Volume Title: High-Skilled Migration to the United States and Its Economic Consequences

Volume Author/Editor: Gordon H. Hanson, William R. Kerr, and Sarah Turner, editors

Volume Publisher: University of Chicago Press

Volume ISBNs: 978-0-226-52552-5 (cloth); 978-0-226-52566-2 (e-ISBN)

Volume URL: http://www.nber.org/books/hans-4

Conference Date: July 10-11, 2016

Publication Date: May 2018

Chapter Title: Understanding the Economic Impact of the H-1B Program on the United States

Chapter Author(s): John Bound, Gaurav Khanna, Nicolas Morales

Chapter URL: http://www.nber.org/chapters/c13842

Chapter pages in book: (p. 109 – 175)

4

Understanding the Economic Impact of the H-1B Program on the United States

John Bound, Gaurav Khanna, and Nicolas Morales

An increasingly high proportion of the scientists and engineers in the United States were born abroad. At a very general level, the issues that come up in the discussion of high-skilled immigration mirror the discussion of low-skilled immigration. The most basic economic arguments suggest that both high-skilled and low-skilled immigrants (a) impart benefits to employers, to owners of other inputs used in production such as capital, and to consumers; and (b) potentially, impose some costs on workers who are close substitutes (Borjas 1999). Evidence suggests, however, that the magnitude of these costs may be substantially mitigated if US high-skilled workers have good alternatives to working in sectors most affected by immigrants (Peri, Shih, and Sparber 2013; Peri and Sparber 2011). Additionally, unlike low-skilled immigrants, high-skilled immigrants contribute to the generation of knowledge and productivity through patenting and innovation, both of which serve to shift out the production possibility frontier in the United States and may also slow the erosion of the US comparative advantage in high tech (Freeman 2006; Krugman 1979).

In this chapter, we study the impact that the recruitment of foreign com-

John Bound is professor of economics and research professor at the Population Studies Center at the University of Michigan and a research associate of the National Bureau of Economic Research. Gaurav Khanna is assistant professor of public policy economics at the University of California, San Diego, School of Global Policy and Strategy. Nicolas Morales is a PhD candidate in economics at the University of Michigan.

We would like to acknowledge the Alfred P. Sloan Foundation for generous research support. We thank Breno Braga, Gordon Hanson, Minjoon Lee, Rishi Sharma, Sebastian Sotelo, Sarah Turner, and seminar participants at the NBER Global Talent SI Conference, University of California, San Diego, University of California, Davis, and the University of Michigan for insightful comments and NE Barr for editorial assistance. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see http://www.nber.org/chapters/c13842.ack.

puter scientists on H-1B visas had on the US economy during the Internet boom of the 1990s. An H-1B is a nonimmigrant visa allowing US companies to temporarily employ foreign workers in specialized occupations. The number issued annually is capped by the federal government. During the 1990s, we observe a substantial increase in the number of H-1B visas awarded to high-skilled workers, with those in computer-related occupations becoming the largest share of all H-1B visa holders (US General Accounting Office 2000). Given these circumstances, it is of considerable interest to investigate how the influx of H-1B visa holders during this period might have affected labor market outcomes for US computer scientists and other US workers, and overall productivity in the economy.

We focus on the period 1994 to 2001 for a number of reasons. During the latter half of the 1990s, the US economy experienced a productivity growth attributable, at least in part, to the information technology (IT) boom, facilitated by the influx of foreign talent (Jorgenson, Ho, and Samuels 2016). At the same time, the recruitment of H-1B labor by US firms was at or close to the H-1B cap during this period, enabling us to treat foreign supply as determined by the cap. Finally, more recent growth of the IT sector in India and changes in the law authorizing the H-1B have complicated the picture since 2001.[1] Nonetheless, in the appendix we show that our model does a reasonable job of predicting employment and wages all the way until 2015.

In earlier work evaluating the impact of immigration on computer science (CS) domestic workers, we constructed a dynamic model that characterizes the labor supply and demand for CS workers during this period (Bound et al. 2015). We built into the model the possibility that labor demand shocks, such as the one created by the Internet boom, could be accommodated by three sources of CS workers: recent college graduates with CS degrees, US residents in different occupations who switch to CS jobs, and high-skilled foreigners. Furthermore, our model assumed firms faced a trade-off when deciding to employ immigrants: foreigners were potentially either more productive or less costly than US workers, but incurred extra recruitment/hiring costs.

The approach we took in that analysis was distinctly partial equilibrium in nature—that is, we focused on the market for computer scientists and ignored any wider impacts that high-skilled immigration might have on the US economy (Nathan 2013). While we believe that approach could be used to understand the impact that the availability of high-skilled foreign labor might have had for this market, it precludes any analysis of the overall welfare impact of the H-1B program in particular, or of high-skilled immigration more generally.

The implications of the model regarding the impact of immigration on the employment and wages of native workers depended on the elasticity of

1. See Khanna and Morales (2015) for a long-run extension of this work that also models the Indian IT sector.

labor demand for computer scientists. As long as the demand curve sloped downward, the increased availability of foreign computer scientists would put downward pressure on the wages for computer scientists in the United States. However, in the case of computer scientists, other factors may affect this relationship. First, even in a closed economy, the contribution of computer scientists to innovation reduces the negative effects foreign computer scientists might have on the labor market opportunities for native high-skilled workers. In addition, in an increasingly global world, US restrictions on the hiring of foreign high-skilled workers are likely to result in greater foreign outsourcing work by US employers. Indeed, if computer scientists are a sufficient spur to innovation, or if domestic employers can readily offshore CS work, any negative effects that an increase in the number of foreign CS workers might have on the domestic high-skilled workforce would be offset by increases in the domestic demand for computer scientists.

In Bound et al. (2015), we used data on wages, domestic and foreign employment, and undergraduate degree completions by major during the late 1990s and early in the twenty-first century to calibrate the parameters of our model to reproduce the stylized facts of the CS market during the analytic period (1994 to 2001). Next, we used the calibrated model to simulate counterfactuals on how the economy would have behaved if firms had been restricted in the number of foreign CS workers they could hire to the 1994 level. Conditional on our assumptions about the elasticity of the demand curve for computer scientists, our simulation suggests that had US firms faced this restriction, CS wages and the number of Americans working in computer science and the enrollment levels in US CS programs would have been higher, but the total number of CS workers in the US would have been lower.

The predictions of our model did not depend on the specific choice we made for noncalibrated parameters, with one important exception: crowd-out in the market for computer scientists depended crucially on the elasticity of demand for their services. Ideally, we would have been able to use exogenous supply shifts to identify the slope of the demand curve for computer scientists, as we use exogenous shifts in demand to identify supply curves. In other contexts, researchers have treated the increase in foreign-born workers in the US economy as exogenous. However, in the current context, immigration law in the United States implies that most of the foreign-born and trained individuals who migrate to the United States to work as computer scientists do so because they are sponsored by US-based firms. Thus, it seems implausible to treat the number of foreign-born computer scientists in the United States as an exogenous increase in supply. In the end, without credible sources of identifying information, we resorted to parametrically varying the elasticity of the demand for computer scientists.

In the current analysis, we take a different track. We interpret the arguments about the potential productivity effects of high-skilled immigrants in terms of models of endogenous technical change. Within the context of a simple general-equilibrium model of the US economy, we link productivity increases in the US economy during the 1990s to increases in the utilization of computer scientists in the economy. This allows us to derive the demand curve for computer scientists.

Within the context of our model, it is possible to understand the effect that the availability of high-skilled foreign workers has on the earnings of both high- and low-skilled workers, the goods available in the economy, and profits in the high-tech sector of the economy. However, our conclusions are dependent both on our modeling choices and on values of our calibrated parameters. For this reason, we do extensive sensitivity analyses to determine which of our conclusions are robust.

A key feature of high-skilled immigrants is that they contribute to innovation. While this point is well understood, we know of no earlier work that has tried to quantify the magnitude of this effect within the context of an explicit model of the US economy. The magnitude of this effect is important because it speaks to the magnitude of any first-order gains to US residents of high-skilled immigration, and because it has a direct influence on the slope of the labor demand curve for close substitutes for high-skilled immigrants.

Our model is limited in a number of important respects. While we allow for endogenous technical change, we incorporate trade in a very stylized manner and do not allow explicitly for outsourcing.[2] As such, we think our model captures relatively short-run effects of H-1B immigration. Although in this sense our model is different from models incorporated in recent work by, for example, Grossman and Rossi-Hansberg (2008) or di Giovanni, Levchenko, and Ortega (2015), we believe that it captures important elements of the current debate about the H-1B program.

We review this literature in detail and describe the market for CS workers in section 4.1. Section 4.2 presents the model we build to characterize the market for CS workers when firms can recruit foreigners. In section 4.3, we describe how we calibrate the parameters of the model and in section 4.4 we run counterfactual simulations where firms have restrictions on the number of foreigners they can hire. Section 4.5 talks about welfare changes under this counterfactual scenario. We conclude with section 4.6, which presents a discussion based on the results of the analysis.

4.1 The Market for Computer Scientists in the 1990s

4.1.1 The Information Technology Boom of the Late 1990s

The mid-1990s marks the beginning of the use of the Internet for commercial purposes in the United States, and a concomitant jump in the

2. Available evidence suggests that outsourcing options were somewhat limited during the 1990s (Liu and Trefler 2008), though it is not clear that this is still true.

Table 4.1 **Immigration and the computer science workforce**

Year	1970 (%)	1980 (%)	1990 (%)	2000 (%)	2010 (%)
Computer scientists as a fraction of workers with a BA/MA	1.68	1.83	3.30	5.66	5.28
Computer scientists as a fraction of STEM college graduates	16.86	23.60	35.99	53.31	54.90
Immigrants as a fraction of BA/MAs	2.10	5.43	6.86	8.41	12.77
Immigrants as a fraction of computer scientists	2.37	7.09	11.06	18.59	27.82
Immigrants as a fraction of other STEM workers	3.63	9.72	10.71	12.69	18.21

Source: US Census (years 1970 to 2000); ACS (2010).

Note: Sample restricted to employed workers with a bachelor's or a master's degree. Definition of computer scientists and STEM workers determined by occupational coding (for details, see the "Details of the Data Used" section of the appendix). Immigrant defined as one born abroad and migrated to the United States after the age of eighteen.

number of Internet users. One indicator of a contemporaneous increase in demand for IT workers is the rise of research and development (R&D) expenditures among firms providing computer programming services and computer-related equipment. Specifically, the share of total private R&D expenditures for firms in these sectors increased from 19.5 percent to 22.1 percent between 1991 and 1998.[3] The entry and then extraordinary appreciation of tech firms like Yahoo!, Amazon, and eBay provide a further testament to the boom in the IT sector prior to 2001.

These changes had a dramatic effect on the labor market for computer scientists. According to the census, the number of employed individuals working either as computer scientists or computer software developers increased by 161 percent between the years 1990 and 2000. In comparison, during the same period, the number of employed workers with at least a bachelor's degree increased by 27 percent and the number of workers in other science, technology, engineering, and math (STEM) occupations increased by 14 percent.[4] Table 4.1 shows that computer scientists as a share of the college-educated workforce and the college-educated STEM workforce was rising before 1990, but increased dramatically during the 1990s. Indeed, by 2000 more than half of all STEM workers were computer scientists. In figure 4.1, panel A, we use Current Population Survey (CPS) data to show a similar pattern, additionally showing that the growth of CS employment started in the second half of the decade—a period corresponding to the dissemination of the Internet.

The Internet innovation affected educational choices as well as employment decisions. We show in figure 4.1, panel B, that the CS share of both all

3. Bound et al. (2015) calculation using Compustat data.

4. Here and elsewhere, our tabulations restrict the analysis to workers with at least a bachelor's degree and use the IPUMS-suggested occupational cross walk. Other STEM occupations are defined as engineers, mathematicians, and computer scientists. For more details see the "Details of the Data Used" section of the appendix.

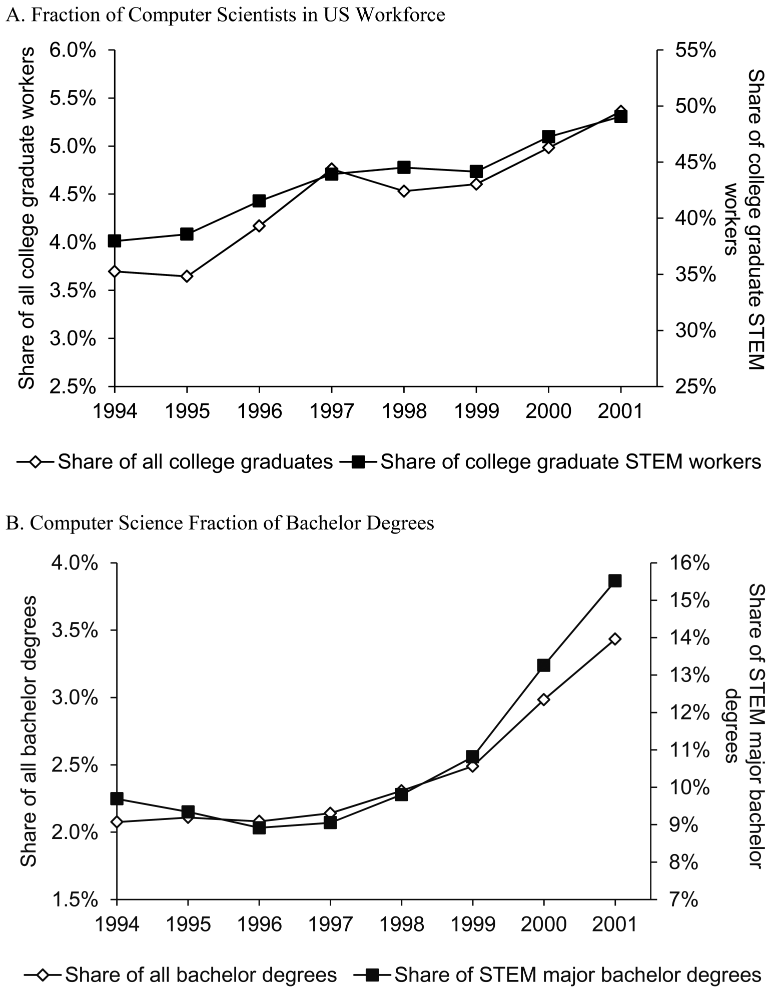


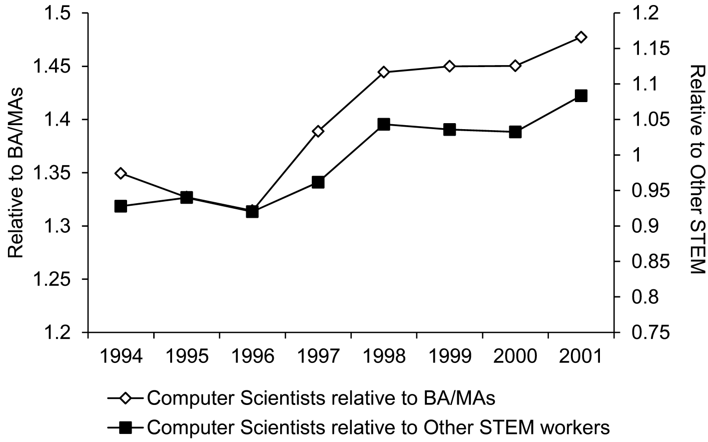
Fig. 4.1 High-skilled immigration and the IT boom

Source: Panels A, C, and D, March Current Population Survey. Panel B, Integrated Postsecondary Education Data System (IPEDS). Panel E, author's calculations updating Lowell (2000).

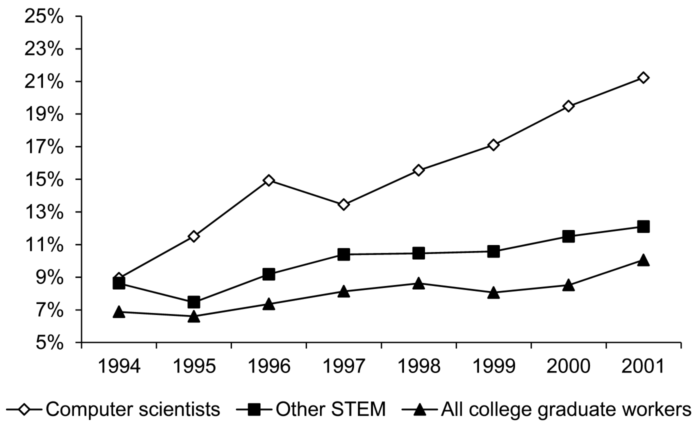
bachelor's degrees and STEM major degrees increased dramatically during this period, in both cases rising from about 2 percent of all bachelor's degrees granted in 1994 to almost 3.5 percent in 2001.

The behavioral response would be different if the boom was only a temporary response to the Y2K bug. The employment and educational evidence, however, suggests that many expected this boom, as a response to technological innovations, to be permanent. Indeed, in 1997, the Bureau of Labor Statistics (BLS) projected a steady increase in CS employment after the turn of the century. More specifically, the BLS predicted that between

C. Earnings of Computer Scientists Relative to Other groups



D. Immigrants as Fraction of Workers by Occupation



E. H-1 Visas

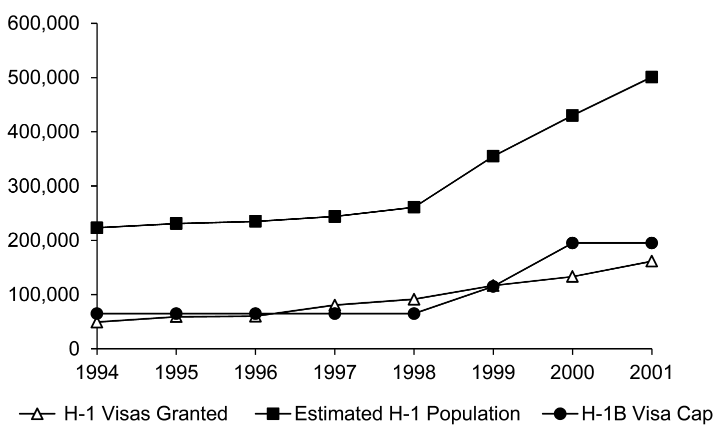


Fig. 4.1 (cont.)

1996 and 2006 "database administrators, computer support specialists, and all other computer scientists" would be the fastest growing occupation and "computer engineers" would be the second fastest in terms of jobs. Furthermore, they predicted that "computer and data processing services" would grow by 108 percent—the fastest growing industry in the country.[5]

In addition to affecting employment and enrollment decisions, there is also empirical evidence that CS wages responded to expanding Internet use. From the census we observe an 18 percent increase in the median real weekly wages of CS workers between 1990 and 2000. The CPS presents similar patterns: starting in the year 1994 we observe in figure 4.1, panel C, that wages of computer scientists increased considerably when compared to both workers with other STEM occupations and all workers with a bachelor's degree. In fact, during the beginning of the 1990s, the earnings of CS workers were systematically lower than other STEM occupations; the wage differential tends to disappear after 1998.

4.1.2 The Contribution of Immigration to the Growth of the High-Tech Workforce

Employment adjustments in the market for computer scientists occurred disproportionately among foreigners during the Internet boom. Evidence for this claim is found in table 4.1 and figure 4.1, panel D, where we use census and CPS data to compare the share of foreign computer scientists to the share of foreign workers in other occupations.[6] In the second half of the 1990s, the foreign fraction of CS workers increased considerably more than both the foreign fraction of all workers with a bachelor's degree and the foreign fraction of all workers in a STEM occupation. In particular, in 1994 the share of foreigners working in computer science was about the same as the share working in other STEM occupations, but later in the decade, during the boom in Internet use, the share of foreigners among all CS workers rose steeply, comprising about 30 percent of the increase in all CS workers during this period.

The growth in the representation of the foreigners among the US CS workforce was fueled by two supply-side developments in this period. First, the foreign pool of men and women with college educations in science and engineering fields increased dramatically (Freeman 2009). In India, an important source of CS workers in the United States, the number of first degrees conferred in science and engineering rose from 176,000 in 1990 to 455,000 in 2000. Second, the Immigration Act of 1990 established the H-1B visa program for temporary workers with at least a bachelor's degree work-

5. Source: BLS Employment Projections http://www.bls.gov/news.release/history/ecopro_082498.txt.

6. Here and elsewhere, we define foreigners as those who immigrated to the United States after the age of eighteen. We believe that this definition is a reasonable proxy for workers who arrived in the United States on nonimmigrant visas.

ing in "specialty occupations" including engineering, mathematics, physical sciences, and business among others.

Firms wanting to hire foreigners on H-1B visas must first file a Labor Condition Application (LCA) in which they attest that the firm will pay the visa holder the greater of the actual compensation paid to other employees in the same job or the prevailing compensation for that occupation, and the firm will provide working conditions for the visa holder that do not adversely affect the working conditions of the other employees. At that point, prospective H-1B nonimmigrants must demonstrate to the US Citizenship and Immigration Services Bureau (USCIS) in the Department of Homeland Security (DHS) that they have the requisite education and work experience for the posted positions. The USCIS may approve the petition for the H-1B holder for a period of up to three years, with the possibility of a three-year extension. Thus foreign workers can stay a maximum of six years on an H-1B visa, though firms can sponsor these workers for a permanent resident visa. Because H-1B visas are approved solely for the applying firm, H-1B foreign workers are effectively tied to their sponsoring company.

Since 1990, when the visa was initiated, the number of H-1B visas issued annually has been capped. The initial cap of 65,000 visas per year was not reached until the mid-1990s, when demand began to exceed the cap. However, the allocation tended to fill each year on a first-come, first-served basis, resulting in frequent denials or delays on H-1Bs because the annual cap had been reached. After lobbying by the industry, Congress raised the cap first to 115,000 for FY1999 and then to 195,000 for FY2000–2003, after which the cap reverted to 65,000. Figure 4.1, panel E, shows the growth in the number of H-1 visas (the H-1 was the precursor to the H-1B) issued 1976–2008, estimates of the stock of H-1 visas in the economy each year, and the changes in the H-1B visa cap.[7]

Through the decade of the 1990s, foreign workers with H-1B visas became an important source of labor for the technology sector. The National Survey of College Graduates shows that 55 percent of foreigners working in CS fields in 2003 arrived in the United States on an H-1B or a student-type visa (F-1, J-1). Furthermore, institutional information indicates a significant increase in the number of visas awarded to workers in computer-related occupations during the 1990s. A 1992 US General Accounting Office report shows that "computers, programming, and related occupations" corresponded to 11 percent of the total number of H-1 visas in 1989, while a

7. The Immigration and Nationality Act of 1952 established the precursor to the H-1B visa, the H-1. The H-1 nonimmigrant visa was targeted at aliens of "distinguished merit and ability" who were filling positions that were temporary. Nonimmigrants on H-1 visas had to maintain a foreign residence. The Immigration Act of 1990 established the main features of H-1B visa as it is known today, replacing "distinguished merit and ability" with the "specialty occupation" definition. It also dropped the foreign residence requirement and added a dual-intent provision, allowing workers to potentially transfer from an H-1B visa to immigrant status.

report from the US Immigration and Naturalization Service (2000) finds that computer-related occupations accounted for close to two-thirds of the H-1B visas awarded in 1999. More specifically, the US Department of Commerce (2000) estimated that during the late 1990s, 28 percent of all US programmer jobs went to H-1B visa holders.

While H-1B visa holders represent an important source of computer scientists, they do not represent all foreigners in the country working as computer scientists. A significant number of such foreigners are permanent immigrants, some of whom may have come either as children or as students. Other foreigners enter the country to work as computer scientists in the United States on L-1B visas, which permit companies with offices both in the United States and overseas to move skilled employees from overseas to the United States. While we know of no data showing the fraction of computer scientists working in the United States on L-1B visas, substantially fewer L-1(A&B) visas are issued than are H-1Bs.[8]

4.1.3 The Impact of Immigrants on the High-Tech Workforce in the United States

Critics of the H-1B program (Matloff 2003) argue that firms are using cheap foreign labor to undercut and replace skilled US workers, although even the fiercest critics do not claim that employers are technically evading the law (Kirkegaard 2005). Rather, they argue that firms skirt the requirement to pay H-1B visa holders prevailing wages by hiring overqualified foreigners into positions with low stated qualifications and concomitant low "prevailing wages." These critics claim that the excess supply of highly qualified foreigners willing to take the jobs in the United States, plus the lack of portability of the H-1B visa, limits the capacity of H-1B workers to negotiate fair market wages.

One way to get a handle on the extent to which H-1B visa holders are being underpaid relative to their US counterparts is to compare foreigners on H-1B visas to those with green cards—an immigrant authorization allowing the holder to live and work in the United States permanently, with no restrictions on occupation. Using difference-in-difference propensity score matching and data from the 2003 New Immigrant Survey, Mukhopadhyay and Oxborrow (2012) find that green card holders earn 25.4 percent more than observably comparable temporary foreign workers. Using log earnings regressions and data from an Internet survey, Mithas and Lucas (2010) find that IT professionals with green cards earn roughly 5 percent more than observationally equivalent H-1B visa holders. Comparisons between green card and H-1B holders are far from perfect. Since many green card holders begin as H-1B visa holders who are eventually sponsored by their employers for permanent residence status, it is reason-

8. See Yeaple (chapter 2, this volume) for a discussion on L-1 and H-1B visas.

able to assume that green card holders are positively selected on job skills. Given this consideration, it is somewhat surprising that the observed green card premium is not larger than this 5 percent.

Perhaps the most compelling work concerning productivity differences between H-1B visa holders and their US resident counterparts comes from a recent paper by Doran, Gelber, and Isen (2015), who analyze H-1B lotteries used in FY2006 and FY2007 to identify the productivity effects on firms of hiring an additional H-1B worker. During these two years, firms that submitted an LCA during the day the H-1B quota was hit would enter a lottery to determine whether they were permitted to hire the additional H-1B worker. Doran, Gelber, and Isen (2015) find that winning the lottery had no effect on subsequent patenting or employment in the affected firm, consistent with the notion that a firm unable to hire a H-1B worker would end up hiring an alternative, equally productive worker.[9]

While there may be no incontrovertible estimate of the productivity (conditional on earnings) advantage of foreign high-skilled labor, simple economic reasons suggest this advantage must exist. US employers face both pecuniary and nonpecuniary costs associated with hiring foreigners. A small GAO survey (US General Accounting Office 2011) estimated the legal and administrative costs associated with each H-1B hire to range from $2,300 to $7,500. Assuming that these workers earn $60,000 per year in total compensation, which would seem to be conservative, this amounts to no more than 2 percent of compensation spread over six years. It seems reasonable to assume that employers must expect some cost or productivity advantage when hiring foreigners, however modest. If not, why would they incur the associated effort and expense?

Whatever the perceived cost or productivity advantages, H-1B critics argue that US employers' use of foreign labor in high-skill jobs either "crowds out" native workers from these jobs or puts downward pressure on their wages. Although, as far as we know, critics of the H-1B program have not yet estimated the magnitude of either of these effects, recent work by economists has started to fill this void. Kerr and Lincoln (2010) and Hunt and Gauthier-Loiselle (2010) provide original empirical evidence on the link between variation in immigrant flows and innovation measured by patenting, finding evidence that the net impact of immigration is positive rather than simply substituting for native employment. Kerr and Lincoln (2010) also show that variation in immigrant flows at the local level related to changes in H-1B flows do not appear to adversely affect native employment and have a small, statistically insignificant, effect on their wages. More recently, Peri, Shih, and Sparber (2014) found positive effects of high-skilled

9. Doran, Gelber, and Isen (2015) point estimates suggest that replacing a US resident with a H-1B holder might raise patenting at small firms by 0.26 percent (95 percent CI −0.42 0.47 percent), implying that the H-1Bs visa holders are no more than 4.7 percent more productive than are US resident workers.

immigrant workers on the employment and wages of college-educated domestic workers.

A potential issue with the analyses of Kerr and Lincoln (2010) and Peri, Shih, and Sparber (2014) is that the observed, reduced-form outcomes may capture concurrent changes in area-specific demand for computer scientists. To circumvent the problem, each paper constructed a variable that is the total number of individuals working on H-1B visas nationally interacted with local-area dependency.[10] However, given the nature of the H-1B visa, the location of immigrants depends, in large part, on the location of employers hiring them. If, because of local agglomeration effects, the IT boom was concentrated in areas of the country that were already IT intensive (such as Silicon Valley), then the measure of local dependency would be endogenous, an issue that Kerr and Lincoln (2010) and Peri, Shih, and Sparber (2014) understand.

Ghosh, Mayda, and Ortega (2014) take a different approach. They match all LCAs with firm-level data on publicly traded US companies, comparing changes in labor productivity, firm size, and profits between 2001 and 2006 for firms that were highly dependent on H-1B labor with firms that were not. They argue that the H-1B-dependent firms would feel more effects than their counterparts from the dramatic drop in the H-1B cap from 195,000 to 65,000 in 2004. And, indeed, they find that, over this period, labor productivity, firm size, and profits all declined more for the H-1B-dependent firms, which they attribute to the loss of the H-1B labor. The concern here is that the firms more dependent in H-1B labor in 2001 would have been systematically different from those less dependent in ways correlated with the change in performance between 2001 and 2006.

In another paper, Peri, Shih, and Sparber (2015) use data on the number of LCAs filed by firms in local (metro) areas during 2007 and 2008 as a measure of potential demand for H-1B workers, and the number of H-1B applications filed by foreigners as their measure of H-1Bs hired. In 2007 and 2008, the number of H-1B applications exceeded the annual quotas, and lotteries were used in awarding visas. The large gap between these two measures represents the unmet demand for skilled foreign workers. Cross-metro-area variation in this variable is due to at least two sources: (a) cross-metro-area demand for foreign high-skilled labor, and (b) truly random fluctuations in the fraction of LCAs picked in the lotteries. While this second source of variation should be truly random, Peri, Shih, and Sparber (2015) find

10. Kerr and Lincoln (2010) and Peri, Shih, and Sparber (2014) hope that the variation in this variable is driven largely by changes in the cap on new H-1B visas that occurred over the last twenty years. That said, it is unclear the extent to which the variation they use is being driven by variation in the visa cap. Because of the dot-com bubble bust in 2000 and 2001, the variation in the H-1B cap is only loosely related to the actual number of H-1Bs issued. What is more, the cap will have different effects across areas, and one can worry about the exogeneity of this variation. In addition, it is hard to imagine that the cap was exogenous to the demand for IT workers.

too little of such variation to reliably identify the net effects of high-skilled labor immigration.

Previous researchers studying the impact of H-1B workers on the US economy have focused on identifying exogenous variation in the number of H-1B workers, typically finding that H-1B workers tend to raise productivity and act as complements to, rather than crowding out, college-educated native workers. However, as these researchers have acknowledged, it is easy to question the validity of the instruments used in these analyses. Rather than using a natural experiment to identify effects, we derive effects from a calibrated model. The model allows us to connect endogenous productivity advances in the IT sector during the 1990s to changes in the demand for CS labor. While the validity of the conclusions that Kerr and Lincoln (2010), Peri, Shih, and Sparber (2014, 2015), and Ghosh, Mayda, and Ortega (2014) depend on the validity of the natural experiments they use to identify effects, our conclusions depend on our model accurately reflecting key features of the US economy. As such, the credibility of our results hinges on the plausibility of our assumptions and/or the robustness of our conclusions to variations in the specific modeling choices we made.

4.2 A Model of the Product and Labor Markets

Our model consists of two major sections. The first is the product market where goods are produced by firms and sold to consumers. The second is the labor market for college graduates, where US workers decide whether to work as computer scientists or in other occupations. Our product market has two sectors: the IT sector and the "other" sector. The IT sector is monopolistically competitive, wherein firms produce different varieties of the same IT good. Firms in the IT sector are heterogeneous in terms of their level of productivity, which is exogenously drawn. Importantly, we include the possibility of endogenous technological change, whereby CS workers' innovation causes the production function to be increasing returns to scale at the aggregate level. All other goods in the economy are produced in the residual "other" sector, which is a perfectly competitive sector with homogeneous firms.

Every period a firm chooses its inputs to maximize profits. Since firms in the IT sector are monopolistically competitive, they have some market power when making these choices. Firms use intermediate inputs from the other sector and labor to produce their output. The labor inputs consist of three types of workers: computer scientists, college-educated non–computer scientists, and non-college-educated workers. In our model, all foreign immigrants are hired as computer scientists. The IT-sector firms are also able to export their products to foreign markets, whereas the US economy imports only non-IT goods. Consumers, on the other hand, choose how much of each good to consume in order to maximize their utility subject to their labor income. Like firms, they make these choices every period, and have no savings.

Building on this setup, we include the labor supply decisions of college graduates. Since human capital investments and career choices have long-term payoffs, US workers in our model are allowed to choose their fields of study and occupations based on the information they have today and their expected payoffs in the future. They are then allowed to switch occupations, by paying a switching cost, when a change occurs in the current or expected payoffs associated with any occupation. Given the labor supply decisions of US workers, the labor supply of immigrants, and the labor demand from firms in each sector, the market clears to determine the equilibrium wages for each type of worker. Equilibrium prices are determined in the product market, where the demand for the two types of goods from consumers meets the supply of these goods from firms.

4.2.1 Product Market

Household Problem

There are X number of consumers in the economy who supply one unit of labor each. Each consumer has the same preferences over the two goods: C_d produced by the IT sector and Y_d, the good produced by the residual sector in the economy. We assume that preferences can be represented by the constant elasticity of substitution (CES) utility function in equation (1).

$$U(C_d, Y_d) = \left[\gamma C_d^{(\sigma-1)/\sigma} + (1-\gamma) Y_d^{(\sigma-1)/\sigma}\right]^{\sigma/(\sigma-1)}. \tag{1}$$

Y_d is assumed to be homogeneous, whereas the IT good C_d is composed of a continuum of varieties (indexed by ν) in the framework introduced by Dixit and Stiglitz (1977):[11]

$$C_d = \left(\int_{\nu\in\Omega} c_{di}^{(\varepsilon-1)/\varepsilon}\, d\nu\right)^{\varepsilon/(\varepsilon-1)}, \tag{2}$$

where Ω is the set of varieties and ε is the elasticity of substitution between the varieties of IT goods. In our analysis, we set the consumption bundle to be the numeraire.[12] In appendix section "Consumer Demand for Goods," we solve for the demand for each good.

Consumers are also workers, and while they have identical consumption preferences, they do not all receive the same labor income as they work in different occupations earning different wages. Furthermore, workers can either be native workers (denoted by a subscript n) or foreign workers (denoted by a subscript F).

11. This setting with one composite and one homogeneous good follows recent papers such as Melitz and Ottaviano (2008), Demidova (2008), and Pfluger and Russek (2013).

12. This means that the ideal price index is normalized to 1: $\{\gamma + (1-\gamma)[(P_c/P_Y)[\gamma/(1-\gamma)]^{\sigma}]\}^{\sigma/(\sigma-1)}/\{P_c + P_Y[(P_c/P_Y)[\gamma/(1-\gamma)]^{\sigma}]\} = 1$.

We outline the details of the labor supply decisions in section 4.2.3, where we discuss how workers choose their field of college majors and occupations over time. The decision of whether to attend college or not is made outside this model. This means that the supply of non–college graduates $\bar{H}$ is exogenous, and so is the total supply of native college graduates $(\overline{L_n + G})$. Those who do get a college degree can choose whether to work as a computer scientists L_n, or in some other occupation that requires a college degree G.

High-skilled immigrants who come in on H-1B visas can do so only if they meet the skill requirements of the visa and only if firms recruit them. As we have mentioned before, during the 1990s immigrants coming in as H-1Bs were increasingly being recruited as computer scientists. For simplicity, we will assume that all recruited H-1Bs are computer scientists L_F.

The size of the labor force in the economy is $X = \bar{H} + \overline{L_n + G} + L_F$ and total income m can be written as the sum of the labor income for the different types of workers plus profits earned by firms in the IT sector (Π) as in equation (3):

$$m = w(L_n + L_F) + sG + r\bar{H} + \Pi, \tag{3}$$

where w is the wage paid to computer scientists, s the wage earned by non-CS college graduates, and r is the wage paid to non–college graduates.

We assume that foreign computer scientists are willing to come and work in the United States at any available wage and are marginally more productive than native computer scientists. Each year the number of immigrants in the economy is capped at a given level $\bar{L}_F$ and because of this small productivity premium the cap always gets exhausted. Native computer scientists face a residual demand curve after all available foreigners have been hired.

One way to think about this assumption in our model is that any extra productivity is almost entirely offset by the recruitment costs of hiring foreigners. Also, due to H-1B restrictions, immigrants get paid the same wage as native computer scientists. In what remains of subsection 4.2.1 we will refer to foreign and native computer scientists as a single group, since from a firm's point of view they are indifferent between hiring the two at the going wage.[13]

Production in the IT Sector

The IT sector produces an aggregate IT good C. There are N monopolistically competitive heterogeneous firms that produce a different variety of this good as shown in equation (2). Following the framework introduced by Hopenhayn (1992) and Melitz (2003), each of these firms will have a different level of productivity. We assume each firm j has a Cobb-Douglas technology in the labor aggregate and intermediate inputs from the other sector as in equation (4):

13. In the data, we see that H-1Bs are almost entirely hired by larger firms. While this is an interesting and suggestive feature of the data, we leave it for future researchers to explore.

$$c_j = \phi_j L_c^{\beta} y_{cj}^{\psi_1} x_{cj}^{1-\psi_1}, \tag{4}$$

where y_{cj} is the amount of intermediate goods from sector Y and x_{cj} is the labor aggregate. Firm technology, $A(\ell_j) = \phi_j L_c^{\beta}$, has an endogenous component L_c^{β} and an exogenous component ϕ_j, which is a productivity draw that varies across firms. The term L_c^{β} captures a technological spillover in the IT sector that depends on the total number of computer scientists employed. Since computer scientists are innovators, we assume that their innovations create spillovers that increase the productivity of all firms in the sector, and this is captured by the β term.

The firm employs all three types of labor available in the economy. We assume that production technology has a nested CES structure.

$$x_j = \left[\alpha^c h_j^{(\tau-1)/\tau} + \left(1-\alpha^c\right) q_j^{(\tau-1)/\tau}\right]^{\tau/(\tau-1)}, \tag{5}$$

where h_j is the number of non–college graduates and q_j is the labor aggregate for college graduates. Here τ is the elasticity of substitution between college graduates and non–college graduates. Due to the nested nature of the CES function, we know that q_j is

$$q_j = \left[(\delta+\Delta)\ell_j^{(\lambda-1)/\lambda} + (1-\delta-\Delta) g_j^{(\lambda-1)/\lambda}\right]^{\lambda/(\lambda-1)}, \tag{6}$$

where ℓ_j is the number of CS workers and g_j the non-CS college graduates employed by firm j. Here λ is the elasticity of substitution between the CS workers and non-CS college graduates.

In equation (4) it is clear that the IT-sector firms have two drivers of technological change. The exogenous component of technology ϕ_j, has been modeled similar to the setup in the trade and the industrial organization literature (Chaney 2008; Hopenhayn 1992; Melitz 2003). The endogenous component of technology, captured by β, depends on the total number of computer scientists hired by the IT sector. These computer scientists innovate and create new technologies, increasing overall firm productivity. Here, we modify the setup used in the literature on economic growth (Acemoglu 1998; Arrow 1962; Grossman and Helpman 1991; Romer 1990).[14]

In the IT sector, the number of potential entrepreneurs is assumed to be fixed and their productivities have a known distribution $\Psi(\phi_j)$ with a positive support over $(0, \infty)$ and an associated density function $\psi(\phi)$. There is a productivity cutoff $\phi = \phi^*$ that captures the productivity level of the firm that breaks even. Therefore, the marginal producing firm earns no profits ($\pi(\phi^*) = 0$). Since profits are an increasing function of the productivity level, the equilibrium ϕ^* determines which firms produce ($\phi_j > \phi^*$) and which

14. Since we do not model economic growth, there are some clear departures from this literature. While many papers assume that the *rate of change* of technology depends on the quantity of a type of labor, we assume the *level* of technology depends on labor. Furthermore, a lot of this literature models a separate R&D sector that sells patents for these technologies—whereas in our model technology is assumed to be nonexcludable.

ones do not ($\phi_j < \phi^*$). The conditional distribution of $\psi(\phi)$ on $[\phi^*,\infty)$ can therefore be written as

$$\mu(\phi) = \begin{cases} \dfrac{\psi(\phi)}{1 - \Psi(\phi^*)}, & \text{if } \phi \geq \phi^* \\ 0, & \text{otherwise} \end{cases}.$$

The productivity distribution $\Psi(\phi_j)$ of entrepreneurs is assumed to be a Pareto distribution, with parameters k and $\phi_{\min}$ such that $\Psi(\phi_j) = 1 - (\phi_{\min}/\phi_j)^k$.

The intuition behind this modeling choice is that whenever economic conditions change, the firms that get pushed into/out of production are the marginal firms (those with ϕ_j closer to ϕ^*), while the larger more productive firms produce regardless. We expect such behavior in the IT sector when we allow more immigrants into the economy. As immigration allows firms to pay lower wages, the marginal firms are the ones that enter into production and large firms capture most of the increase in profits. For a given mass of potential producers, N_e, the total number of firms that produce can be written as in equation (7):[15]

$$N = \left(1 - \Psi(\phi^*)\right) N_e. \tag{7}$$

Such a model follows an approach to market entry closer to Chaney (2008) rather than the original Melitz (2003) model where the potential pool of entrants is not fixed.[16]

The firm's problem therefore boils down to maximizing profits by choosing the amount of labor inputs. If they choose to produce, they pay an upfront fixed cost of production f, which is in terms of the cost of the non-IT good P_Y (equation [8]). Each firm is a monopolist for their own variety and faces a demand curve as in equation (A.2).

$$\max_{\ell_j, g_j, h_j, y_{cj}} \pi_j = \phi_j P_c C^{1/\varepsilon} c_j^{(\varepsilon-1)/\varepsilon} - w\ell_j - sg_j - rh_j - P_Y y_{cj} - P_Y f. \tag{8}$$

The first-order conditions from this exercise determine the labor demand from the IT sector for each type of labor. Total labor hired by this sector is denoted by the subscript c, and aggregate employment of each type of worker can be expressed as L_c, G_c, and H_c.

15. While our model of firm entry does not have dynamic implications, Waugh (chapter 6, this volume) provides a more extensive treatment of the potential effects of skilled immigration on firm entry and exit dynamics.

16. In the original Melitz setting there are a number of potential entrants who have to pay an additional fixed cost f_e to get a productivity draw, and once they know their productivity they produce if $\phi_j > \phi^*$. New entrants in this model can be both high and low productivity and end up driving expected net profits to zero. Di Giovanni, Levchenko, and Ortega (2015) think of the case with a fixed pool of potential producers as the short run, where the number of varieties available only changes through the entry and exit of marginal firms, having small effects on aggregate welfare.

Production in the Non-IT Sector

The non-IT sector produces good Y and is assumed to be perfectly competitive. We assume the representative firm in this sector has a Cobb-Douglas constant returns-to-scale technology over intermediate inputs from the other sector and the labor aggregate.

$$Y = C_y^{\psi_2} X_y^{1-\psi_2}, \tag{9}$$

where again C_y represents intermediate inputs from the IT sector and X_y the labor aggregate. This sector also employs the three types of labor denoted by subscript Y. Therefore, X_y can be written as

$$X_y = \left[\alpha^y H_y^{(\tau-1)/\tau} + \left(1-\alpha^y\right) Q_y^{(\tau-1)/\tau}\right]^{\tau/(\tau-1)}. \tag{10}$$

Again, by the nested CES assumption, Q_y can be represented by

$$Q_y = \left[\delta L_y^{(\lambda-1)/\lambda} + (1-\delta) G_y^{(\lambda-1)/\lambda}\right]^{\lambda/(\lambda-1)}. \tag{11}$$

This sector is less intensive in computer scientists than the IT sector. To capture this, we model the intensity of CS workers to be higher in the IT sector (the incremental share is captured by Δ in equation [6]), and allow the computer scientists in the IT sector to have an additional impact on the technology in the firm (captured by β). Both sectors have the same elasticity of substitution between college and non–college graduates (τ) and between computer scientists and non-CS college graduates (λ).

The representative firm in the non-IT sector has to therefore solve the following maximization problem:

$$\max_{L_y, G_y, H_y, C_y} \Pi_y = P_y C_y^{\psi_2} X_y^{1-\psi_2} - wL_y - sG_y - rH_y - P_c C_y. \tag{12}$$

The first-order conditions determine the demand for the intermediate inputs and the different types of labor in this sector. Together with the demand for labor from the IT sector, we can then derive the aggregate labor demand for each worker. Section 4.2.3 describes the supply of the different types of workers, and section 4.2.4 describes the equilibrium, where we also detail how the labor demand curve shifts over time given the technological boom in the 1990s.

4.2.2 Trade with the Rest of the World

The US economy trades both IT goods and the other goods with the rest of the world (W). Information technology firms export final goods to consumers in other countries, whereas US consumers import the other goods from the rest of the world.[17]

17. While we do not explicitly model outsourcing decisions, we do allow for the fact that imported goods in the other sector can be used as intermediate goods in production for the IT sector.

We assume consumers in the rest of the world (W) have the same utility function as US consumers:

$$U_W(C_W, Y_W) = \left[\gamma_W C_W^{(\sigma-1)/\sigma} + (1-\gamma_W) Y_W^{(\sigma-1)/\sigma}\right]^{\sigma/(\sigma-1)}. \tag{13}$$

Since the United States is the only producer of IT goods, foreign consumption is equivalent to US exports of IT goods. Imports into the United States from the rest of the world are represented by Y_{IM}. For convenience we assume trade is balanced, implying that the value of imports must equal the value of exports:

$$P_c C_W = P_y Y_{IM}. \tag{14}$$

Here we assume that the United States is the only producer of IT. Even though Freeman (2006) stresses how high-skilled immigration may help the United States maintain its comparative advantage in IT, we may expect that immigration policy affects IT production elsewhere in the world, especially via the diffusion of knowledge. Khanna and Morales (2015) draw up a general-equilibrium model of both the United States and India—the other major producer of IT—to study how the H-1B program affects production, human capital accumulation, and labor market welfare for agents in both countries. The possibility of migrating to the United States induces students and workers in other countries to accumulate CS-specific human capital, and return migrants help facilitate the diffusion of technology. Over time, in the latter half of the first decade of the twenty-first century, India becomes the major exporter of IT, eroding the US's comparative advantage. Khanna and Morales (2015) can be thought of as a long-run extension of our current work, with consistent implications for the period of study here (the 1990s).

4.2.3 Labor Supply of US Computer Scientists

The firms' decision problem determines not only the product market equilibrium but also the demand curves for the different types of labor. To describe the workers' decisions, we develop a dynamic model of labor supply that captures the choices made in deciding a field of study in college and occupational choices later in life. The model builds on previous work by Freeman (1975, 1976) and Ryoo and Rosen (2004), and closely follows the setup of Bound et al. (2015). While Bound et al. (2015) was a partial-equilibrium model that studied the decisions made between CS and STEM occupations for a given labor demand elasticity, we extend it to a general-equilibrium framework that includes all types of labor and rigorously models the firm's decision to derive the labor demand curve that the workers face as well.

While we model the decisions to choose a field of study for US workers who attend college, we do not explicitly model the decision to attend college in the first place. This is because we assume that changes in wages for CS-related occupations do not greatly affect the college-going decision for students. The supply of workers who have only a high school degree $\bar{H}$ is therefore assumed to be the same whether or not there were changes in the number of foreign computer scientists in the labor market. Therefore, the total supply of US workers with a college degree $(\overline{L_n + G})$ is also assumed to be fixed. However, we do model the decisions of these college-educated workers as they make choices between majoring in CS degrees or other degrees and then their occupation choices in each year of their lives until retirement.

In our model, there are three potential sources of CS workers. First, there are those who earn CS bachelor's degrees from US institutions and join the workforce only after they finish college. Second, there are college-educated US residents working in other occupations who can switch into computer science, but must pay costs to switch occupations. Third, there are foreigners who are being recruited on temporary work visas.

Given that most foreign workers that come on H-1Bs are computer scientists, we model computer science as the only profession that they get hired into. There are therefore two sources of non-CS college-educated workers—those who graduate with any degree that is not computer science and those that switch from CS work to non-CS work by paying the switching cost.

We model US college graduates as maximizing their lifetime utility by making two types of decisions. When they are twenty years old they choose their field of study in college, which influences their initial occupation at graduation. From ages twenty-two to sixty-five, they choose between working as a computer scientist or in another occupation. All individuals have rational, forward-looking behavior and make studying and working decisions based on the information available in each period.

The labor demand curve derived from the firms' decision problem discussed in the previous sections shifts out yearly due to productivity shocks. These shifts help identify the labor supply parameters and trace out the labor supply curve.

Field of Study Decision

In our model students choose their field of study when they are undergraduate juniors. Equation (15) captures this decision. At age twenty, a student i draws idiosyncratic taste shocks for studying computer science or another field: η_i^{cs} and η_i^{o}, respectively. This student has expectations about the prospects of starting a career in each occupation after graduation (age twenty-two), which have values V_{22}^{cs} and V_{22}^{o}, respectively. Given this information, an individual chooses between pursuing computer science or a different choice of major at the undergraduate level.[18]

18. We are assuming that students decide their major after the end of their second year in school. Bound et al. (2015) experiment with a four-year time horizon and doing so made little qualitative difference.

Worker utility is a linear function of their tastes and their career prospects in each sector and they discount their future with an annual discount factor ρ. Additionally, there is an attractiveness parameter θ_o for studying in a field that is not computer science that all students experience. This parameter may be negative if, on average, students prefer studying computer science

$$\max\left\{\rho^2 \mathbb{E}_t V_{22}^{cs} + \eta_i^{cs}, \rho^2 \mathbb{E}_t V_{22}^{o} + \theta_o + \eta_i^{o}\right\}. \tag{15}$$

We assume that the individual taste parameters η_i^{cs} and η_i^{o} are independently and identically distributed and for $d = \{cs, o\}$, can be defined as $\eta_i^d = \sigma_0 v_i^d$, where σ_0 is a scale parameter and v_i^d is distributed as a standard Type I Extreme Value distribution. This assumption allows the decisions of agents to be formulated in aggregate probabilities, and is therefore commonly used in dynamic discrete choice models (Rust 1987; Kline 2008). We describe the probability of enrollment in degrees in the "Labor Supply Derivations" section in the appendix.

One crucial parameter for how studying choices are sensitive to different career prospects is the standard deviation of taste shocks. Small values of σ_0 imply that small changes in career prospects can produce big variations in the number of students graduating with a CS degree.

Occupational Choice

The field of study decisions determine if an individual enters the labor market at age twenty-two, either as a computer scientist or in a different occupation. However, individuals can choose to switch occupations between the ages of twenty-two and sixty-five. At the start of each period, individuals use the information at hand and choose their occupation in order to maximize the expected present value of their lifetime utility.

Switching occupations, however, is costly for the worker, and these costs vary with age. This is because workers have occupation-specific human capital that cannot easily be transferred across occupations (Kambourov and Manovskii 2009). The occupational switching costs are modeled as a quadratic function of a worker's age, allowing for the fact that it becomes increasingly harder to switch occupations as workers get older.[19]

Like in the college-major decision, we assume that workers have linear utility from wages, taste shocks, and career prospects.[20] The value functions of worker i at age a between twenty-two and sixty-four at time t if she starts the period as a computer scientist or other occupation are therefore going to be

$$V_{t,a}^{cs} = \max\left\{w_t + \rho \mathbb{E}_t V_{t+1,a+1}^{cs} + \varepsilon_{it}^{cs},\ s_t - \zeta(a) + \rho \mathbb{E}_t V_{t+1,a+1}^{o} + \varepsilon_{it}^{o} + \theta_1\right\}, \tag{16}$$

19. While our model has no general human capital accumulation and wages do not vary with the age of a worker, the implications of the model would still hold if individuals expect similar wage-growth profiles in each occupation.

20. Wages must be totally consumed in that same year and workers cannot save or borrow.

$$(17) \quad V^{o}_{t,a} = \max\left\{w_t - \zeta(a) + \rho\mathbb{E}_t V^{cs}_{t+1,a+1} + \varepsilon^{cs}_{it}, s_t + \rho\mathbb{E}_t V^{o}_{t+1,a+1} + \varepsilon^{o}_{it} + \theta_1\right\},$$

where $\zeta(a) = \zeta_0 + \zeta_1 a + \zeta_2 a^2$ is the monetary cost of switching occupations at age a, and θ_1 is the taste attractiveness parameter for not working as a computer scientist experienced by all workers. Finally, all workers retire at age sixty-five and their retirement benefits do not depend on their career choices. Therefore, at age sixty-five workers face the same decision problem without consideration for the future.

As in the college-major decision problem, we will assume that taste shocks are independently and identically distributed and for $d = \{cs, o\}$ can be defined as $\varepsilon^d_{it} = \sigma_1 v^d_{it}$ where σ_1 is a scale parameter and v^d_t is distributed as a standard Type I Extreme Value distribution.

The standard deviation of the taste shocks, the sector-attractiveness parameter, and the cost of switching occupations will affect the sensitivity of occupational switching to changes in relative career prospects. Since individuals are forward looking, the working decisions depend upon the equilibrium distribution of their career prospects. We describe the probabilities of employment, occupational switching, and the expected value of future prospects in the "Labor Supply Derivations" section in the appendix.

Labor Supply of Foreign Computer Scientists

We model high-skilled foreign workers as only being hired as computer scientists, since during the 1990s a majority of H-1Bs were hired into this occupation. By 2001, more than 21 percent of all computer scientists were born abroad and immigrated after the age of eighteen (March CPS). We assume that high-skilled foreigners have a perfectly elastic labor supply curve to the United States, since the wage that a computer scientist could obtain in countries like India or China, for instance, is substantially lower than it is in the United States (Clemens 2013). This wage premium creates a large queue of foreigners ready to take jobs in the United States. There is, however, an institutionally imposed cap on the total number of H-1Bs that restricts the number of foreign computer scientists each year.

Institutional requirements also force firms to pay foreigners the prevailing US wage. We assume that the additional costs of recruiting foreigners offset the productivity advantage that foreigners may have over their US counterparts. During the 1990s, a large fraction of the CS workers coming from abroad were on H-1B visas. Given that this was a period when the H-1B cap was usually binding, and given our assumption that foreign and domestic CS workers are effectively identical, we treat the quantity of foreign CS workers coming to the United States as exogenous.

4.2.4 Equilibrium

Equilibrium in each period can be defined as a set of prices and wages $(P_{ct}, P_{Yt}, w_t, s_t, r_t)$, quantities of output and labor $(C^*_t, Y^*_t, C^*_{dt}, C^*_{yt}, C^*_{Wt},$

Y^*_{dt}, Y^*_{ct}, Y^*_{1Mt}, L^*_{nt}, K^*_{Ft}, G^*_t; H^*_t), number of firms (N_t), and the productivity cutoff (ϕ^*_t) such that[21]

- Consumers in the United States and the rest of the world maximize utility by choosing C_t and Y_t taking prices as given, and choose their college major and occupations taking wages as given.
- Firms in both sectors maximize profits taking wages and aggregate prices as given.
- In the IT sector, the firm with productivity ϕ^*_t gets zero profits. All firms with $\phi_{jt} > \phi^*_t$ produce, while those with $\phi_{jt} < \phi^*_t$ do not.
- Output and labor markets clear. The equations for the market-clearing conditions are in the "Market-Clearing Conditions" section of the appendix.

Native college graduates face the decision of whether to work as computer scientists or in some other occupation that requires a college degree. This decision is no longer static, but has an intertemporal dimension that requires the definition of the dynamic equilibrium in the labor market for college graduates. As in Bound et al. (2015), this equilibrium is characterized by the system of equations ([15]–[17]) and a stochastic process Z_t. In the "Labor Supply Derivations" section of the appendix we characterize further equations, including future expectations, and the dynamic supply of colleges and workers.

A unique equilibrium is pinned down each period by an aggregate labor demand curve for US computer scientists relative to other college graduates that comes from the product market model.

Even though this labor demand curve from the two sectors has no closed-form solution, we will express it as in equation (18), a setup that will prove to be useful for the calculations in the following sections.

$$(18) \qquad \frac{L_{nt}}{G_t} = Z_t + \Upsilon\left(\frac{w_t}{s_t}\right),$$

where $\Upsilon(w_t/s_t)$ is a baseline-relative demand curve that depends on the relative wage; Z_t is a shifter that can be thought of as a combination of the productivity shocks from the IT boom that shifts out the relative demand for computer scientists every year and the cap of foreign computer scientists $\bar{L}_F$ that shifts in the relative demand curve every period; and Z_t is assumed to follow a random walk process with high persistence such that

$$(19) \qquad Z_t = 0.999Z_{t-1} + 0.001\bar{Z} + \xi_t,$$

where $\bar{Z}$ is the steady-state value of Z_t and ξ_t is an i.i.d. shock.[22]

The equilibrium in the labor market can be expressed by a mapping from

21. Note that we have introduced a t subscript to each of the variables to denote that there is a different equilibrium for each time period.

22. We assume workers consider both the technological progress from the IT boom as well as the increase in immigrants to be a series of highly persistent shocks.

the state variables: $s = \left\{R_t, L_{n,t-1}^{22}, \ldots, L_{n,t-1}^{64}, G_{t-1}^{22}, \ldots, G_{t-1}^{64}, Z_{t-1}\right\}$ and exogenous productivity shock ξ_t to the values of L_{nt}, w_t, G_t, s_t, and $\mathbf{V_t}$, the vector of career prospects at different occupations for different ages, that satisfies the system of equations for labor supply as well as each period's relative demand curve.

4.3 Calibration

We calibrate the parameters of our model in order to determine how welfare changes due to immigration. We have a total of twenty-five parameters: σ, ε, γ, γ_W, ψ_1, ψ_2, β, α_c, α_y, τ, λ, δ, Δ, k, $\phi_{\min}$, N_e, and f from the product market, and σ_0, σ_1, θ_0, θ_1, ζ_0, ζ_1, ζ_2, and ρ from the US college graduates labor market. We focus on the period 1994–2001 that corresponds to the IT boom and when the H-1B cap was mostly binding.

In order to calibrate the different parts of the model, we follow a sequential approach. First, we calibrate the parameters in the product market assuming total labor supply of L_t, G_t, and H_t are fixed (i.e., ignoring the choice of native workers between L_t and G_t). What makes this possible in our model is the fact that adjustment costs imply that the stock of the different types of labor are fixed in the very short run. This approach is akin to the approaches taken by Freeman (1975, 1976) and Ryoo and Rosen (2004) in their modeling of adjustments on the labor market for scientists.

In the next step we use the calibrated parameters to derive the aggregate labor demand curve for computer scientists relative to other college graduates for every year. As a third step, we use the predicted shifts in labor demand to calibrate the parameters of the labor supply curve of different types of college graduates. Finally, we use the calibrated labor supply curve, labor demand curve, and product-demand parameters to calculate welfare under the economy where immigration is encouraged via the H-1B program and the counterfactual scenario where immigration is restricted.

4.3.1 Product Market Calibration

We calibrate the parameters of the product market to match different features of the data as explained in the following subsections. The details of the data we use, including sources and definitions of the different sectors and occupations, can be found in the "Details of the Data Used" section of the appendix.

The model is calibrated separately for each year between 1994 and 2001. While some parameters are assumed to be constant over time, others change in order to capture structural changes in the economy. Particularly, the production-function parameters (α_{ct}, α_{yt}, δ_t, Δ_t, ψ_{1t}, and ψ_{2t}) will be recalibrated every year to capture the technological change that affects the two sectors during this period. This can be thought of as describing the skill-biased technological change over this period, since the share of labor cost

Table 4.2 **Calibrated parameters from the product market**

Time-invariant parameters									
		σ	1.00			k	2.62		
		ε	3.20			N_e	0.25		
		τ	1.70			f	1.07–1.24		
		β	0.23			$\phi_{\min}$	1		
Time-varying parameters									
		1994	1995	1996	1997	1998	1999	2000	2001
γ		0.042	0.046	0.050	0.052	0.054	0.055	0.055	0.054
γ_W		0.014	0.015	0.015	0.016	0.014	0.015	0.016	0.013
ψ_1		0.522	0.524	0.525	0.524	0.523	0.521	0.517	0.513
ψ_2		0.055	0.054	0.054	0.053	0.053	0.052	0.052	0.051
α_c		0.438	0.432	0.427	0.419	0.410	0.401	0.395	0.390
α_y		0.502	0.494	0.486	0.479	0.473	0.468	0.465	0.463
	$\lambda = 1$	0.053	0.055	0.059	0.063	0.067	0.069	0.073	0.072
δ	$\lambda = 2$	0.224	0.227	0.233	0.240	0.248	0.253	0.262	0.260
	$\lambda = 4$	0.395	0.398	0.401	0.405	0.414	0.420	0.430	0.428
	$\lambda = 1$	0.217	0.215	0.215	0.218	0.225	0.237	0.249	0.270
Δ	$\lambda = 2$	0.174	0.168	0.153	0.147	0.146	0.157	0.158	0.175
	$\lambda = 4$	0.073	0.066	0.048	0.039	0.036	0.046	0.042	0.057

Notes: σ: elasticity of substitution between C and Y; ε: elasticity of substitution across IT varieties; τ: the elasticity of substitution between college graduates and non–college graduates; β: the technological spillover of computer scientists in IT; f: fixed cost of production; N_e: mass of potential producers; k and $\phi_{\min}$: distribution and scale parameters from the Pareto distribution; γ: distributional parameter of domestic CES utility; γ_W: distributional parameter of foreign CES utility; ψ_1, ψ_2: production-function parameters for intermediate inputs in IT and the other sector, respectively; α_c, α_y: distributional parameter for non–college graduates in the IT and other sector production function; δ: distributional parameter for computer scientists in both sectors; Δ: distributional parameter for computer scientists in IT; λ: elasticity of substitution between CS and non-CS college graduates.

that these sectors spend in computer scientists is increasing over time. The utility parameters γ_t and γ_{Wt} are also allowed to shift over time to capture changes in local and foreign consumer preferences toward the IT sector. A summary of all calibrated parameters in the product market can be found in table 4.2.

Domestic Utility Function Parameters

The three parameters in the consumer utility function are σ, ε, and γ_t; σ is the elasticity of substitution between the composite IT good C and the good Y. We calibrate this parameter using the ratio of first-order conditions of goods Y and C from the consumer's utility maximization problem: $[\gamma/(1-\gamma)](C/Y)^{-1/\sigma} = P_c/P_Y$.

This relationship can be reformulated as

(20)
$$\log\left(\frac{C}{Y}\right) = -\sigma\log\left(\frac{1-\gamma}{\gamma}\right) - \sigma\log\left(\frac{P_c}{P_Y}\right).$$

We estimate σ using a regression of the relative quantity index on the relative price index. We use data from the Bureau of Economic Analysis (BEA) industry-specific price and quantity indices.[23] The BEA data allows us to distinguish prices and quantities in the IT sector, and all the other sectors in the economy. The coefficient of this regression is statistically indistinguishable from $\sigma = 1$. Given the plausibly exogenous technological change during the period that drives down prices, we use this estimate as our main specification and proceed using a Cobb-Douglas utility specification. We also run a series of robustness checks running the results for different values of σ that are summarized in the "Sensitivity Analysis" section of the appendix.

The elasticity across IT varieties, ε, is calibrated using the markup condition that comes from the IT firms' profit-maximization condition (equation [21]). We follow an approach similar to Gaubert (2015) and match average value added to cost ratios for the IT sector. The data for this is again taken from the BEA's annual industry accounts that report value added, as well as costs like compensation to employees and taxes. For a marginal cost $MC(c_i)$, the price markup can be used to determine the value of ε:

(21)
$$p_i = \frac{\varepsilon}{\varepsilon - 1} MC(c_i).$$

We calibrate $\varepsilon = 3.26$. Bernard et al. (2003) calculate a value of 3.8 for all US plants, whereas Broda and Weinstein (2006) find a value of 2.2 for varieties of "automatic data processing machines and units." Since our estimates lie within this region, we believe them to be reasonable. We show that our results are robust to other reasonable values of this parameter in the "Sensitivity Analysis" section of the appendix.

We calibrate the distribution parameter γ_t to match the share of expenditures in the IT good (using equation [22]). Again we use data from the BEA on industry-specific gross domestic product (GDP) of IT as a share of total GDP.[24]

(22)
$$\frac{P_c C}{m} = \frac{P_c}{P_c + [((1-\gamma)/\gamma)P_c]^{\sigma}}.$$

23. The BEA price indices methodology can be found here: http://www.bea.gov/national/pdf/chapter4.pdf and http://www.bls.gov/opub/hom/pdf/homch17.pdf. The specific methodology for personal computers and peripheral equipment are detailed at http://www.bls.gov/cpi/cpifaccomp.htm, where they discuss adjusting for quality as well. While they do adjust for quality differences, we may still underestimate quality changes in IT (Gordon 1990), which would affect our estimate of β. We do a rigorous sensitivity analysis for different values of β.

24. For all time-varying parameters that are matched to shares observed in the data we run a regression of the raw share on a linear and quadratic time trend to recover the time invariant parameters. We then predict the share using those coefficients and calibrate the parameters to match the predicted shares.

We calibrate γ_t conditional on the equilibrium prices, the share of consumption of the IT good γ and the calibrated value of σ. For the Cobb-Douglas specification we just use the share of IT-industry GDP to total domestic GDP. As already discussed, γ_t is time varying in order to capture potential changes in consumer preferences over time for the IT good relative to the rest of the goods in the economy. Table 4.2 shows how γ_t steadily rises from 0.042 at the start of the period to 0.052 by the year 2001.

Foreign Utility Function

Consumers from the rest of the world are assumed to have the same utility function as consumers in the United States. While we assume the elasticity of substitution σ is the same for both countries ($\sigma = 1$), the distribution parameter γ_{tW} is selected to match the share of consumption of the rest of the world for US IT products. We use the share of exports in IT to US GDP and the relative size of the US economy to the rest of the world to pin down this parameter. Again, we allow this parameter to change over time to capture potential changes in preferences for consumers abroad.

Production-Function Parameters

The elasticity of substitution between high school and college graduates (τ) and between computer scientists and other college graduates (λ) are assumed to be time invariant and equal across sectors. To calibrate τ we follow several influential papers that provide estimates for this parameter such as Katz and Murphy (1992), Card and Lemieux (2001), and Goldin and Katz (2007) and set $\tau = 1.7$, which is an average of their estimates.[25] We present our results for a range of values of λ (1, 2, and 4) that correspond to aggregate relative labor demand elasticities of 1.02, 1.99, and 3.98. Ryoo and Rosen (2004) estimate aggregate relative demand elasticities that lie between 1.2 and 2.2 for engineers, which are included in the range of values we use.

To calibrate the value of β, the technological spillover from total CS in the IT sector, we look at the relationship between the price decline in IT and the increase in total CS working in the sector. We use the aggregate CS in IT equilibrium condition that gives us a relationship between prices of IT and total labor in CS as in equation (23):

$$(23)\quad \log P_c = \mho(w_t, s_t, r_t) - \frac{1}{\varepsilon}\log C_t - \psi_1 \frac{\varepsilon - 1}{\varepsilon}\log P_y + \frac{(1-\beta(\varepsilon-1))}{\varepsilon}\log L_c.$$

We run the regression of $\log(P_c)$ on a linear and quadratic time trend, the log of quantity of IT good, the log price of the other good, and the log of

25. Katz and Murphy (1992) find 1.41, Card and Lemieux (2001) find estimates between 2 and 2.5, and Goldin and Katz (2007) find 1.64. Strictly speaking, these numbers refer to the elasticity of substitution between college- and non-college-educated labor in the US economy, while our parameter is sector specific. The aggregate elasticity involves both within- and between-sector components. However, our simulations suggest that setting $\tau = 1.7$ produces an aggregate elasticity indistinguishable from 1.7 to the first digit.

total computer scientists in IT. The time trend aims to capture fluctuations in the wages of the different types of workers over time. The calibrated value of β is 0.233. Effectively, this procedure attributes all of the total factor productivity (TFP) change to the increase in computer scientists working for the IT sector, while in reality there are several other factors that also affect technical progress in IT. As a result, our estimates will tend to overestimate the impact of computer scientists on technological change. Our estimate is quite close to the Peri, Shih, and Sparber (2014) estimates of changes in TFP attributable to the total number of STEM workers. In the "Sensitivity Analysis" section of the appendix, we explore the sensitivity of our results to our estimate of β.

The production-function parameters α_{ct}, α_{yt}, δ_t, Δ_t, ψ_{1t}, and ψ_{2t} are calibrated separately every year to reflect the skill-biased technological change the two sectors face during the period. This allows us to capture that, increasingly, firms in both sectors spend a higher share of their expenditures on college graduates.

The shares of expenditures on non–college graduates in both sectors are matched to the observed share of labor income for each year in the March CPS. Here we define the shares observed in the data as $\vartheta_{t,C,H}$ and $\vartheta_{t,Y,H}$, such that

$$\vartheta_{t,C,H} = \frac{\alpha_{ct}\bar{H}_{ct}^{(\tau-1)/\tau}}{\alpha_{ct}\bar{H}_{ct}^{(\tau-1)/\tau} + (1-\alpha_{ct})\bar{Q}_{ct}^{(\tau-1)/\tau}}, \tag{24}$$

where $\bar{H}_{ct}$ and $\bar{H}_{ct}$ are the quantities observed in the CPS for each sector. We analogously calibrate α_{yt} using the shares observed in the data ($\vartheta_{t,Y,H}$, $\bar{H}_{yt}$, and $\bar{Q}_{ct}$).

In both sectors we have the parameter δ_τ that is the distribution parameter associated with computer scientists. We calibrate this parameter to match the relative wage of CS to other college graduates (w_t/s_t). The IT sector has a higher share of CS than the other sector, so we calibrate the parameter Δ to match the share of total labor expenditure spent in CS by the IT sector in a manner similar to our approach for calibrating α_{yt} and α_{ct}.

In table 4.2 we can see how skill-biased technological change in the economy changes these parameters over time; δ_t steadily increases over this period as both sectors want to hire more computer scientists. The values of α_{yt} and α_{ct} steadily decrease for both sectors, showing that they spend more of their income on college graduates than on high school graduates. Parameters associated with the intermediate inputs from another sector (ψ_{1t}, ψ_{2t}) are calibrated using the share of intermediate inputs from other sectors relative to the GDP, which we obtain from the BLS input-output tables.

Entry into Production in the IT Sector

There are four parameters related to the entry decision and productivity distribution in the IT sector. The number of firms in the sector depend on f, the fixed cost of production, and N_e, the mass of potential producers. The Pareto distribution parameters k and ϕ_{min}, determine the productivity levels of these firms. All these parameters are assumed to be time invariant.

We calibrate f to match the average firm size in the IT sector observed in the data for the steady-state year 1994. In order to do this we use information on the number of firms and total employment in the IT sector from the US Census Bureau's Statistics of US Businesses (SUSB).[26] In 1994 we calibrate f to match the ratio of total employees and number of firms in the data for the IT sector. The calibrated values for f are 1.24, 1.14, and 1.07 (for λ values of 1, 2, or 4, respectively). For the rest of the years we allow the number of firms N_t to adjust endogenously, as the profits from production change over time.

The parameter N_e is calibrated using information on establishment entry and exit.[27] We look at the total number of establishments over 500 employees in 2001 and calibrate the ratio of $(N_t/N_e) = (N_{94}/N_{01})$. Given that N_t in 1994 is used to calibrate f, we get the rescaled $N_e = 0.25$.[28]

The Pareto distribution parameter k is set to match the standard deviation of logarithm of US domestic plant revenues. Following Demidova (2008), we use the simulation reported by Bernard et al. (2003) of 0.84. In our model the standard deviation of $Ln(p_i c_i)$ is $(\varepsilon - 1)/k$, so given our value of $\varepsilon = 3.2$ we get a value of $k = 2.62$. The scale parameter ϕ_{min} is related to the choice of units in which to measure productivity, so we follow the convention in the literature and normalize it to 1.

Total Quantity of Labor

To calibrate the product market parameters, we use the total quantities observed in the data for each occupation type $\bar{L}_t$, $\bar{G}_t$, and $\bar{H}_t$ as if they were exogenously given. We normalize the US working population from the March CPS in 1994 to 100, and then allow for the population in our model to grow at the same rate as the growth in the US population. The shares of each type of worker are set equal to those observed in the data each year, which allows us to know the total number of college- and non-college-graduate workers, as can be seen in table 4.3.

26. This information comes from the 1992 Statistics of US Businesses (SUSB), http://www.census.gov/econ/susb/. Since the information was only available for 1992 and 1997–2012, we use the figures for 1992 as a proxy for 1994.

27. We get information on entry and exit of establishments in the IT sector by year from the Business Dynamics Statistics (http://www.census.gov/ces/dataproducts/bds/). Entry and exit was only available for establishments, not firms when looking at specific industries.

28. Other papers such as Demidova (2008) and Melitz and Redding (2015) use the exit rate to calibrate parameters related to fixed cost of production and entry, but unlike us calibrate the slightly different Melitz (2003) model. The strategy we use is somewhat different as we have a fixed pool of potential entrants.

Table 4.3 **Normalized population and growth as observed in the data**

Year	X_t	L_{tF}	L_{tn}	G_t	H_t
1994	100.00	0.13	0.99	24.30	74.59
1995	101.18	0.16	1.02	24.85	75.16
1996	103.31	0.19	1.12	25.61	76.39
1997	105.25	0.24	1.20	26.26	77.55
1998	107.35	0.26	1.27	27.06	78.76
1999	109.12	0.31	1.30	27.85	79.67
2000	110.95	0.37	1.35	28.71	80.52
2001	111.77	0.40	1.37	29.51	80.49

Note: Total working population as shown in the CPS is normalized to 100 in 1994. For subsequent years we allow total population to grow at the same rate as the working population in the United States. The shares of each type of occupation are then used to calculate the total number of workers in each category.

4.3.2 Deriving the Labor Demand Curve

Once we calibrate the product market parameters we are able to derive a labor demand curve for computer scientists relative to other college graduates. Such a demand curve does not have a closed-form solution that comes directly from the model, so we derive it by first changing the relative values of $\overline{L}_t \,/\, \overline{G}_t$ that we feed into the model and then calculating the predicted value of w_t/s_t. We run this exercise only for the steady-state year, 1994, and calculate w_t/s_t for different values of $\overline{L}_t \,/\, \overline{G}_t$ that ranges between 0.04 and 0.07.[29] We then fit a second-order polynomial to get a closed-form solution of the relative labor demand curve.[30]

The elasticity of labor relative demand for computer scientists to other college graduates depends crucially on the parameter λ. We derive the labor demand for our three values of λ and get what we call the baseline labor demand curve as in equation (25), calculated using the calibrated model for the steady-state year 1994:[31]

$$\frac{\widehat{L_t}}{G_t} = \hat{\Upsilon}\left(\frac{w_t}{s_t}, \lambda\right). \tag{25}$$

For the remaining years we allow the demand curve to shift for two reasons. First, to capture the innovation taking place in the economy. This exogenous technological change is captured by the time-varying parameters

29. Relative total computer scientists to other college graduates in the data is 0.0406 in 1994 and goes up to 0.0466 in 2001. We therefore capture more than the range of possible values in the data.

30. The second-order polynomial perfectly predicts the model with a $R^2 = 1$. We experiment with higher-order polynomials to fit the labor demand curve and our results do not change.

31. The elasticity of the derived labor demand curve is very close to the value of λ, more specifically 1.015, 1.99, and 3.98 for λ equal to 1, 2, and 4, respectively.

of the production functions. Second, the demand curve shifts to capture the relative changes in the stock of college graduates to non–college graduates, which is determined outside of the model.

We can calculate the labor demand shifter Λ_t as in equation (26). This shifter applies to the total demand of computer scientists relative to other college graduates, including both native and foreign computer scientists.

$$\hat{\Lambda}_t = \frac{L_t}{G_t} - \hat{\Upsilon}\left(\frac{w_t}{s_t}\right). \tag{26}$$

As a last step, in order to use the variation in the demand curve to trace out the relative supply curve for *native* computer scientists only, we subtract the relative number of foreign computer scientists each year to derive to the total demand shifter Z_t as presented in equation (18). As a reminder, we treat the quantity of foreign CS workers coming to the United States as exogenous since the H-1B cap was binding throughout this period. Given that we assume foreign CS workers are willing to work at any wage and are slightly more productive than natives, they get hired first until they exhaust the H-1B cap, while native workers face a residual labor demand curve. The total shifter $Z_t = \hat{\Lambda}_t - \left(\bar{L}_{tF} / G_t\right)$ allows us to write the labor demand for native CS relative to other college graduates as in equation (27):

$$\frac{L_{nt}}{G_t} = Z_t + \hat{\Upsilon}\left(\frac{w_t}{s_t}\right). \tag{27}$$

In the steady state, $\hat{\Lambda} = 0$ and $\bar{Z} = -\left(\bar{L}_{94,F}/\bar{G}_{94}\right)$.

4.3.3 Calibrating Labor Supply

On the labor supply side of the model, we have eight parameters that need to be calibrated—$\{\sigma_0,\theta_0,\sigma_1,\theta_1,\zeta_0,\zeta_1,\zeta_2,\rho\}$. Of these, we pick the annual discount rate to be $\rho = 0.9$, and calibrate the other parameters to match the data. In our model we assume the total quantities of non–college graduates $\bar{H}_t$, native college graduates $\left(\overline{L_n + G}\right)_t$, and foreign computer scientists $\bar{L}_{tF}$ are determined outside the model.

In the way we set up the model, changes in lagged degree attainment, employment, and wages are driven by the exogenous technology shocks that shift out the demand curve for the different types of labor over this decade. As the demand curve shifts, it traces out the labor supply curve for workers. The technological developments that drive these shifts in the labor demand are assumed to not affect the parameters of the workers' labor supply decisions.

We use data on relative wages, employment, lagged degree attainment, and age shares to calibrate the remaining seven parameters. The first three series compare computer scientists to non-CS college-graduate workers. For example, relative wages compare the wages for CS workers with wages for non-CS college graduates. To do this, we use the March CPS. Details of the sample used in the data and specific variable definitions can be found in the "Details of the Data Used" section of the appendix.[32]

We simultaneously match wages, employment, and the share of US CS workers that are young (between age twenty-two and forty) in 1994 and 2001.[33] We also match relative degrees in computer science for 1994, 1997, and 2001. The series we use from the data are as follows:[34]

1. $L_{n,t}/G_t$ = (US computer scientists)/(Non–CS college educated US workers) for $t = \{1994, 2001\}$.
2. w_t/s_t = (Median weekly wages for computer scientists)/(Median weekly wages for non–CS educated) for $t = \{1994, 2001\}$.
3. $q^{cs}_{t+2} \,/\, q^{o}_{t+2}$ = [US computer science college degrees awarded (lagged 2 years)]/[US non–CS college degrees awarded (lagged 2 years)] for $t = \{1994, 1997, 2001\}$.
4. $\text{age}_t^{22,40}$ = (US computer scientists with age between 22 and 40)/(US $\text{CS}^{22,40}$ + US $\text{CS}^{41,65}$) for $t = \{1994, 2001\}$.

To simultaneously find parameter values that solve the model under these data restrictions, we use a Nelder-Mead simplex method. While the system uses all the data at the same time, there is strong intuition behind the identification of each parameter. For example, the relative degree-attainment data should help identify the taste parameters for field of major decisions (σ_0 and θ_0) as well as the fixed cost of switching occupations (ζ_0), whereas the relative employment data should help pin down the occupation-specific tastes (σ_1 and θ_1). The age shares in CS employment help identify the occupation switching cost parameters that depend on age (ζ_1 and ζ_2).

Labor Supply Calibration Results

Figure 4.2 shows the data used and the model fit from this exercise. The figures report both the path of the variables of interest predicted by the model and the CPS data we use for these series. We match two extreme years (1994 and 2001) for employment and wages and three years (1994, 1997, and 2001) for lagged degree attainment, and the remaining years plotted are an out-of-sample test of our method. The years in between (1995 to 2000) include years where there were observed changes to immigration laws, and other potentially structural changes that may make it difficult for the data to fit perfectly.

32. We exclude imputed wages and multiply top-coded values by 1.4. Bollinger and Hirsch (2007) show that including imputations can lead to biased results, whereas the top-coding adjustment is standard in the literature (Lemieux 2006). We smooth the raw data over three-year moving averages as follows: $X_{t,\text{smooth}} = (1/3)(X_{t-1,\text{raw}} + X_{t,\text{raw}} + X_{t+1,\text{raw}})$.

33. Given that in our labor supply model we impose all cohorts are the same size, we normalize the number of computer scientists of a given age group dividing by the total number of college graduates in that age group before calculating the age shares.

34. We have an exactly identified system as we use nine data moments to recover ten parameters $\{\sigma_0, \theta_0, \sigma_1, \theta_1, \zeta_0, \zeta_1, \zeta_2\}$, and two implied values of technology in the years we match the wage/employment data $\{A_{94}, A_{01}\}$.

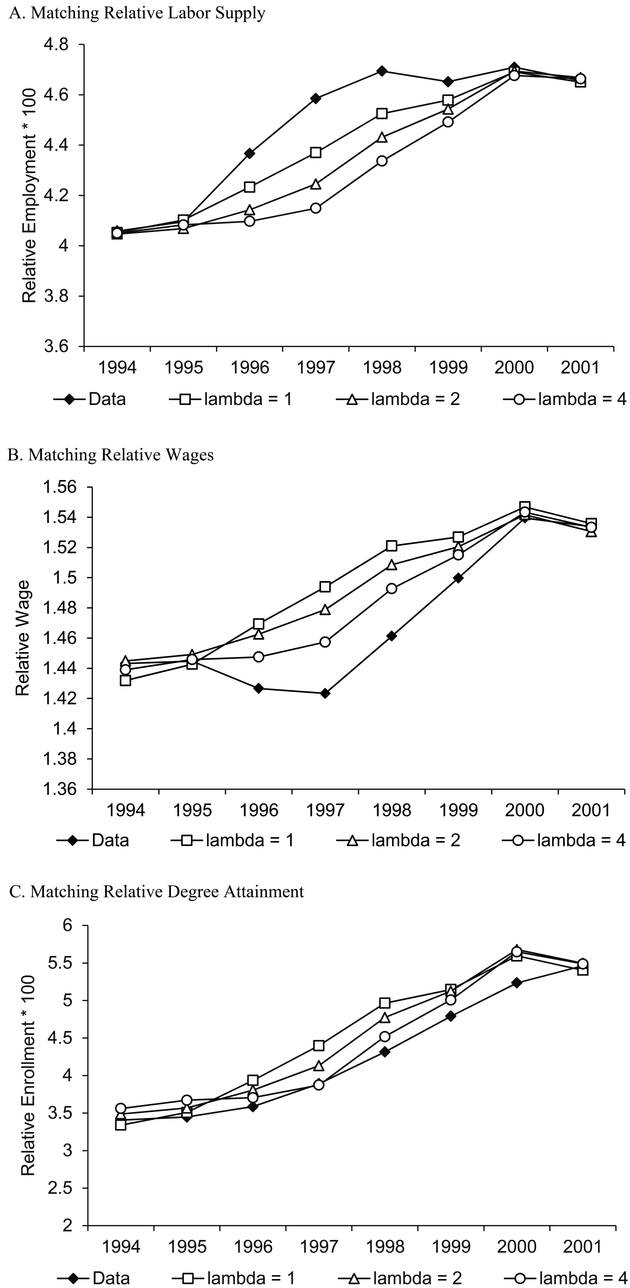


Fig. 4.2 Calibrating labor supply parameters

Notes: In the calibration exercise, the years 1994 and 2001 were used to match the data for employment and wages, whereas the years 1994, 1997, and 2001 were used to match the data on degree attainment (lagged two years). The years in between are an out-of-sample test. Wage and employment data come from the March CPS, whereas degree data is from IPEDS. See the "Details of the Data Used" section of the appendix for more details and the "Extended Out-of-Sample Tests (until 2015)" section of the appendix for a longer-run view of the out-of-sample tests in later years.

Table 4.4 **Labor supply calibrated parameters**

		Calibrated value		
Parameter	Description	$\lambda = 1$	$\lambda = 2$	$\lambda = 4$
σ_0	Std. dev. of study-area taste shocks	0.0141	0.0215	0.0217
σ_1	Std. dev. of occupation taste shocks	0.9420	0.8887	0.9282
θ_0	Mean taste for not studying CS	−0.1341	−0.1072	−0.1362
θ_1	Mean taste for not working in CS	2.1766	1.8627	1.9278
η_1	Sector-switching cost (constant)	0.3265	0.4059	0.4145
η_2	Sector-switching cost (linear)	0.0307	0.0529	0.0488
η_3	Sector-switching cost (quadratic)	−0.0001	−0.0007	−0.0006

The employment series in figure 4.2, panel A, and the wage series shown in panel B, fit well at the start and end of the period, but it misses some years in between, particularly because it can't match the dip in wages that occur after 1994 and the simultaneous spike in employment in that same period. Last, the lagged degree-attainment series can be seen in figure 4.2, panel C, and matches the data relatively well.

In figure 4A.1 in the appendix we extend this exercise to later years, and study how well our calibrated parameters match the data in the first decade of the twenty-first century. We do a good job of matching wages and employment in this out-of-sample exercise, but overpredict enrollment in computer science for the years after 2004.

Table 4.4 presents the values of the calibrated parameters for the different values of λ. On average, we can see that there is a mean taste for not working in CS occupations, which is consistent with the wage differential seen across CS and non-CS work.

These calibrated parameters allow us to trace out the labor supply curve for computer scientists relative to non-CS college-educated workers. In order to do this, we use the model setup and the parameters and vary the relative wage to measure the response in relative quantities of labor. This derives the relative supply curve that we then use in the labor market to find the equilibrium wage.[35]

4.3.4 Endogenous Variables during the IT Boom

The calibration exercise so far helps us identify the parameters in the model that govern the trends in the endogenous variables over time. We can study these trends to understand how our model predicts what is happening at the time of the IT boom and the influx of foreign computer scientists. Given the solution of the model in each period, we study how prices and wages, employment by occupation and sector, and quantities produced change over time.

35. Our estimated relative labor supply elasticities lie between 1.96 and 2.48.

While US workers were more likely to work in CS occupations over time, the fraction of foreigners in CS work was increasing at a yet faster rate. Also consistent with the trends seen in the data for this period, the wage for computer scientists increases faster than the wages in other occupations. This IT boom overall leads to an increase in consumption of the IT good and a fall in prices of the IT good, which benefits consumers.

Figure 4.3, panel A, shows how the ratio of US computer scientists to non-CS college graduates (L_{US}/G) evolves over this period according to our model for the different values of λ. During the time of the boom, this ratio increases from about 0.040 to 0.047 for $\lambda = 2$, as more and more US workers shifted into CS work. At the same time, there was an increasing share of foreigners in CS occupations—the ratio of foreign-to-US computer scientists ($L_{Foreign}/L_{US}$) more than doubled from about 0.13 in 1994 to about 0.29 in 2001.

Our model predicts that over this period IT-sector employment grew faster than employment in the other sector, and most of this was driven by hiring in CS occupations (figure 4.3, panel B). The ratio of employment in IT to non-IT sectors over time $(L_C + G_C + H_C)/(L_Y + G_Y + H_Y)$ increases over this period, highlighting the importance of the IT boom in employing more workers. At the same time, with the influx of foreign computer scientists, the intensity of CS workers in the IT sector eventually increases. This can be seen in the series that plots the ratio of CS to non-CS workers in the IT sector $L_C/(G_C + H_C)$ in figure 4.3, panel B. The overall growth in the IT-sector employment, therefore, was skewed toward CS employment.

While employment for CS workers and the IT-sector workers as a whole was increasing over this period, we can also study how the relative wages for these types of workers change. Figure 4.3, panel C, plots the CS wage relative to the non-CS college-graduate wage (w/s) and relative to the non-college-graduate wage (w/r). Consistent with the data, the model predicts that wages for computer scientists increase at a faster rate than wages for the other types of workers.

The boom in the IT sector increased overall production and consumption for IT goods. Figure 4.3, panel D, shows how relative consumption (C/Y) increases and relative prices (P_c/P_y) fall over this period as the supply of IT goods from firms increases. The reduction in the price of IT goods will affect overall consumer utility as laid out by the model, and the following section will discuss how we calculate utility for the different types of workers and the owners of firms.

4.4 Counterfactuals

In order to isolate the impacts of high-skilled immigration on the various endogenous variables and on worker welfare, we conduct a counterfactual exercise. In the exercise we restrict the stock of immigrants to be constant

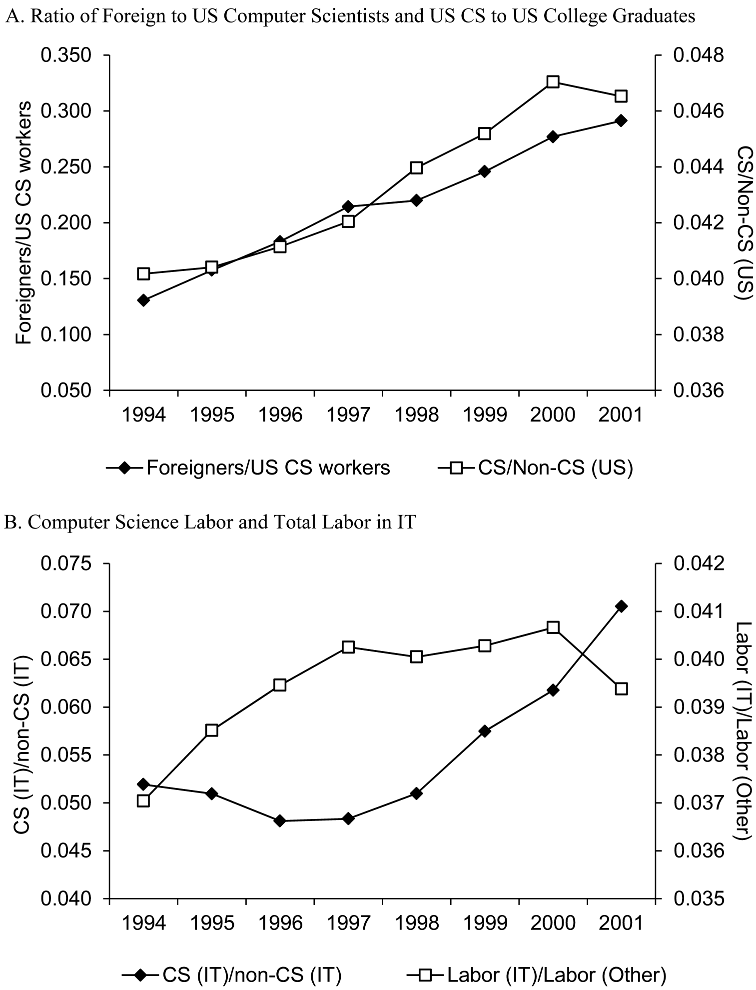


Fig. 4.3 Endogenous variables over time

Notes: Model predictions for ratio of endogenous variables over time:

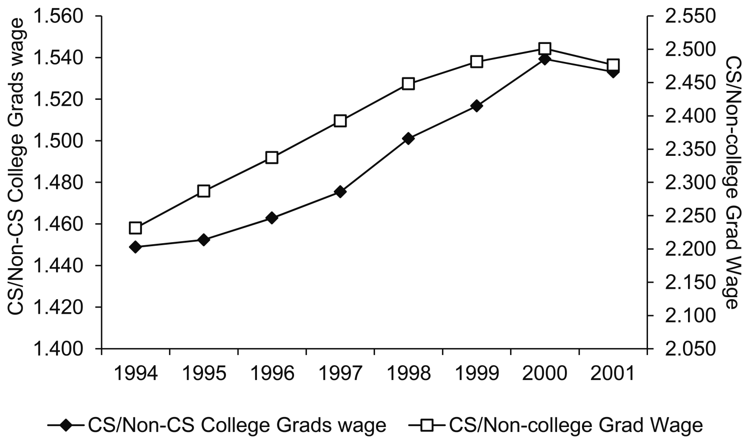
1. Foreign CS workers to US CS workers ($L_{\text{Foreign}}/L_{\text{US}}$), and for US CS workers to all US college-graduate workers ($L_{\text{US}}/(L_{\text{US}} + G)$).

2. CS labor to non-CS labor in the IT sector ($L_c/(G_c + H_c)$) and the total labor in IT relative to total labor in the other sector ($(L_c + G_c + H_c)/(L_y + G_y + H_y)$).

3. CS wage relative to non-CS college-graduate wage (w/s) and the CS wage relative to non-college-graduate wage (w/r).

4. Relative prices for the IT good (P_c/P_y) and relative consumption (C/Y).

C. Relative wages for CS workers



D. Relative Price and Quantity of the IT Good

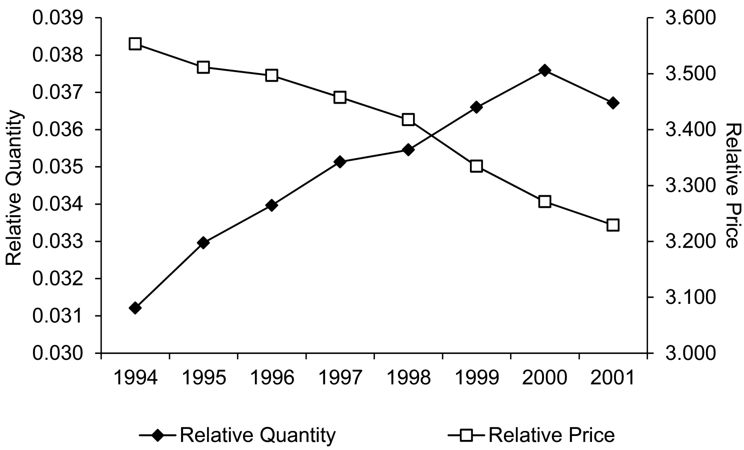


Fig. 4.3 (cont.)

at the 1994 level, and subject the economy to the same innovation shocks that were experienced during this period. Using the identified parameters, we can then trace out what happens to all the endogenous variables over this period in a situation where the stock of immigrants is fixed.

We use the notation "open" to refer to the real scenario under the H-1B regime, and "closed" to the counterfactual of restricted immigration. We can then define any endogenous variable, x_s, under the two scenarios $s = \{\text{open,closed}\}$. For example, $L^{\text{US}}_{\text{open}}$ is the number of US computer scientists in the "real" scenario under which high-skilled immigration is encouraged via the H-1B program, and all CS workers earn a wage w_{open}. In contrast,

L^{US}_{closed} and w_{closed} are the employment of US computer scientists and wages for all computer scientists in the counterfactual scenario where the stock of foreigners is restricted to its 1994 level.

4.4.1 Employment and College Degrees in Computer Science

Figure 4.4, panel A, describes the restriction under the counterfactual exercise. It shows how, under the real scenario where the economy is open to H-1B immigration, there is an increase in the stock of foreign computer scientists, whereas under the counterfactual scenario where the economy is closed, the stock of foreign computer scientists is restricted to the 1994 level.

How this restriction affects the stock of US computer scientists in our model can be seen in figure 4.4, panels B and C. Over this period there is an increase in the total number of computer scientists when we allow for immigration, but the number of US computer scientists actually decreases with respect to the closed economy every year as the number of immigrants increases. In 2001, the number of US computer scientists was between 6.1 percent and 10.8 percent lower under the open than in the closed economy (table 4.5). These numbers imply that for every 100 foreign CS workers that enter the United States, between thirty-three to sixty-one native CS workers are crowded out from computer science to other college-graduate occupations.

When the economy is open to immigration under the H-1B program, some US computer scientists switch over to non-CS occupations, shifting out the supply of these workers. This can be seen in figure 4.4, panel D. While over time there has been a rapid increase in the number of non-CS college-educated workers, this increase would have been lower if the number of foreign CS workers were restricted. In fact, the growth rate between the open and closed economies plotted in figure 4.4, panel D, mirrors the decrease in panel C as US workers switch from CS to non-CS occupations.

Since students in our model choose their college major in their junior year, a change in the wages for computer scientists will affect these choices. Under the open economy scenario, the fraction of CS degrees in 2001 would be between 1.3 and 2.6 percentage points lower than in the closed economy, as can be seen in figure 4.4, panel E.

4.4.2 Wages

Over the period of study, wages grew for CS workers, but this growth would have been higher if immigration was restricted (figure 4.5, panel B). An influx of foreign CS workers depresses the CS wage, and shifts some US workers into non-CS occupations. At the end of the decade, our model implies wages for CS workers would have been between 2.6 percent and 5.1 percent lower under the open economy (table 4.5).

With an increase in the foreign CS workforce, college-educated US CS workers shift into non-CS occupations, and this tends to lower the non-CS wage. At the same time, however, as the equilibrium amount of total

CS workers increases, so does the marginal product of non-CS college-educated workers. This increases the demand for non-CS workers, and tends to increase their wage, making the net effect positive (figure 4.5, panel C). Overall, table 4.5 shows an increase in the non-CS wage due to immigration of about 0.04 percent–0.28 percent in 2001. As expected, both the changes in CS wage and the non-CS wage for college graduates are sensitive to what value of λ we choose, but qualitatively our results do not change across specifications.

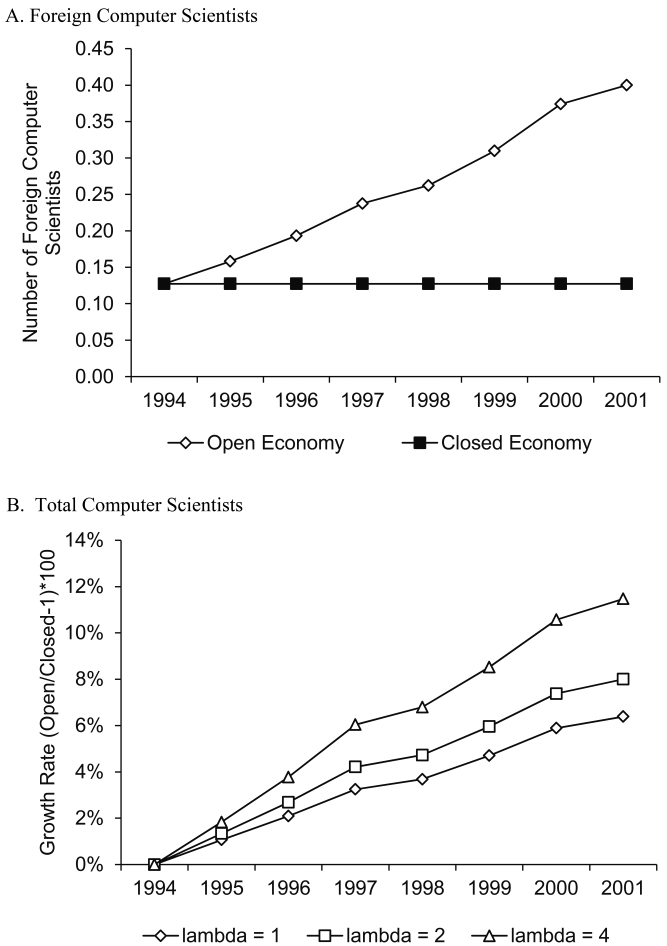


Fig. 4.4 Employment under the real and counterfactual scenarios

Note: The closed economy is where immigration is restricted to the 1994 levels, whereas in the open economy the stock of immigrants grows according to the data. Total size of the workforce is normalized to 100 in 1994.

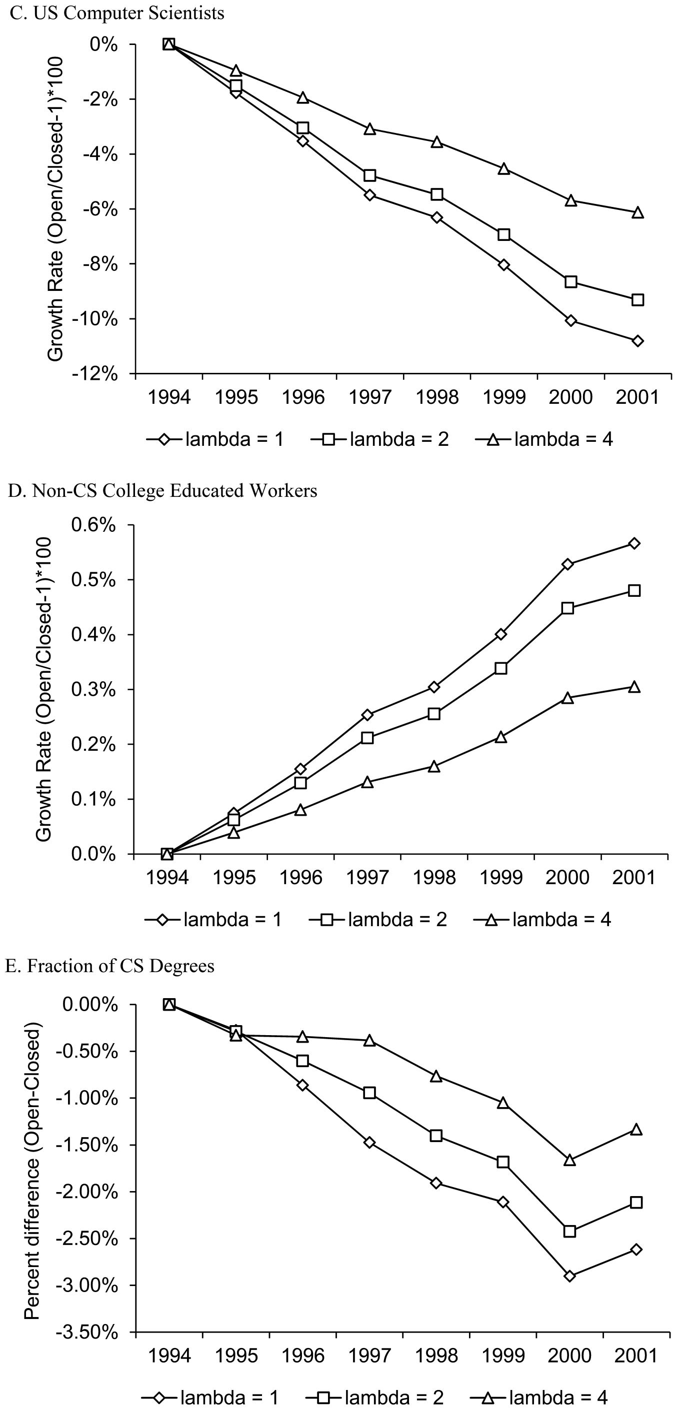

C. US Computer Scientists
Growth Rate (Open/Closed-1)*100
0%
-2%
-4%
-6%
-8%
-10%
-12%
1994 1995 1996 1997 1998 1999 2000 2001
lambda = 1
lambda = 2
lambda = 4
D. Non-CS College Educated Workers
Growth Rate (Open/Closed-1)*100
0.6%
0.5%
0.4%
0.3%
0.2%
0.1%
0.0%
1994 1995 1996 1997 1998 1999 2000 2001
lambda = 1
lambda = 2
lambda = 4
E. Fraction of CS Degrees
Percent difference (Open-Closed)
0.00%
-0.50%
-1.00%
-1.50%
-2.00%
-2.50%
-3.00%
-3.50%
1994 1995 1996 1997 1998 1999 2000 2001
lambda = 1
lambda = 2
lambda = 4

Fig. 4.4 (cont.)

Table 4.5 **Percent changes when allowing immigration (2001)**

	λ = 1	λ = 2	λ = 4
Relative price	−1.86	−1.85	−2.42
Relative quantity	1.89	1.89	2.48
Number of firms	0.50	0.51	0.56
Wage computer scientists	−5.13	−3.47	−2.57
Wage college graduates, non-CS	0.28	0.10	0.04
Wage non–college graduates	0.43	0.44	0.52
Total employment in CS	6.39	8.00	11.47
US computer scientists	−10.81	−9.32	−6.12
College graduates, non-CS	0.57	0.48	0.30

Notes: Percent changes are calculated using the endogenous variables from the closed and open economy. For each year we consider the situation of going from a closed to an open economy (allowing immigration), that is $[(X_{\text{open}}/X_{\text{closed}}) - 1] \times 100$. Results shown for different values of λ and only look at year 2001.

Since the labor supply of non–college graduates is assumed to be fixed and inelastic, only changes in the demand for non–college graduates determine the difference in their wages under the real and counterfactual scenarios. When the economy is open to immigration, the equilibrium number of total college graduates employed increases due to immigration. This raises the marginal product of non-college-graduate labor, and shifts out the demand for non-college-graduate workers, raising the overall wage for non–college graduates (figure 4.5, panel D). Under the open economy, wages for non–college graduates would have been between 0.43 percent and 0.52 percent higher by the end of this period (table 4.5).[36]

4.4.3 Prices, Output, and the Entry of Firms

While high-skilled immigration affected both employment and wages, it also affects overall output and prices of the different goods produced in the economy. These changes will affect overall consumer welfare, and also the profits accruing to firm owners.

Over the period of study, relative prices of IT goods were falling steadily, and some of this fall can be attributed to the increase in CS employment due to immigration. Figure 4.6, panel A, and table 4.5 show how under the open economy, prices would have been between 1.9 percent and 2.4 percent lower in 2001.

At the same time, the relative consumption of IT goods was increasing, and this increase would have been lower without the growth in the foreign workforce (figure 4.6, panel B). Immigration also raises the profits of firms who can now hire relatively cheaper labor, and this causes new firms to enter

36. Since the non-college-graduate workforce is a lot larger than the CS workforce, the relative shift in wages is a lot lower compared to the CS wage.

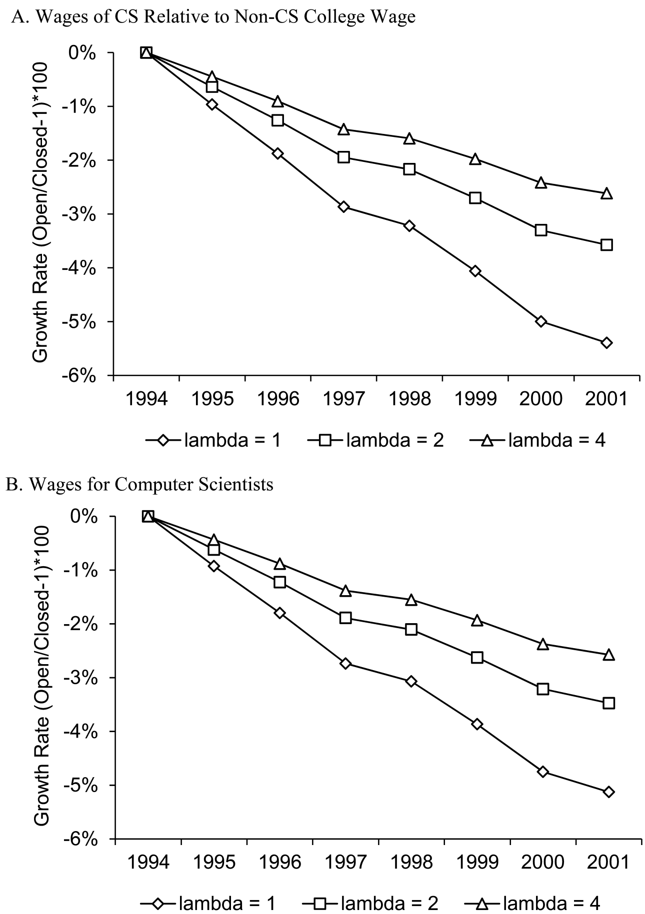


Fig. 4.5 Wages under the real and counterfactual scenarios

Notes: The closed economy is where immigration is restricted to the 1994 levels, whereas in the open economy, the stock of immigrants grows according to the data. All monetary values are in units of the numeraire (the consumption bundle).

the IT sector. Figure 4.6, panel C, shows how by allowing immigration, the number of IT firms would be higher. At the end of this period, there would be between 0.50 percent and 0.56 percent fewer IT firms if immigration was restricted (table 4.5).

4.5 Welfare

Using our estimated parameters and counterfactual exercises, we can measure the overall economic impacts on the different agents in the economy due to the increase in the number of foreign computer scientists. In order to

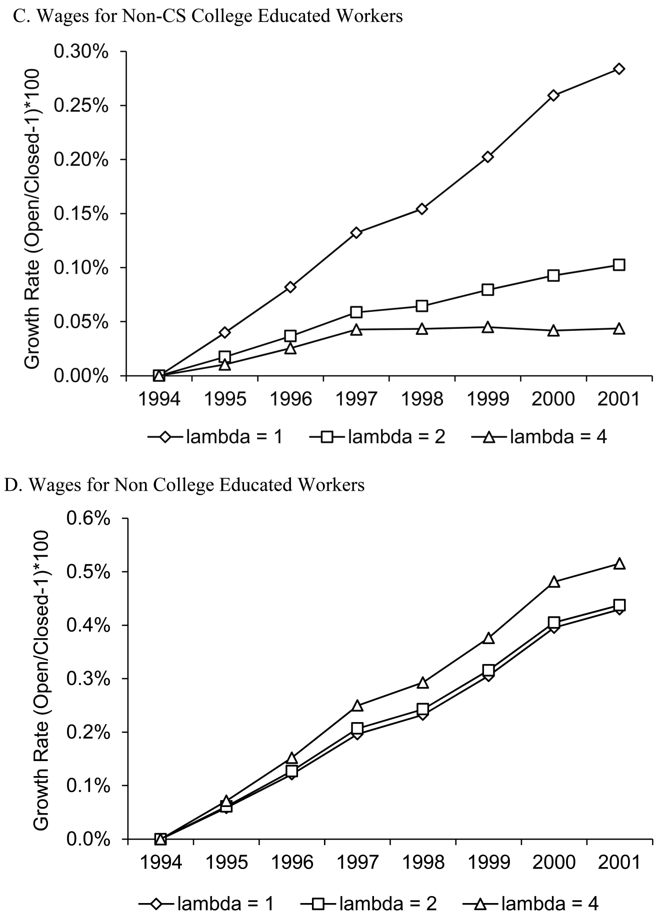


Fig. 4.5 (cont.)

compare losses and benefits and the distributional consequences of immigration, we look at the welfare of all types of workers and the owners of firms.

4.5.1 Calculating Welfare

Calculating Worker Welfare

Given the structure of our CES utility function, we can calculate consumer welfare as a function of the income of each type of agent. For a given income level m_i, the indirect utility of the agent is just the product of his income and the ideal price index. However, since the ideal price index is the numeraire, indirect utility is just the income of each type of worker: $V_i(m_i) = m_i$. We then compare the welfare of individuals under the two

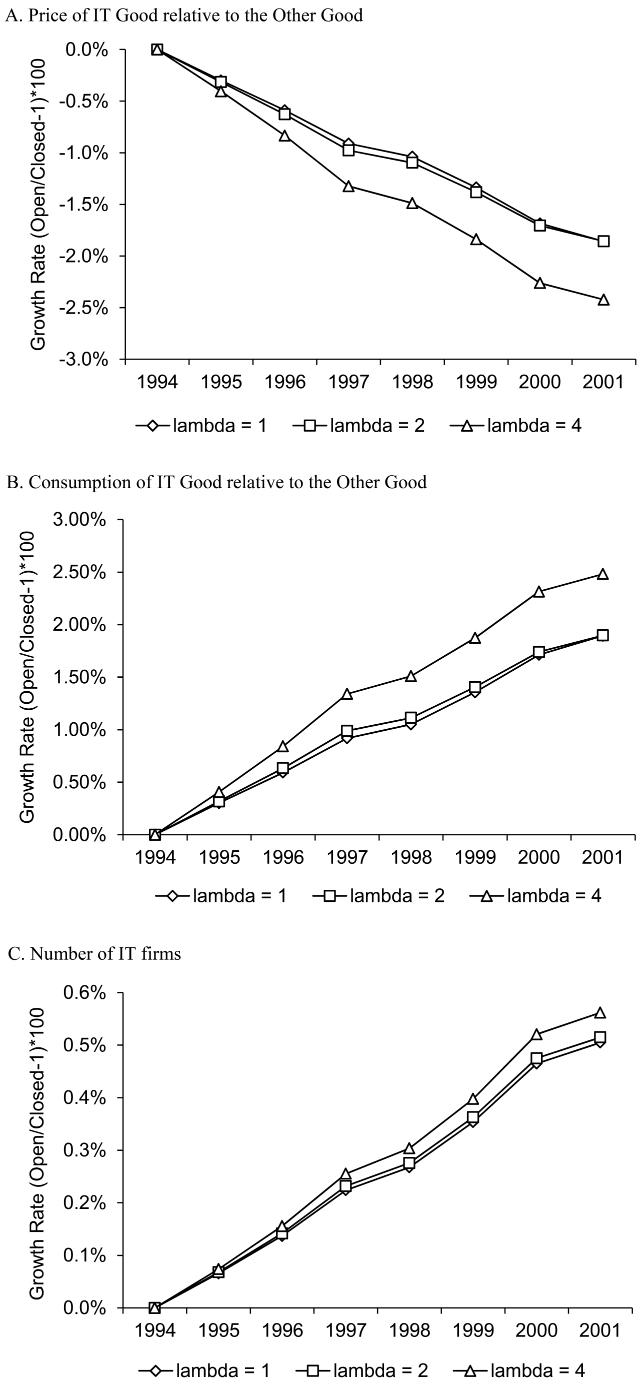


Fig. 4.6 Output and prices under the real and counterfactual scenarios

Notes: The closed economy is where immigration is restricted to the 1994 levels, whereas in the open economy, the stock of immigrants grows according to the data. Prices are in units of the numeraire (the consumption bundle).

scenarios: (a) the real scenario, where high-skilled immigration is encouraged under the H-1B program; and (b), the counterfactual scenario where the stock of immigrants is restricted to the 1994 level. For all welfare calculations we will only be focusing on welfare changes for those individuals who are US born, ignoring the changes in welfare for migrant computer scientists.

Workers are divided into four groups: those who are computer scientists and stay in CS occupations in the presence of immigration, those who are CS workers but switch to non-CS work because of immigration, those who were non-CS college graduates even before there was immigration, and those who are non–college graduates. We then proceed to calculate welfare changes in two different ways: percent utility changes and compensating variation.

Our model shows that when there is an influx of foreign computer scientists, the equilibrium wage for CS workers falls and pushes some native college-educated computer scientists into non-CS work. As the equilibrium number of hired computer scientists increases, the marginal product, and hence the demand for other types of workers, will also increase, tending to push up their wages. The wage for non-college-educated workers and non-CS college-educated workers unambiguously rises for all specifications of λ.

For those that stay in their occupation groups under both real and counterfactual scenarios, we can calculate the percent utility changes by just looking at the percent change in the wage for each group (e.g., the percent change in utility for the computer scientists that stay in CS occupations under the presence of immigration is just the percent change in w between the open and closed economy). For computer scientists that switch to non-CS occupations when we allow for immigration, we use information from both the utility change for the CS workers that stay and the change for those that were always non-CS college graduates.

By knowing the form of the indirect utility function, we can also calculate how much income we must compensate different types of workers who lose from immigration. This compensating variation (CV) depends on the indirect utility calculated at the original prices P_c and original income levels m_i, and compares it to a scenario with new prices and income $(P_c' m_c')$. A useful feature of the compensating variation is that we can scale up the results using total labor income in the US economy from the data to measure how much workers should be compensated (in USD) if immigration restrictions were imposed. Given that the ideal price index is our numeraire, we can write the compensating variation as $CV = m_i - m_i'$.

The number of computer scientists who stay in CS occupations even in the presence of immigration is L_{open}. Their overall change in income in the presence of increased immigration is therefore given by $(w_{\text{closed}} - w_{\text{open}})L_{\text{open}}$. When there is immigration, non-college-graduate workers benefit from the rise in wages that is caused by the increase in their marginal product. The increase in income for this group is therefore $r_{\text{open}} - r_{\text{closed}}\bar{H}$.

Similarly, the number of non-CS college-educated workers who were always in these other occupations is given by G_{closed}. Their overall change in income is given by $(s_{\text{closed}} - s_{\text{open}})G_{\text{closed}}$. Given that we find the wages for non-CS college-educated workers to be lower in the presence of immigration, there is a loss in income to these workers due to immigration.

Last, for the group of workers who switch from CS to non-CS work in the presence of immigration, we must take into account their switching costs and change in utility because of different tastes in each occupation. The marginal worker who switches experiences a different loss in utility than the inframarginal worker. The overall change in terms of income equivalent for this group of workers can be approximated by $(1/2)(L_{\text{closed}} - L_{\text{open}})[(s_{\text{closed}} - s_{\text{open}}) + (w_{\text{closed}} - w_{\text{open}})]$.[37]

Calculating Profits

In our model, firms in the perfectly competitive residual sector earn no profits. In the monopolistically competitive IT sector, however, only the marginal firm earns 0 profits. In the current setup we follow Chaney (2008), where there is an underlying mass of firms that already know their entrepreneurial capabilities and choose whether to produce or not given their productivity. There is, therefore, free entry into the production decision that drives the profit for the marginal producing firm down to zero.

For the firms in the IT sector, the marginal producing firm has a productivity ϕ^*, and a profit $(\phi^*) = 0$. Using the notation highlighted in section 4.2.1, we know that the average profit for producing firms can be represented by

$$\begin{aligned}(28)\qquad \int_{\phi^*}^{\infty} \pi(\phi)\mu(\phi)\,d\phi = {} & \int_{\phi^*}^{\infty} PC^{1/\varepsilon}c_j^{(\varepsilon-1)/\varepsilon}\mu(\phi)\,d\phi - w\int_{\phi^*}^{\infty} l_j\mu(\phi)\,d\phi \\ & - s\int_{\phi^*}^{\infty} g_j\mu(\phi)\,d\phi - r\int_{\phi^*}^{\infty} h_j\mu(\phi)\,d\phi - f.\end{aligned}$$

The total profits are then the average profits times the number of firms $N = (1 - \Psi(\phi > \phi^*))N^e$, where N^e is the number of total potential producers in the sector.

We can also calculate profits for different types of firms using the features of this distribution. For example, we know that the cutoff productivity will change across the regimes where there is immigration and there is not. In the presence of immigration, firm profits will rise and allow newer firms to enter on the margin.[38] This then allows us to calculate the profits for the new entrants and the incumbent firms separately. Let ϕ^*_{open} and ϕ^*_{closed} be the cutoff values of productivity under each regime. The new firms that enter

37. The intuition for this expression is the following: a CS worker who switches experiences a change in welfare that equals the change in CS wage up to the relative wage that induces them to switch. From that point on, the additional change in welfare will equal the change in the wages of non-CS college graduates. We assume that for minor changes in wages the demand curve can be approximated linearly.

38. Alternatively, in the Melitz (2003) framework of the model, firms will enter at any point of the distribution.

when there is immigration will have a productivity $\phi_j \in [\phi^*_{\text{open}}, \phi^*_{\text{closed}}]$. Whereas the incumbents have a productivity $\phi_j \in [\phi^*_{\text{closed}}, \infty)$. These cutoffs, therefore, change the limits of integration and the conditional distribution functions.

The marginal distribution for the incumbents is determined by

$$\mu_{\text{closed}}(\phi) = \begin{cases} \dfrac{k\phi^{-(k+1)}}{1 - \Psi\left(\phi^*_{\text{closed}}\right)}, & \text{if } \phi \geq \phi^* \\ 0, & \text{otherwise} \end{cases}$$

with $\Psi(\phi^*_{\text{closed}}) = 1 - (1/\phi^*_{\text{closed}})^k$.

The total profits to incumbents is then these average profits times the number of incumbents: $N_{\text{incumbent}} = (1 - \Psi(\phi_{\text{closed}}))N^e$. The total profits for new entrants is simply the difference between the profits for incumbents and the total profits for all firms in the open economy scenario.

Such an exercise can also be done to derive the profits for the firm in any percentile. For example, the firm in the 90th percentile has a productivity $\phi_{90} = (1/0.1)^{1/k}$. Since the number of firms above the 90th percentile is simply $N_{90} = 0.1N_e$, we can derive the profits for these firms in the scenario with and without immigration.

4.5.2 Welfare Changes Due to Immigration

The changes to the welfare of workers in this economy depends on the changes in income and the prices due to immigration. Figure 4.7, panels A, C, and E, show how much workers, under a regime of restricted immigration, need to be compensated to maintain the same level of utility as they had in the open economy. These numbers have been translated into 1999 USD. Overall worker welfare is higher under immigration, and the amount of the compensating variation rises steadily between 1994 and 2001. The compensating variation for all workers in 2001 is between \$8.2 and \$10.9 billion, depending on the value of λ.

This overall increase in utility due to immigration, however, hides a lot of distributional changes. Figure 4.7, panels A, C, and E, split up the workers into four groups: (a) those who stay in CS occupations even after immigration, (b) those who switch from CS to non-CS, (c) college graduates who were always non-CS, and (d) non–college graduates. As these panels show, US computer scientists are negatively affected by immigration, while other workers gain. The positive effect for college graduates gets partly offset by the mobility of the college educated across occupations, where computer scientists switching to non-CS occupations depress the wage. The losses for computer scientists and the gains for non-CS college graduates get closer to zero when the ease of substitution between CS and non-CS college graduates gets higher. On the other hand, the compensating variation for non–college

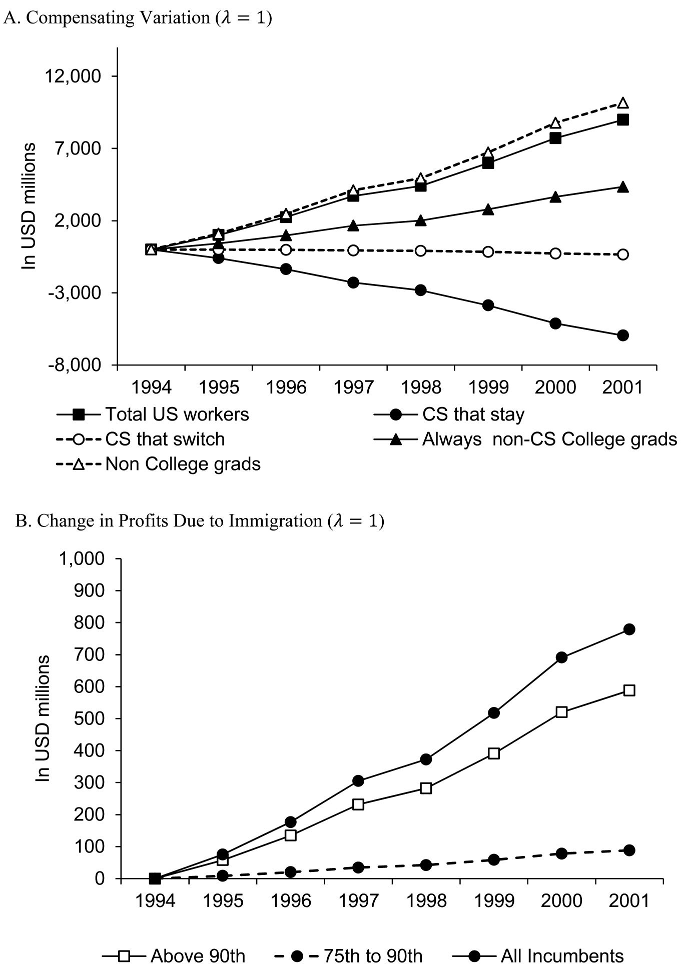
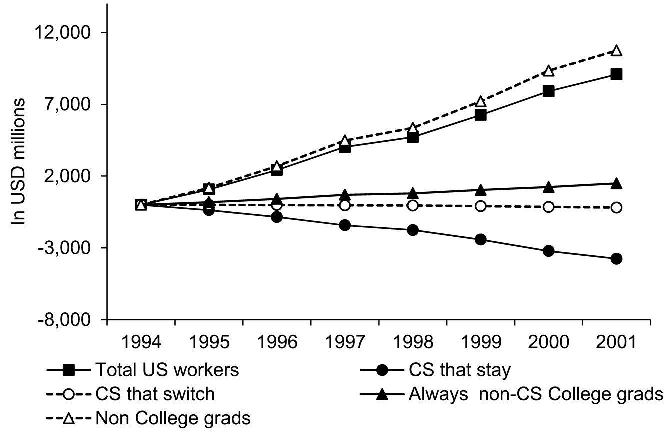


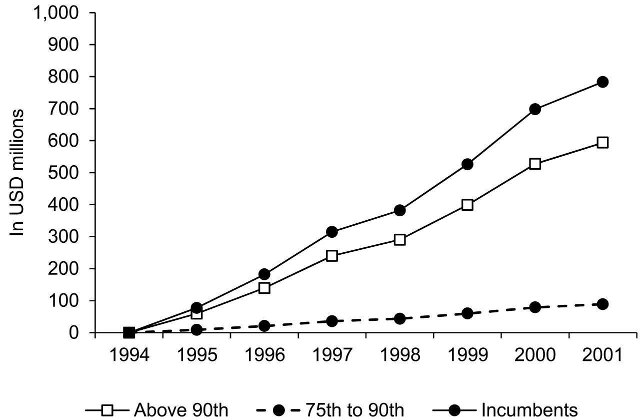
Fig. 4.7 Welfare changes due to immigration

Notes: The closed economy is where immigration is restricted to the 1994 levels, whereas in the open economy, the stock of immigrants grows according to the data. Compensating variation in this scenario is how much the workers must be compensated if immigration is restricted to the 1994 level. Compensating variation and profits are in millions of 1999 USD. The scaling up to USD was done using CPS data for the total amount of labor income across each year separately. Panels A, C, and E split up the workers into four groups—(a) those who stay in CS occupations even after immigration, (b) those who switch from CS to non-CS, (c) college graduates who were always non-CS, and (d) non–college graduates. Panels B, D, and F split up the firms into three different categories—(a) "all incumbents" are only the firms that still produce when immigration is restricted. Among these incumbents, the (b) "above 90th percentile" firms are those that have a productivity level that is above the 90th percentile in the productivity distribution, and similarly (c) "75th to 90th percentile" firms have a productivity level that lies between the 75th and 90th percentiles of the Pareto productivity distribution.

C. Compensating Variation ($\lambda = 2$)



D. Change in Profits Due to Immigration ($\lambda = 2$)



E. Compensating Variation ($\lambda = 4$)

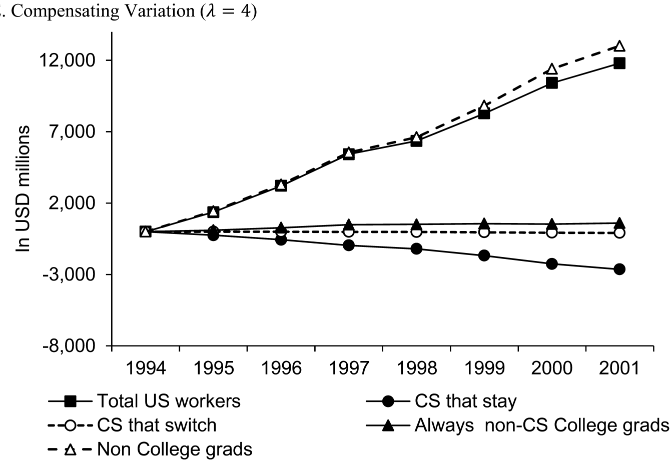


Fig. 4.7 (cont.)

F. Change in Profits Due to Immigration ($\lambda = 4$)

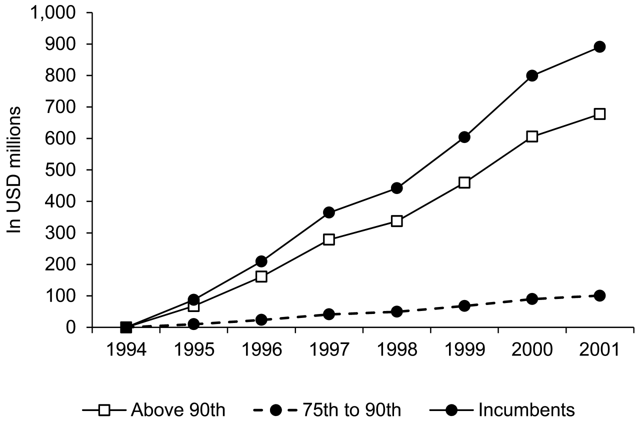


Fig. 4.7 (cont.)

Table 4.6 **Percent change in utility when allowing for immigration and compensating variation**

	Percent change in utility			Compensating variation (million USD)		
	λ = 1	λ = 2	λ = 4	λ = 1	λ = 2	λ = 4
All US workers	0.20	0.21	0.27	8,204	8,290	10,904
All college graduates	−0.12	−0.16	−0.14	−1,955	−2,453	−2,110
Computer scientists that stay	−5.13	−3.47	−2.57	−5,951	−3,752	−2,631
Computer scientists that switch	−2.48	−1.71	−1.27	−348	−189	−85
Non-CS college graduates that stay	0.28	0.10	0.04	4,344	1,488	606
Non–college graduates	0.43	0.44	0.52	10,159	10,743	13,014

Note: We compare utility changes when going from a closed to an open economy, so percent changes are calculated for each year and subgroup as $[(V_{open}/V_{closed}) - 1] \times 100$, where V is indirect utility for that specific group. Compensating variation figures are expressed in million USD.

graduates increases when λ increases. Table 4.6 summarizes the utility percent changes from allowing immigration and compensating variation for 2001, corroborating the idea that there are significant distributional effects from increased immigration.

While workers as a whole benefit from more immigration, firms make higher profits, too. In figure 4.7, panels B, D, and F, the firms are split up into three different categories: (a) "all incumbents" are only the firms that still produce when immigration is restricted; among these incumbents, the (b) "above 90th percentile" firms are those that have a productivity level that is above the 90th percentile in the productivity distribution; and similarly, (c)

Table 4.7 **Percent change in profits when allowing for immigration (2001)**

	$\lambda = 1$		$\lambda = 2$		$\lambda = 4$	
	Share of profits	Percent change	Share of profits	Percent change	Share of profits	Percent change
All firms	—	0.61	—	0.62	—	0.70
All incumbent firms	100	0.61	100	0.62	100	0.70
90th–100th percentile	84.82	0.54	85.09	0.55	85.05	0.62
75th–90th percentile	9.64	0.71	9.59	0.73	9.60	0.82
< 75th	5.54	1.45	5.32	1.51	5.36	1.66

Notes: Columns titled "share of profits" show the share of profits among all incumbents by firm size for 2001 in the open economy. We compare profit changes when going from a closed to an open economy, so percent changes in aggregate profits for each year and subgroup are calculated as $[(\Pi_{open}/\Pi_{closed}) - 1] \times 100$. Percentiles are defined using the Pareto distribution we are assuming for productivities in the market. Rows 2–5 only consider incumbent firms (those that operate under the open and closed economy); row 1 shows the growth rate between open and closed taking into account the marginal firms that start producing in the open economy. Results shown for different values of λ and only look at year 2001.

"75th to 90th percentile" firms have a productivity level that lies between the 75th and 90th percentiles. Profits for all firms are increasing over this period, and most of the profits are captured by the firms in the top 10 percent of the productivity distribution. While we believe there is considerable heterogeneity in the profits firms receive as a result of the H-1B program, it is important to note that the distribution of profits in the model is determined by our assumption on the Pareto distribution of firm productivities. In 2001, the aggregate profits in the IT sector were between \$0.78 and \$0.89 billion (1999 USD), and between \$0.59 and \$0.68 billion went to the firms that had a productivity level above the 90th percentile. Table 4.7 summarizes the changes in profits for the different values of λ; overall profits increase between 0.61 percent and 0.70 percent in 2001 when allowing for immigration.

4.5.3 Alternative Modeling Specifications

We analyze how two particular features of the IT sector in our model affect our results. The first is our assumption of monopolistic competition and the existence of different varieties in IT products. This makes the IT sector smaller than the perfectly competitive optimal size. An increase in the number of immigrants, and therefore workers, will expand this sector and lead to welfare gains. At the same time, as more firms enter, the increase in varieties benefits consumers as well. The second nonstandard feature of our model is the presence of technological spillovers driven by innovation by computer scientists. An increase in the CS workforce due to immigration leads to more innovation and has an additional impact on overall production, lowering prices and increasing welfare for consumers.

In table 4.8 we compare the monopolistically competitive model with a traditional perfectly competitive setup, and also shut down the presence of technological spillovers to study how our results change. In moving from a perfectly competitive to a monopolistically competitive model, the welfare changes due to immigration are roughly similar. There is a slightly larger welfare gain due to immigration in the monopolistically competitive model both in the absence or the presence of technological spillovers. Shutting down the possibility of technological spillovers, however, has a larger impact on the gains from immigration. In the absence of spillovers, $\beta = 0$, the overall gains to worker utility is only between 0.02 percent and 0.03 percent, whereas the spillovers $\beta = 0.23$ increase these gains to about 0.21 percent. How the results change with other values of β is discussed in the "Sensitivity Analysis" section of the appendix. Therefore, while the monopolistic-competition assumption does not affect worker welfare much, the presence of technological spillovers does.

One advantage of the monopolistically competitive setup is that it allows us to get a measure of how firm profits are affected by immigration. The profit numbers should be interpreted with caution, however, since our framework implies that profits are simply a fixed proportion of total revenues. Nonetheless, given that IT firms spend a substantial amount of funds in lobbying Congress to raise the H-1B cap, it is reasonable to believe that firms stand to benefit from an influx of high-skilled immigrants.

Importantly, our model includes the labor supply decisions of college-educated US workers. This allows students and workers to move out of immigrant-intensive fields and occupations when there is an influx of high-skilled workers from abroad. The negative effects on CS workers are mitigated as US CS workers switch to non-CS jobs, and fewer students graduate with CS degrees. However, since CS workers are also innovators, the

Table 4.8 **Percent change in utility—perfect competition versus monopolistic competition in the IT sector**

	Perfectly competitive		Monopolistic competition	
	$\beta = 0$	$\beta = 0.23$	$\beta = 0$	$\beta = 0.23$
All US workers	0.02	0.20	0.03	0.21
All college graduates	−0.34	−0.16	−0.34	−0.16
Computer scientists that stay	−3.76	−3.58	−3.64	−3.47
Computer scientists that switch	−1.90	−1.76	−1.85	−1.71
Non-CS college graduates that stay	−0.08	0.10	−0.08	0.10
Non–college graduates	0.25	0.43	0.26	0.44

Note: We compare utility changes when going from a closed to an open economy, so percent changes are calculated for each year and subgroup as $[(V_{open}/V_{closed}) - 1] \times 100$, where V is indirect utility for that specific group. In the perfectly competitive cases, we assume that the IT sector is no longer under monopolistic competition. The $\beta = 0$ refers to the case where there is $\lambda = 0.23$.

Table 4.9 **Changes in profits and income for different labor supply specifications**

	Percent change in income/profits		Compensating variation/change in profits (million USD)	
	Baseline	Inelastic supply	Baseline	Inelastic supply
All US workers	0.21	0.46	8,290	17,798
All college graduates	−0.16	0.01	−2,453	225
Computer scientists that stay	−3.47	−7.51	−3,752	−8,467
Computer scientists that switch	−1.71	—	−189	—
Non-CS college graduates that stay	0.10	0.60	1,488	8,692
Non–college graduates	0.44	0.72	10,743	17,572
Profits	0.61	0.94	783	1,197

Note: The baseline case is when we apply our full labor supply model for college graduates. The inelastic case shows what happens when workers are not allowed to change occupations or degree-choice decisions. All specifications use a value of $\lambda = 2$, $\sigma = 1$, and $\beta = 0.23$. Dollar values for compensating variation and profits are in millions of 1999 USD. The scaling up to USD was done using CPS data for the total amount of labor income. Changes in income for different worker groups and profits are calculated as $[(X_{open}/X_{closed}) - 1] \times 100$.

economy as a whole no longer benefits as much from technological improvements when US workers leave CS occupations. In table 4.9, we compare our baseline model that allows for labor supply decisions to an inelastic supply model where US students and workers are no longer allowed to change their decisions in the face of high-skilled immigration. Immigration has even more negative impact on US CS workers when we restrict adjustments on the labor supply side. Since workers can no longer switch into non-CS occupations, the increase in labor supply from abroad depresses CS wages and hurts CS workers the most. On the other hand, welfare in the economy as a whole increases since there are more computer scientists and hence more innovators.

4.6 Discussion

Isolating the impacts of high-skilled immigration is challenging in the absence of credible instruments that exogenously vary the share of foreign workers. Nonetheless, given the rapidly increasing share of immigrants in the skilled labor force, it is an important issue to examine. We develop a general-equilibrium model of the US economy, calibrated using data from 1994 to 2001, to estimate how the increasing share of foreign high-skilled workers affects the welfare of different types of workers, firms, and consumers. We do so by examining the welfare of US natives under a counterfactual scenario where we restrict the fraction of immigrants to their 1994 levels.

While our conclusions depend on the specifics of our model, we believe them to be reasonable. As long as the supply curve of US workers is not infinitely elastic, and we believe that evidence indicates rather conclusively that it is not, the availability of high-skilled foreign immigrants will shift out the supply of high-skilled workers in the US economy. However, as long as the demand curve for high-skilled workers is downward sloping, the influx of foreign high-skilled workers will both crowd out and lower the wages of US high-skilled workers. As a result, output in the high-skill-intensive sector of the economy will rise, but will rise less than if the crowd-out effects were negligible. The fact that high-skilled workers contribute to innovation tends to mute such crowd-out effects, but our results suggest such effects are not nearly large enough to fully compensate for the crowd-out.

Overall, our results suggest that high-skilled foreign workers contribute to the well-being of the typical US consumer, mainly through the assumption that these workers contribute to innovation at the same rate as US high-skilled workers. Indeed, under our calibrations, accounting for foreign workers' effect on innovation, the gains to consumers are an order of magnitude larger than gains excluding this effect. At some level, this is hardly surprising. While simple models of the impact of immigration on native welfare suggests the immigrant surplus is second order (Borjas 1999), if the immigrants shift out the production possibility frontier, their effect will be first order.

In our model, immigration also raises profits in the IT sector. While the magnitude of these gains depends on the markup in the IT sector, as long as there is a markup, which we consider safe to assume, high-skilled immigrant labor raises IT sector profits. It is then no surprise that Bill Gates and other IT executives lobby in favor of increasing quotas for high-skilled immigrants.

Although our results suggest that the introduction and expansion of the H-1B program in the 1990s brought gains to both US consumers and IT-sector entrepreneurs, we also found indications of losses for US computer scientists and potential computer scientists. Recent work (Peri and Sparber 2009, 2011) has emphasized the importance of immigration affecting the occupational choice of US natives. Our results tend to support the importance of this view. Indeed, our estimates suggest that high-skilled immigration has had a significant effect on the choices made by US workers and students.

Researchers (e.g., Peri and Sparber 2011) have emphasized that high-skilled immigrants have the potential for opening up opportunities for US workers—someone who might otherwise have been an engineer or computer scientist now becomes a manager. We have no doubt that this is true and, in a primitive way, we have built this into our model. The influx of skilled immigrants induces some college graduates to leave computer science and raises the productivity of non-CS college graduates. Still, for many college graduates who entered or might have entered the CS field, their options have been curtailed.

Our model is far too simple to allow for policy evaluations of alternatives to our current system of high-skilled immigration. However, we note that our model (and simple economic reasoning) suggests that high-skilled immigration does tend to crowd out US workers to some extent. We suspect that allowing essentially unlimited immigration of high-skilled workers by, for instance, awarding green cards to all foreign students attending US colleges and universities would have dramatic effects on the US labor market. Not all of these would be positive.

In the end we want to emphasize the limitations of our work. While our focus is on how the influx of foreign workers affects the United States, we recognize that US policy on high-skilled immigration has profound effects on both labor-sending countries and other countries that produce in the IT sector. Also, our analysis is constrained to the 1990s, whereas in the long run, US immigration policy is likely to affect the position of the United States in the world economy. We leave exploration of these issues to future research.

Appendix

Additional Model Details

Consumer Demand for Goods

Given the consumer utility functions described in the "Household Problem" section, it is possible to write the price index P in the form of equation (A.1):

(A.1) $$P_c = \left(\int_{v\in\Omega} p_i^{1-\varepsilon}\, dv\right)^{1/1-\varepsilon}.$$

Consumers maximize utility in equation (1) subject to a budget constraint $m = P_c C_d + P_Y Y_d$, where m is total income. The utility-maximizing first-order condition for a given variety is therefore

(A.2) $$\left(\frac{c_{di}}{C_d}\right)^{-(1/\varepsilon)} = \frac{p_i}{P}.$$

We can then write the demand for aggregate goods as a function of prices, total income m and the parameters γ and σ.

(A.3) $$C_d = \frac{m}{P_c + P_Y\{[(1-\gamma)/\gamma](P_c/P_Y)\}^{\sigma}}$$

(A.4) $$Y_d = \frac{m\{[(1-\gamma)/\gamma](P_c/P_Y)\}^{\sigma}}{P_c + \{[(1-\gamma)/\gamma](P_c/P_Y)\}^{\sigma}}.$$

Labor Supply Derivations

In order to determine the labor supply of US-born workers, we use the setup described in section 4.2.3. First, we study the probability of students enrolling in CS degrees. Given the distributional assumptions and equation (15), it follows that the probability (q_t^{cs}) that a student graduates with a CS degree can be written in logistic form:

$$q_t^{cs} = \left[1 + \exp\left(-\left(\rho^2 \mathbb{E}_{t-2}[V_{22}^{cs} - V_{22}^{o}] - \theta_o\right) / \sigma_0\right)\right]^{-1}. \tag{A.5}$$

This setup allows us to map the graduating probability described above to employment. Let $\left(L_t^a + G_t^a\right)$ be the number of college graduates with age a in time period t, then the number of graduates with a CS degree in year t is represented by $R_t = q_t^{cs}\left(L_t^{22} + G_t^{22}\right)$.

Next, we derive the occupational choice decisions based on the setup in the "Occupational Choice" section. Defining $q_{t,a}^{dD}$ as the probability that a worker at age a between twenty-two and sixty-four moves from occupation d to occupation D, it follows from the distributional assumptions that the probability of workers switching from computer science to other occupations, and vice versa, can be represented as

$$q_{t,a}^{o,cs} = \left[1 + \exp\left(-\left(w_t - s_t - \zeta(a) - \theta_1 + \rho \mathbb{E}_t\left[V_{t+1,a+1}^{cs} - V_{t+1,a+1}^{o}\right]\right) / \sigma_1\right)\right]^{-1} \tag{A.6}$$

$$q_{t,a}^{cs,o} = \left[1 + \exp\left(-\left(s_t - w_t - \zeta(a) + \theta_1 + \rho \mathbb{E}_t\left[V_{t+1,a+1}^{o} - V_{t+1,a+1}^{cs}\right]\right) / \sigma_1\right)\right]^{-1}. \tag{A.7}$$

Here we can see that the switching probabilities depend upon both the current wage differential and expected future career prospects in each occupation. The standard deviation of the taste shocks, the sector-attractiveness parameter, and the cost of switching occupations will affect the sensitivity of occupational switching to changes in relative career prospects.

Since individuals are forward looking, the working decisions depend upon the equilibrium distribution of their career prospects. Under the extreme value errors assumption, we can use the properties of the idiosyncratic taste shocks distribution to derive the expected values of career prospects (Rust 1987). The expected value functions for an individual at age a between twenty-two and sixty-four working as a computer scientist or in another occupation are respectively

$$\begin{aligned} \mathbb{E}_t V_{t+1,a+1}^{cs} = \sigma_1 \mathbb{E}_t[\varpi + ln\{&\exp((w_{t+1} + \rho \mathbb{E}_{t+1} V_{t+2,a+2}^{cs}) / \sigma_1) \\ &+ \exp((s_{t+1} - \zeta(a) + \theta_1 + \rho \mathbb{E}_{t+1} V_{t+2,a+2}^{o}) / \sigma_1)\}] \end{aligned} \tag{A.8}$$

$$\begin{aligned} \mathbb{E}_t V_{t+1,a+1}^{o} = \sigma_1 \mathbb{E}_t[\varpi + ln\{&\exp((s_{t+1} + \theta_1 + \rho \mathbb{E}_{t+1} V_{t+2,a+2}^{o}) / \sigma_1) \\ &+ \exp((w_{t+1} - \zeta(a) + \rho \mathbb{E}_{t+1} V_{t+2,a+2}^{cs}) / \sigma_1)\}]. \end{aligned} \tag{A.9}$$

where gamma $\varpi \cong 0.577$ is the Euler's constant and the expectations are taken with respect to future taste shocks.

Given this setup we can use the occupational-switching probabilities to derive the aggregate employment in each sector. Since we allow workers at age twenty-two to also pay the switching costs and get their first job in an occupation that is different from their field of study, the number of computer scientists at age twenty-two is a function of the number of recent graduates with a CS degree and the occupational-switching probabilities:

$$L_{nt}^{22} = (1 - q_{t,22}^{cs,o})R_t + q_{t,22}^{o,cs}[(L_{nt}^{22} + G_t^{22}) - R_t] \tag{A.10}$$

$$G_t^{22} = (1 - q_{t,22}^{o,cs})[(L_{nt}^{22} + G_t^{22}) - R_t] + q_{t,22}^{cs,o}R_t \tag{A.11}$$

where R_t is the number of recent graduates with a CS degree, and $(L_{nt}^{22} + G_t^{22}) - R_t$ is the number of college graduates with any other degree. Similarly, the supply of computer scientists at age a from twenty-three to sixty-five is a function of past employment in each occupation and the switching probabilities:

$$L_{nt}^{a} = \left(1 - q_{t,a}^{cs,o}\right)L_{n,t-1}^{a-1} + q_{t,a}^{o,cs}\left[G_{t-1}^{a-1}\right] \tag{A.12}$$

$$G_t^{a} = \left(1 - q_{t,a}^{o,cs}\right)G_{t-1}^{a-1} + q_{t,a}^{cs,o}\left[L_{n,t-1}^{a-1}\right], \tag{A.13}$$

where L_{nt}^{a} is the exogenous number of workers in computer science at age a in time period t, and G_t^{a} is the number of workers at age a working in other occupations.

The aggregate domestic labor supply of computer scientists and other workers is the sum across all ages:

$$L_{nt} = \sum_{a=22}^{a=65} L_{nt}^{a} \tag{A.14}$$

$$G_t = \sum_{a=22}^{a=65} G_t^{a}. \tag{A.15}$$

Here we can see that the labor supply in each occupation depends on past employment, new college graduates, and on wages through the occupational-switching probabilities.

Market-Clearing Conditions

The following equations describe the market-clearing conditions for the labor and output markets. Total consumer expenditure equals labor income plus firm profits (equation [A.16]):

$$P_{tc}C_{dt}^{*} + P_{yt}Y_{dt}^{*} = m = w_t(L_{nt}^{*} + L_{Ft}^{*}) + s_tG_t^{*} + r_tH_t^{*} + (\Pi_t + P_{yt}fN_t). \tag{A.16}$$

Total quantity produced in the IT sector equals domestic consumer demand, intermediate inputs in the other sector, and exports (equation [A.17]):

$$N_t^{\varepsilon/(\varepsilon-1)}\left(\int_{\phi_t^*}^{\infty} c_{it}^{(\varepsilon-1)/\varepsilon}\mu(\phi)\,d\phi\right)^{\varepsilon/(\varepsilon-1)} = C_t^{*} = C_{dt}^{*} + C_{yt}^{*} + C_{Wt}^{*}. \tag{A.17}$$

Total quantity produced in the other sector, net of inputs, equals domestic consumer demand and intermediate inputs in the other sector (equation [A.18]):

$$C_y^{*\psi_2}X_y^{*1-\psi_2} = Y_t^{*} = Y_{dt}^{*} + Y_{Ct}^{*} + fN_t^{*} - Y_{IMt}^{*}. \tag{A.18}$$

Trade in goods is balanced:

$$P_{ct}^{*}C_{Wt}^{*} = P_{yt}^{*}Y_{IMt}^{*}. \tag{A.19}$$

Given that the supply of non–college graduates is inelastic $\bar{H}_t$, and the demand comes from both sectors, their labor market clears as in equation (A.20):

$$\bar{H}_t = H_{ct}^{*} + H_{yt}^{*}. \tag{A.20}$$

Total labor supply for college graduates (CS and non-CS) is fixed, such that total demand for college graduates has to be equal to total supply in each period (equation [A.21]):

$$\overline{L_{nt} + G_t} + \bar{L}_F = L_t^{*} + G_t^{*} = L_{ct}^{*} + L_{yt}^{*} + G_{ct}^{*} + G_{yt}^{*}. \tag{A.21}$$

Details of the Data Used

This study draws on a variety of data sets. Our descriptive statistics in table 4.1 rely on the IPUMS census from 1970 to 2000. We restrict the sample to employed workers. We use the IPUMS-suggested occupation crosswalk and define computer scientists as computer systems analysts, computer scientists, and computer software developers with at least a bachelor of arts (BA) degree. We define foreigners as either naturalized citizens or noncitizens who immigrated after the age of eighteen. For early census years, the year of immigration is only available in ranges. In order to construct a precise year of immigration value for workers in those samples, we choose to select a random value within the year range for each individual.

Data on earnings, domestic employment, and foreign employment used in the calibration 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 eighteen. Earnings are deflated to 1999 dollars, and top-coded values are multiplied by 1.4.

In our analysis we drop imputed earnings. In order to identify these 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. The database also contains ten Census Bureau flags that identify a small fraction (less than 1 percent) of earnings as allocated. Over the period under study, around 26 percent of earnings were allocated. This fraction of imputations varies over time—between 19.14 percent (in 1994) and 29.47 percent (in 2003). These numbers are consistent with Bollinger and Hirsch (2007), who find that between 1998 and 2006, the nonresponse rate was about 20 percent. The small difference in our numbers arises both from using a different sample (restricted to those with a bachelor of arts or master of arts degree) and because nonresponse is not the only reason the CPS imputes earnings.

In order to define workers in computer science, we use the occupational codes 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: "055 electrical and electronic," "064 computer systems analysts and scientists," and "229 computer programmers."

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

In descriptive statistics, we compare the CS workforce to STEM workers. The 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 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 US Census Bureau's Statistics of US Businesses (SUSB). In these data sets we define the IT sector as the subsectors of "computer and electronic product manufacturing," "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.

Extended Out-of-Sample Tests (until 2015)

In section 4.3.3 we describe how we calibrate our labor supply parameters for the period 1994–2001. We matched observed moments of relative wages, employment, and enrollment for three years, and performed an out-of-sample test for the years in between, as shown in figure 4.2. A natural question to ask is whether our calibrated parameters are able to predict movements in key data series for the years after 2001 as an additional out-of-sample test of our model.

In figure 4A.1, we perform such an out-of-sample analysis. To do this, we constructed the relative labor supply shocks Z_t from equation (27) using information on the relative wages, relative native employment, and relative foreign employment for the period 2002–2015 together with the relative labor demand curve for the base year 1994. In a second step, we fed those shocks into our labor supply model using the parameters calibrated for the

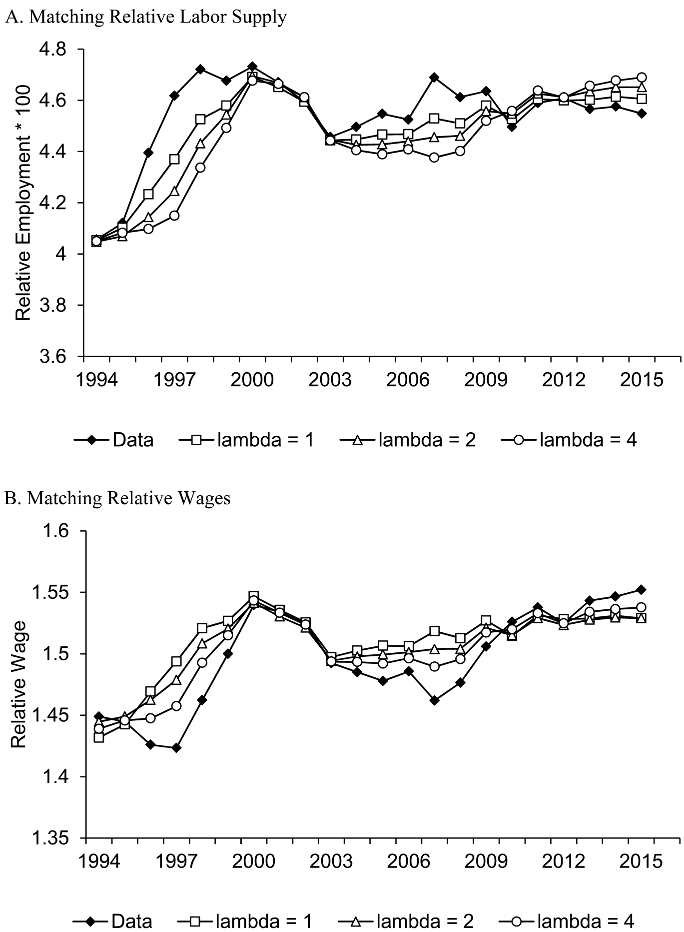


Fig. 4A.1 Employment, wages, and enrollment (1994 to 2015)

Source: Wage and employment data come from the March CPS, whereas degree data is from IPEDS.

Notes: In the calibration exercise, the years 1994 and 2001 were used to match the data for employment and wages, whereas the years 1994, 1997, and 2001 were used to match the data on degree attainment (lagged two years). The years in between, and after 2001, are an out-of-sample test. See the "Details of the Data Used" section of the appendix for more details about the data.

C. Matching Relative Degree Attainment

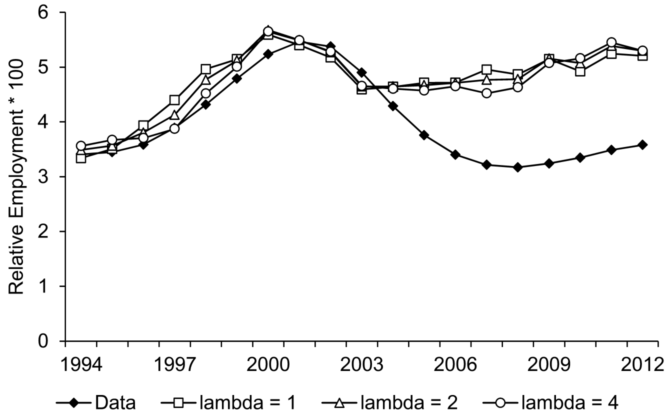


Fig. 4A.1 (cont.)

1994–2001 period to observe how relative wages, employment, and enrollment series are predicted by our model for the post-2001 period. As can be seen in figure 4A.1, we consistently predict employment and wages for the post-2001 years, but overestimate enrollment rates for the years 2005 onward.

Sensitivity Analysis

We check how sensitive our results are to variations in key parameters and specifications of the model. So far we have presented all our results for three different values of λ. Despite slight differences in the magnitudes of changes in income and profits, the results are qualitatively similar across different values of this parameter. For simplicity, we fix λ = 2 when doing our sensitivity analysis on all the other parameters of the model.

First we look at how sensitive our results are to variations in the elasticity of substitution between the IT good and the non-IT good, represented by the parameter σ. As we see in table 4A.1, the more elastic the relative product-demand curve, the larger the income increase for all US workers is when we allow for immigration. This is consistent with economic intuition, since a higher value of σ implies that consumers are more willing to substitute consumption from non-IT to IT goods. When we allow for immigration, the larger number of computer scientists in the economy increases the size of the IT sector and consumers shift into consuming more IT goods. Profits for IT firms rise and workers that are complements to CS workers are better off for higher values of σ. Furthermore, since IT production drives technological change, as and when more resources are devoted to IT for higher values of σ, the price of IT goods falls, increasing overall welfare. Overall, our qualitative results are similar for different values of σ with the

Table 4A.1 Changes in profits and income for different elasticities of substitution between IT and non-IT good in consumer utility

	Percent change in income/profits			Compensating variation/change in profits (million USD)		
	$\sigma = 1$	$\sigma = 2$	$\sigma = 5$	$\sigma = 1$	$\sigma = 2$	$\sigma = 5$
All US workers	0.21	0.25	0.41	8,290	32,760	102,943
All college graduates	−0.16	−0.11	0.07	−2,453	7,060	34,637
Computer scientists that stay	−3.47	−3.65	−3.43	−3,752	−3,360	−1,471
Computer scientists that switch	−1.71	−1.78	−1.61	−189	−127	48
Non-CS college graduates that stay	0.10	0.16	0.34	1,488	10,547	36,061
Non–college graduates	0.44	0.48	0.63	10,743	25,700	68,305
Profits	0.62	1.66	7.70	783	2,106	9,160

Note: All specifications use a value of $\lambda = 2$ and $\beta = 0.23$. Dollar values for compensating variation and profits are in millions of 1999 USD. The scaling up to USD was done using CPS data for the total amount of labor income. Changes in income for different worker groups and profits are calculated as $[(X_{\text{open}}/X_{\text{closed}}) - 1] \times 100$.

only difference being that for high values of σ the population of all college graduates is better off due to immigration.

We also vary the technological change parameter and check how sensitive our results are to reasonable values of β. Our calibrated value of β is 0.233, and we redo our results for values of 0, 0.1, and 0.5. Table 4A.2 shows how the compensating variation and profits change as we vary β. Immigration is benefecial for higher values of β for all types of US workers and firms, since a larger CS workforce increases the gains from technology. Firms directly benefit from higher output, whereas consumers benefit from lower prices as the value of β rises. Overall, however, our qualitative results are similar across the different β levels. The only qualitative difference is that for the scenario where there is no technological progress ($\beta = 0$), the subpopulation of non-CS college graduates is worse off when there is immigration. This happens because without the spillover of aggregate computer scientists, the positive effect they had for being complements to computer science gets smaller and is offset by the lower wages caused by computer scientists switching occupations.

Last, we vary the elasticity of substitution between varieties of the IT good ε across a reasonable range. In section 4.5.3, we discuss results for the baseline case where the goods are perfect substitutes and all IT firms are similar. In table 4A.3, we see that as we lower the elasticity of substitution ε from a value of 3.2 to 2, the overall welfare gains from immigration are enhanced. While close substitutes in the labor market are worse off for smaller values of ε, complements are better off. The overall impacts, however, are similar both qualitatively and quantitatively.

In other results, we test to see whether using a Melitz-style (2003) model

Table 4A.2 **Changes in profits and income for different values of the technological spillover parameter**

	Percent change in income/profits				Compensating variation/change in profits (million USD)			
	$\beta = 0$	$\beta = 0.1$	$\beta = 0.233$	$\beta = 0.5$	$\beta = 0$	$\beta = 0.1$	$\beta = 0.233$	$\beta = 0.5$
All US workers	0.03	0.10	0.21	0.41	1,051	4,150	8,290	16,522
All college graduates	−0.34	−0.26	−0.16	0.05	−5,275	−4,066	−2,453	758
Computer scientists that stay	−3.64	−3.57	−3.47	−3.27	−3945	−3,862	−3,752	−3,530
Computer scientists that switch	−1.85	−1.79	−1.71	−1.54	−205	−198	−189	−171
Non-CS college graduates that stay	−0.08	0.00	0.10	0.31	−1,125	−6	1,488	4,458
Non–college graduates	0.26	0.33	0.44	0.64	6,326	8,216	10,743	15,764
Profits	0.43	0.51	0.62	0.82	554	653	783	1,047

Note: All specifications use a value of $\lambda = 2$ and $\sigma = 1$. Dollar values for compensating variation and profits are in millions of 1999 USD. The scaling up to USD was done using CPS data for the total amount of labor income. Changes in income for different worker groups and profits are calculated as $[(X_{\text{open}} / X_{\text{closed}}) - 1] \times 100$.

Table 4A.3 **Changes in profits and income for different values of the elasticity of substitution across varieties in consumer utility**

	Percent change in income/profits		Compensating variation/change in profits (million USD)	
	$\varepsilon = 0$	$\varepsilon = 3.2$	$\varepsilon = 0$	$\varepsilon = 3.2$
All US workers	0.26	0.21	10,272	8,290
All college graduates	−0.11	−0.16	−1,641	−2,453
Computer scientists that stay	−3.78	−3.47	−4,000	−3,752
Computer scientists that switch	−1.83	−1.71	−181	−189
Non-CS college graduates that stay	0.18	0.10	2,540	1,488
Non–college graduates	0.50	0.44	11,913	10,743
Profits	0.67	0.62	615	783

Note: All specifications use a value of $\lambda = 2$, $\sigma = 1$, and $\beta = 0.23$. Dollar values for compensating variation and profits are in millions of 1999 USD. The scaling up to USD was done using CPS data for the total amount of labor income. Changes in income for different worker groups and profits are calculated as $[(X_{\text{open}} / X_{\text{closed}}) - 1] \times 100$.

of entry significantly affects our conclusions and found out that it does not. In our baseline model, there is an underlying fixed number of potential entrepreneurs that always know their level of productivity ϕ_j, a setup closer to Chaney (2008). An alternative setup is one where firms do not initially know their level of productivity ϕ_j, but must pay a fixed sunk cost f_e to draw their level of productivity from the known distribution. Firms wish to pay this cost as long as their expected profits are positive. As more firms draw and produce, expected profits fall until they are zero. Once a firm draws their productivity ϕ_j, they may choose to pay another fixed cost f and produce if $\phi_j > \phi^*$. This setup is closer to the one described by Melitz (2003).[39]

We may expect these setups to have different implications for firm profits and overall welfare. In our baseline model, an increase in immigration tends to increase firm profits, and the marginal firms enter into producing goods. Since all firms already know their productivity level, the more productive firms always produce. The new entrants are therefore firms that have a productivity level in the immediate neighborhood of ϕ^*. The overall increase in productivity, therefore, is small since the new entrants are firms that have relatively low productivity. Furthermore, the increase in profits is almost entirely captured by the larger firms. In the alternative Melitz (2003) framework, entry may happen at any part of the productivity distribution. When the expected profits rise because of immigration, new entrants may potentially draw very high levels of productivity and enter at the extreme tails of the distribution. The overall increase in productivity is higher, which drives down the price of the IT good and increases consumer utility. Furthermore the new entrants capture all the increase in profits, whereas the profits for the incumbents do not change. Compared to the baseline model, we find that the change in welfare due to immigration is higher under the Melitz entry setup both because of higher aggregate profits and consumer utility. This is because the new firms that enter the industry are not just the firms with marginal productivity, but could also be firms with very high levels of productivity. These firms have higher profits and drive down the output prices more. Qualitatively, however, all our results stay the same across the two models.

References

Acemoglu, D. 1998. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics* 113 (4): 1055–89.

Arrow, K. J. 1962. "The Economic Implications of Learning by Doing." *Review of Economic Studies* 29 (3): 155–73.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum. 2003. "Plants and Productivity in International Trade." *American Economic Review* 93 (4): 1268–90.

Bollinger, C., and B. Hirsch. 2007. "How Well Are Earnings Measured in the Current Population Survey? Bias from Nonresponse and Proxy Respondents." For the North American Summer Meetings of the Econometric Society, Duke University,

39. To see a discussion about these two models in the context of immigration models, see di Giovanni, Levchenko, and Ortega (2015).

June 21–24. http://www2.gsu.edu/~ecobth/ES_Bollinger-Hirsch_Response_Bias_Apr07.pdf.
Borjas, G. 1999. “The Economic Analysis of Immigration.” In *Handbook of Labor Economics*, vol. 3A, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier.
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, part 2): S187–223.
Broda, C., and D. Weinstein. 2006. “Globalization and the Gains from Variety.” *Quarterly Journal of Economics* 121 (2): 541–85.
Card, D., and T. Lemieux. 2001. “Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis.” *Quarterly Journal of Economics* 116 (2): 705–46.
Chaney, T. 2008. “Distorted Gravity: The Intensive and Extensive Margins of International Trade.” *American Economic Review* 98 (4): 1707–21.
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.
Demidova, S. 2008. “Productivity Improvements and Falling Trade Costs: Boon or Bane?” *International Economic Review* 49 (4): 1437–62.
di Giovanni, J., A. Levchenko, and F. Ortega. 2015. “A Global View of Cross-Border Migration.” *Journal of the European Economic Association* 13 (1): 168–202.
Dixit, A. K., and J. E. Stiglitz. 1977. “Monopolistic Competition and Optimum Product Diversity.” *American Economic Review* 67 (3): 297–308.
Doran, K., A. Gelber, and A. Isen. 2015. “The Effect of High-Skilled Immigration Policy on Firms: Evidence from H-1B Visa Lotteries.” NBER Working Paper no. 20668, Cambridge, MA.
Freeman, R. B. 1975. “Supply and Salary Adjustments to the Changing Science Manpower Market: Physics, 1948–1973.” *American Economic Review* 65 (1): 27–39.
———. 1976. “A Cobweb Model of the Supply and Starting Salary of New Engineers.” *Industrial and Labor Relations Review* 29 (2): 236–48.
———. 2006. “Does Globalization of the Scientific/Engineering Workforce Threaten US Economic Leadership?” In *Innovation Policy and the Economy*, vol. 6, edited by Adam B. Jaffe, Josh Lerner, and Scott Stern. Cambridge, MA: MIT Press.
———. 2009. “What Does Global Expansion of Higher Education Mean for the US?” NBER Working Paper no. 14962, Cambridge, MA.
Gaubert, C. 2015. “Firm Sorting and Agglomeration.” Working Paper, University of California, Berkeley.
Ghosh, A., A. M. Mayda, and F. Ortega. 2014. “The Impact of Skilled Foreign Workers on Firms: An Investigation of Publicly Traded US Firms.” IZA Discussion Paper no. 8684, Institute for the Study of Labor.
Goldin, C., and L. F. Katz. 2007. “Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing.” *Brookings Papers on Economic Activity* 2 (2007): 135–65.
Gordon, R. 1990. *The Measurement of Durable Goods Prices*. National Bureau of Economic Research Monograph. Chicago: University of Chicago Press.
Grossman, G. M., and E. Helpman. 1991. “Quality Ladders in the Theory of Growth.” *Review of Economic Studies* 58 (1): 43–61.
Grossman, G. M., and E. Rossi-Hansberg. 2008. “Trading Tasks: A Simple Theory of Offshoring.” *American Economic Review* 98 (5): 1978–97.
Hopenhayn, H. A. 1992. “Entry, Exit, and Firm Dynamics in Long Run Equilibrium.” *Econometrica* 60 (5): 1127–50.

Hunt, J., and M. Gauthier-Loiselle. 2010. "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics* 2 (2): 31–56.
Jorgenson, D. W., M. S. Ho, and J. D. Samuels. 2016. "Education, Participation and the Revival of US Economic Growth." NBER Working Paper no. 22453, Cambridge, MA.
Kambourov, G., and I. Manovskii. 2009. "Occupational Specificity of Human Capital." *International Economic Review* 50 (1): 63–115.
Katz, L. F., and K. M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107 (1): 35–78.
Kerr, W. R., and W. F. Lincoln. 2010. "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention." *Journal of Labor Economics* 28 (3): 473–508.
Khanna, G., and N. Morales. 2015. "The IT Boom and Other Unintended Consequences of Chasing the American Dream." Working Paper no. 460, Center for Global Development.
Kirkegaard, J. 2005. "Outsourcing and Skill Imports: Foreign High-Skilled Workers on H-1B and L-1 Visas in the United States." Working Paper no. 05-15, Washington, DC, Peterson Institute for International Economics.
Kline, P. 2008. "Understanding Sectoral Labor Market Dynamics: An Equilibrium Analysis of the Oil and Gas Field Services Industry." Cowles Foundation Discussion Paper no. 1645, Cowles Foundation for Research in Economics, Yale University.
Krugman, P. 1979. "A Model of Innovation, Technology Transfer, and the World Distribution of Income." *Journal of Political Economy* 87 (2): 253–66.
Lemieux, T. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review* 96:461–98.
Liu, R., and D. Trefler. 2008. "Much Ado about Nothing: American Jobs and the Rise of Service Outsourcing to China and India." NBER Working Paper no. 14061, Cambridge, MA.
Lowell, B. Lindsay. 2000. "H-1B Temporary Workers: Estimating the Population." CCIS Working Paper no. 12, Center for Comparative Immigration Studies, University of California, San Diego. http://escholarship.org/uc/item/4ms039dc#page-1.
Matloff, N. 2003. "On the Need for Reform of the H-1B Non-Immigrant Work Visa in Computer-Related Occupations." *University of Michigan Journal of Law Reform* 36 (4): 815–914.
Melitz, M. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71:1695–725.
Melitz, M., and G. Ottaviano. 2008. "Market Size, Trade, and Productivity." *Review of Economic Studies* 75:295–316.
Melitz, M., and S. Redding. 2015. "New Trade Models, New Welfare Implications." *American Economic Review* 105 (3): 1105–46.
Mithas, S., and H. C. Lucas. 2010. "Are Foreign IT Workers Cheaper? US Visa Policies and Compensation of Information Technology Professionals." *Management Science* 56 (5): 745–65.
Mukhopadhyay, S., and D. Oxborrow. 2012. "The Value of an Employment-Based Green Card." *Demography* 49:219–37.
Nathan, M. 2013. "The Wider Economic Impacts of High-Skilled Migrants: A Survey of the Literature." IZA Discussion Paper no. 7653, Institute for the Study of Labor.
Peri, G., K. Shih, and C. Sparber. 2013. "STEM Workers, H1B Visas and Productivity in US Cities." Working Paper, University of California, Davis, and Colgate University.
———. 2014. "Foreign STEM Workers and Native Wages and Employment in US Cities." NBER Working Paper no. 20093, Cambridge, MA.

———. 2015. "Foreign and Native Skilled Workers: What Can We Learn from H-1B Lotteries?" NBER Working Paper no. 21175, Cambridge, MA.

Peri, G., and C. Sparber. 2009. "Task Specialization, Immigration, and Wages." *American Economic Journal: Applied Economics* 1 (3): 135–69.

———. 2011. "Highly Educated Immigrants and Native Occupational Choice." *Industrial Relations* 50 (3): 385–411.

Pfluger, M., and S. Russek. 2013. *Firms in the International Economy—Firm Heterogeneity Meets International Business*. Cambridge, MA: MIT Press.

Romer, P. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5, part 2): S71–102.

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. "The Engineering Labor Market." *Journal of Political Economy* 112 (S1): S110–40.

US Department of Commerce. 2000. "Digital Economy 2000." Technical report, Washington, DC, US Department of Commerce.

US General Accounting Office. 2000. "H-1B Foreign Workers: Better Controls Needed to Help Employers and Protect Workers." Technical Report GAO/HEHS 00-157, Washington, DC, US General Accounting Office.

———. 2011. "H-1B Visa Program: Reforms Are Needed to Minimize the Risks and Costs of Current Program." Technical Report GAO/HEHS 00-157, Washington, DC, US General Accounting Office.

US Immigration and Naturalization Service. 2000. "Characteristics of Specialty Occupation Workers (H 1B)." Technical report, Washington, DC, US Immigration and Naturalization Service.