High Skilled Immigration and the Market for Skilled Labor: The Role of Occupational Choice

JIE MA*

April 20, 2018

Abstract

n recent years, immigration rates have increased dramatically among the most highly skilled workers. How does this inflow affect labor market outcomes among highly skilled native-born workers? I estimate a general equilibrium model in which individuals adjust to skilled immigration by changing occupations and investing into human capital differently. Moreover, I estimate the demand functions for native and immigrant workers and find that skilled immigrants and natives are imperfect substitutes in some occupations and are complements in others. Counterfactual exercises indicate that even large inflows of foreign skilled workers have limited impacts on domestic workers. In particular, the skill rental rates for native science and engineering workers would have been approximately 2% higher if firms were not able to hire more foreigners than they did in 1994. On the other hand, had the U.S. workers been constrained to remain in their original occupations, the adverse impacts of foreign labor competition would be more severe. When natives' occupational choices respond to immigration, the negative effects are diffused. The extent to which this occupational mobility helps to absorb the immigration shock depends not only on the substitution elasticity in the directly affected occupations, but also on the demand elasticity of native labor in the destination occupations where natives move to.

JEL Codes: F22, J61, J23, J24

*Department of Economics, McGill University, 855 Sherbrooke Street West Montreal, QC, H3A 2T7. (email: jie.ma2@mail.mcgill.ca) I am indebted to Fabian Lange, Theodore Papageorgiou, Markus Poschke for their patience, guidance, encouragement support, and deep generosity of time. I would also like to thank Christopher Ferrall for his support and for developing the wonderful software Niqlow. Thanks to the conference/seminar participants at the Barcelona GSE Summer Forum, EALE, IZA, CEA, McGill University, HEC Montreal, and Colgate university for their useful comments. All errors are my own.

1 Introduction

Contrary to popular perception, many of the immigrants to the U.S. in the last decades were highly skilled. Between 1990 and 2010, the number of skilled immigrants residing in the U.S. rose by about 4.8% annually.[1] Today 16% of the U.S. workers with a bachelor's education are immigrants. The inflow of immigrants has furthermore been unevenly spread. 25% of computer scientists and electronics engineers are immigrants, but only 6% of those working in the legal professions are. Basic economic arguments (Borjas, 1999) suggest that such an unbalanced and sizable flow of migrants might have substantial and detrimental effects on natives with similar skills working in the same professions. Given the empirical distribution of skilled migrants across occupations, we expect any labor market effects to be largest among U.S. workers in the science, technology, engineering, and mathematics (STEM) occupations.[2] However, with a few exceptions (Peri et al., 2015; Hanson and Slaughter, 2015; Bound et al., 2015), the literature so far has neglected the question of how skilled immigrants has affected native born workers in the STEM fields. How do natives' human capital decisions react to skilled immigration? Would they have specialized in different occupations?

A comprehensive answer to these questions is crucial to understand the economic impacts of skilled immigrant inflows. And the answer requires multiple key inputs. First, we need to estimate the demand for skilled workers across occupations. Second, we need to understand how the native-born workers choose occupations and whether occupational switching can serve as a pressure-valve mitigating and diffusing any consequences of the inflows of skilled immigrants. In particular, I build a general equilibrium model focusing on dynamic occupational choices of native workers with the following key features: (a) skilled natives and immigrants are allowed to be imperfect substitutes or complements, and substitutability or complementarity can vary across occupations; (b) workers are heterogeneous in terms of their multi-dimensional innate abilities; (c) workers accumulate occupation specific human capital through learning-by-doing, and the occupation specific human capital is partially transferable across occupations.

In this paper, the proposed and estimated general equilibrium model is used to identify the wage impacts of skilled immigration in the STEM fields over the last two decades, taking into consideration of occupational adjustments by natives. The general equilibrium approach departs from the recent literature and allows me to address the following main points. First, I estimate the labor demand functions across occupations which are general enough to allow the possibility that native and foreign workers could be complements. In other papers in the literature (Bound et al., 2015; Llull, 017a; Dustmann et al., 2012), the common assumption presented is that native and immigrant

[1] Source: Migration Policy Institute

[2] The U.S. Immigration and Customs Enforcement list disciplines including: Physics, Actuarial Sciences, Chemistry, Mathematics, Applied Mathematics, Statistics, Computer Science, Computational Science, Psychology, Biochemistry, Robotics, Computer Engineering, Electrical Engineering, Electronics, Mechanical Engineering, Industrial Engineering, Information Science, Civil Engineering, Aerospace Engineering, Chemical Engineering, Astrophysics, Astronomy, Optics, Nanotechnology, Nuclear Physics, Mathematical Biology, Operation Research, Neurobiology, Biomechanics, Bioinformatics, Acoustical Engineering, Geographic Information Systems, Atmospheric Sciences, Software Engineering, Econometrics, etc.

labor of the same type are perfect substitutes.[3] This assumption of perfect substitutability is not innocuous. It discusses the impacts of immigration only within the score of potential competitions. Allowing the possibility of complementarity integrates the potential beneficial side into the story to avoid the overestimation of negative effects. Second, by explicitly modeling occupational choices by natives, I can quantify and correct the biases in the estimated effects of immigration on wages introduced by ignoring native workers' labor market adjustments. Another common assumption in the literature is the fixed native labor supply in the narrowly defined labor markets (Borjas, 1999; 2003; Llull, 017b; Ottaviano and Peri, 2012). Failing to account for native workers' adjustments may lead to a substantial bias in the estimated wage impacts of skilled immigration. Using the estimated model, I identify that a non trivial fraction of native workers adjust their occupational choices in response to foreign labor competition, which mitigates any initial impacts through diffusion. And third, the equilibrium model with multiple occupations allows me to evaluate the impacts of a selective immigration policy based on occupations that would not be possible otherwise.

The two-sector equilibrium model (Computer Science occupations (CS) v.s. Other-STEM occupations) builds on Lee (2005) and Lee and Wolpin (2006).[4] The supply side of the model extends the structure of Keane and Wolpin (1994; 1997). Native workers live from age 22 to 65 and make yearly forward-looking decisions on their occupational choices. Native workers differ in their innate abilities for working in either sector.[5] Agents choose their occupations according to their comparative advantage (Roy, 1951). This comparative advantage evolves over time because human capital accumulates throughout the life-cycle. Learning-by-doing on the job leads to accumulation of occupation-specific human capital which is only partially transferable across occupations. Switching occupation is costly in the current model due to the loss of previous accumulated human capital. In their human capital investment decisions, natives make forecasts about future wages, which depend on future inflows of skilled immigration and future sector specific productivity shocks. Individuals are assumed to be able to perfectly foresee future career prospects when making decisions.[6] Skilled foreign workers are allowed to accumulate human capital through learning-by-doing as well. However, the lack of portability of the H-1B visa across occupations restricts their occupational mobility after entering the U.S. labor market. As a result, the high skilled immigrants can not reoptimize their occupations as their domestic peers.[7] In this model, the supply of foreign worker is directly taken from the data and is exogenously determined by the US immigration policies (H1-B visa regulations). The features of the H1-B visa effectively separate the foreign labor supply from the labor demand, which provides plausible exogenous variations to identify the occupational specific production functions. See Section

[3] Llull (017a) allows for some flexibility by defining labor in skill units. However, after adjusting for the observed characteristics, native and immigrant labor are perfectly substitutable.

[4] Computer Science occupations include computer systems analysts, computer scientists, computer software developer.

[5] The different dimensions of workers' innate abilities are allowed to be freely correlated.

[6] This is not a rational expectation model. The perfect foresight assumption introduces potential bias into the model. Later in the paper, I discuss the impacts of this assumption in detail.

[7] The H-1B is a non-immigrant visa in the United States under the Immigration and Nationality Act, section 101(a)(15)(H). It allows U.S. employers to temporarily employ foreign workers in specialty occupations. It is the primary source that skilled immigrants entering the U.S. labor market.

2 for more details.[8] On the labor demand side, I assume a Constant Elasticity of Substitution (CES) production function for each occupation, which use domestic and immigrant labor to produce outputs. This approach flexibly allows for the possibilities of imperfect substitution even complements between immigrant and native workers in production, and moreover allows the substitutability or complementarity to vary across occupations. Even though, the literature has not reached a consensus on what the effect of immigration on a defined labor market, most of the papers agree on the existence of asymmetric effects across different workers (Ottaviano and Peri, 2012; Piyapromdee, 2015; Llull, 017a; Dustmann et al., 2012). The flexible specification of the occupational specific production functions incorporates the asymmetric effects of skilled immigration. Incorporating the asymmetric effects is crucial to correctly estimate native occupational responses because it correctly describes the changes in relative wages and the adjustment incentives. Uneven inflows of skilled immigrants can change market prices for different occupations differently, and thus generate incentives for natives to adjust their human capital decisions. The general equilibrium framework with multiple occupations is a crucial feature of the model because it links the immigration-induced labor supply shocks with the native workers' labor market adjustment through changes in relative wages.

I fit the model to U.S. data from the Current Population Survey (CPS), the American Community Survey (ACS), and the Panel Study of Income Dynamics (PSID). The model is identified by combining the structural assumptions with the exogenous variation in the flow of skilled immigrants.[9] I then use the estimated parameters to quantify the effect of skilled immigration on labor market outcomes.

My estimation results add to the previous papers (Peri and Sparber, 2009; Ottaviano and Peri, 2012; Manacorda et al., 2012) and find complementarity between native and immigrant workers at the finer occupational level. The substitutability or complementarity does vary across occupations. My estimates suggest that skilled native and foreign labor are imperfect substitutes in the CS occupations, but are complements in the other-STEM occupations. Then, I explore the ONET data to provide some evidence for the micro-structure to the imperfect substitution or complementarity between skilled natives and immigrants. Task specialization within occupations is a potential source that can explain my findings.

With the estimated parameters, I first want to quantity the effect of skilled immigration on the STEM occupations. In order to do so, I consider a counterfactual economy in which the total foreign skilled labor supply is kept constant to its 1994 level, also holding the occupation mix of immigrants fixed. This experiment attempts to study the impact of immigration policies on total quantity restriction, such as the variations in the H-1B cap. Natives are allowed to freely switch occupations according to the changes in market conditions. The results of this exercise deviate from existing papers and find a very limited impact on native workers even for large inflows of skilled immigrants. Wages of skilled domestic workers on average increase by 2.41% due to this highly

[8] It is reasonable that one raises the concern with the endogeneity of immigration policies. The legislation responds to economic shocks with a lag. Historically, we observe that the U.S. business environment causes the demand for H1-B visas to vary year by year, but the H1-B caps which restrict the maximum number of new visas typically do not vary. This lagged response provides enough variations needed to identify the production functions.

[9] The full model is estimated in two steps following Jeong, Kim and Manovskii's (2015).

restrictive counterfactual policy. When natives are allowed to respond to immigration by switching occupations, the labor supply shocks induced by inflows of skilled immigrants are mitigated and diffused, and redistribution effects are partially arbitraged out.

One thing to note is that a universal restriction on the number of skilled immigrants across occupations is suboptimal in the counterfactual world. The economy can perform significantly better with selective immigration policies that favor occupations where complementarity exists between natives and immigrants. In the real world, the point-based immigration system adopted by the Canadian government and the Optional Practical Training (OPT) period used by the United States Citizenship and Immigration Services (USCIS) provide workers in certain occupations and fields of study with better access to a country's labor markets. Thus, in the second set of counterfactual analysis, I manipulate the occupation mix of skilled immigrants to replicate the effect of a selective immigration policies. The foreign labor in the other-STEM occupations is allowed to increase. Optimizing occupation mix of skilled immigrants achieves a Pareto improvement, which benefits all natives compared to a simple overall cap. Workers in the occupations where complementarity exists experience a direct increase in wages due to inflows if skilled immigrants, while others experience a positive spillover effect as a result of native workers' occupational mobility.

I then use the estimated model to assess the importance of its key features: the substitutability between domestic and foreign workers, the general equilibrium setting with multiple occupational groups, and lastly the human capital decisions by heterogeneous agents. First, my estimates indicate that domestic and foreign workers are imperfect substitutes in the CS occupations and are complements in the other-STEM occupations. If one instead falsely assumes that immigrants and natives are perfect substitutes, then one will overestimate the costs that immigrants impose on skilled natives. Second, a partial equilibrium model, unlike the general equilibrium, will ignore any wider impact of foreign computer scientists on other occupations. The quantitative results of this paper demonstrate that a non-trivial number of domestic CS workers switch to the other-STEM occupations. This implies that the inflows of skilled immigrants in the CS occupations affect workers in the other-STEM occupations as well. A partial equilibrium model will miss this adjustment effect. Failing to account for this adjustment may lead to a substantial bias in the estimation of wage effects across occupations. Third, unobservable heterogeneity is an important determinant of both individual wages and occupational choices (Keane and Wolpin, 1997). Shutting down heterogeneity in labor will substantially increase occupational mobility. As a result, a model without heterogeneity in labor will overestimate the natives' sensitivity in terms of occupational mobility.

In the model, the occupational adjustment margin in general mitigates and diffuses any initial impacts of skilled immigrants. Results also suggest a significant heterogeneity cross individuals in their occupational adjustments to immigration. I use the estimated model to predict individual's option value of free occupational mobility. The computational exercise suggests that the occupational adjustment margin is more valuable for younger workers when faced with foreign labor competition. Younger workers are willing to pay up to $40,000 to remove a temporary restriction on occupational mobility. If natives were forced to stay in their current occupations permanently where increas-

ing foreign labor competition is expected, they would on average require more than $ 100,000 of compensation variations to maintain the same level of lifetime utility. Human capital is occupation specific and is only partially transferable cross occupations in the current setting; thus, earlier mistakes in human capital decisions will have long lasting effects. I also find that the valuation of free occupational mobility is higher if no one else in the economy is endowed with the same option. In a Roy world, native workers who do not have strong comparative advantages in either sector have a larger margin to adjust and gain more from exercising the option of free occupational mobility.

This paper contributes to a growing body of literature that studies the impacts of skilled immigration on U.S. labor market outcomes (Kerr and Lincoln, 2010; Hunt and Gauthier-Loiselle, 2010; Hunt, 2011; 2013; Borjas and Doran, 2012; Moser et al., 2014; Bound et al., 2015).[10] Motivated by the empirical evidence documented in Peri and Sparber (2011),Peri et al. (2015), and D'Amuri and Peri (2014), the main focus of this paper is on the domestic workers' occupational adjustments in response to skilled immigration. Adding this adjustment margins relaxes the assumption of fixed labor supply like those in Borjas (2003), Card (2009), Ottaviano and Peri (2012), or Llull (017b). It contributes by showing that the estimated impacts of skilled immigrants differ substantially depending on whether we capture the labor market equilibrium adjustments. In the paper, I identify a non-trivial portion of domestic workers switching occupations in response to inflows of skilled immigrants. It highlights the importance of accounting for natives occupational adjustments when quantifying the wage impact of immigration. Llull (017a) study three adjustment mechanisms of native workers: education, occupation, and labor participation. The occupational adjustments in this paper have a very similar flavor. This paper also departs from previous papers in two additional dimensions. The first one is that it uses flexible enough production functions that allow natives and immigrants to be imperfect substitutes or even complements. This enlarges the scope within which we discuss the impact of skilled immigration.Bound et al. (2015) assume natives and immigrants are perfect substitutes. With a normal downward sloping labor demand curve, their model only captures half of the story. And the second one is that it introduces the occupational dimension on the demand side, which allows the elasticity of substitution between natives and immigrants varying across occupations. This accurately captures the asymmetric effects of skilled immigration across occupations and generates correct incentives to occupational mobility. This additional dimension allows the model to evaluate the effect of a selective immigration policy.

This paper also adds to another strand of the literature studying the native's internal migration in response to immigration. In the literature, Card (2001), Borjas (2006), and Piyapromdee (2015), model the special equilibrium responses to immigration in a static framework. These papers estimate the wage effect of immigration across different local labor market and find mixed evidence of this adjustment margin.[11] The role of geographic mobility in these paper is analogous to the

[10] Regression analysis in the literature has found no clear evidence of crowd-out of native employment, and in some cases has found crowd-in.The literature studying the human capital externalities of skilled immigrants has found that immigration through H-1B program leads to large positive impacts on innovation measured by the number of patent being filed in the U.S.

[11] Card (2005) finds that native workers are rather insensitive in terms of geographic mobility. Borjas (2006) shows that native migration can substantially reduce the negative wage impacts of immigration. Piyapromdee (2015)builds a spatial equilibrium model in which she finds that the extent to which the geographic mobility reduces the adverse

occupational mobility here: both geographic and occupational mobility act as pressure valves that mitigates the wage effects of immigration through diffusion in equilibrium.

There are very few structural papers studying natives' occupational response to foreign competition. The study of Bound et al. (2015) and Llull (017a) are the only ones to my knowledge falling into this category. Bound et al. (2015) utilize a calibrated model to analyze the employment and wage adjustment of native computer scientists. They assume a decreasing returns technology in which domestic and foreign computer scientists are perfect substitutes. The partial equilibrium setting in their study focuses only on the market for computer scientists and misses the wider impacts that high-skilled immigration might have on the U.S. economy. Llull (017a) proposes a rich model incorporating the educational, occupational, and participation adjustments of natives to immigration, which mitigate initial effects on wages and inequality through diffusion as the mechanism I emphasize in this paper. Llull (017a) also identifies a significant heterogeneity in these adjustments both across individuals and across different margins. In this paper, I focus on the relevant occupational adjustment margin and add more structures to the production functions. Besides allowing skilled immigrants and natives to be complements to include the potential benefit side of the story, the occupation specific production functions are flexible enough to capture the heterogeneity in the elasticity of substitution. Adding the occupational dimension to the labor demand side and letting the production functions vary across occupations capture the asymmetric impacts of immigration across different occupations. This feature is crucial because it provides us new insights into the effects of changing the immigrant occupational composition and the consequences of selective immigration policies, like the OPT in the U.S., which has not been studied in the previous papers. Analogously to Llull (017a), I also explore the heterogeneity in adjustments but mainly across individuals.

The paper is structured as follows. Section 2 describes the OPT and the H-1B visa program in the U.S. Section 3 specifies the model and Section 4 discusses the data, identification and estimation procedure. Section 5 presents the results. Section 6 shows the counterfactual experiments. I discuss some potential model specification issues and some limitation of the current model in Section 7 and conclude in Section 8.

2 Relevant Immigration Policies and Impacts

The primary visa program through which skilled immigrants enter the U.S. is the H-1B program.[12] The H-1B visa program for temporary workers in specialty occupations was established by the Immigration Act of 1990. H-1B visas require applicants to have at least a bachelor's degree or its

impacts in local labor markets depends on the substitutability between different types of labor and the local labor market composition.

[12]Other temporary worker visas, similar to H-1Bs, are the L-1 and TN visa. These programs are less than 10% of the size of the H-1B program for high skilled workers and contain institutional features that limit the firms' ability to use them to circumvent the H-1B quota. The Department of Homeland Security has argued that limited substitution exists across the H-1B and L-1 visas. Neither visa category has shown substantial increases after the H-1B cap was dramatically reduced in 2004.

equivalent.It effectively restricts the annual flow of skilled immigrants.[13]

The H-1B temporary visa is noteworthy in the context of this paper because it not only enables highly skilled immigrants to work in the U.S., but it creates a binding contract between a particular worker and a sponsoring firm.[14] The process begins with a sponsoring firm filing a Labor Condition Application (LCA) to USCIS for a prospective employee. Once the application is approved, it allows a foreign skilled worker to stay a maximum of six years on an H-1B visa. If the worker is unable to adjust the status of their visa into one that allows permanent residence by the time their visa period expires, the H-1B visa holder must leave the country. An important consequence of this sponsorship is that foreign workers are tied to their sponsoring firms, which to a large extent prevents immigrants from switching occupations.[15] This feature simplifies the analysis in the paper. Upon entering the U.S. labor market, skilled immigrants are effectively tied to one particular occupation due to the binding contract. Given the lack of portability of the H-1B visa, foreign workers are very insensitive to changes in wages across occupations. Therefore, the occupational mobility (self-selection) of skilled immigrants after entering the U.S. market is not a concern in this paper. However, even immigrants face formidable barriers to occupational mobility, they accumulate human capital through learning-by-do as natives.

Since 1990, the United States has capped the number of H-1B visas that are granted each year.[16] The annual cap has fluctuated over the years, and the policy debate typically focuses on whether the cap should be increased. During the early 1990s, the initial cap was set at 65,000 visas per year. When first introduced, the cap was rarely reached. By the mid-1990s, the allocation was based on a first come first served principle, resulting in frequent denials or delays on H-1Bs because the annual quota was usually exhausted within a short period of time. The USCIS then instituted a lottery system to randomly select qualified petitions. Figure 1 shows the changes in the H-1B visa cap and the estimated population of H-1B holders. The initial cap of 65,000 visas was increased to 115,000 in 1999 and to 195,000 in 2001. The cap then reverted to 65,000 in 2004. The cap has since remained unchanged. The cap is binding recently and the chance of getting an H-1B visa is less than 1. In 2016, the probability of winning the lottery was less than 40% at its 'historical low'. The annual flow of foreign skilled workers into the U.S. is effectively restricted by the H-1B cap, and the temporary visa only allows visa holders to stay for six years. Based on these two facts, I argue that the stock of skilled immigrants is also inelastically supplied and policy driven, which is the plausible exogenous variation required to identify the labor demand. One maybe concern about the endogeneity of immigration policies in the U.S. However, over years, we see binding caps and

[13]The specialty occupations are defined as requiring theoretical and practical application of a body of highly specialized knowledge in a field of human endeavor including, but not limited to, architecture, engineering, mathematics, physical sciences, social sciences, medicine and health, education, law, accounting, business specialties, theology and arts.

[14]In the LCA's for H-1B workers, the employer must attest that the firm will pay the non-immigrant the greater of the actual compensation paid to other employees in the same jobs or the prevailing compensation for that occupation, and the firm will provide working conditions for the foreign worker that do not cause the working conditions of the other employees to be adversely affected.

[15]The H-1B allows visa holders to switch employers but the new job has to match the original ones in terms of title, requirements, and background.

[16]There are exemptions for foreigners who work at universities and non-profit research facilities.

news about high tech executives lobby to expand the H-1B program while the cap has not been increased for more than 10 years after it was abruptly cut down by two thirds in 2004.

The H-1B visa cap places a restriction on the number of skilled immigrants, while the Optional Practical Training program (OPT) favors foreign workers with special training and who work in specific occupations. OPT is a period during which undergraduate and graduate students with F-1 status, who have completed or have been pursuing their degrees for more than nine months, are permitted by the USCIS to work for a certain period of time on a student visa. STEM occupations have a total OPT length of 36 months, which is two years longer than other non-STEM occupations. When OPT expires, if students fail to acquire a valid working visa, they have to either leave the country or enroll into another educational program. Longer OPT length means multiple visa application opportunities. This greatly increases the chance of actually getting a temporary working visa. As a result, a noteworthy portion of H-1B beneficiaries work in the STEM occupations, especially the computer science related occupations (see the occupational composition of H-1B beneficiaries in Figure 2).[17] Figure 3 plots the fraction of skilled immigrants in three different groups using March CPS data. The bottom flat line is the fraction of immigrants in the high-skilled labor force. The proportion of foreign workers mildly increased until 2001 and then stabilized afterwards, consisting approximately 12% of the high-skilled labor force. The proportion of foreign workers in the CS occupations is persistently higher than the other-STEM occupations, both of which are higher than the non-STEM occupations. Policies like the OPT, which favor the STEM occupations, are responsible for this pattern. One clear advantage of an equilibrium model with multiple occupations, like the one proposed in this paper, is that it is useful to evaluate the effects of changes in immigration composition and the consequences of selective immigration policies.

3 An Equilibrium Model of Dynamic Labor Supply and Demand of STEM Workers

To analyze the effects of skilled immigration on native workers in different occupations, I extend the static Roy model to dynamic general equilibrium settings. I estimate the model with the U.S. data, and later in the paper, I used the estimates to quantify the effect of inflows of skilled immigrants in the STEM occupations, accounting for the occupational choices by native workers. The equilibrium framework of this paper models the natives human capital decision explicitly in response to immigration, relaxing the common assumption in the literature that native labor supply in a narrowly defined market is fixed. In the model, all agents work in the STEM occupations, which are the occupations that receive large inflows of skilled immigrants over the last two decades. Natives and immigrants are two different types of labor in production. Whether they are substitutes or complements in the production, I directly estimate it from the data rather than making restrictive assumptions. The previous cross-skill cell analysis in the literature (Borjas, 2003; Card, 2009; Otta-

[17] H-1B Beneficiaries are workers renewing their H-1B visa as well as newly arrived workers

viano and Peri, 2012; etc.) has not brought a consensus on what the effect of immigration on wages because the result is sensitive to assumptions on elasticities of substitution between natives and immigrants. The production functions in the paper are flexible enough to allows both substitution and complement.

3.1 Labor Supply

3.1.1 Native Career Decisions and Labor Supply

Natives enter the labor market at age 22 with a bachelor's degree. At the beginning of each period between age 22 and 65 (the exogenous retirement age), individuals choose $d \in \{cs, ncs\}$, working either as a computer scientist ($d = cs$) or in the other-STEM occupations ($d = ncs$) to maximize their expected present value of their lifetime utility.[18]

Individual Human Capital Formation, Wages, and Preferences

An individual enters the labor market with full knowledge of his or her own innate ability, which is modeled as a realization from a bivariate normal distribution.[19]

$$\epsilon = \begin{pmatrix} \epsilon^{cs} \\ \epsilon^{ncs} \end{pmatrix} \sim N\ (\mu,\ \Sigma)$$

The initial ability endowments are occupation specific, e.g., ϵ^{cs} denotes the individual innate ability to work as a computer scientist. The two dimensions of the innate ability are allowed to be freely correlated. I again estimate the correlation from data without making any restrictive assumption. This ϵ is permanent and persistent heterogeneity. I decide to explicitly model the innate ability because the unobservable heterogeneity is an important determinant of both individual wages and occupational choices (Keane and Wolpin, 1997). In the identification part, I will discuss in detail how the variations in occupational employment shares across birth cohorts and over time, and the variations in relative wages across years provide sufficient conditions for identifying the innate ability distribution, which is an application of Heckman and Honore (1990).

Individuals accumulate occupation specific human capital when engaged in a productive activity (working in one occupation) through learning-by-doing. This occupation specific human capital is only partially transferable across occupations(Keane and Wolpin,1994; 1997).

[18] Adding extra occupation is very costly. Each additional occupation implies an extra choice, an additional experience variable, an additional dimension in the innate ability distribution, more than 15 additional parameters to estimate, an additional equilibrium skill price, and an additional expectation process. As noted in section 7, I show that the current occupational choices and the current classification that capture the most relevant occupations and discuss the potential biases when ignoring other adjustment margins.

[19] Even though I call this innate ability, it actually captures more than unobserved ability. Since the educational decisions and investments prior to work are not explicitly modeled, they can be captured by this 'heterogeneous ability'. Llull (017a) explicitly models the educational adjustments and finds the inflow of low skilled immigrants causes the primary-age native males to increase their education by three years.

The human capital evolves endogenously with age a based on individual occupational choices. The occupation specific human capital depends on occupational tenures (x_a^{cs}, x_a^{ncs}) and general work experience x_a, where $x_a = x_a^{cs} + x_a^{ncs}$.[20]

$$H_a^{cs} = \exp\{\alpha_1 x_a^{cs} + \alpha_2 x_a^{ncs} + \alpha_3 x_a^2 + \alpha_4 x_a^3 + \epsilon^{cs}\} \tag{1a}$$

$$H_a^{ncs} = \exp\{\alpha x_a^{cs} + \alpha_6 x_a^{ncs} + \alpha_7 x_a^2 + \alpha_8 x_a^3 + \epsilon^{ncs}\} \tag{1b}$$

The function 1a and 1b are specified in a consistent fashion with an optimal human capital investment framework and the corresponding Mincer earning equations (Ben-Porath, 1967; Heckman et al., 2006). The log-wage is a quadratic function of general labor market experience which is consistent with the standard Ben-Porath human capital accumulation model with a linearly declining rate of investment on-the-job. The specification departs from the standard Mincer equations by the following two features. First, I add a third-order term in experience to improve the fit relative to Mincer's original specification (Murphy and Welch, 1990; Card, 1999), specifically to improve the fit to the observed flattening of the life-cycle earnings profiles. Second, this specification introduces different returns to occupation tenures in different occupations. Kambourov and Manovskii (2009) provide sufficient evidence to support the occupational specificity of human capital, which in this model is captured by the separate returns to tenures in different occupations. One feature of the occupational specificity of human capital is that it accounts for the implicit switching cost of occupational mobility. With α_1 greater than α_2, workers with longer tenures in the other-STEM occupations would experience a wage loss when switching to the CS occupations. Switching results in a wage penalty due to the loss of the previously accumulated occupation specific human capital. The nature of human capital accumulation also introduces asymmetric switching costs among individuals and across different cohorts, which generates heterogeneous adjustments across individuals. Education is absent in the human capital function because this paper studies a homogeneous educational group, high skilled workers. With little variation in years of education, empirically it is impossible to disentangle returns to education from the skilled rental rate Π_t^s.

The labor market is competitive with no search friction. Workers are paid their marginal product. Thus, wages are determined by the product of current equilibrium rental rates (Π_t^{cs} and Π_t^{ncs}) and individual occupation specific human capital (H_a^{cs} and H_a^{ncs}), which gives the standard Mincer equation. Π_t^s is determined in equilibrium, as discussed below.

$$\begin{aligned} W_{a,t}^s &= \Pi_t^s H_a^s \\ s &\in \{cs, ncs\} \end{aligned} \tag{2}$$

The market is assumed to be complete. Individuals can fully insure against risks, so no precautionary saving is required. Agents, hence, can be modeled as risk neutral. Adding saving decision here is unnecessary. When making educational decisions and certain types of occupational decisions

[20] There is no constant in the H_a because the constant is not separately identifiable from the equilibrium rental rates. In this specification, I assume the initial ability distribution is constant across years.Yearly variations in wages of new entrants are mainly attributed to changes in returns to human capital.

(becoming self-employed or an entrepreneur), which require some initial capital investments, saving decisions and financial constraints are indispensable components Castro and Ševčík (2016). Because financial constraints would distort agent's human capital decisions in that case. However, since no initial capital requirements are needed for taking a new occupation in the current model, trimming the saving decision keeps the model simple without introducing a bias.

And agents derive flow utility from wages and age specific idiosyncratic taste shocks η_a.[21]

$$u^{cs}_{a,t} = W^{s}_{a,t} + \eta_a \tag{3a}$$

$$u^{ncs}_{a,t} = W^{ncs}_{a,t} \tag{3b}$$

Taste shocks only appear in the flow utility of the CS occupations. One could have the taste shocks appearing in both flow utility functions. However, the two taste shocks would not be separately identifiable. It is the difference between these two taste shocks that are relevant to individual occupational choices. Equivalently, I choose to model it as a shock only in the CS occupations.

As shown in equation (4), taste shocks are transitory and are independent draws from a family of normal distributions whose variance changes along with age.

$$\begin{aligned} \eta_a &\sim \mathbb{N}\ (0,\ \sigma^2_{\eta_a}) \\ \sigma^2_{\eta_a} &= \sigma^2_{\eta} \exp(\gamma a) \end{aligned} \tag{4}$$

I expect γ to be negative because empirically more experienced individuals tend to switch occupation less often. Kambourov and Manovskii (2008) document a declining occupational mobility profile along with age. This parsimonious specification of age varying taste shocks together with the occupation specific human capital would potentially improve the model fit.

Individual Occupational Choices

Followed the notation of Lee and Wolpin (2006), let $\Omega_{a,t}$ be the vector of state variables at age a and time t, variables known that determine the remaining expected present value of lifetime utility. Given the structure of the model, the state space at any age a includes the current equilibrium skill rental rates ($\Pi_t = (\Pi^{cs}_t,\ \Pi^{ncs}_t)$), the future career prospects ($\Pi_t(e)$), the current occupation tenures ($x^{cs}_a,\ x^{ncs}_a$), innate ability (ϵ), and the realization of taste shocks (η_a). The evolution of state space $\Omega_{a,t}$, especially the way how natives form expectations, is discussed in the next section.

Given the information set $\Omega_{a,t}$, a native worker solve the following dynamic programming problem.

[21] I also estimate the model with log flow utility $u = \log(c) + \eta$, and the alternative specification doesn't make qualitative changes to the results.

The bellman equations of the two alternative value functions are as follows.

$$V^{cs}(\Omega_{a,t}) = \quad W^{cs}_{a,t} + \eta_a + \beta \mathbb{E} V(\Omega_{a+1,t+1}|d = cs, \Omega_{a,t}) \tag{5a}$$

$$V^{ncs}(\Omega_{a,t}) = \quad W^{ncs}_{a,t} + \beta \mathbb{E} V(\Omega_{a+1,t+1}|d = ncs, \Omega_{a,t}) \tag{5b}$$

Where $\mathbb{E}()$ indicates expectation, β is the subjective discount factor. This finite horizon dynamic discrete problem (DDP) is solved by backward recursion. The decision problem stops after retirement at age 65. To initiate this iteration, I specify the ending value functions for age 65 as

$$V^{cs}(\Omega_{65,t}) \quad = \quad W^{cs}_{65,t} + \eta_{65} \tag{6a}$$

$$V^{ncs}(\Omega_{65,t}) \quad = \quad W^{ncs} \tag{6b}$$

In each period, native workers choose the greater of the alternative value functions.

$$V(\Omega_{a,t}) \quad = \quad max\ \{V^{cs},\ V^{ncs}\} \tag{7}$$

Evolution of the State Variable

The state space of a native worker at age a and time t as stated before is

$$\Omega_{a,t} = \{a, x^{cs}_a, x^{ncs}_a, \epsilon, \eta_a, \Pi_t, \Pi_t(e)\}$$

where $\Pi_t(e)$ represents the future career prospects.

The evolution of age a, innate ability ϵ, and taste shock η_a is trivial. The innate abilityϵ is permanent heterogeneity; taste shocks η_a are transitory and idiosyncratic. Occupational tenures x^{cs}_a and x^{ncs}_a evolve endogenously. If a native worker is in the occupation s ($d = s$), this individual accumulates one additional year of occupational tenure according to the following rule, which has a return in the future and changes individual comparative advantages as well.

$$x^s_{a+1} = x^s_a + \mathbf{1}(d = s).$$

The current skill rental rates Π_t are equilibrium outcomes, which will be discussed in detail later in the equilibrium part. In order to make their decisions, individuals also need to forecast the future career prospects $\Pi_t(e)$.

Let $F(a+1, x^{cs}_{a+1}, x^{ncs}_{a+1}, \epsilon, \eta_{a+1}, \Pi_{t+1}, \Pi_{t+1}(e)|\Omega_{a,t})$ denote the distribution of these state variables in the next period conditional on the current state, I can split this joint distribution into three parts

$$\begin{aligned} &F(a+1, x^{cs}_{a+1}, x^{ncs}_{a+1}, \epsilon, \eta_{a+1}, \Pi_{t+1}, \Pi_{t+1}(e)|\Omega_{a,t}) \\ &\qquad = F^{\eta_{a+1}}(\eta_{a+1}) \quad F(a+1, \epsilon|a, \epsilon) \quad F(x^{cs}_{a+1}, x^{ncs}_{a+1}, \Pi_{t+1}, \Pi_{t+1}(e)|\Omega_{a,t}) \end{aligned} \tag{8}$$

Equation 8 implies that the processes for taste shocks, permanent heterogeneity and age are independent of the processes for occupational tenures and career prospects.

Expectation Forecasting the evolution of future skill rental rates

$$\{\Pi_\tau(e)\}_{\tau=t+1}^{\infty} = \{\Pi_\tau^{cs}(e), \Pi_\tau^{ncs}(e)\}_{\tau=t+1}^{\infty}$$

is a complex task because it depends on future sector specific productivity shocks, total inflows of skilled immigrants, and occupational composition of immigrants as well. Under rational expectations (Lee and Wolpin, 2006), the process of future career prospects should be the one that individuals make the best possible forecast conditional on all the available information in the current period.Thus, it requires one to specify all the above processes. Moreover, it also requires one to feed the agents equilibrium choices into the expectation, which implies the expectation processes to coincide with the realized processes. To make the model model tractable and computationally feasible, I assume that agents form a deterministic expectation, similar to a perfect foresight model.

$$\Pi_\tau^s(e) \quad = \quad \hat{\Pi}_\tau^s(o) \; \forall \tau > t \; s \in \{cs, ncs\} \tag{9}$$

where $\hat{\Pi}_\tau^s(o)$ denote the evolution of the observed market skill rental rates that are directly measured using the CPS data. In this current model, I assume that the evolution of $\hat{\Pi}_\tau^s(o)$ are perfect anticipated by the workers. Workers form expectations about future career prospects according to the evolution because it summarizes all relevant information needed for individual occupational choices. The career prospects reflect all relevant information including the future processes of the sector specific productivity shocks, flows of immigrants, cohort sizes, and the composition of immigrants.

This assumption implies that

$$F(x_{a+1}^{cs}, x_{a+1}^{ncs}, \Pi_{t+1}, \Pi_{t+1}(e)|\Omega_{a,t}) = F(x_{a+1}^{cs}, x_{a+1}^{ncs}, \Pi_{t+1}|\Omega_{a,t})F(\Pi_{t+1}(e)|\Pi_t(e))$$

The deterministic expectation rule reduces the computational burden but also introduces a bias. It is relevant to have some idea how far are my current estimates away from the rational expectation model. I answer this question without directly estimating the full rational expectation model. Instead, I consider two computationally feasible versions: (i) the perfect foresight model of which the estimated parameters are reported in this paper. Agents are assumed to have full information about the career prospects. (ii) the myopic model, where agents have essentially static expectations about the future skill rental rates. They incorrectly assume a steady state and that the current skill rental rates would last forever. In this myopic model, agents would be surprised by the arrival of new shocks and modify their expectation on a less frequent basis (the MIT shocks). These two versions are two extreme cases how agents would from expectation. In the perfect foresight model, agents have full information about the future processes, while in the myopic model agents have no information. The more realistic rational expectation would fall between these two polar extremes. Based on my estimation results, the estimates of two polar extremes are quantitatively close, which implies that the current estimates are not far from the rational expectation model.

Aggregate Native Labor Supply

Individual labor supply in efficiency units differs due to the heterogeneity in individual human capital (H_a^{cs} and H_a^{ncs}). To get the total labor supply for each occupational group, I first aggregate the labor supply for age group a, $NS_{a,t}^s$.

$$NS_{a,t}^s = \iint \mathbb{I}^s(\Omega_{a,t}) H_a^s(\ x_a^{cs},\ x_a^{ncs},\ \epsilon) \mathrm{d}F(x_a^{cs},\ x_a^{ncs},\ \epsilon|a,t) \mathrm{d}F(\eta_a) \tag{10}$$

$\mathbb{I}^s(\Omega_{a,t})$ is the indicator variable that occupation s is chosen at age a and year t. For the age group a in year t, there is a joint distribution of the innate ability and the occupational tenures $F(x_a^{cs},\ x_a^{ncs},\ \epsilon|a,t)$ which incorporates all relevant information about the entire history of skill rental rates, taste shocks and expectations of career prospects. Jointly with the distribution of the current taste shock $F(\eta_a)$, $F(x_a^{cs},\ x_a^{ncs},\ \epsilon|a)$ determines the aggregate labor supply for age group a.

Since the cohort population size also varies, to compute the aggregate labor supply, NS_t^s, I assign each cohort aggregate labor supply ($NS_{a,t}^s$) a weight proportional to his or her birth cohort size, $w_{a,t}$. As a result, the aggregate labor supply in occupation s at time t is

$$NS_t^s \quad = \quad \sum_{a=22}^{a=65} w_{a,t} NS_{a,t}^s \tag{11}$$

Where the weight, which is proportional to the cohort population size, is measured using the CPS data

$$w_{a,t} = \frac{N_{a,t}}{\sum_{i=22}^{i=65} N_{i,t}}.$$

The model predicts that the fraction of natives working in the CS occupations in age group a and year t has the following expression,

$$P_{a,t}^{cs} = \iint \mathbb{I}_{cs}(\Omega_{a,t}) \mathrm{d}F(x_a^{cs},\ x_a^{ncs},\ \epsilon|a,t) \mathrm{d}F(\eta_a). \tag{12}$$

3.1.2 Immigration Labor Supply

Immigrants accumulate occupation specific human capital through on-the-job learning by doing as with their native counterparts. However, they are faced with massive mobility barriers across occupations due to the lack of H-1B portability discussed in the previous section. Immigrants are high insensitive to changes in the market prices because of the massive mobility barriers. In this model, immigrants are allowed to accumulate occupation specific human capital but are not allowed to freely switch occupations. The supply of immigrants in each occupational group is assumed to be perfectly inelastic, and the actual quantity is regulated by the variations in the H-1B cap. The model takes the supply of immigrants directly from the data. In the CPS and the ACS data, we observe annual incomes for all full-time full year skilled workers as well as their citizenship status.

The skill rental rate $\hat{\Pi}_t^{s\star}$ paid to foreign labor is measured directly by using the annual income of new foreign entrants. New entrants have no occupation specific tenure ($x_a^{cs} = 0$, $x_a^{ncs} = 0$), and the mean of innate abilities is normalized to 0.

$$log(W_{i,t}^s) = log(\hat{\Pi}_t^{s\star}) + \epsilon_i^s$$

Let $\overline{log(W_{i,t}^s)}$ denote the mean log wage of the new foreign entrants. One can compute the skill rental rate for foreign labor directly $\hat{\Pi}_t^{s\star} = exp(\overline{log(W_{i,t}^s)}\ )$. For each skilled immigrants, his or her labor supply in efficiency units is back out by $H_{i,t}^s = \frac{W_{i,t}^s}{\hat{\Pi}_t^{s\star}}$. The total foreign labor supply in occupation s (M_t^s) is aggregated across the immigrant population. Taking the foreign labor supply directly from the data is less restrictive than it seems at first glance. It is flexible enough to take into account new foreign entrants and the human capital accumulation of the previous generations of immigrants.

To sum up the labor supply side, the model mainly focuses on the occupational choices by native workers in response to foreign labor competition. When working, individuals are paid their marginal products, and they accumulate occupation specific human capital, which maps into future wages. Forward-looking individuals could be interested in an occupation that pays a lower contemporaneous wage if the experience provides a high enough return in the future. Occupation decision, hence, affect and are affected by future career prospects, which are impacted by, but not limited to, inflows of immigrants. The skill rental rates Π_t are equilibrium outcomes that channel the effect of immigration towards native occupational choices and wages. If immigration put negative pressure on one occupational group s, there is a strong incentive for natives to switch to the other occupation. Likewise, it may also increase the wage level in one occupational group through task specialization, which can make this occupation more attractive.

3.2 Aggregate Production Function and the Demand for Labor

In each occupational group, there is an aggregate firm that produces a single output, combining two types of labor, native labor N_t^s and foreign labor M_t^s using a Constant Elasticity of Substitution (CES) technology. Capital is separable from labor. Representative firms solve static profit maximization problems in every period.

Firms use production technologies which are occupation specific. The profit maximization problem of the representative firm in the occupational group s is:

$$\max_{\{N_t^s, M_t^s\}} Z_t^s((1-\delta^s)(N_t^s)^{\rho^s} + \delta^s(M_t^s)^{\rho^s})^{\psi^s/\rho^s} - \Pi_t^s N_t^s - \Pi_t^{s\star} M_t^s \tag{13}$$

The production function presented in Equation (13) includes a CES aggregate of domestic and immigrant labor. Parameter$\delta^s$is connected to the factor shares, and ψ^s describes the curvature of the production function (return to scale), which is also closely related to the demand elasticity of skilled labor. Parameter ρ^s relates to the substitutability between the two types of labor. All the above parameters will be estimated from the data. The production function is general enough to allow two types of labor being imperfect substitutes or even complements. The sector specific aggregate productivity shock Z_t^s is assumed to be a log stationary process which allows for business cycle

fluctuations around the mean.

As noted above, I abstract from physical capital in the occupation specific production function. This is because of the data availability. There is no aggregate capital measure available for the STEM occupations. Due to the abstraction of physical capital, the interpretation of sector specific shock Z_t^s deviates from the conventional total factor productivity (TFP). One should view Z_t^s as a combination of the TFP and physical capital adjustments. The consensus in the literature is that overall estimated effects of immigration are non-negligible if physical capital does not react to immigration, and virtually zero if capital fully adjusts. In this model, physical capital is assumed to adjust with some sluggishness to immigration, and the actual adjustment is incorporated in the estimated Z_t^s process.

Equation (13) differs from the CES technology popularized in the immigration literature in the following three aspects: (1) the production functions are occupation specific, (2) the CES technologies are flexible enough to allow for imperfect substitutability and even complementarity between natives and immigrants and (3) the substitutability or complementary can vary across occupations to accurately capture the incentives to occupational by natives.

Adding occupations to the labor demand side is important. Occupational adjustment (specialization) is an important adjustment mechanism by natives to react to the labor market competition induced by immigrants (Peri and Sparber, 2009). In this paper, I endogenize the occupational choices by natives which requires the production functions or the labor demand side to accurately generate the wage heterogeneity across occupations. The skill-cell approach in Borjas (2003) and Ottaviano and Peri (2012) captures the heterogeneous impacts of immigration across skill cells. However, even within each cell, there is a massive distributional effect of immigration. Occupational specific production functions generate wage heterogeneity, which allows the model to quantify the heterogeneous effects among natives within a skill cell. Additionally, previous papers (Llull, 017a; Dustmann et al., 2012) note that immigrants downgrade upon entry into the destination country. As a result, they do not compete with the natives that share the same observable skills, but, instead, with those that work in the same jobs. As indicated above, occupation is a more accurate measurement of labor competition.

The current production functions display sufficient flexibility that allows for both imperfect substitutability and complementarity. Instead of making some disputable assumption, I estimate the substitutability directly from the data. Natives and immigrants may be imperfect substitutes in production because of their different skills. And it is important to account for this imperfect substitution (Ottaviano and Peri, 2012). The flexibility of the production functions expands the scope within which we discuss the impacts of immigration in the literature. It goes beyond imperfect substitution(Ottaviano and Peri, 2012; Bound et al., 2015; Llull, 017a). Inflows of skilled immigrants can benefit natives in one occupation if natives and immigrants are complements in the production function. Limiting the analysis to perfect or imperfect substitution with a normal downward sloping labor demand curve focuses the discussion exclusively on the potential losses of immigration. The negative effects of immigration on wages and inequality would mechanically be overstated.

The substitutability of native and immigrants can vary across occupations. The current general equilibrium model can correct the estimation bias introduced by the assumption of fixed native labor supply by explicitly accounting for occupational adjustments by natives. Allowing for the occupation specific production functions captures the asymmetric effects of immigration across occupations, generates further wage heterogeneity across occupations, which allows later to quantify the distributional effects of immigration.

The FOCs with respect to native labor deliver the implicit demand functions for native skilled labor.

$$\Pi_t^s = Z_t^s \psi^s (1-\delta^s)((1-\delta^s) + \delta^s (\frac{M_t^s}{N_t^s})^{\rho^s})^{\psi^s/\rho^s - 1} (N_t^s)^{\psi^s - 1} \tag{14}$$

The native labor demand ND_t^s is implicitly determined by equation (14). As in Equation (14), I write the native skill rental rates as a function of foreign labor supply M_t^s. (marginal product) for natives goes up when foreign labor increases whenever $\psi^s > \rho^s$. This is the case when natives and immigrants are complements in production. Not allowing for possible complementarity would automatically put downward pressure on native wages whenever there is an inflow of immigrants, which rules out the possible benefits of immigration by construction.

3.3 Equilibrium

A dynamic general equilibrium can be characterized by a system of equations representing the agent's labor supply decision (value functions, choice functions, agent's expectation), the firm's labor demand decision (demand functions, technology process), and market clear conditions. In particular, the equilibrium skill rental rate series $\{\Pi_t\} = \{\Pi_t^{cs}, \Pi_t^{ncs}\}$ in this model has to satisfy the following conditions:

1. Based on the skill rental rates $\{\Pi_t\}$ and future career prospects $\{\Pi_t(e)\}$, the native labor supply NS_t is the aggregation (equation 10 - 11) of individual labor supply decision, which is fully described by equation 5 - 8.

2. Firms maximize profit given skill rental rates $\{\Pi_t\}$. The labor demand ND_t is determined as in equation 14.

3. $\{\Pi_t\}$ clears the skill markets in every period.

$$NS_t = ND_t \ for \ all \ t$$

where $NS_t = \{NS_t^{cs}, NS_t^{ncs}\}$ and $ND_t = \{ND_t^{cs}, ND_t^{ncs}\}$.

4 Data, Identification and Estimation Method

4.1 Data

To estimate the model, I fit simulated moments to statistics computed with micro-data obtained from three data sources: the Current Population Survey (CPS), the American Community Survey (ACS), and the Panel Study of Income Dynamics (PSID). The fitted data moments include information on choice probabilities, gross occupational mobility, mean and the variance of wages conditional on occupational choices. The cohort sizes and career prospects are directly estimated from the CPS.

I follow the measure of immigrants as in Bound et al. (2015). Immigrants are defined as individuals born abroad and migrated to the US after the age of 18. I explore the longitudinal features of the PSID to get information on the age profile of gross occupational mobility between the two mutually exclusive occupational choices. I follow the method proposed by by Kambourov and Manovskii (2008) which uses the Retrospective Occupation-Industry Supplemental Data Files to correct for classification errors in occupation coding. The pattern of gross occupational mobility found is very similar to the mobility pattern found in Kambourov and Manovskii (2008) at the one digit level. Occupational mobility declines sharply as people age.

There is another aspect of the data that are worth discussing in more detail here. There is a discrepancy between the model and the data. It occurs because I fit a life-cycle model to repeated cross-sectional data. The model presented in the previous section is a life cycle occupational choice model, where the targeted moments are confounded with cohort effects as well. Individuals at age 65 in the year of 2000 entered the labor market in 1956. They faced very different market conditions when making their educational and earlier occupational choices. Consequently, their human capital investment decisions are drastically different from other birth cohorts, i.e., workers age 65 in the year of 2015. A single cohort model is incapable of capturing the empirical data patterns.

To address the issue, in the estimated version of the model, I explicitly model multiple birth cohorts. Different cohort groups share identical innate ability distribution, and their human capital is accumulated in a similar fashion, but they experience different market conditions. To resolve the inconsistency, it is necessary to solve for the optimal career decision rules for different birth cohorts, i.e., the cohort-specific set of value functions. Specifically, different birth cohorts solve life-cycle choice problems subject to different market conditions and career prospects. In Appendix E, I show how to combine the simulation results of different cohorts to match the empirical moments presented below.

4.2 Estimation by Simulated Method of Moment

I estimate parameters of the model by minimizing the distance of simulated moments to their empirical counterparts. Simulated statistics are obtained from simulating the behavior of cohorts of 5000 natives. Each simulated cross-section includes 5000 observations, which are weighted using data on cohort sizes and immigrant population sizes. The moments matched describe the gross occupa-

tional mobility, occupation employment share, and wage distributions over the life cycle and across cohorts. During the estimation process, I combine different data sources, which all have different sample sizes. The standard asymptotic results don't apply here. In order to get the correct inference on the estimates, I follow the modification of the standard asymptotic results proposed by Görlach (2017). The new asymptotic distribution explicitly addresses the issue with multiple samples. See Appendix D for details.

4.2.1 Choice of Moments

The data moments to be matched are as follows where a is between 22 to 65 and t covers the period from 1994 to 2013:

- *Age Profile of Occupation Employment Share*
  $p_{a,t}$ =proportion of age a native STEM workers working in the CS occupations in year t.
- *Conditional Wage Distribution*
  1. First moments: the mean wage of occupation s, in age group a and year t, $\bar{W}^s_{a,t}$.
  2. Second moments: the variance of wages of occupation s, in age group a and year t, $var^s_{a,t}$.
- *Age Profile of Gross Occupational Mobility*
  Mob_a =fraction of age a workers switching between the CS occupations and the other-STEM occupations.

4.2.2 Estimation Procedure

The parameter space Θ in the model can be naturally divided into two subsets $[\Theta^s|\Theta^d]$. Θ^scontains parameters that determine the native labor supply, including parameters governing human capital formation, individual preferences, and ability heterogeneity. Θ^d includes parameters entering the occupation-specific production functions. The supply and the demand side are treated as two separate parts. The key elements connecting these two subsets are the equilibrium skill rental rates. I use a two-step estimation procedure similar to the one proposed by Jeong et al. (2015) that separates the supply side estimation from that of the demand side. The two-step estimation is less efficient, but it significantly reduces the computational requirements.

I assume that fundamentals of the native labor supply have remained constant. The innate ability distribution, preferences, and human capital production function remain unaltered. Variations in the conditional wage distribution and occupational employment share are attributed to changes in skill rental rates.[22] The first stage selects the fundamental parameters governing labor supply side to match natives' conditional wage distributions, occupational employment shares, and gross occupational mobility patterns. The estimation of the first stage delivers the measure of effective labor

[22] In the traditional demand and supply framework, the labor supply curve is fixed over the recent two decades. Any skill rental rate variation is simply movements along the supply curve. Consequently, the labor supply side is well identified.

supply at the individual level, which later aggregate to obtain the labor supply ($\hat{N}_t^s$).

Next, in the second stage, I combine the first stage outputs ($\hat{N}_t^s$) with observed quantities ($\hat{\Pi}_t^s$,$\hat{\Pi}_t^{s\star}$and $\hat{M}_t^s$) to estimate the production functions using maximum likelihood. The production parameters are identified using the time variation in skill rental rates and in aggregate labor quantities. The partially exogenous supplies of skilled immigration provide variations needed to identify Θ^s.

4.2.3 Recover the Career Prospects

As mentioned in the model section, agents can perfect anticipate the future career prospects in both occupations, which is directly measured using the CPS data. The career prospects$\hat{\Pi}_\tau^s(o)$ are recovered using the method known in the human capital literature as the flat spot method (Heckman et al., 1998). This method is based on the fact that most optimal human capital investment models have the feature that at some point in the working life-cycle, optimal net investment is zero (Bowlus and Robinson, 2012). Human capital of a given cohort over those years is constant. For a cohort in the flat spot area of their human capital profile, any changes in wages purely reflect changes in skill rental rates. By applying the flat spot method, the future career prospects are directly identified from the CPS data.

Figure 4 plots the evolution of skill rental rates in both occupations. Note these two series are subject to normalization. For the illustration purpose, the skill rental rates in 2010 are normalized to unity in both occupations in Figure 4. The detailed data cleaning and smooth techniques are discussed in Appendix C.

4.3 Identification

In this section, I illustrate how some of the crucial parameters are identified.

The identification of the innate ability can be considered as an application of Heckman and Honore (1990) in a dynamic setting. In their discussion, under the normality assumption, the ability distribution can be identified by using a single cross-sectional data. The cross-sectional data need to contain information about occupational choices and conditional wage distributions. Closely related to Heckman and Honore (1990), I make a distributional assumption about the innate ability (a bivariate normal distribution) and explore the variations provided by natives' occupational choices and their conditional wage distributions. In Heckman and Honore (1990), the sufficient condition for identification is one single cross-sectional data set. In this paper, I use repeated cross-sectional data for 20 years and explore additional time variations. The time variations of skill rental rates are particularly important which provide me an additional source of identification. Since the underlying innate ability distribution is assumed to be time invariant, any variation in skill rental rates over years directly maps to the observed data variations in the employment shares of new entrants across occupations, which in turn reveals the general shape of the ability distribution. Sufficient variations

in skill rental rates over years are desirable. The current study covers the period from 1994 to 2013, which provides sufficient variations in skill rental rates because of the Internet boom and bust occurred within these years. Sufficient variations in skill rental rates and the resulting large shifts in occupational employment shares explore the ability distribution extensively.

In order to identify the labor supply side, I make somewhat restrictive assumptions about the fundamentals of the native labor supply side, e.g. the innate ability distribution is time invariant. How realistic are these assumptions? We don't observe the occupation specific innate ability, otherwise, we can directly measure and control for it. Thus, I can only provide some indirect evidence to support this assumption. I check three ability or skill measures of the American youth: (1) the cognitive ability, (2) the non-cognitive ability, and (3) the social ability, across the National Longitudinal Survey of Youth (NLSY) 79 cohort and the NLSY97 cohort. To make the ability measures comparable across two data sets, I apply the method proposed by Altonji et al. (2012). I plot the cross sample comparison of these three measures in Figure 11. For all three skill measures, the two distributions for different cohorts are quite similar, but there is some limited suggestion that the NLSY97 cohort is a bit stronger in all three dimensions than the NLSY79 cohort. The above analysis shows some evidence that the ability or skill distributions are quite similar across cohorts over 20 years. If there were any distributional difference across cohorts, it is likely that the difference is relatively limited. Ignoring the time variations in the ability distributions, the current model attributes all the observed changes in employment shares to the changes in skill rental rates. This would potentially imply overestimation of the price variations, thus underestimating the factors that drive persistence of occupational decisions in the model, and eventually overestimating the economy's overall ability to absorb shocks induced by immigrants.

The identification of the human capital functions comes jointly from the variations in occupational employment shares when the market skill rental rates change over time and the life-cycle wage dynamics. Specifically, when the rental rates change, the net flow of workers across occupations provides information about the transferability of occupational specific human capital, i.e., the relative magnitudes of returns to different occupational tenures. In the case, where the other-STEM occupations value more individual occupational tenure in the CS occupations, the increasing skill rental rate for the other-STEM occupations would cause a greater net flow of native workers from the CS occupations to the better alternative. Not only we expect to observe a greater aggregate net flow, we also expect the age composition of switchers tilts more towards the more experienced workers. In addition to the time variations, within one year when the skilled rental rates are constant, wage differences among workers attribute to different returns to occupation tenures and to the total labor market experience. Consequently, the income age profiles are informative in terms of the shape of the human capital function.

The identification of the taste shock parameters mainly comes from the gross occupational mobility profile. The age specific taste shocks are identified to explain the residual variations in the gross occupational mobility, the remaining variations after purging of the effects of the occupational specific human capital.

For the labor demand side, I estimate all the parameters once using maximum likelihood. In this part, I illustrate step by step how the demand parameters are identified in regression style for illustration purpose. The inputs of demand side estimation are the skill rental rates and the equilibrium labor quantities in efficiency units for both native and foreign workers. For native workers, the skill rental rates $\hat{\Pi}_t^s$ and the aggregate labor supply $\hat{N}_t^s$ are either directly estimated or aggregated using estimates from the individual occupational choices, whereas skill rental rates $\hat{\Pi}_t^{s\star}$ and aggregate labor supply $\hat{M}_t^s$ of immigrants are directly measured from data. Given the CES functional form, the share parameter δ^s and the parameter related to substitutability ρ^s can be identified by exploring time variations between the relative quantity and the relative rental rate. The regression of how δ^s and ρ^s are identified is illustrated by the following estimation equation 15

$$\log(\frac{\hat{\Pi}^s}{\hat{\Pi}^{s\star}})_t \quad = \quad \log(\frac{1-\delta^s}{\delta^s}) + (\rho^s - 1)\log(\frac{\hat{N}^s}{\hat{M}^s})_t + \mu_t \tag{15}$$

To identify the return to scale parameter ψ^s and the technology process Z_t^s, I impose one additional identification assumption on the Z_t^s process. I assume Z_t^s is a log stationary process. Let $\bar{z}^s$ and Δ_t^s denote the long-run average of $\log(Z_t^s)$ and the deviation from the long-run average respectively. $\bar{z}^s$ and ψ^s are separably identifiable, where $\hat{y}_t^s$ and $\hat{x_t}^s$ are functions of known parameters and quantities.[23]

$$\log(\hat{y}_t^s) \quad = \quad (\bar{z}^s + \log\psi^s) + \psi^s \hat{x_t}^s + \Delta_t^s \tag{16}$$

Once we identify $\bar{z}^s$ and ψ^s, the technology process $Z_t^s = \exp(\bar{z}^s + \Delta_t^s)$ is identified observation by observation as well. In all the counterfactual experiments, representative firms are assumed to know the technology processes in both occupational groups.

5 Results

In this paper, I fixed the annual discount factor to the value 0.95 which is within the reasonable range in literature.

5.1 Estimation Results

Table 2 presents the estimates of Θ^s, the parameters from labor supply side. All parameters are statistically significant at 1% level. For the parameters related to returns to occupational tenure, the first year of the CS tenure augments CS human capital by about 7.56% with little attenuation in the rate of increase at higher years of experience. The first year of the other-STEM tenure increases the

[23] where $\hat{x_t} = \frac{1}{\hat{\rho}^s}\log[(1-\hat{\delta}^s)(\hat{N}_t^s)^{\rho^s} + \hat{\delta}^s(\hat{M}_t^s)^{\rho^s}]$ and $\hat{y}_t = \frac{\hat{\Pi}_t^s}{(1-\hat{\delta}^s)(\hat{N}^s)^{\hat{\rho}^s}}$.

other-STEM skill by 8.51%.[24] Both occupational groups value tenures in the other occupations but to a lesser extent. An additional year of the CS (the other-STEM) tenure augments the other-STEM (the CS) skill by less than 6.95% (7.60%). The value of estimates implies that the two occupations are close alternatives measured by the transferability of occupation specific human capital. This suggests that when facing foreign labor competition, native CS workers are very likely to choose the other-STEM occupations as their alternative occupations to counteract the negative impacts. Transferability of occupation specific human capital varies across occupational groups. In (Keane and Wolpin, 1997), the white-collar sector is found to discount tenures in the blue-collar occupations a lot. This is not the case with the occupational groups in this paper.

The two dimensions of the innate abilities are mildly negatively correlated with a correlation coefficient equal to -0.17. The negative sign indicates that those natives who have a talent for working in the CS occupations are less efficient workers when employed in the other-STEM occupations. Different ability endowments (initial comparative advantages) of natives lead them to different choices of occupations. The innate ability (permanent heterogeneity) makes individuals more likely to persist in their choices. The negatively correlated abilities restrict the extent to which natives can resort to occupational adjustments when faced with foreign labor competition. However, with a small negative correlation coefficient (-0.17), the ability barrier of occupational mobility is limited, which again suggests that the CS and the other-STEM occupations are close alternatives. Permanent heterogeneity is an important determinant of individual occupational choices and income inequality. Omitting such heterogeneity could lead to a substantial bias in other fundamental parameters. This would potentially underestimate cross-experience effects and overestimate returns to occupational tenures, which eventually would distort the occupational adjustments by natives and bias the estimated impact of immigration. As an aggregate of a larger set of occupations, the other-STEM occupational group has a larger variance. For both occupational groups, there exists substantial heterogeneity in innate ability across individuals, which is an important determinant of both the income inequality and individual occupational choices (Keane and Wolpin, 1997). One standard deviation increase in the innate ability is associated with a 23% and 35.5% increase in annual wages for the CS and the other-STEM occupations respectively.

One of the most important differences between the production function in Equation (13) and the nested-CES production function used in the immigration literature (Borjas, 2003; Ottaviano and Peri, 2012; Llull, 017a; Bound et al., 2015) is that Equation (13) allows for complementarity between natives and immigrants. Moreover, the substitutability and complementarity are also allowed to vary across occupations. The estimates of the labor demand side are presented in Table 3. First, both occupations display decreasing return to scale with estimated values between 0.58-0.6.[25] The

[24]Holding everything else being constant, Kambourov and Manovskii (2008) find smaller returns to occupation tenures. They find one year of occupational tenure is associated with a 2.4% - 4% increase in wage. The major reason for the difference is that in Kambourove and Manovskii's (2008) specification they also include industry tenures, job tenures, and general labor market experiences. In this model, without job and industry tenure, occupational tenure should account more for individual's wage growth.

[25]The occupation specific production function displays decreasing return to scale. This is because there is no capital in my production. The estimated return to scale parameter is equivalent to the labor share in U.S. According to the BLS, the labor share is approximately 0.58 in the year of 2016.

more interesting result is about the substitutability between foreign and native labor across occupations. For the CS occupations, immigrants and natives are imperfect substitutes with an elasticity of substitution $\frac{1}{1-\rho} = 6.9$, suggesting that the negative impacts of skilled immigrants would be very limited in the CS occupations. This estimates are consistent with the ones in Card (2009) and Piyapromdee (2015) in the literature. Unlike the previous papers, I find the existence of complementarity in the other-STEM occupations. Papers that assume only perfect substitutability or some extent of imperfect substitutability would by construction eliminate this possibility. Bound et al. (2015) assume that native and foreign workers are perfect substitutes in a decreasing return to scale production function. Skilled immigrants in their specification, by construction, would crowd out native workers and impose negative effects on wages of native workers. The production parameters of the other-STEM occupations in this paper draw attention to the potential benefit of skilled immigrants, which has been understudied in the literature. Inflows of foreign workers to the other-STEM occupations would result in wage gains rather than wage losses for native workers, and consequently generate a crowding-in effect rather than a crowding-out effect on native employment.

5.1.1 Possible Source of Complementarity

Estimates in this paper indicate that skilled natives and immigrants are imperfect substitutes in the CS occupations but are complements in the other-STEM occupations. The substitutability clearly varies across occupations, which would have drastically different policy implications. But before moving to the counterfactual exercises, it is interesting to understand what causes the cross occupational difference first.

Workers are heterogeneous and are equipped with different skills while jobs have different task contents or work activities. Peri and Sparber (2009; 2011) find that high skilled immigrants specialize in occupations demanding quantitative and analytical skills, whereas their native born counterparts specialize in occupations requiring interactive and communicative skills. They find strong evidence supporting occupation specialization of native and foreign workers. In this paper, I propose a similar argument but push it into a finer level. There exists possible task specialization of foreign and native workers even within an occupation. I rely on O*NET data to provide some preliminary evidence supporting my argument.

O*NET provides a comprehensive list of work activity or task measures, which can be classified into four major categories, information input, mental processes, work output, and interacting with others. Different occupations involve different work activities and to different intensity. As shown in Figure 12, the CS occupations involve intensively one major type of work activities, the mental processes. The other-STEM occupations, e.g., the chemical engineer, involve more diverse types of work activities in the task space. These occupations require workers to interact more with others and to spend more time on information processing. When more types of tasks or work activities are intensively involved in one occupation, this leaves plenty room for task specialization of labor with different comparative advantages. Inflows of skilled immigrants in one occupation, who have

comparative advantages in performing quantitative and analytical tasks, will cause skilled natives to reallocate their task supply toward performing interactive and communicative tasks.[26] Task specialization within one occupation shifts out the production frontier and gives rise to the observed production complementarity.[27]

The remaining parameters of the model, which are crucial for the model to fit individual choices, wages and occupational mobility observed in the data, are also reasonable and in line with the literature. As expected, the variance of the age specific taste shocks decreases with age, indicating that occupation specific human capital alone can not capture the sharp decline in the gross occupational mobility profile.

5.2 Sample Fit

5.2.1 In-sample Fit

Figure 5 - 6 depict a snapshot of the model fit in the year 2000. The model fit of the log income profile is remarkable. It captures the curvature of income profiles at the beginning and also the flat spot in the latter part of one's career. The third polynomial in human capital function helps to improve the model fit of the age profile.[28] The variance of log wage presents the U shape pattern which has already been documented many times in the literature. There is no mechanism in the current model that generates this U-shape variance profile. As a result, the model only fits the level rather than the age profile of wage dispersion. Adding learning to comparative advantages can help generate the increasing variance of wages with labor market experience (Gibbons and Waldman, 1999; Papageorgiou, 2014). Adding learning to the current model would generate more gross occupational mobility. Along with on-the-job learning by doing, this extra randomness in individual occupational choices generated by information friction would mute the impact of the permanent ability heterogeneity and reduce the persistent of occupation choices. Adding learning to the current model would amplify the effectiveness of occupational adjustments by natives.

[26] Besides the outward shift of production frontier, the reallocation of task supply could bring additional benefits to skilled native workers.According to Deming (**?**), the U.S. labor market increasing rewards social and communicative skills.

[27] I am not able to formally test the existence of task specialization within one occupation. In the publicly available survey data, there is occupation code at the individual level. But there is no detailed measure about what types of tasks being performed and what is the corresponding intensity at the individual level. The O*NET only provides task measures at occupation level.

[28] I also estimate another version of this basic model. Rather than specifying a human capital function with a third order polynomial, in the other vision, I specify a random accumulation process for occupation specific tenure. In that process, occupation tenures are accumulated in a stochastic way that the probability of acquiring an additional unit of tenure is a decreasing function of age. When agents spend one period in occupation s ($d_s = 1$), individuals randomly accumulate occupation specific tenures according to the following rule.

$$x'_s = \begin{cases} x_s + 1 & p = \exp(-\gamma_s x_s) \\ x_s & 1 - p \end{cases}$$

My estimates indicate that the accumulation process approximately stops around age 45-50 with minor differences across occupations. This is consistent with the empirical findings of the flat spot area in income profiles. The random accumulation process also captures the curvature of the income age profiles.

The model explicitly incorporates the cohort effect. With multiple birth cohorts, the model is able to match the native employment shares across occupations over age. The parsimonious taste shock specification with the occupation specific human capital generates a sharp decrease in gross occupational mobility over age which outperforms a model with only occupation specific human capital and fits better to the observed patterns.

For the remaining 19 years, the model fits the corresponding targeted moments well. Figures vary slightly from year to year. But there is no qualitative difference. Overall this simple model is able to capture the patterns of natives' conditional wage distributions, employment shares, and their gross occupational mobility.

5.2.2 Out of Sample Fit

The previous section provides some evidence that the model fit of the targeted moment dimension is good. And it is reassuring to explore the out of sample fit of the model presented in this paper. This section presents some additional exercises that provide further validation of the model.

First, I take the ratio of the two equilibrium skill rental rates $\frac{\pi_{cs,t}}{\pi_{ncs,t}}$ and plot this time series on the left side of Figure 7. I put the Nasdaq composite index on the right side in Figure 7. The equilibrium relative skill rental rate basically reproduces the key features of the Nasdaq index, which are not directly targeted. The relative skill rental rate series peaked around 2000 before the dot-com bust hit. After the bust, it was gradually recovering until 2007 when the financial crisis occurred. Recently it has stayed on an uphill track for about 6 years. The pattern of relative skill rental rate mimics closely the Nasdaq composite index. The correlation coefficient between these two series is very high (0.81).

Next, taking the changes in the relative skill rental rate as given, which are the fundamentals that drive natives' choices of occupations and fields of study, how sensitive are the natives in response to these price variations? In Figure 8, I plot the share of 22-year-old native workers who choose to be a computer scientist from 2000-2013 using the ACS with the model predicted relative skill rental rate over the same time. There is no such strong correlation. But individual occupational choices are closely related to their formal decisions on fields of study in college. This is especially true for the STEM workers since the STEM occupations require years of specialized training. College graduates have chosen their fields of study before they formally enter the labor market. As a result, when I move the relative skill rental rate series ahead by 3 years (1997-2010), there appears this strong positive correlation (see Figure 9). This time lag suggests that it is the current market prices, rather than the expectation, that matters more for individual's major choices.

After the initial moves (choosing fields of study), natives remain alert to price variations. However, switching occupation becomes less favorable and less frequent in the latter part of one's career. Agents accumulate occupation specific human capital through on-the-job learning by doing, and thus the implicit switching cost is higher for the more experienced workers. My model predicts

that the increasing switching cost profile makes the more experienced workers less sensitive to price changes. Figure 10 plots the relative skill rental rate and the employment shares of three different birth cohorts. There are two features emerged in Figure 10. First, there is a level effect that cohort groups who enter the market initially facing higher relative skill rental rate $\frac{\pi_{cs}}{\pi_{ncs}}$ maintain a higher employment share in the CS occupations throughout the subsequent periods. The 1978 cohort graduated from college in the year of 2000, the peak of the Internet boom. The CS employment share of this cohort stays unambiguously higher than the 1986 cohort who entered the labor market at the lowest point of the relative skill rental rates. This is consistent with the conclusion of Kahn (2010) that there exist some long-term labor market consequences of graduating from college in a bad economy. Here, graduating at different phases of the industry cycle has a persistent effect on individuals occupational choices. Second, the correlation coefficients between the employment share of the CS occupations and the relative skill rental rate vary across birth cohorts. The correlation coefficients across birth cohorts are presented in Table4. The correlation coefficient is higher for younger workers, 0.89 for the 1986 cohort as opposed to 0.41 for the 1978 cohort. Younger workers are subjected to less mobility friction measured by the amount of occupation specific human capital (implicit switching cost), are more sensitive to changes in market prices, and are more able to counteract the initial impacts of immigrants.

To sum up, in this section, I first show that equilibrium skill rental rates capture the salient features of the relevant market over the period of interest. Furthermore, I show that workers, especially inexperienced ones, can re-optimize their occupational choices to counteract the impacts of immigrants. Occupational mobility is indeed the relevant adjustment margin that helps to mitigate and diffuse any impact of foreign labor competition. Ignoring this adjustment by natives will introduce a substantial bias into the estimated impact of immigrants.

6 Counterfactual Exercises

I use the estimated model to simulate different counterfactual economies from 1994-2013 to evaluate the labor market effects of immigration, to study different types of immigration policies, and to study the key features of the current model. The exercises consist of hitting the calibrated model with the same sector specific technological shocks. Specifically, conditional on the estimated production parameters $\{\psi^s, \delta^s, \rho^s\}$ and the equilibrium skill rental rate $\{\Pi_t^s\}$ and the equilibrium labor supply $\{M_t^s, N_t^s\}$, I am able to recover values of the Z_t^s process during the period of 1994-2013:

$$Z_t^s = \frac{\Pi_t^s}{\psi^s(1-\delta^s)((1-\delta^s)+\delta^s(\frac{M_t^s}{N_t^s})^{\rho^s})^{\psi^s/\rho^s-1}(N_t^s)^{\psi^s-1}}$$

In all the following exercises, I assume that the economy is hit by these series of sector specific shocks. And these processes are policy invariant. Moreover, each agent is hit by the same realization of taste shocks that are also policy or exercises invariant to purge of the additional randomness in individual occupation choices.

6.1 Policy Relevant Exercises

In this section, to evaluate the labor market effects of immigration, I compare baseline simulations with simulations of counterfactual economies without the massive increase in skilled immigrants over the past twenty years. Specifically, I explore the effects of two different types of immigration policies: (1) a cap on total skilled immigrants, (2) a selective immigration policy that optimizes the occupational mix of immigrants. The outcomes of interest are the equilibrium skill rental rates and employment share of native skilled workers. In these exercises, the native workers can freely switch their occupations.

6.1.1 A Cap on Total Skilled Immigrants v.s. Selective Immigration Policy

In the first counterfactual exercise, I simulate a counterfactual economy in which the stock of foreign CS workers is fixed at its 1994 level whereas the stock of skilled immigrants in the other STEM occupations follows its actual observed path.[29] The purpose of this exercise is to assess that to what extent the rapid growth in the recruitment of foreign computer scientists affected outcomes of native workers in the STEM occupations.

Figure 13 depicts the counterfactual exercise and its results. In the top right panel, I present the resulting impact on the skill rental rate for the CS workers. In this counterfactual economy, the skill rental rate for the CS occupations would be higher. As illustrated in the bottom left panel, the native labor supply in the CS occupations would also be higher. One advantage of a general equilibrium model with multiple occupations is that it enables us to study broader effects of foreign computer scientists. For instance, we can evaluate the impacts of foreign computer scientists on the wages of native workers in the other-STEM occupations. In the second row of Table 5, I compute these effects.[30] Had the foreign CS workers been fixed at its pre-boom level, the skill rental rate in the CS occupations would have increased by 2.52%. The cap placed on the total labor supply of foreign CS workers would also benefit natives working in the other-STEM occupations. The skill rental rate would have increased slightly by 1.37% for the other-STEM occupations. This spill-over effect is primarily attributed to the occupational reoptimization of native workers. When experiencing less competition from immigrants in the CS occupations, the immediate result is that the skill rental rate for natives would increase in the CS occupations. Skilled natives who once didn't have any comparative advantage working as a computer scientist now would find it beneficial to switch to the CS occupations. This leads to an increase in the native labor supply in the CS occupations, a decrease in the native labor supply and consequently an increase in the equilibrium rental rates in the other-STEM occupations. The simulation result confirms the proposed channel above. On average the native labor supply in the CS occupations would grow by 5.24%, while the native labor supply in the other-STEM occupations would reduce only by 2.81%. The asymmetricity in labor supply changes across occupations implies that those switchers are only marginally better working

[29] Bound et al. (2015) adopt the same counterfactual settings as this one.

[30] The number reported is a 20-year average of percentage changes between counterfactual data and the real data.

in the other-STEM occupations in the actual economy.

Bound et al. (2015) do similar counterfactual exercises and find that rental rates increase by 2.8-3.8% while native labor supply increases by 7%-13.6% in the CS occupations. I find instead a more muted effect. According to my simulations, the skill rental rates increase by 2.5% and 1.4% in the CS and in the other STEM occupations respectively. Three major modeling factors in the current model setting are responsible for this limited effect. First, native and foreign workers are imperfect substitutes in the CS occupations. The lack of substitutability limits the extent of foreign labor competition. As a result, the model in this paper only generates small price variations in the CS occupations and an even smaller price change in the other-STEM occupations. Second, natives are heterogeneous in terms of their productivity working in different occupations (occupation specific human capital). Those who would switch in the counterfactual economy are not as productive as always-takers in the CS occupations. Labor supply measured in efficiency units accounts explicitly for individual productivity difference. Thus we see the native labor supply in efficiency units only increases by about 5 % in this counterfactual economy as opposed to 11.3% when measured by the number of workers. In addition, the occupation specific human capital in the current model adds some additional friction to occupational mobility, which would restrict the natives' mobility responses.

The second experiment, presented in the first row of Table 5, reduces the labor supply for both the CS and the other-STEM occupations to their 1994 levels. Skill rental rates for both occupations would still increase, but the magnitudes are smaller compared to the first counterfactual. The skill rental rates would increase by 2.51% and 1.22% for the CS and the other-STEM occupations respectively. Note that there are in total less foreign workers in the current counterfactual economy compared with the previous one. Surprisingly, native workers are instead worse off compared to the first counterfactual economy. As illustrated in the labor demand side, the skill rental rate (marginal product) for native workers in the other-STEM occupations is

$$\Pi_t^{ncs} = Z_t^{ncs}\psi^{ncs}(1-\delta^{ncs})((1-\delta^{ncs}) + \delta^{ncs}(\frac{M_t^s}{N_t^s})^{\rho^{ncs}})^{\psi^{ncs}/\rho^{ncs}-1}(N_t^{ncs})^{\psi^{ncs}-1}$$

The skill rental rate is an increasing function of foreign labor supply if natives and immigrants are complements in production($\psi^{ncs} > \rho^{ncs}$). My estimates confirm the existence of complementarity in the other-STEM occupations. Allowing more foreign labor supply in this occupational group where complementarity between foreign and native labor exists actually would benefit all native workers.

The difference results in the previous counterfactual exercises help to make the following two points. First, it is important to take a general equilibrium approach when assessing the impacts of immigration. If the occupation mobility mechanism identified by this paper is a viable channel and native workers are sensitive in terms of occupational mobility, the occupations where native workers move to matter and need to be modeled explicitly. Furthermore, to have a comprehensive evaluation of the wage impacts of immigration, the indirect impact or the spillover effect on the destination occupations should also be properly measured.

Second, the previous two experiments also shed some light on the welfare impacts of two types of immigration policies, the overall cap on quantity and the selective immigration policy based on

occupations and fields of study. One example of the selective immigration policy is the point-based system employed by Canada that grants more entries to workers in special occupations. Another example of this type is the OPT period in the US which functions in a similar manner that grants more entries to workers in the STEM occupations. On the other hand, the H-1B program resembles an overall cap that controls the total number of skilled workers entering the U.S. labor market. The previous simulation indicates that a selective immigration policy based on occupations and fields of study have the potential to outperform the overall cap and to achieve higher welfare for natives. Optimizing occupation mix of immigrants and directing skilled immigrants towards occupations where complementarity exists would benefit incumbent natives in these occupations by shifting outwards the production frontier. Outsiders can also share the gains through labor mobility.

6.2 Mechanism Related Exercises

6.2.1 Explore the Demand Parameters

The occupation specific production function is crucial in terms of identifying the impacts of skilled immigration and providing policy suggestions. This raises the question of how robust these results are to variations in how the production function is specified. In this part, I explore different counterfactual economies in which the labor demand parameters differ in various ways from the basic model estimates. I conduct two sets of exercises in this section. First, I use my estimates of the other-STEM occupations, but change the parameters of the CS occupations. For comparability, I use similar parameters as those in Bound et al. (2015). In Bound et al. (2015), immigrants and native workers are assumed to be perfect substitutes ($\rho = 1$), and skilled immigrants are more productive than their native-born counterparts (equation 17). The comparable share parameter δ takes a value of 0.52. I then explore different curvature (return to scale) parameters ($\psi = 0.75,\ 0.50,\ 0.25$), which closely relate to different demand elasticities. Later, I explore to what extent the variations in the parameters determining the magnitude of immigration impacts.

$$Y^{cs} = Z_t((1-\delta^{cs})N_t^{cs} + \delta^{cs}M_t^{cs})^{\psi^{cs}} \tag{17}$$

The production function in equation (17) assumes that natives and immigrants are perfect substitutes. With $\psi^{cs} < 1$, this function implies a downward sloping labor demand in the CS occupations. By construction, the function form generates a negative wage impact of immigrants and a crowding out effect on native employment in the CS occupations. Quantitatively how strong this crowding out effect is depends on the demand elasticity. This demand elasticity depends on the return to scale parameter ψ^{cs}

$$\eta^{cs} \quad = - \quad \frac{1}{1-\psi^{cs}}. \tag{18}$$

With a sudden rise of foreign skilled workers in the CS occupations, the less elastic the labor demand is, the greater the negative wage effects would be. On the contrary, if skilled labor demand is more elastic, an increase in immigrant labor supply can be better absorbed with a minimal crowding out effect. When value of ψ^{cs} varies, the demand elasticity varies from -1.3 to -4.0, which are within the plausible range in the literature.

One clear limitation of the production function in equation (17) is that it rules out the possibility that foreign computer scientists might be, in fact, have a positive effect on wages of natives. The assumption of perfect substitutability would lead to an overestimation of the negative impact of immigration. The results from the counterfactual economies discussed above are presented in Panel A of Table 6.

In a separate set of counterfactual exercises, for which the results are shown in Panel B of Table 6, natives and immigrants are assumed to be perfect substitutes in both occupation groups. As the above exercise, I explore different demand elasticities for skilled labor. For all the counterfactual exercises, I hit the economy with the same recovered sector specific technology shocks, but firms cannot increase the recruitment of foreign computer scientists above its 1994 level.

When comparing results in the column where $\psi = 0.5$ from Table 6 to the baseline results in Section 6.1, it is obvious that in both panels A and B, the magnitudes of impacts induced by skilled immigration are greater.[31] This is because skilled immigrants and natives are considered to be perfect substitutes in the CS occupations in both panels. Perfect substitution amplifies the potential negative impacts of high-skilled immigrants.

Comparing results within each panel, as ψ increases, the labor demand becomes more elastic. With immigration shocks of the same size, the more elastic labor demand is capable of better absorbing the supply shock induced by increasing immigrants, resulting in a smaller wage drop. The smaller wage drop, in turn, generates a very limited crowding out effect.

Another important message learned from Table 6 is by comparing the results in the same column across panels. The two panels differ in the production technology used in the other-STEM occupations. The cross-panel differences reiterate the point made earlier - to be able to better assess the impacts of skilled immigration on the U.S. labor market, a general equilibrium model is more appropriate. When workers move across occupations in response to immigration shocks, any potential impacts would be mitigated and diffused. The effectiveness of the diffusion mechanism relies also on the labor demand of the alternative occupations. In the current context, the more elastic the native labor demand is in the other-STEM occupations, the better ' buffer ' the economy provides against immigration shocks. With a large immigration inflow in the CS occupations, native workers react by switching to the other-STEM occupations where the market price becomes more favorable. The more elastic the native labor demand is in the other-STEM occupations, the smaller skill rental rate reduction would be caused. A small reduction in skill rental rates indicates that more native workers who now lose their comparative advantages in the CS occupations can find better ' shelters

[31] $\psi = 0.5$ is comparable to results in sector 6 because my estimates $\psi^{cs} = 0.58$ implies a similar labor demand elasticity in the CS occupations.

' (with more favorable market conditions). In addition, more natives leaving the CS occupations further neutralizes the negative shocks by maintaining wages in that sector at a higher level.

The demand elasticity of skilled natives in the other-STEM occupations has the following expression.[32]

$$\eta_N^{STEM} = \frac{1}{\frac{\mathrm{d}\Pi}{\mathrm{d}N}\frac{N}{\Pi}} = \begin{cases} \frac{1}{[(\rho-1)+\frac{(1-\alpha)(\psi-\rho)N^\rho}{\alpha M^\rho+(1-\alpha)N^\rho}]} & \rho \neq 1 \\ \\ \frac{1}{\psi-1}\frac{(1-\alpha)N+\alpha M}{(1-\alpha)N} & \rho = 1 \end{cases} \tag{19}$$

In Table 7, I evaluate the above expression using the quantities in 1994. In column 1, the labor demand of the other-STEM occupations is more elastic in the lower panel. In this column, the counterfactual economy in the lower panel can better ' absorb ' the immigration-induced supply shocks. The quantitative results confirm the initial conjecture. For the immigration shocks of the same size, the economy in the lower panel observes more occupational mobility responses. The native labor supply grows by 5.99% in the lower panel as opposed to 4.74% in the upper panel. However, wage fluctuations are attenuated in the lower panel. Changes in skill rental rates of both occupational groups are smaller in the lower panel. The same argument also applies to the second column. In the third column, the labor demand of the other-STEM occupations in the lower panel now is less elastic. As a result, the prices in the lower panel are more fragile and more volatile to immigration shocks. Fewer natives are able to re-optimize their occupations. The interaction between the CS and the other-STEM occupations is important when ones intend to provide a comprehensive assessment of the impacts of skilled immigration over the past two decades.

In summary, there are two main points learned from the counterfactual experiments in this part. First, to precisely evaluate the impacts of skilled immigration, accurate production parameters are necessary. The crucial parameters in the flexible CES production function are the substitution elasticity (ρ^s) and return to scale parameters (ψ^s). Second, a general equilibrium model capturing the interaction between the CS and the other-STEM occupations is more appropriate. We have seen that the interaction has a non-negligible impact on natives' mobility responses. The extent to which the potential negative impacts are diffused relates to the demand elasticity in alternative occupations where native workers move to.

6.2.2 Heterogeneous Effects and Valuations of Occupational Mobility

The equilibrium effects on wages described above summarize a variety of heterogeneity. Individual affect the equilibrium wage by changing their occupational decisions differently. This paper empha-

[32] The formula that uses inverse demand function to derive the demand elasticity is valid when the inverse demand function is monotone. The local monotonicity holds at least for those points where I compute various elasticities.

sizes that the occupational mobility is an important adjustment margin that has been understudied in the literature. And there is a wide variety of heterogeneous adjustments among natives. Who are those natives that are more likely to switch occupations to counteract the labor supply shocks induced by immigrants? Economically, how important is the occupational adjustment for different individuals? Does an agent's valuation of occupational mobility vary with individual characteristics? To answer questions like this, in the following exercises, I consider the occupational mobility as an option and then quantify the individual valuation of this option using compensating variations (CV). CV is the dollar amount agents require in order to maintain the same level of lifetime utility when they are constrained to remain in their original occupations regardless of the market conditions.

Over the Internet boom (1994-2000), a period with a massive inflow of immigrants in the STEM occupations, native workers are forced to remain in their original occupations. With the help of large variations in skill rental rates during this period, I first identify individuals who would exercise this option if the mobility restriction was removed, those marginal workers who benefit from occupational mobility.

Figure 14 indicates the fraction of individuals in different age groups that make different occupational choices when the mobility restriction is removed. These numbers are computed using a simulation of 20000 native workers.

When facing with a massive inflow of immigrants, More than 80% of workers under the age of 40 (also see Table 8) change their occupational choices. Younger workers who just enter the labor market are most likely to exercise this option. About 35% of young workers age 22 to 26 would switch to a different occupation in response to the increasing foreign labor supply. However, adjustments by the more experienced workers are almost nonexistent. Only less than 2% of the more experienced workers approaching the end of their career would exercise this option. Very few experienced native workers find it beneficial to switch. Their occupational specific human capital in one occupation becomes predominant that the high implicit switching cost prevents them from enjoying the gains from occupational mobility. This result is natural that younger workers are relatively more flexible who can undo the immigration shocks themselves by re-optimizing their career path, but the experienced workers are left vulnerable in such a case due to their lack of ability to adjust for the new market conditions.

Economically speaking, how valuable is the option of occupational mobility? Even with a temporary restriction (only six years) on individual occupational mobility, individuals are harmed in the sense that their lifetime utility would be adversely affected. The temporary restriction is especially costly for younger workers, whose cost is estimated to be more than $45,000. Human capital is occupation specific. Early work experiences and human capital investment have long lasting effects. If young workers are permanently forced to stay in occupations where an increase in foreign competition is expected, they would demand more than $100,000 to compensate for not being able to re-optimize their occupational choices.

To fill in with more details, I compute the average CVs for identified switchers for each age group. The average CV decreases, as shown in Figure 15. The CV drops from more than $45,000 for new

entrants to almost zero for workers who are about to retire. Overall, it is more costly for younger workers to stick to their previous occupations when market conditions become less favorable. Young workers have a longer career path. Career concerns about choosing the ' right ' occupations play a big role here.

There is a plenty of heterogeneity in individual option values. Within one age group, in the two-dimensional space of occupation specific human capital, workers located along the relative rental rate line require for high expected CVs. For age group a, given the two dimensional human capital distribution $H_a = \begin{pmatrix} H_a^{cs} \\ H_a^{ncs} \end{pmatrix}$, the expression for expected CVs is

$$\mathbb{E}CV(H_a) = \mathbb{E}(CV|mob = 1, H_a)P(mob = 1|H_a) + \mathbb{E}(CV|mob = 0, H_a)P(mob = 0|H_a) \tag{20}$$

where $mob = 1$ if one would switch occupations when the market condition changes. When $mob = 0$, agents would not exercise the option, which implies $\mathbb{E}(CV|mob = 0, H_a) = 0$.

In Appendix F, I derive general expressions to calculate the expected CVs for native workers. I then apply these formulas to compute the distribution of expected CVs for new entrants when foreign labor supply changes. As shown by the contour map in Figure 16, marginal individuals who do not have dominant advantages in any of these occupations place a higher value on the option. Along the relative skill rental rate line, workers are the most sensitive to changes in market prices, exercise their option of switching most frequently. These workers at the margin place the highest valuation on the occupational mobility.

7 Discussion

This section addresses some potential issues about the current model specification.

In the current setting, I consider a binary choice model within the STEM domain over two mutually exclusive options, working in the CS or the other-STEM occupations. This framework can be extended to include a broader set of occupations at a more detailed level without much substantial modification of the basic model. However, I choose to focus this paper only on the STEM occupations because as shown in the introduction the STEM occupations have been experiencing the massive inflows of skilled immigrants in the U.S. labor market. And to identify the parameters, I use information on conditional wage distributions of narrowly defined age groups. The sample size would be very small when we look at the more detailed occupational groups. It is difficult to look at occupation choices at a more detailed level given the current data availability. With larger administrative data sets, such as the Danish Integrated Database for Labor Market Research (IDA) (Foged and Peri, 2016), the basic model can be extended to include more detailed occupations.

With the available data, the non-STEM occupations could be a potential third choice. First, as noted before, adding extra occupation is very costly. Each additional occupation implies an extra

choice, an additional experience variable, an additional dimension in the innate ability distribution, more than 15 additional parameters to estimate, an additional equilibrium skill price, and an additional expectation process. Besides, whether to include the non-STEM occupations depends on how frequently do workers switch between the STEM and the non-STEM occupations. Would the non-STEM occupations be a good alternative where native workers move to? The task contents of the STEM occupations differ a lot from those of the non-STEM occupations. As a result, the STEM occupations require years of special training. The implicit switching cost can be very high moving cross the STEM boundary.

To formally address the concern above, I use the linked monthly CPS data from 1994-2014 to explore individual's occupational mobility patterns cross the STEM boundary. The data clean technique proposed by Moscarini and Thomsson (2007) is applied here to get rid of the possible classification errors. Table 9 panel A indicates that if the non-STEM workers switch occupations, almost all of them, approximately 94%, would switch to another non-STEM occupations. However, for the STEM workers, conditional on switching occupations, 62.8% of them move to the non-STEM occupations. This is to some extent due to the fact that managerial occupations are classified as the non-STEM occupations. Moving to managerial positions within the same industry is instead occupational upgrading. Both the causes and implications of this vertical occupational mobility are different from the horizontal mobility discussed in this paper according to comparative advantages. Once the occupational mobility toward managerial positions in the same industry is adjusted and treated as mobility within the same occupational group, the binary choice captures more than 2/3 of the total switches for the STEM workers.

The previous results talk about the gross occupational mobility. Figure 17 plots the employment shares (the net occupational mobility) of the STEM occupations and of the CS occupations. The employment share of the STEM occupations in the U.S. is stable, about 4% over the past two decades, which indicates that no structural or systematic changes occurred during this period that makes the STEM occupations more favorable for natives. However, the employment share of the CS occupations increases by nearly 2/3 over the same period. These patterns suggest that the STEM occupations as a whole do not become more attractive to native workers; the CS occupations become more popular among native STEM workers.

The previous adjustment and discussion about employment share only partially alleviate the concerns. However, still 1/3 of the STEM switchers move to the non-STEM occupations even after the adjustment. This option is excluded in the model, so is the labor market attachment decision (Llull, 017a). When the negative wage impacts of immigration in the STEM occupations are large enough, some marginal workers would be discouraged and leave the STEM occupations. If the disincentive is sufficiently large, they might even adjust their labor market participation decisions as well. Omitting these relevant margins could lead to a selection bias of the wage impacts of immigration on the STEM occupations and an overestimation of the imperfect substitutability in the labor demand side. Furthermore, by restricting the analysis only to the STEM occupations, this paper

ignores the wider impacts of skilled immigrants through diffusion to the non-STEM occupations. The third occupational choice and the labor force participation decision would introduce additional distribution effects. Natives who adjust along these margins would be better off compared with the baseline in this paper by switching to a better alternative. With more adjustment margins by native, the more complete model would imply that the economy has better abilities to mitigates the initial impacts of immigration. However, native workers who are in the destination or receiving occupations would face with some downward pressure on their wages. The total impact is not clear in this case. The empirical question is which of the two confronting distributional effects dominates the other. And it also depends crucially on the production technology in the third occupation.

8 Conclusion

Immigration is a major policy issue and concern in the U.S. Most of the previous literature has focused on the low-skilled immigration. In this paper, I focus on the high-skilled immigrants who play a qualitatively and quantitatively important role in the U.S. labor market recently. This paper estimates a labor market general equilibrium model to quantify the wage effects of skilled immigration on the STEM occupations. I estimate demand for skilled labor across occupations and explicitly model the native workers' occupational mobility as an adjustment margin in response to foreign labor competition.

Despite the public concerns, my results indicate that a large inflow of skilled immigrants has limited impacts on natives with similar skills and who are working in the same professions. For some occupations, e.g., the other-STEM occupations, even complementarity exists. Increases in the supply of foreign labor actually have positive impacts on natives' welfare. Even when native and foreign labor directly compete with each other in some occupations (the CS occupations), native workers are not perfectly substitutable. Furthermore, native workers optimize their occupational choices. The occupational mobility acts as a pressure-valve that mitigates and diffuses any impacts of skilled immigration.

The estimation and counterfactual exercises in the paper deliver the following points. First, a selective immigration policy based on fields of study and occupations can outperform an overall cap on total immigrants. Optimizing the occupation mix of skilled immigrants and channeling skilled migrants to occupations where complementarity exists would benefit natives more compared to an overall cap. The success of the policy hinges on the accurate estimates of occupation specific production functions. Second, the estimates provide us with some useful insights into the underlying mechanism. I emphasize in this paper that the general equilibrium framework is important. It allows us to analyze the overall welfare impacts of skilled immigration. Even when some occupations don't directly receive skilled immigrants, they are relevant in the sense that they would affect the economy's overall ability to absorb immigration shocks. Last, individuals value the option of occupational mobility. Had native workers been constrained to remain in their original occupations

when the market condition changes, their lifetime utility would be adversely affected. This restriction would be particularly costly for younger workers.

While the model incorporates mobility between different types of occupations according to comparative advantages, a model explicitly taking into consideration of the vertical occupational movement is also relevant and desirable. D'Amuri and Peri (2014) and Peri and Sparber (2011) document that natives tend to upgrade their jobs in response to low-skilled immigration. They take on more complex and communication-intensive tasks and leave low skilled manual tasks to immigration. This protects them from the direct foreign competition. Llull (017a) studies the educational upgrading by natives as a way to counteract the immigration shocks. In the context of high-skilled immigration, do inflows of foreign skilled labor push more natives to managerial positions? By moving up the job ladder, natives potentially could gain more. Does the speed of vertical mobility differ across occupations? If it does, what drives the differential? Adding this margin provides a more comprehensive assessment of the impacts of skilled immigration.

References

Altonji, J. G., Bharadwaj, P., and Lange, F. (2012). Changes in the characteristics of american youth: Implications for adult outcomes. *Journal of Labor Economics*, 30(4):783–828.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of political economy*, 75(4, Part 1):352–365.

Borjas, G. J. (1999). The economic analysis of immigration. *Handbook of labor economics*, 3:1697–1760.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, pages 1335–1374.

Borjas, G. J. (2006). Native internal migration and the labor market impact of immigration. *Journal of Human resources*, 41(2):221–258.

Borjas, G. J. and Doran, K. B. (2012). The collapse of the soviet union and the productivity of american mathematicians. Technical report, National Bureau of Economic Research.

Bound, J., Braga, B., Golden, J. M., and Khanna, G. (2015). Recruitment of foreigners in the market for computer scientists in the united states. *Journal of Labor Economics*, 33(S1 Part 2):S187–S223.

Bowlus, A. J. and Robinson, C. (2012). Human capital prices, productivity, and growth. *The American Economic Review*, pages 3483–3515.

Card, D. (1999). The causal effect of education on earnings. In *Handbook of labor economics*, volume 3, pages 1801–1863. Elsevier.

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

Card, D. (2005). Is the new immigration really so bad?*. *The Economic Journal*, 115(507):F300–F323.

Card, D. (2009). How immigration affects us cities. *making cities Work: prospects and policies for Urban America*, pages 158–200.

Castro, R. and Ševčík, P. (2016). Occupational choice, human capital, and financial constraints.

D'Amuri, F. and Peri, G. (2014). Immigration, jobs, and employment protection: evidence from europe before and during the great recession. *Journal of the European Economic Association*, 12(2):432–464.

Dustmann, C., Frattini, T., and Preston, I. P. (2012). The effect of immigration along the distribution of wages. *Review of Economic Studies*, 80(1):145–173.

Foged, M. and Peri, G. (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, 8(2):1–34.

Gibbons, R. and Waldman, M. (1999). A theory of wage and promotion dynamics inside firms. *The Quarterly Journal of Economics*, 114(4):1321–1358.

Görlach, J.-S. (2017). Borrowing constraints, migrant selection, and the dynamics of return and repeat migration. *Work*.

Hanson, G. and Slaughter, M. (2015). High-skilled immigration and the rise of stem occupations in us employment. In *Education, Skills, and Technical Change: Implications for Future US GDP Growth*. University of Chicago Press.

Heckman, J. J. and Honore, B. E. (1990). The empirical content of the roy model. *Econometrica: Journal of the Econometric Society*, pages 1121–1149.

Heckman, J. J., Lochner, L., and Taber, C. (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of economic dynamics*, 1(1):1–58.

Heckman, J. J., Lochner, L. J., and Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. *Handbook of the Economics of Education*, 1:307–458.

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

Hunt, J. (2013). Does the united states admit the best and brightest computer and engineering workers. Technical report, Rutgers University working paper.

Hunt, J. and Gauthier-Loiselle, M. (2010). How much does immigration boost innovation? *American Economic Journal: Macroeconomics*, 2(2):31–56.

Jeong, H., Kim, Y., and Manovskii, I. (2015). The price of experience. *American Economic Review*, 105(2):784–815.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, 17(2):303–316.

Kambourov, G. and Manovskii, I. (2008). Rising occupational and industry mobility in the united states: 1968–97*. *International Economic Review*, 49(1):41–79.

Kambourov, G. and Manovskii, I. (2009). Occupational specificity of human capital*. *International Economic Review*, 50(1):63–115.

Keane, M. P. and Wolpin, K. I. (1994). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte carlo evidence. *The Review of Economics and Statistics*, pages 648–672.

Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of political Economy*, 105(3):473–522.

Kerr, W. R. and Lincoln, W. F. (2010). The supply side of innovation: H-1b visa reforms and us ethnic invention. Technical report, National Bureau of Economic Research.

Lee, D. (2005). An estimable dynamic general equilibrium model of work, schooling, and occupational choice. *International Economic Review*, 46(1):1–34.

Lee, D. and Wolpin, K. I. (2006). Intersectoral labor mobility and the growth of the service sector. *Econometrica*, 74(1):1–46.

Llull, J. (2017a). Immigration, wages, and education: A labour market equilibrium structural model. *The Review of Economic Studies*.

Llull, J. (2017b). The effect of immigration on wages: exploiting exogenous variation at the national level. *Journal of Human Resources*, pages 0315–7032R2.

Manacorda, M., Manning, A., and Wadsworth, J. (2012). The impact of immigration on the structure of wages: theory and evidence from britain. *Journal of the European economic association*, 10(1):120–151.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, pages 995–1026.

Moscarini, G. and Thomsson, K. (2007). Occupational and job mobility in the us*. *The Scandinavian Journal of Economics*, 109(4):807–836.

Moser, P., Voena, A., and Waldinger, F. (2014). German jewish émigrés and us invention. *The American Economic Review*, 104(10):3222–3255.

Murphy, K. M. and Welch, F. (1990). Empirical age-earnings profiles. *Journal of Labor economics*, 8(2):202–229.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, 10(1):152–197.

Papageorgiou, T. (2014). Learning your comparative advantages. *Review of Economic Studies*, 81(3):1263–1295.

Peri, G., Shih, K., and Sparber, C. (2015). Stem workers, h-1b visas, and productivity in us cities. *Journal of Labor Economics*, 33(S1 Part 2):S225–S255.

Peri, G. and Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, 1(3):135–69.

Peri, G. and Sparber, C. (2011). Highly educated immigrants and native occupational choice. *Industrial Relations: a journal of economy and society*, 50(3):385–411.

Piyapromdee, S. (2015). The impact of immigration on wages, internal migration and welfare.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, 3(2):135–146.

Appendix A

Table 1: Fraction of Immigrants

Year	Immigrants as a fraction of Skilled Worker	Immigrants as a fraction of Computer Scientists	Immigrants as a fraction of Other-STEM workers
1970	2.10%	2.37%	3.63%
1980	5.43%	7.09%	9.72%
1990	6.89%	11.06%	10.71%
2000	8.41%	18.59%	12.69%
2010	12.77%	27.82%	18.21%

Table 2: Estimates of Native Labor Supply

Perfect Foresight Expectation

Coeff	Computer Science		Other STEM	
	val.	std. err	val.	std. err
CS Exp.	0.0756	(0.0012)	0.0760	(0.0016)
Other STEM Exp.	0.0695	(0.0013)	0.0851	(0.0011)
Total Exp^2 /100	-0.2009	(0.0020)	-0.2028	(0.0024)
Total Exp^3 /1000	0.0124	(0.0014)	0.0141	(0.0010)
		Covariance Matrix		
unobs. Heterogeneity	0.0561	(0.0060)		
	-0.0143	(0.0037)	0.1263	(0.0066)
		Taste Shock		
Trend	-0.0998		(0.0067)	
Variance	26.975		(2.9435)	

Table 3: Estimates of Industry Production Function

Perfecto Foresight Expectation

Coeff	Computer Science		other-STEM	
	val.	std. err	val.	std. err
Share	0.4956	(0.1152)	0.4229	(0.1311)
Rho	0.8542	(0.1592)	0.4914	(0.1217)
Return to Scale	0.5863	(0.1770)	0.6085	(0.2108)

Occupation Specific Shock Process

Coeff	Computer Science		other-STEM	
	val.	std. err	val.	std. err
Autoregressive term	0.7287	(0.0774)	0.1925	(0.2126)
Constant	1.2406	(0.3501)	3.7004	(0.9736)
St. dev. of innovations	0.0325	(0.0114)	0.0411	(0.0121)

Table 4: Correlation Coefficient

	1978 Cohort	1982 Cohort	1986 Cohort
Relative Efficient Price	0.41	0.60	0.89

Table 5: Fixed Immigrant Labour Supplies at Their 1994 Level

	$\Delta\Pi_{cs}$	$\Delta\Pi_{ncs}$	Δ^N_{CS}	Δ^N_{ncs}
M_{cs} Fixed & M_{ncs} Fixed	2.41%	1.22%	5.49%	−2.96%
M_{cs} Fixed & M_{ncs} Old Path	2.52%	1.37%	5.24%	−2.81%

Table 6: Summary Results from Counterfactual Simulation With Different Production Parameters

Panel.A $\rho_{cs} = 1$ While other-STEM sector Using Estimated Production Function				
	ψ_{cs}=0.75	ψ_{cs}=0.50	ψ_{cs}=0.25	Basic
$\Delta\Pi_{cs}$	2.28%	3.44%	3.70%	2.51%
$\Delta\Pi_{ncs}$	1.24%	2.00%	2.01%	1.37%
Δ^N_{cs}	4.74%	5.57%	7.66%	5.24%
Δ^N_{ncs}	−2.55%	−2.90%	−4.07%	−2.81%
Panel.B Same Production Function in Both CS and other-STEM sector (ρ=1)				
	ψ=0.75	ψ=0.50	ψ=0.25	Basic
$\Delta\Pi_{cs}$	2.01%	3.15%	4.61%	2.51%
$\Delta\Pi_{ncs}$	0.72%	1.66%	2.55%	1.37%
Δ^N_{cs}	5.99%	6.91%	7.11%	5.24%
Δ^N_{ncs}	−3.31%	−3.73%	−3.90%	−2.81%

Table 7: Elasticities of Counterfactual Simulation With Different Production Parameters

Panel.A $\rho_{cs} = 1$ While other-STEM sector Using Estimated Production Function						
	ψ=0.75		ψ=0.50		ψ=0.25	
	CS	NCS	CS	NCS	CS	NCS
Elasticity	-4.47	-2.12	-2.23	-2.12	-1.49	-2.12
Panel.B Same Production Function in Both CS and other-STEM sector (ρ=1)						
	ψ=0.75		ψ=0.50		ψ=0.25	
	CS	NCS	CS	NCS	CS	NCS
Elasticity	-4.47	-4.48	-2.23	-2.24	-1.49	-1.49

Table 8: Age Composition of Switchers

Age Group	Fraction as % of Total Switcher	Accumulative %
$22-26$	34.63%	34.63%
$27-21$	27.83%	62.46%
$32-36$	19.23%	81.69%
$37-41$	9.52%	91.21%
$42-46$	4.76%	95.97%
$47-51$	2.45%	98.41%
$52-65$	1.59%	100%
total	100%	

The age composition is computed using simulation of 20000 native workers.

Table 9: Composition of Monthly Occupational Mobility

A: Unadjusted Composition

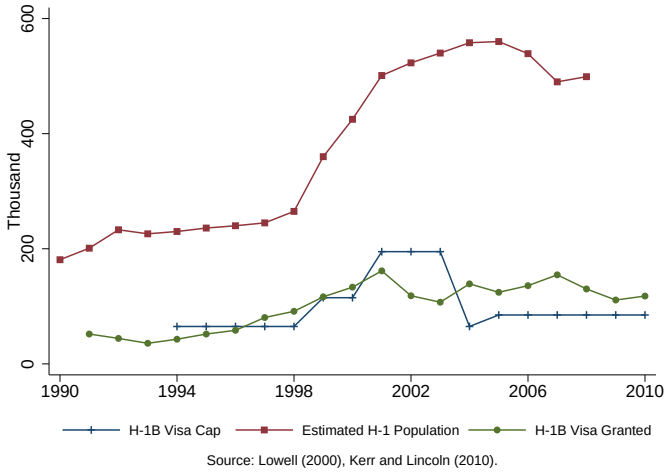
	STEM	Non-STEM
STEM	37.2%	62.8%
Non-STEM	7.4%	92.6%

B: Adjusted Composition

	STEM	Non-STEM
STEM	66.7%	33.3%
Non-STEM	7.4%	92.6%

The monthly occupation switch probability is computed using linked montly CPS data from 1994-2013. The row corresponding to occupational groups in month t. The column corresponding to occupational groups in month t+1.

Figure 1: H-1B Petition Cap and Estimated H-1B Population



Note: Population stock is constructed using estimations of inflows (visa granted) and outflow (deaths, permanent residency, or emigration) of H-1B workers. In later years, the number of visa granted could exceed the visa cap due to exemptions for foreigners who work at universities and non-profit research facilities.

Figure 2: Occupations of H-1B Worker Beneficiaries in 2010

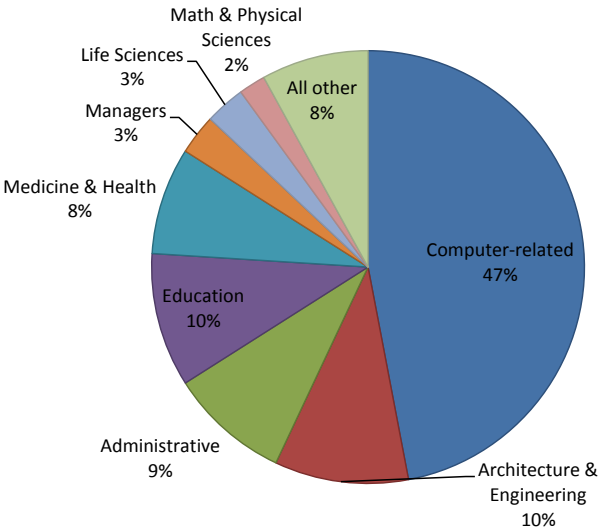


Figure 3: Fraction of Immigrants

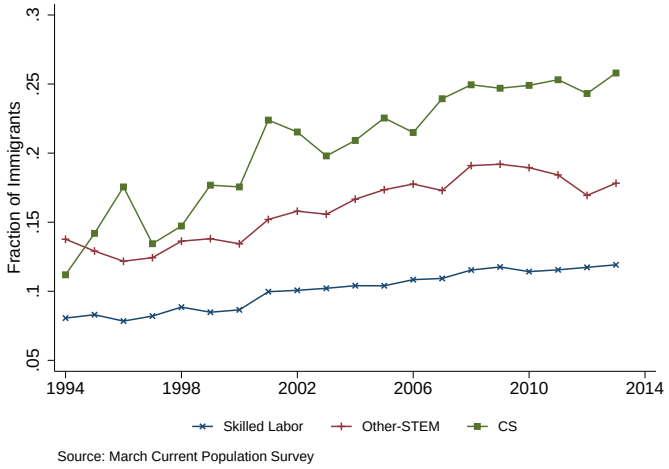


Figure 4: Evolution of Skill Rental Prices

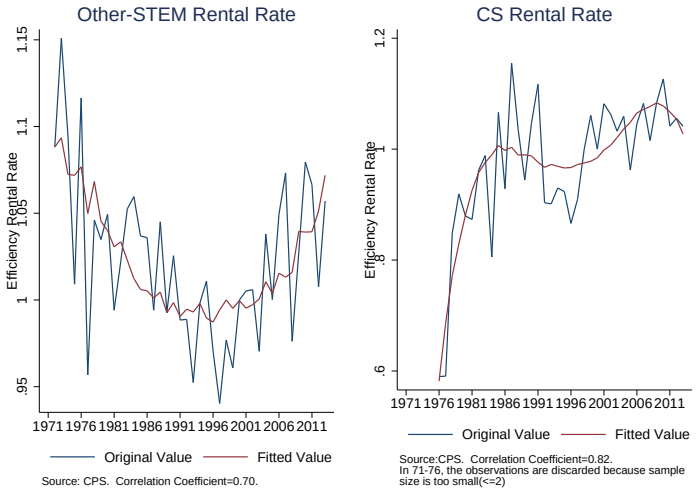
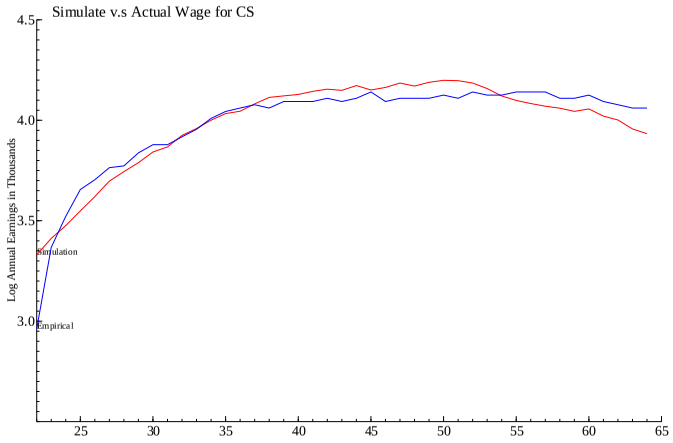
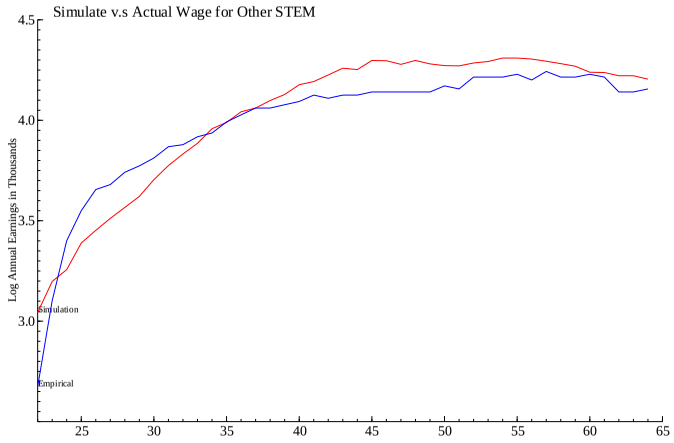


Figure 5: Wage Distribution Fit (Year 2000)

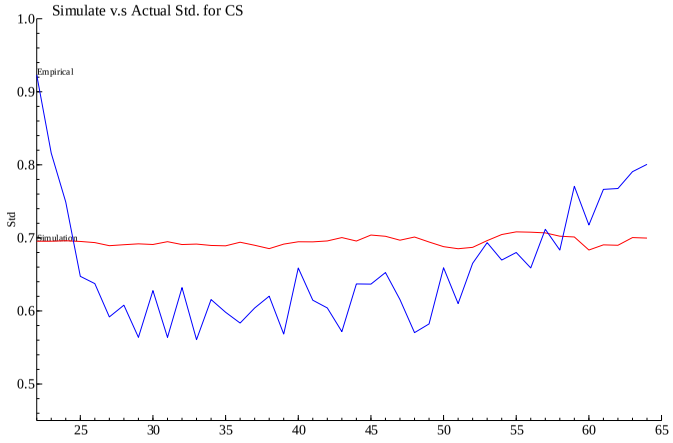
(a) Log CS Wage Profile



(b) Log Other-STEM Wage Profile



(c) Std of CS



(d) Std of Other-STEM

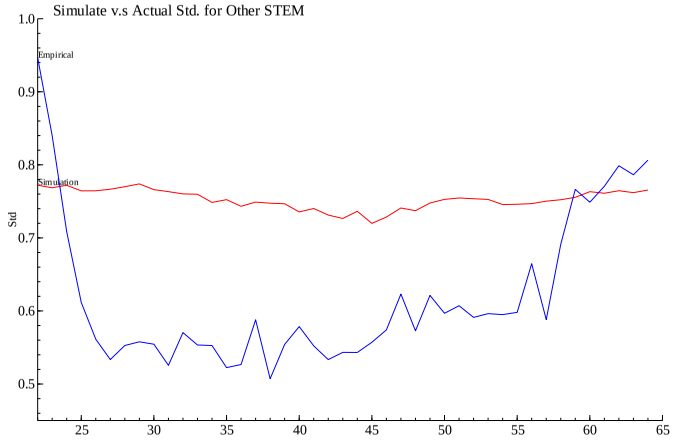
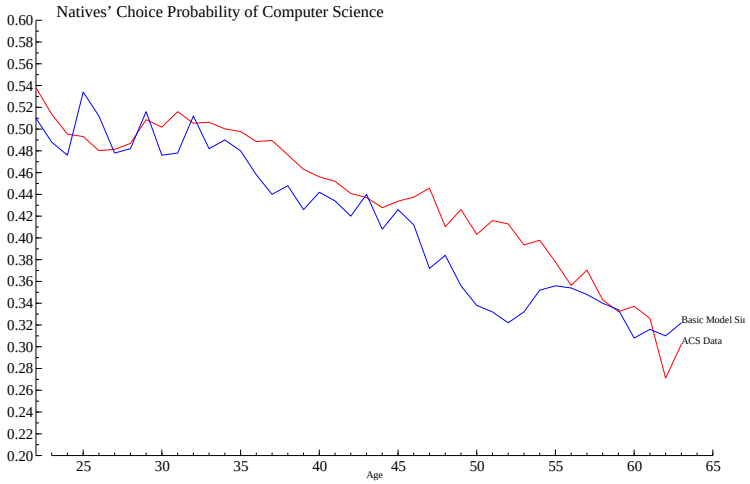


Figure 6: Occupational Choices and Mobility Fit (2000)

(a) Fraction of Native CS Workers



(b) Gross Occupational Mobility

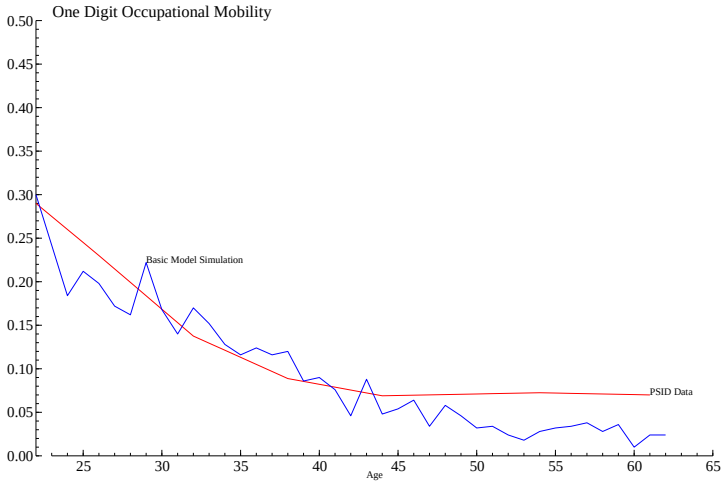
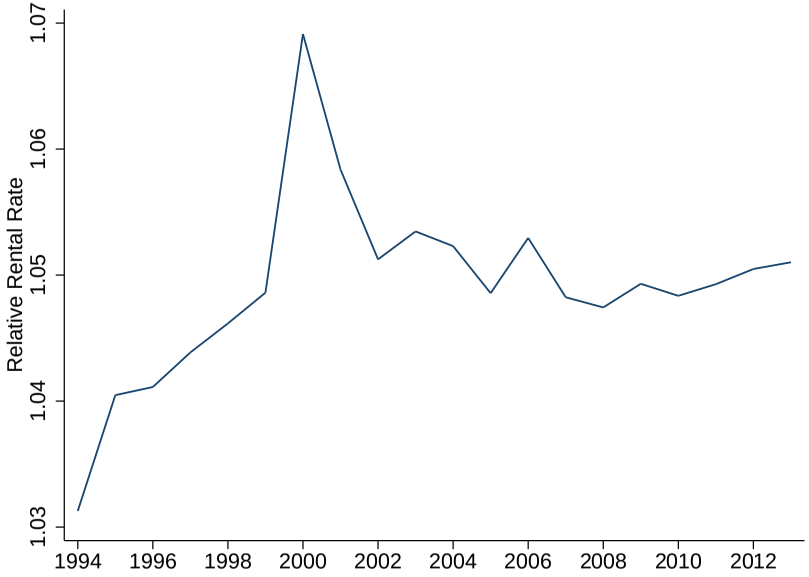


Figure 7: Model Predicted Prices v.s. Nasdaq Index

(a) Model Predicted Relative Prices



(b) Nasdaq Composite Index

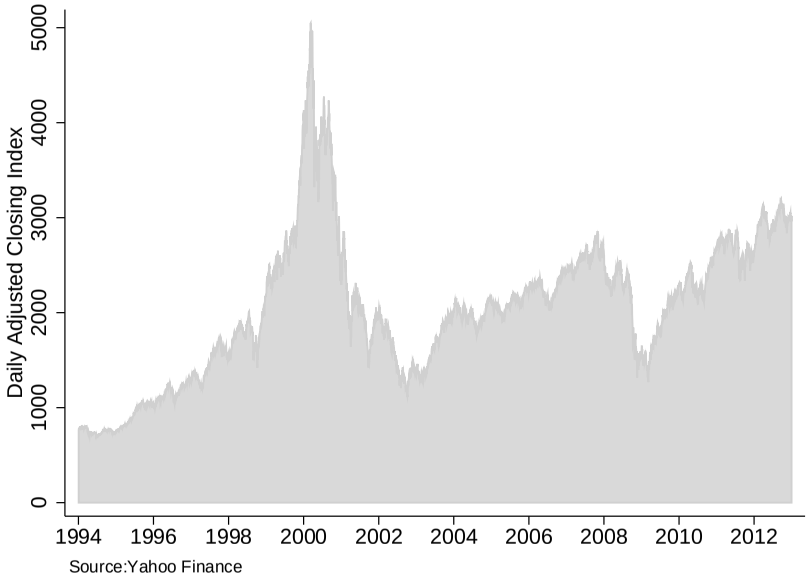


Figure 8: Response of New Entrants

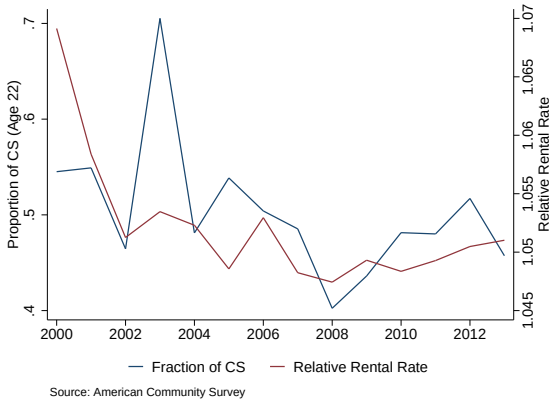


Figure 9: Lagged Response of New Entrants

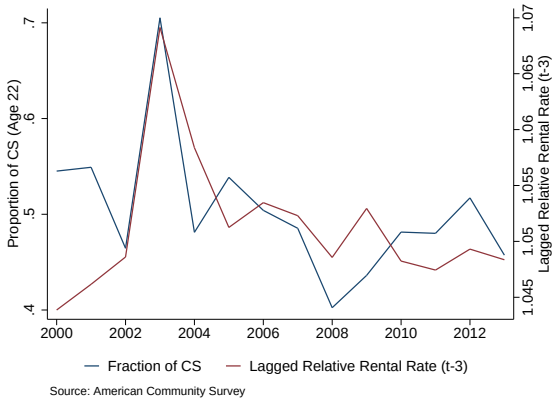


Figure 10: Cohort Response

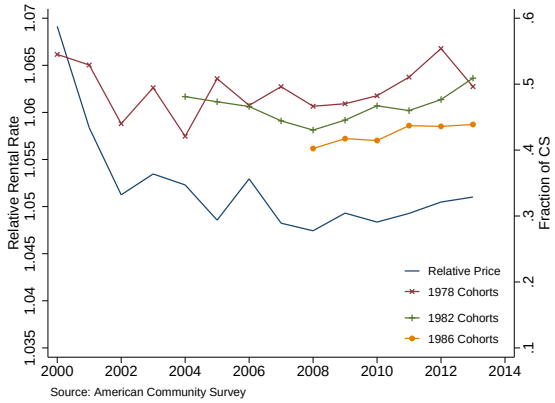
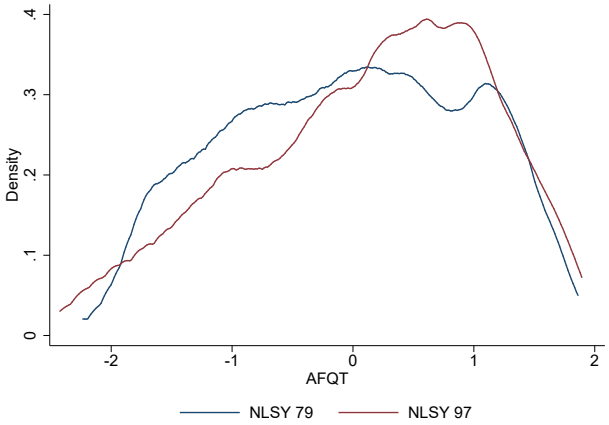
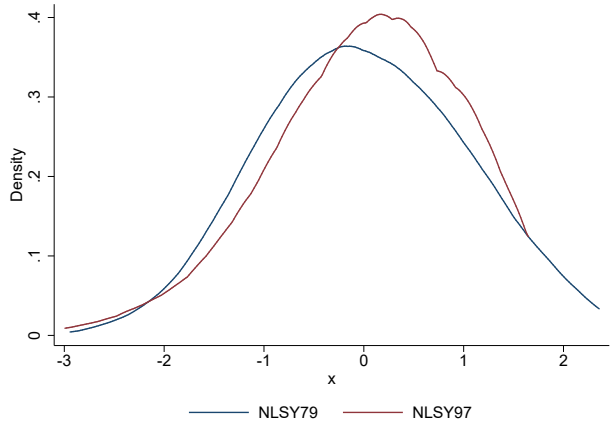


Figure 11: Skill Measure Across Cohorts

(a) Cognitive Ability (AFQT)



(b) Non-Cognitive Ability



(c) Sociability Ability

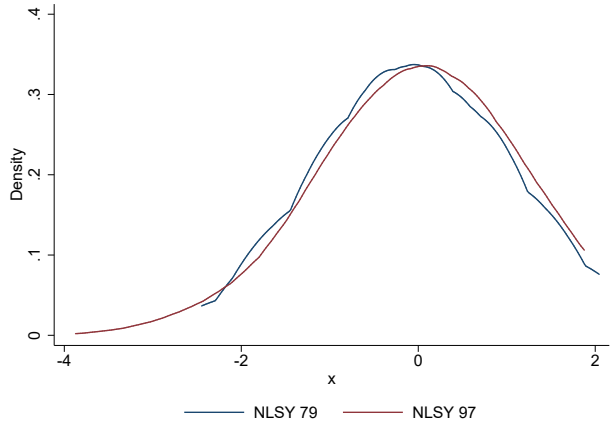


Figure 12: Task Diversity

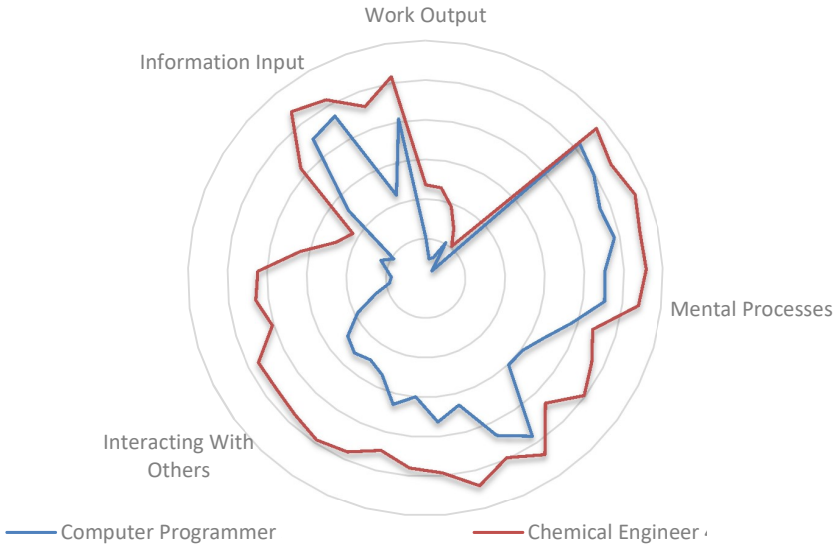


Figure 13: Fixed Foreign Computer Scientists at 1994 Level

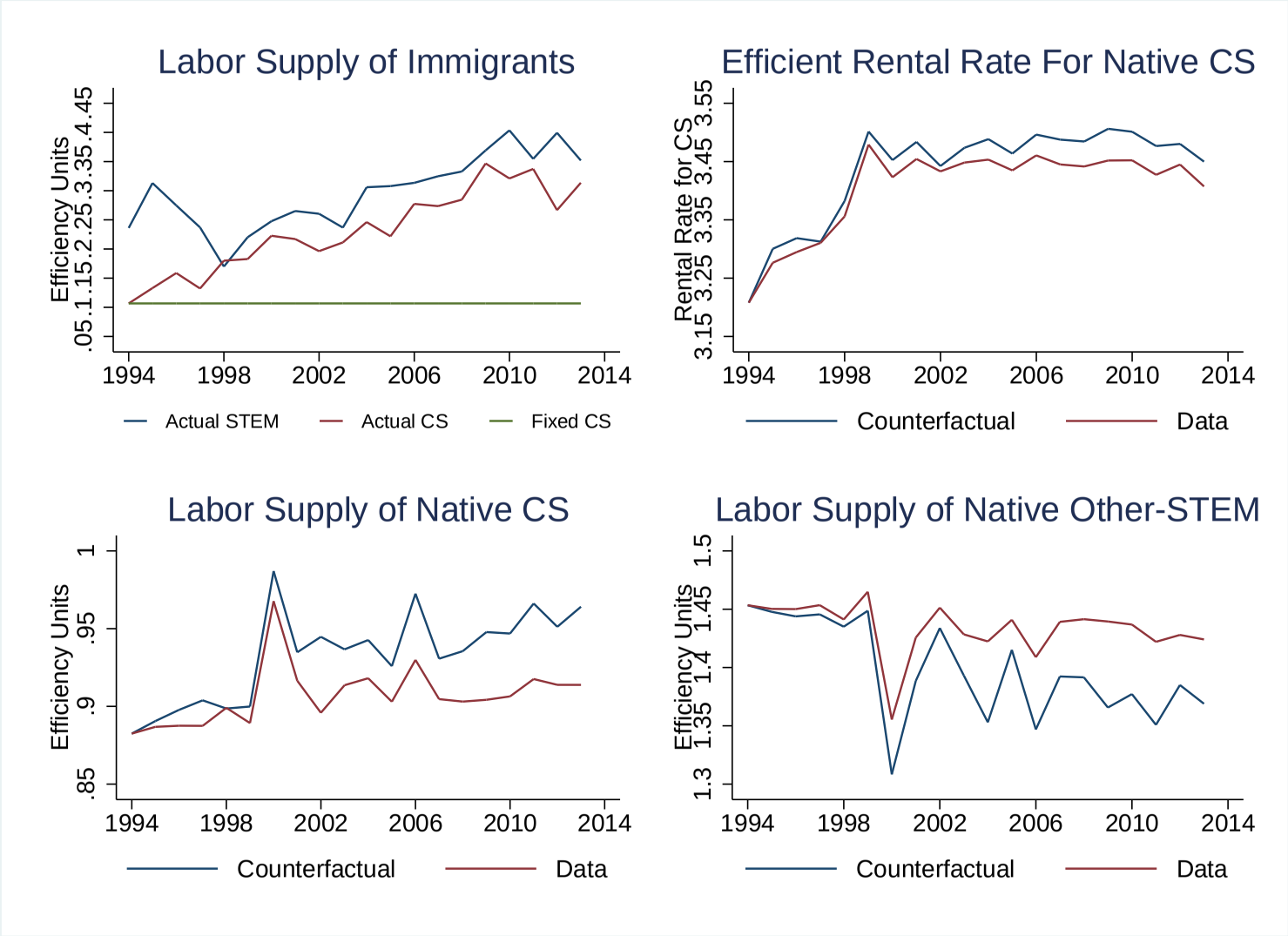


Figure 14: Fraction of Switcher Over Age

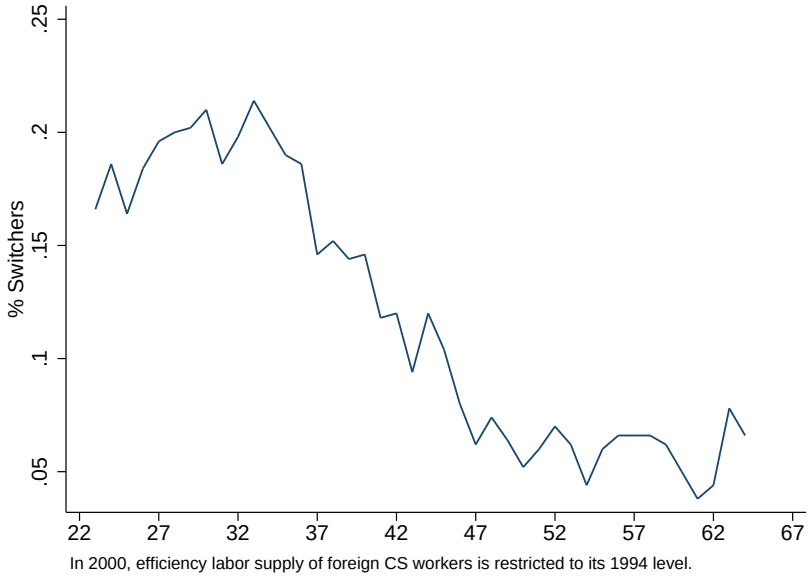


Figure 15: Age Profile of Average CV Required by Switchers

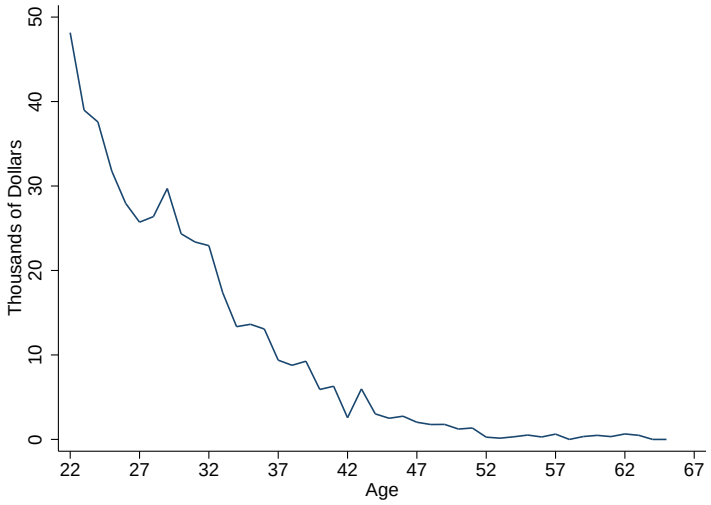


Figure 16: Contour Map of Expected CVs For Workers Aged 22

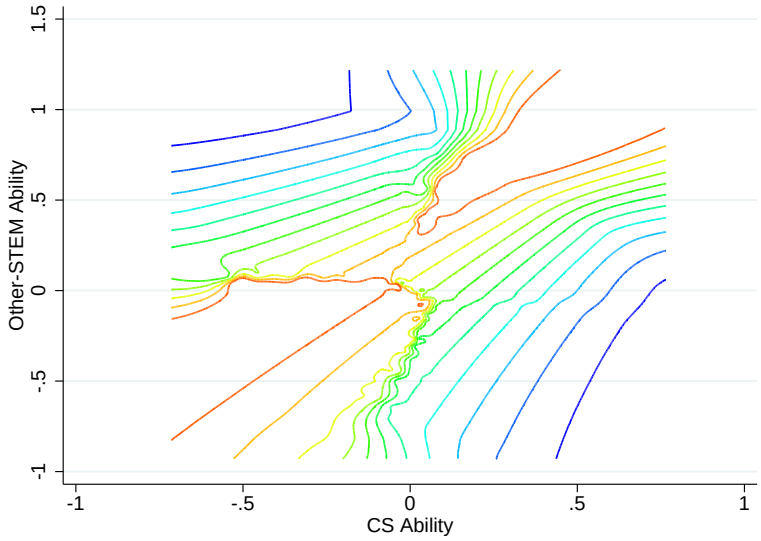
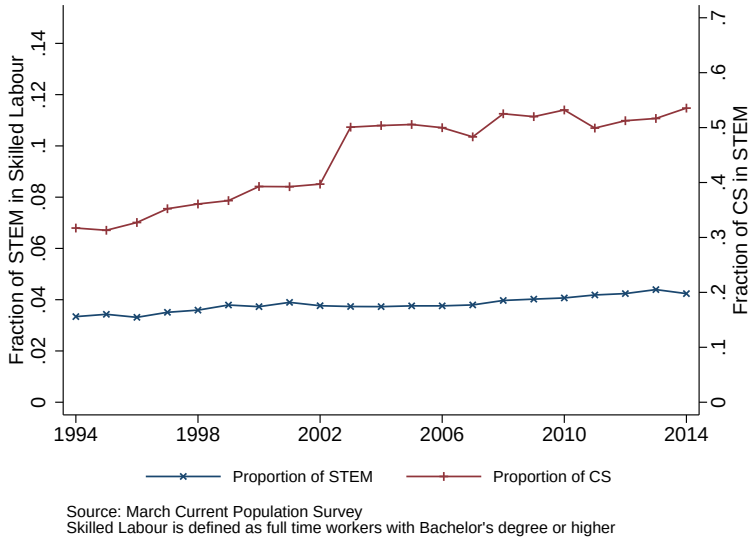


Figure 17: Employment Share of STEM and CS Workers



Appendix B

Detailed Description of Data Cleaning

For the data cleaning, I first restrict the analysis applying only to the skilled labor force, defined as those who have a Bachelor's degree or higher and are currently in the labor force. The status ' in labor force ' is defined as currently at work, having jobs not at work, in armed force, unemployed with experience and unemployed without experience. Because the hour choice is omitted in the discrete choice model, I further restrict the sample to full-time full-year workers whose total hours worked (the product of usual hour worked per week and usual weeks worked) exceed 1500 per year to better match the model. For the income wage data, I first use CPI index suggested by IPUMS website to deflate the income in 1999 dollar, and top-coded values are multiplied by 1.4. The hourly wage rate is calculated following the standard approach, dividing income wage by total hour worked. Then the hourly wage rate is employed to deal with possible outliers. Individuals with hourly wage rate lower than 7 dollars[33] and higher than 200 dollars are discarded. I use the variable ' year of immigration ' to differentiate immigrant and native workers. If a worker migrates to U.S. older than age 18, they are considered as foreign workers. To define consistently two occupational groups, I use the IPUMS suggested occupation crosswalk (OCC1990) and define CS workers as computer system analysts, computer scientists, and computer software developers in OCC1990.

The efficiency rental rates paid to skilled immigrants are measured by the average annual income of foreign new entrants in each occupational group. There are two major ways to define new entrants. First, in the model, I assume skilled workers enter the labor market after graduating from college at average age 22. The average annual income of 22-year-old foreign computer scientists is treated as the measure of the skill rental rate. I try different measures by varying the age range, such as the range from age 22 to 24 which is the normal range of college graduation. Another way to define the new entrants is to use another variable: reason not at work last year. For foreign workers aged 22 to 30, if their answer to the previous question is ' at school ' then they are classified as new entrants. The second measure suffers from the small sample problem because most of the answers are not missing.

Appendix C

Recover Skill Rental Rates Series Using Flat Spot Method

The estimation follows the method proposed by Bowlus and Robinson (2012). I made the assumption of competitive labor markets for each human capital type. The wages for any individual i of a particular occupation are given by

$$W_{i,t}^{s} = \Pi_{t}^{s} H_{i,t}^{s}.$$

[33] The level of federal minimum wage

This implies that within each occupation the change in wages between t and $t+1$ is given by

$$\frac{W^s_{t+1}}{W^s_t} = \frac{\Pi^s_{t+1}}{\Pi^s_t}\frac{H^s_{i,t+1}}{H^s_{i,t}}.$$

Therefore, the change in skill rental rates is given by

$$m^s_t = \frac{\Pi^s_{t+1}}{\Pi^s_t} = \frac{W^s_{t+1}}{W^s_t}\frac{H^s_{i,t}}{H^s_{i,t+1}}.$$

The flat spot method estimates the change in skill rental rates by restricting estimation to observations where human capital levels do not change over time, i.e., where $\frac{H^s_{i,t}}{H^s_{i,t+1}} = 1$. As a result, the changes in observable wages are equivalent to the changes in skill rental rates. For the choice of flat spot region, I also follow Bowlus and Robinson (2012) and choose 51-62 as the suitable range for college graduates.

Solution Method Using $\{\frac{\Pi^s_{t+1}(o)}{\Pi^s_t(o)}\}$

Using the flat spot method, I obtain the series of observed skill rental rates $\{m^s_t\} = \{\frac{\Pi^s_{t+1}(o)}{\Pi^s_t(o)}\}$. Now I assume the evolution of the skill rental rates for both occupations is fully anticipated by agents.

$$\Pi^s_\tau(e) = \hat{\Pi}^s_\tau(o) = m^s_t m^s_{t+1}.., m^s_\tau \Pi^s_t, \ \forall \tau > t$$

This relationship yields a sequence of skill rental rates that can be written solely in terms of Π^s_t. The value of Π^s_t is determined by equating labor supply and demand in period t.

Appendix D

Asymptotic Distribution of the SMM Estimators with Multiple Samples

The asymptotic distribution of the SMM estimators used in this paper follows results of Gorlach (**?**).

The criterion function to be minimized has the following general functional form:

$$\begin{aligned} M(\theta) &= D(\theta)'WD(\theta) \\ &= (m^d - m^s(\theta))'W(m^d - m^s(\theta)) \end{aligned}$$

Where m^d denotes the data moments, $m^s(\theta)$ is the simulated moments using the model. θ is the parameters of interest and W is any weighting matrix.

Important assumptions need to be made in addition to the regularity assumptions required by the usual asymptotic theory of M-estimators:

Additional Assumption 1: Different samples used are drawn independently. This implies that any cross-sample moments are zeros and the weighting matrix W will be block diagonal.

Additional Assumption 2: The sample size N_ζ of the dataset ζ used increases at a proportional rate

$$\lim_{\substack{N \to \infty \\ N_\zeta \to \infty}} (N_\zeta / N) = \lambda_\zeta$$

where $N = \sum_\zeta N_\zeta$ and $0 < \lambda_\zeta < \infty$ for all samples ζ, which means that none of the samples is irrelevant relative to the others.

Additional Assumption 3: The simulated sample size N_ζ^s increase at a rate such that

$$\lim_{\substack{N_\zeta \to \infty \\ N_\zeta^s \to \infty}} (N_\zeta / N_\zeta^s) = n_\zeta$$

with $0 < n_\zeta < \infty$ for all sample ζ.

The application of the central limit theorem then yields the asymptotic distribution for the parameter $\hat{\theta}$:

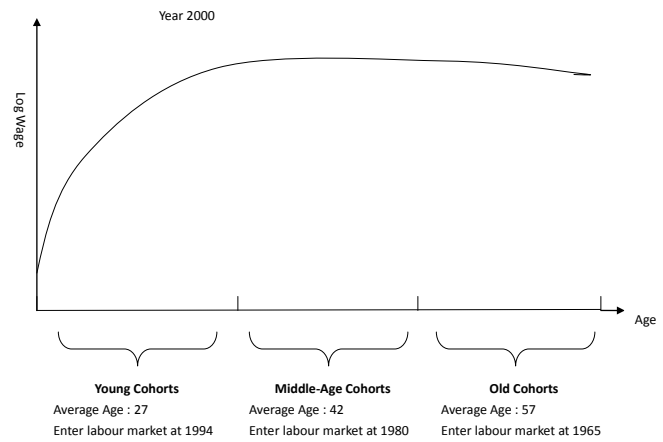
$$\begin{aligned}
\sqrt{N}(\hat{\theta} - \theta) \quad \to \quad & N(0, (\frac{\partial D'}{\partial \theta}(\hat{\theta}) W \frac{\partial D}{\partial \theta'}(\hat{\theta}))^{-1} \\
& (\Sigma N(1 + n_\zeta) \frac{\partial D'_\zeta}{\partial \theta}(\hat{\theta}) W_\zeta var(m_\zeta^d - m_\zeta^s(\hat{\theta})) W_\zeta' \frac{\partial D_\zeta}{\partial \theta'}(\hat{\theta})) \\
& (\frac{\partial D'}{\partial \theta}(\hat{\theta}) W \frac{\partial D}{\partial \theta'}(\hat{\theta}))^{-1}
\end{aligned}$$

Appendix E

Addressing the Cohort Effects

I assume that individuals form deterministic expectations of future skill rental rates. Individuals solve the occupational choice problem under the deterministic path of skill rental rates. However, different birth cohorts enter the labor market at different point of time and use different phases of the deterministic path to form their expectation. Therefore, it is necessary to solve the optimal occupation decision rule for each birth cohort groups separately, i.e., for the cohort specific set of value functions.

Figure E.1: Illustration of the Perfect Foresight Model



I assume each year the native labor force consists three birth-cohort groups: young, middle-aged and old workers.[34] The reason that I only consider three cohort groups is simply for the computational feasibility. For each group, the average birth year is computed and assigned as a group characteristic. Taking the old cohort group in the year 2000 as an example, workers in this birth cohort group are on average 57 years old who entered the labor market in 1965. Different birth cohort groups solve different dynamic choice problems since they experience different phases of the industrial development, form different expectations, and invest in occupation specific human capital differently. In a particular year, the synthetic income age profile of these three birth cohort groups and the employment share age profiles are computed to better match the cross-sectional data used. Figure E.1 takes the year 2000 as an example to show graphically how to construct the synthetic income age profile using the simulated data from three birth cohort groups. All other model specifications remain unchanged as described in the basic model.

Appendix F

Derive Expression for Expected CV

By the law of iterated expectation, the computation of expected CVs breaks down into two parts.

$$\mathbb{E}CV(H_a) = \mathbb{E}(CV|mob = 1, H_a)P(mob = 1|H_a) + \mathbb{E}(CV|mob = 0, H_a)P(mob = 0|H_a)$$

[34] Each group of workers has an age span of 14-15 years. For example, workers aged between 22 to 36 are considered to be young workers, while workers aged 37-51 and aged 52-65 are considered as middle-age and old workers respectively.

Assume that the skill rental rates change from $\begin{pmatrix} \Pi_{old}^{cs} \\ \Pi_{old}^{ncs} \end{pmatrix}$ to $\begin{pmatrix} \Pi_{new}^{cs} \\ \Pi_{new}^{ncs} \end{pmatrix}$. For an individual age a with occupation specific human capital $\begin{pmatrix} H_a^{cs} \\ H_a^{ncs} \end{pmatrix}$, I derive the expected CV given his or her human capital H_a. Let d_{old} and d_{new} denote this individual's occupational choices under old and new market conditions respectively. $mob = 1$ if occupational mobility is desirable when market conditions change. mob takes value 1 either in the case $\{d_{old} = ncs,\ d_{new} = cs\}$ or$\{d_{old} = cs,\ d_{new} = ncs\}$. When $mob = 0$, individuals would not exercise their options of occupational mobility anyway, which implies $\mathbb{E}(CV|mob = 0, H_a) = 0$. This simplify the above formula

$$\mathbb{E}CV(H_a) \quad = \quad \mathbb{E}(CV|mob = 1, H_a)P(mob = 1|H_a) \tag{21}$$

Since$\{mob = 1\} = \{d_{old} = ncs, d_{new} = cs\} \cup \{d_{old} = ncs, d_{new} = cs\}$, use again the law of iterated expectation.

$$\begin{aligned} &\mathbb{E}(CV|mob = 1, H_a) = \\ &\quad \mathbb{E}(CV|d_{old} = ncs, \quad d_{new} = cs, mob = 1, H_a)P(d_{old} = ncs, d_{new} = cs|mob = 1, H_a) \\ &\quad +\mathbb{E}(CV|d_{old} = cs, \quad d_{new} = ncs, mob = 1, H_a)P(d_{old} = cs, d_{new} = ncs|mob = 1, H_a) \end{aligned} \tag{22}$$

$\{d_{old} = ncs, d_{new} = cs\}$ and $\{d_{old} = ncs, d_{new} = cs\}$ are two mutually exclusive events. This implies

$$P(mob = 1|H_a) = P(d_{old} = ncs, d_{new} = cs|H_a) + P(d_{old} = ncs, d_{new} = cs|H_a).$$

$$P(d_{old} \quad = \quad ncs, d_{new} = cs|mob = 1, H_a) = \frac{P(d_{old} = ncs, d_{new} = cs|H_a)}{P(mob = 1|H_a)} \tag{23}$$

Substitute equation (22) and (23) into (21).

$$\begin{aligned} \mathbb{E}(CV|H_a) = \mathbb{E} \quad &(CV|d_{old} = ncs, d_{new} = cs, H_a)P(d_{old} = ncs, d_{new} = cs|H_a) \\ +\mathbb{E} \quad &(CV|d_{old} = cs, d_{new} = ncs, H_a)P(d_{old} = cs, d_{new} = ncs|H_a) \end{aligned} \tag{24}$$

By definition,

$$P(d_{old} = ncs, d_{new} = cs|H_a) = P(V_{old}^{cs} < V_{old}^{ncs}, V_{new}^{ncs} < V_{new}^{cs}|H_a)$$

The value functions derived in the model part are as follows.

$$\begin{aligned} V_{old}^{cs} \quad &= \quad \Pi_{old}^{cs} H_a^{cs} + \eta_a + \beta \mathbb{E} V_{old}^{'}(d_{old} = cs) \\ V_{old}^{ncs} \quad &= \quad \Pi_{old}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{old}^{'}(d_{old} = ncs) \end{aligned}$$

$d_{old} = ncs$ when $V_{old}^{cs} < V_{old}^{ncs}$. This implies

$$\eta_a^{old} < (\Pi_{old}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{old}^{'}(d_{old} = ncs)) - (\Pi_{old}^{cs} H_a^{cs} + \beta \mathbb{E} V_{old}^{'}(d_{old} = cs)).$$

By the same computation, $d_{new} = cs$ implies

$$\eta_a^{new} > (\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new} = ncs)) - (\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new} = cs)).$$

The taste shocks under the old and new skill rental rates are independent draws from the same normal distribution.

$$\begin{aligned} P(d_{old} = \quad ncs, d_{new} = cs|H_a) &= P(V_{old}^{cs} < V_{old}^{ncs}|H_a) P(V_{new}^{ncs} < V_{new}^{cs}|H_a) \\ &= \Phi(\tfrac{(\Pi_{old}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{old}^{'}(d_{old}=ncs)) - (\Pi_{old}^{cs} H_a^{cs} + \beta \mathbb{E} V_{old}^{'}(d_{old}=cs))}{\sigma_{\eta_a}}) \\ &\Phi(\tfrac{(\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=cs)) - (\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=ncs))}{\sigma_{\eta_a}}) \end{aligned} \tag{25}$$

Then, the expected CV required by individuals to compensate for restricted occupational mobility is given by the following equation

$$\begin{aligned} \mathbb{E}(CV|d_{old} = ncs, d_{new} \quad &= cs, H_a) = \mathbb{E}((\Pi_{new}^{cs} H_a^{cs} + \eta_a + \beta \mathbb{E} V_{new}^{'}(d_{new} = cs)) \\ &-(\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new} = ncs))|d_{new} = cs, H_a) \end{aligned} \tag{26}$$

Taste shocks are normally distributed $\eta_a \sim N(0, \sigma_{\eta_a}^2)$.

$$\begin{aligned} \mathbb{E}(\eta_a|\eta_a > \quad &(\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new} = ncs)) - (\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new} = cs))) \\ &= \sigma_{\eta_a} \frac{\phi(\frac{(\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=ncs)) - (\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=cs))}{\sigma_{\eta_a}})}{1 - \Phi(\frac{(\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=ncs)) - (\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=cs))}{\sigma_{\eta_a}})} \end{aligned} \tag{27}$$

Substitute equation (27) into (26).

$$\begin{aligned} &\mathbb{E}(CV|d_{old} = ncs, d_{new} = cs, H_a) \\ &= (\Pi_{new}^{cs} H_a^{cs} + \quad \beta \mathbb{E} V_{new}^{'}(d_{new} = cs)) - (\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new} = ncs)) \\ &+ \sigma_{\eta_a} \quad \frac{\phi(\frac{(\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=ncs)) - (\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=cs))}{\sigma_{\eta_a}})}{1 - \Phi(\frac{(\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=ncs)) - (\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=cs))}{\sigma_{\eta_a}})} \end{aligned} \tag{28}$$

The similar computation is applicable to the other case.

$$
\begin{aligned}
&P(d_{old} = cs, d_{new} = ncs | H_a) \\
&\quad = \quad \Phi\left(\frac{(\Pi_{old}^{cs} H_a^{cs} + \beta \mathbb{E} V_{old}^{'}(d_{old}=cs)) - (\Pi_{old}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{old}^{'}(d_{old}=ncs))}{\sigma_{\eta_a}}\right) \\
&\qquad \Phi\left(\frac{(\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=ncs)) - (\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=cs))}{\sigma_{\eta_a}}\right)
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
&\mathbb{E}(CV | d_{old} = cs, d_{new} = ncs, H_a) \\
&\quad = (\Pi_{new}^{ncs} H_a^{ncs} + \quad \beta \mathbb{E} V_{new}^{'}(d_{new} = ncs)) - (\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new} = cs)) \\
&\qquad + \sigma_{\eta_a} \quad \frac{-\phi\left(\frac{(\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=cs)) - (\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=ncs))}{\sigma_{\eta_a}}\right)}{\Phi\left(\frac{(\Pi_{new}^{cs} H_a^{cs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=cs)) - (\Pi_{new}^{ncs} H_a^{ncs} + \beta \mathbb{E} V_{new}^{'}(d_{new}=ncs))}{\sigma_{\eta_a}}\right)}
\end{aligned} \tag{30}
$$

ϕ and Φ are the pdf and the cdf function for the standard normal distribution respectively.