176
	(0.261)	(0.239)
College Graduates		0.122
		(0.00559)
Constant	0.0245	0.00364
	(0.0180)	(0.0165)
Observations	2,432	2,432
Adjusted R-squared	**0.0386**	**0.195**

Baseline Correlations with Shares

Notes: Relationship between baseline shares of Indian in US occupations (in 1994) and other correlates (in 1994). 'P(male)' is the likelihood that the individual is male, 'Family worker' is if the person works in a family-run enterprise, and 'Domestic worker' is if the worker does domestic work (without pay).

Table A3: US Demand Affects Supply in India: Pre-Trends Falsification

	2SLS Emp in India	2SLS Emp in India	2SLS Emp in India	2SLS Log(Emp in India)
Emp Indians in US	0.0240	0.103	0.115	5.26e-06
	(0.0376)	(0.0704)	(0.111)	(4.63e-06)
Bartik control	0.00249	-0.119***	-0.120***	5.62e-07
	(0.0188)	(0.0364)	(0.0364)	(2.17e-06)
Instrument	Non-Indian Migrants by Occupation	US Natives by Occupation	Non-Indian Migrants Migrants	Non-Indian Migrants
Observations	1,444	1,344	1,344	1,442
R-squared	0.003	0.099	0.097	0.014
F stat	63.16	34.58	164.4	63.14
Elasticity	0.019	0.045	0.050	0.031
SE	(0.030)	(0.31)	(0.049)	(0.027)

Notes: Tables show 2SLS regressions for the impact of employment by occupation of Indians in the US on the employment by occupation of Indians in India. Sample restricted to all non-CS occupations. 'Bartik control' controls for local labor demand demand shocks. Standard errors clustered by occupation-year. The main specification can be found in Table 1.

Table A4: US Demand Affects Supply in India: non-CS Occupations

	2SLS Emp in India	2SLS Emp in India	2SLS Emp in India	2SLS Log(Emp in India)
Emp Indians in US	1.358***	0.879***	0.697***	2.30e-05***
	(0.331)	(0.247)	(0.174)	(4.47e-06)
Bartik control	0.0607***	0.0224***	0.0220***	-2.07e-07
	(0.0228)	(0.00832)	(0.00832)	(1.90e-07)
Instrument	Non-Indian Migrants	US Natives by Occupation	Non-Indian Migrants by Occupation	Non-Indian Migrants
Observations	4,463	4,463	4,463	4,459
R-squared	0.129	0.136	0.136	0.194
F stat	1044	2782	2121	2120
Elasticity	0.433	0.534	0.42	0.192
SE	(0.105)	(0.15)	(0.106)	(0.037)

Notes: Tables show 2SLS regressions for the pre-trend tests of the Indian occupation response to US shocks. For the pre-trends we consider employment between 1987 and 1994. 'Bartik control' controls for local labor demand demand shocks. Standard errors clustered by occupation-year. The main specification can be found in Table 1.

Table A5: Migration, Computer Scientists and Patenting by Industry: Robustness

	Log(Patents)	Log(Patents)	Log(New Patents)
CS workers	3.99e-06**	5.53e-06**	5.90e-06*
	(1.89e-06)	(2.27e-06)	(3.10e-06)
Observations	250	250	250
R-squared	0.971	0.968	0.329
Number of Industries	25	25	25
Additional Controls	No. of Firms	None	No. of Firms
F stat	12.54	11.11	12.54
Elasticity	0.162	0.225	0.240
SE	(0.076)	(0.092)	(0.126)

Notes: The main specification can be found in Table 4. Here, the specifications 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.

Table A6: 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 ten 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 from 1.7.

Table A7: Compensating Variation in 2010 By Worker

	Total Welfare (USD mn)		**Welfare per Migrant (USD)**	
	US	**India**	**US**	**India**
Always CS	-1221	-207.2	-11253	-1909
Switchers from CS to non-CS	-29.59	-49.06	-272.6	-452.0
Always non CS	-585.1	396.1	-5391.5	3650
Non college	1937	31.8	17844	293.5
Total welfare non-migrants	100.6	171.7	927	1582
Welfare of migrants	5477		50470	
Total welfare (including migrants)	5749		52979	

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 A8: Effect of Migration Over Time

	1995	2000	2005	2010
Wages				
US CS	-0.04%	-0.17%	-0.60%	-0.89%
US non CS college grad	0.00%	0.00%	-0.02%	-0.02%
US non college grad	0.00%	0.02%	0.06%	0.10%
India CS	-3.89%	-13.0%	-13.2%	-9.7%
India non CS college grad	0.20%	0.53%	0.50%	0.69%
India non college grad	0.06%	0.20%	0.14%	0.17%
Occupational Choice				
US CS (native & immigrant)	0.77%	1.31%	1.30%	1.21%
US CS native	-0.41%	-1.37%	-3.21%	-4.72%
India CS	16.44%	38.8%	34.8%	31.2%
US non CS college grad	0.01%	0.05%	0.13%	0.20%
India non CS college grad	-0.09%	-0.41%	-0.57%	-0.74%
IT production				
US IT output	1.05%	1.41%	0.70%	-0.48%
India IT output	27.48%	10.83%	5.80%	15.07%
World IT output	1.52%	2.04%	1.74%	2.80%
US IT price	-0.28%	-0.49%	-0.85%	-1.03%
India IT price	-4.93%	-8.87%	-9.06%	-9.75%
Welfare				
Welfare of US natives	0.002%	0.006%	0.004%	0.002%
Welfare of migrants	79.95%	73.39%	69.99%	52.29%
Welfare in India	0.01%	0.06%	0.09%	0.08%
Combined welfare	0.02%	0.06%	0.10%	0.12%

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 A9: Effect of Migration: Robustness to Parameters

	Baseline	**τ=3**	**θ=4**
Occupational choice			
US CS (native & immigrant)	1.21%	0.78%	0.22%
US CS native	-4.72%	-5.07%	-5.77%
India CS	31.23%	30.45%	29.76%
Welfare			
Welfare of US natives	0.00%	0.00%	0.00%
Welfare in India	0.08%	0.08%	0.07%
Combined welfare	0.12%	0.12%	0.11%

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), and θ (the dispersion parameter of the Frechet distribution).

Table A10: Brain Drain vs Brain Gain: Welfare

	Baseline	**No occupational choice in India**	**No migration but reallocation**	**With return migration**
US welfare				
Always CS	-0.89%	-0.65%	-0.36%	-1.51%
Always non CS	-0.03%	-0.01%	-0.01%	0.00%
Non college	0.10%	0.08%	0.01%	0.10%
India welfare				
Always CS	-9.8%	4.23%	-16.4%	-3.35%
Always non CS	0.54%	-0.14%	0.61%	0.15%
Non college	0.02%	-0.08%	0.07%	-0.03%
Total welfare				
US natives	0.002%	0.01%	-0.01%	0.02%
Welfare in India	0.08%	-0.05%	0.14%	0.02%
Combined (with migrants)	0.12%	0.10%	0.03%	0.12%

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'). 'With return migration' introduces another aspect of 'brain gain' – that of returning migrants, but also inhibits 'brain gain' by lowering the initial benefits of migrating if workers return.

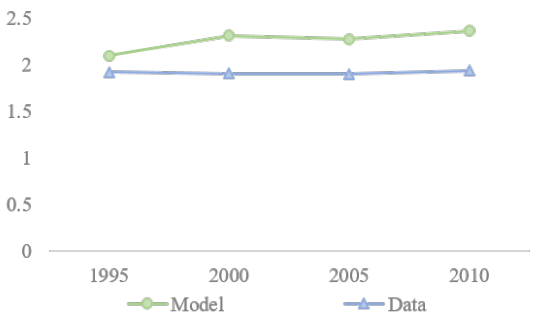
B Out-of-Sample Tests of Model Fit

Figure A1: Model Fit: Wages in India and the US

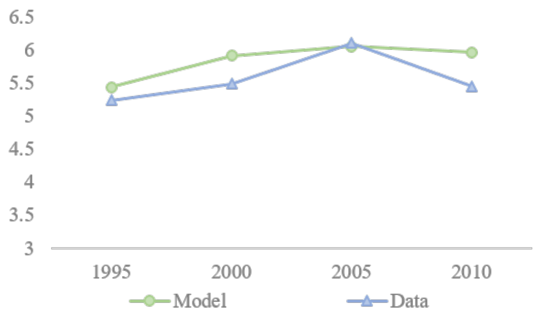
(a) Relative Wage for Computer Scientists: US to India

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

(c) College Wage Premium: US



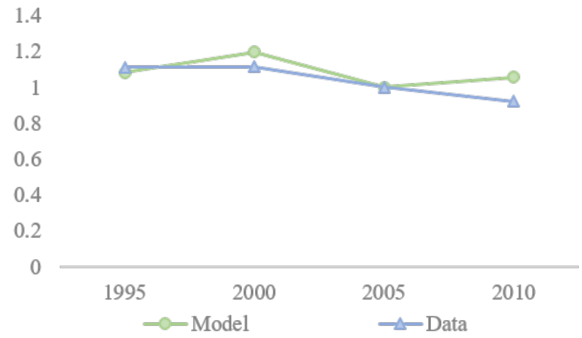
(d) College Wage Premium: India



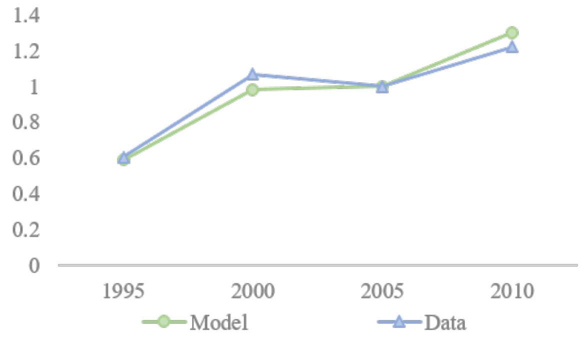
Figures plot the simulated model output and the actual data for the endogenous variables of interest. Top panels show the relative wages for CS and non-CS college graduates between the US and India. Bottom panel shows the college wage premium within each country. The college wage premium is defined as the weighted average of college graduate wages to non-college graduate wages. For data sources please refer to Data Appendix C

Figure A2: Model Fit: Prices, Production and Labor in the IT sector

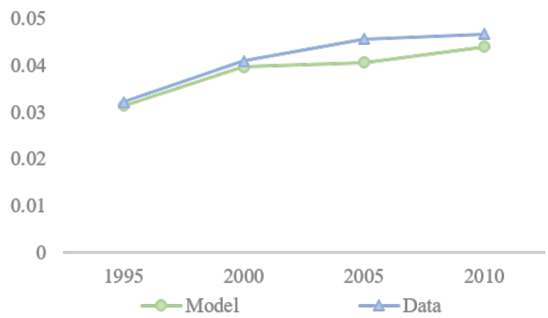
(a) Price Index IT US



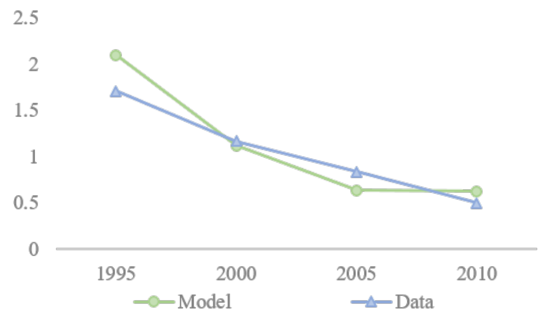
(b) IT Output US



(c) IT share in CS US



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



Figures plot the simulated model output and the actual data for the endogenous variables of interest. Top panels show IT prices and output in the US, where the indices are normalized to be 1 in 2005. Bottom panel shows the fraction of all CS workers that work in IT, and the relative fraction between US and India. For data sources please refer to Data Appendix C

C Details of the Data Used

C.1 US Data

Data on earnings, domestic employment and foreign employment used in the model-solving procedure, and in the descriptive figures come 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. Earnings are deflated to 1999 dollars, and top-coded values are multiplied by 1.4. We drop imputed earnings. In order to identify imputed values, we use a methodology similar to (Bollinger and Hirsch (2007)). From the IPUMS database we use the qinclongj and qincwage variables, and from the NBER database we use the FL665 flag to identify imputations. We also use ten Census Bureau flags that identify a small fraction of earnings as allocated. Over our period around 26% of earnings were allocated. This fraction varies over time (between 19.14% in 1994 and 29.47% in 2003). These numbers are consistent with (Bollinger and Hirsch (2007)) who find that between 1998 and 2006, the non-response rate was 20%. The small difference in our numbers arises from using a different sample (restricted to those with BA/MA degree) and because non-response is not the only reason the CPS imputes earnings.

In order to define workers in Computer Science we use the occupational codes and the crosswalk given the categories in the CPS Outgoing Rotation Group (CPS-ORG) data set. The occupational coding in the CPS-ORG up to 2002 uses the 1990 Census definition. We consider as Computer Scientists those under the occupational titles of: "064 Computer systems analysts and scientists" and "229 Computer programmers". For the years 2000-2 the CPS-ORG reports codes using both the 1990 Census definition and the 2000 Census definition. This allows us to create a crosswalk where we weight the 2000 occupational codes by the 2 occupational categories in the 1990 Census.

College enrollment data is based on Integrated Post-secondary Education Data System (IPEDS) Completions Survey. It consists of bachelor's degrees awarded by the NSF population of institutions. We consider enrollment in computer science and electrical engineers as the number of degrees awarded in these fields lagged by 2 years. For 1994 and 1995, enrollment in electrical engineering was not available by native and foreign students but only shown together with all engineering degrees. We impute the data for these two years by looking at the average growth in electrical engineering for 1996-2002.

STEM occupations are defined as engineers, computer systems analysts and computer scientists, computer software developers, operations and systems researchers and analysts, actuaries, statisticians, mathematicians and mathematical scientists, physicists and astronomers, chemists, atmospheric and space scientists, geologists, physical scientists n.e.c., agricultural and food scientists, biological scientists, foresters and conservation scientists, and medical scientists.

We use data on the prices, quantities, costs and value added from the Bureau of Economic Analysis (BEA) since this source allows us to look into data for specific industry groups. Data on firm entry and exit comes from the Business Dynamic Statistics (BDS), and the 1992 Census' Statistics of U.S. Businesses (SUSB). In these data sets we define the IT sector as the sub-sectors of "Publishing industries, except Internet (includes software)," "Data processing, Internet publishing, and other information services" and "Computer systems design and related services" according to the NAICS 2002 classification. The Non IT sector is defined as all other sectors in the economy.

C.2 Trade Data

Information on imports, exports and RoW consumption of IT from the US and India come from the OECD Trade in Value Added statistics. 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. The data is only available for 1995, 2000, 2005 and 2008-2011.

C.3 India Data

Data on earnings, employment by occupations and sector, and age-shares by occupation comes from the National Sample Survey (NSS). We use the Employment / Unemployment surveys from rounds 50 through 66, which cover 1994 through 2010 with gaps in between. NSS is a nationally representative survey used by many researchers. It is 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. CS are defined as "systems analysts, programmers, and electrical and electronic engineers" based on the National Occupational Codes (NOC). We use the earnings data for the primary occupation only. The IT sector is restricted to be "software" (code 892 in the National Industrial Classification).

There are various sources for education data, the most comprehensive of which is the Ministry of Human Resources and Development that records number of degrees and universities by type of degree (for example, engineering degrees). We combine this 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), 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).

D Estimation Details

D.1 Gravity Equation Parameters

To get our estimates for trade costs and technology we take logs of Equation 23 and get Equation 33 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 $T^s_{t,k}(\xi^s_{t,k})^{-\theta}$.

First, we parametrize the trade costs as in Equation 32. 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 $v_{t,k,b}$.

$$log(d^s_{t,k,b}) = dist_{k,b} + border_{k,b} + CU_{t,k,b} + RTA_{t,k,b} + exp^s_{t,k} + v^s_{t,k,b} \tag{32}$$

We ran a separate equation for each sector and year to capture the differential evolution in technology and trade costs over time. 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.

$$\begin{aligned} log\left(\frac{EX^s_{t,k,b}}{EX^s_{t,b,b}}\right) = \underbrace{log\left((T^s_{t,k}(\xi^s_{t,k})^{-\theta}\right) - \theta exp_{t,k}}_{\text{Exporter fixed effect}} - \underbrace{log\left((T^s_{t,b}(\xi^s_{t,b})\right)^{-\theta}}_{\text{Importer fixed effect}} \\ -\theta(dist_{k,b} + border_{k,b} + CU_{t,k,b} + RTA_{t,k,b} + v_{t,k,b}) \end{aligned} \tag{33}$$

The distance variable $dist_{k,b}$ is a group of 6 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, maximum), and 0 otherwise.

We estimate equation 33 separately by sectors Y and C. Distance, currency union and regional trade agreement data comes from CEPII gravity database and data on industry specific trade flows and and GDP from the OECD. We run the regression separately for years 1995, 2000, 2005 and 2010 as those are the only years we have good data on trade flows for the IT sector for a large number of countries. From the estimates of equation 33 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 drop 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 get the 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. We calculate a weighted average of trade costs and $\hat{\Xi}^s_{t,k}$ based on country-sector GDP to get to the average trade costs and $\hat{\Xi}^s_{t,k,b}$ in RoW (as a single region) we will feed into the model.