Immigrant Wage Growth in the United States: The Role of Occupational Upgrading

Rebecca Lessem[*] and Carl Sanders[†]

August 20, 2019

Abstract

Immigrants to the United States routinely take jobs below their skill qualifications because of barriers to entering occupations. We use a structural model of immigrant job choice to quantify the benefits of potential policies to promote entry into suitable occupations. We estimate the model using longitudinal labor market data on immigrants to the US. Our counterfactual results show that eliminating barriers to occupational entry would lead to only a small earnings increase for the average immigrant in our sample, but a substantial earnings increase for the most highly skilled immigrants.

JEL Codes: J31, J15, J62

[*]Tepper School of Business, Carnegie Mellon University. rlessem@andrew.cmu.edu

[†]Department of Economics, Washington University in St. Louis. carlsanders@wustl.edu

Thanks to seminar participants at Kansas State University, the London School of Economics, Montana State University, New York University, University College London, University of Virginia, the Society for Economics Dynamics, the Society of Labor Economics, the North American Summer Meeting of the Econometric Society, and the Southern Economics Association for helpful comments. Financial support from the National Center for Border Security and Immigration at the University of Arizona is gratefully acknowledged.

1 Introduction

Immigrants to the United States commonly take jobs below their true skill levels because of barriers to entry into occupations. Over time, immigrants may move up the occupational ladder and find jobs that match their skill levels, but at the cost of foregone wages and lost productivity. Some policy interventions aim to match new immigrants with jobs that fit their skill levels rather than simply leaving them on the bottom rungs of the job ladder.[1] However, in the US, these programs are typically small in scale and run by non-profits rather than the government.[2] If there are immigrants who need training to navigate occupational barriers in the US labor market, the shortage of existing programs to help them may be a missed opportunity. However, these programs have costs, and while their potential benefits for the wage growth and labor market assimilation of US immigrants are likely positive, they have not been quantitatively evaluated.

In this paper, we quantify the benefits for US immigrants of facilitating entry into occupations at their skill levels. Previous work has shown that occupational upgrading is responsible for a large portion of immigrants' wage growth. Eckstein and Weiss (2004) and Weiss et al. (2003) demonstrate the importance of both firm and occupational transitions for the wage growth of highly-skilled Russian immigrants to Israel. We build on this work by quantitatively evaluating the potential benefits for both earnings and occupational attainment of policies that reduce the remaining barriers to occupational mobility. To do this, we construct a model of occupational search and estimate it using labor market and demographic data from the New Immigrant Survey (NIS), a survey of new permanent US residents. Consistent with prior work, we find that occupational mobility is an important component of immigrants' wage growth, but we find only small average wage gains from policies that reduce remaining occupational frictions. Our small estimated average returns to reducing occupational barriers mask wide dispersion in the returns to eliminating these frictions. We find the largest effects for the most highly skilled immigrants, suggesting that policies aiming to reduce barriers to occupational entry have significant distributional effects.

In our model, immigrants make occupational choices over their careers in the US. Their wages are a function of observable and unobservable skills, labor market experience, and their current occupation. Every period, workers either remain at their previous occupation, receive shocks into and out of the unemployment pool and the labor force, or get an outside offer from another occupation, drawn from a distribution that depends on their skills. Given the jobs avail-

[1]For example, the Express Entry program in Canada attempts to find appropriate jobs for any prospective immigrants to ease their entry into the domestic labor market.

[2]A typical example of these sorts of programs is the Community Refugee and Immigration Services, a small non-profit in central Ohio which helps immigrants with job searching and interview skills.

able to them, workers choose their career path of jobs to maximize expected wages. We parameterize observed immigrant skill levels as a function of the individual demographics available in the NIS, which include measures such as English skills and type of US entry visa that are often not available in standard Census-based data sets. We show that the offer distributions are non-parametrically identified and estimate the model by simulated maximum likelihood.

Using the model estimates, we perform counterfactuals to quantify the effect of removing occupational frictions on immigrants' wages across their careers. In the first counterfactual, we begin each worker's US career in the same occupation they worked in their home country before migrating to the US. In this counterfactual, average wages increase by 18% at entry, and after 10 years the gain from the counterfactual is around 2.5%, as immigrants typically catch up to or surpass their home country occupation with occupational mobility in the US. In the second counterfactual, we estimate how immigrants' wages would evolve if the immigrants were immediately placed in their model-predicted long run job, effectively skipping many jobs on the way up the job ladder. In this scenario, wages increase by a substantial 25% on average at entry, but the wage gains decrease quickly over time, to around 2.8% after 10 years in the US. Neither of these counterfactuals are directly implementable policy options; however, we see them both as plausible upper bounds on the impact of the removal of barriers to occupational mobility.

The long-run effects of the counterfactuals are small for the average immigrant, but this small average effect masks significant heterogeneity across skill groups. For example, considering only immigrants who come from the top 10% of the highest-paying occupations in their home countries, the home-job counterfactual raises wages by 38% at entry, which is a much larger effect than for the average immigrant. The positive relationship between pre-immigration skills and the wage benefits from eliminating occupational frictions shows that labor market assistance for immigrants may have significant distributional consequences. On one hand, if countries want to assimilate high-skilled immigrants quickly to potentially boost innovation (in line with the findings from Hunt and Gauthier-Loiselle (2010)), reducing occupational frictions can have a significant effect on the most-skilled immigrants; on the other hand, the policy will not help the immigrants in the most need.

The previous literature on immigrant wage growth using US data has generally been focused on documenting the existence and extent of wage assimilation between immigrant and native workers: Immigrants start out in the US earning lower wages than comparable natives, but the gap falls with increasing experience in the US. Chiswick (1978), Borjas (1985), and LaLonde and Topel (1992) document assimilation using cross-sectional data from the US Census, and Duleep and Dowhan (2002) and Lubotsky (2007) do so using longitudinal data from Social Security Administration records. The results of these studies differ in the specifics depending on the data set and time-frame, but they all document the general phenomenon of wage assimilation.[3] They do not, however, go beyond documenting the existence and extent of wage assimilation. Here, we consider the role of a particular set of policies that have been suggested to help immigrants speed up the process of assimilating, and we calculate the average benefits we could expect to see, as well as how these benefits are distributed across different immigrants.

Our primary contribution is quantifying the remaining occupational barriers that immigrants face in the US. There is a small group of papers that study the importance of current rates of occupational upgrading for immigrants, but they typically focus on documenting the existence of occupational upgrading rather than evaluating how useful policies to quicken upgrading would be. As mentioned above, Eckstein and Weiss (2004) and Weiss et al. (2003) look at the role of firm and occupational transitions for the wage growth of a non-representative sample of highly-skilled Russian immigrants to Israel. de Matos (2012) shows reduced-form evidence on immigrants moving to more productive firms over time using linked employer-employee data from Portugal. Imai et al. (2018) use Canadian data to show that home country occupation predicts immigrants' wage growth, but the authors do not explicitly consider occupational upgrading within Canada or quantify the effects of home country occupation on immigrants' wage path. In the US context, Akresh (2008) documents the degree of occupational downgrading when immigrants enter the US and the resulting occupational upgrading after arrival.

Overall, our findings show that there is a role for additional policies that remove the barriers to occupational mobility, but mostly for high-skilled immigrants. Since we find these barriers have little impact on low and medium-skilled immigrants, the difficulties these groups face in the labor market are not due to an inability to find the right job, but come from another source. For example, they may face persistent discrimination in the labor market, immigrants' outcomes may be due to differences in training and education, or immigrants' skills may not be a good match for the US labor market. Finding policies that can help all immigrants ease their transition into a new country will require consideration of other mechanisms that limit wage growth and occupational attainment.

2 Data and Descriptive Statistics

2.1 Data Sources

The New Immigrant Survey (NIS) program conducted in-person surveys of a random sample from immigrants granted US Legal Permanent Resident (LPR) status between May and Novem-

[3] There is some argument over whether recent cohorts are still seeing wage growth in the US; see Borjas and Friedberg (2009).

ber 2003, becoming (colloquially) "green card" holders.[4] The green card recipients, along with their spouses, were primarily interviewed at some point between May and November at the location where the LPR documentation was sent. The immigrants responded to questions about demographics, labor market outcomes, and their migration history. All of the information is self-reported.

The NIS contains a number of demographic and labor market measures not typically available in Census-based data sets. The personal demographics we use are gender, year of birth, education, year of entry in the US, home country, type of US entry visa, home country occupation, and English skills. Respondents report their wage, occupation, and firm tenure at their first and current US job. We use this information to construct a panel of wages and occupations for immigrants in the US.

In the data, we see when a person entered the US, and the year they started their first job in the US labor market. A limitation to the NIS data is that, in the case when a person is not working when they enter the US, we do not know whether they were unemployed or out of the labor force prior to the first job. In addition, since the NIS only asks about the first and current US job, we are always missing any information on jobs between the first job and the current job. These missing data are endogenous with respect to occupational upgrading, since a worker who moves jobs often will have more missing jobs than a worker who never moves jobs. In the estimation of the model, we will deal with these missing data issues. We treat other forms of missing data (e.g., no wage reported for some jobs) as exogenous.

The NIS reports immigrants' English skills (self-classified as poor, fair, good, or excellent) at the time of the survey. We use English skills grouped into "low" (poor and fair) and "high" (good and excellent) as part of our skill measures in estimation. However, these skills are only reported at the time of the survey, which may not necessarily reflect a person's English skills when they entered the US. Language acquisition after immigration likely plays an important role in occupational and wage growth. Cohen-Goldner and Eckstein (2008) and Berman et al. (2003) find that language skill acquisition has high returns in wages, although Berman et al. (2003) only sees this effect in high-skill occupations. To account for this, in our analysis we allow for immigrant English skills to change over time and estimate the English skill acquisition process.

To create our sample, we include LPR recipients who were living in the US at the time of the interview. We drop immigrants with under one year durations in the US, due to the limited amount of information they provide about occupational transitions. We also drop observations with missing demographic data, with most of the cuts coming from people without home occupation information. See Appendix A for the sample creation details.

[4] See Jaso et al. (2006) for more details on the NIS.

Characterization of occupations is key for our analysis. We use data on three reported occupations: a person's occupation in their home country, their initial occupation in the US, and their occupation in the US at the time of the survey. These occupations are coded in the NIS using 3-digit 2000 Census occupational codes, with about 400 unique codes. Without aggregation, there are far too many occupational cells relative to individual observations to perform inference. To avoid this issue, we characterize occupations by calculating the average wage across all workers in each occupation in the Current Population Survey (CPS) in the year that a person started a given job.[5] This procedure takes the average hourly wage of each occupation as a proxy for the productivity of the job, which we call the *job quality* in the remainder of the paper.[6] In our interpretation, a worker who moves to a higher-quality job moves up the occupational ladder.[7]

2.2 Descriptive Statistics

Table 1 shows sample summary statistics from the NIS. The average immigrant is around 38 years old, and the sample is about 57% male. The average immigrant had been in the US for 2.9 years at the end of their first job. On average, at the time of the survey, an immigrant had been in the US for 8.7 years. About 65% of the sample has had some schooling beyond high school, and around 41% of the sample reports high English skills at the time of the survey.

One of the unique features of the NIS data is that we know the visa status for each immigrant when they entered the US. Since the sample consists of LPR recipients, we know that they all eventually received a green card, but there is some heterogeneity in their visa status at entry. About 77% of the sample entered using a valid visa, meaning that about 23% of our sample

[5] An alternative but similar method is to calculate the percentile of the wage distribution for each occupation, as in Autor and Dorn (2009). Early versions of this paper used this method, and the qualitative results were similar.

[6] One concern with using the starting average wages of a job as a measure of job quality is that workers may forecast that some jobs (e.g. manufacturing) will have falling wages in the future, so current average occupation wage is not a good measure of actual job quality. Allowing for forward-looking behavior such as this would make our model below dynamic and invalidate our ability to identify model parameters. To understand the magnitude of this concern, we have also run our estimation and counterfactuals defining job quality as the average occupation-level wage in the *final* period a worker holds their job. These results are essentially identical to what we show in this paper and available on request, helping to alleviate the concern that long-term changes in the job quality distribution bias our results.

[7] Another factor that could affect occupation transitions when moving to the US is the license requirements of a job, as licensing requirements can act as a barrier to entry into occupations even for domestic workers, as in Friedman and Kuznets (1954). Licensing requirements are not restricted to just high skill occupations, and can exist in lower skill occupations such as cosmetology (Kleiner, 2000). To address this issue, we collected data (from the O*NET website) on whether a license is required for a given job and use this as a control variable. It did not affect our results in early specifications so we do not include this in our final analysis.

entered as an undocumented immigrant.[8] We also know the type of visa that each immigrant received, with 36% of the sample moving to the US on a visa sponsored by an employer. Immigrants who obtained employer sponsorship likely had a job offer before moving to the US, so we expect them to be higher-skilled workers and to suffer less of a drop in their job quality after moving to the US. Most of the remainder of those with valid visas entered on family reunification visas.

While this sample is representative of the households of LPR recipients, it is not representative of all US immigrants, because it does not contain information on immigrants who never apply for LPR status or those who apply and are not granted a green card. We expect the sample selection to bias our results towards higher wages and workers in higher-skilled occupations relative to a representative sample of US immigrants. Most obviously, it takes both time and money to apply for and obtain a green card. A second concern is that immigrants who are unsuccessful in the US are likely under-represented in the pool of LPR applicants and recipients, since they will be more likely to leave the US for their home country. Lubotsky (2007) emphasizes that failing to consider migrants returning home can bias wage assimilation estimates upwards. After estimating the model, we find that the returns to the counterfactuals we consider are higher for the most skilled immigrants, who may also be the least likely to return to their home countries. Therefore we expect that the returns to the counterfactuals are lower for those who return to their home countries and our estimated average effects would be smaller in a sample that also included return migrants.

To understand the extent to which our sample differs from the overall population of immigrants in the US, we calculated basic summary statistics on the sample of all immigrants in the 2003 Current Population Survey. Of individuals who were born abroad, the average age is 37, 45% have attended college, and 57% are male. The average age and gender composition of the NIS sample are similar to those of the overall immigrant population, but the NIS sample has a significantly higher percentage of immigrants with some college education (65% vs. 45%). We expect this upwards bias in observable skills to be matched with an upward bias in unobserved ability. Even though our data are not a representative sample of all US immigrants, representative data sets lack many pre-immigration characteristics that are included in the NIS, which we find to be significantly correlated with the returns to reducing occupational frictions. Our references to the "average" immigrant (from the NIS) should be interpreted as reflecting a migrant to the US with skills slightly above those of the average US immigrant.

[8] We know that this group eventually receives visas, but we do not know when this happens. Additionally, visa status at entry is self-reported, so there is no way to know to what extent invalid entries to the US are being under-reported.

2.3 Occupational Upgrading

In this subsection, we describe how immigrants in our sample moved up the occupational ladder with time in the US and how this relates to their pre-immigration demographic characteristics and labor market experiences. Figure 1 shows the distribution of job qualities for the final occupation in the home country, initial occupation in the US, and current occupation in the US; the sample is split by education level and home country labor market experience. Across the board, we see an increase in the mass of immigrants working in low-quality jobs when they first move to the US. However, fewer migrants work in those low quality jobs at their current job in the US than at their initial US job, indicating upward mobility with time in the US. Comparing panels (a) and (b), which split the sample by education, we see that the shift from lower- to higher-quality jobs between the initial and current US job is much more pronounced for people with high levels of education. Panels (c) and (d) of Figure 1 split the sample based on years of home country work experience. Immigrants with large amounts of home experience suffer a larger drop in their job quality after moving to the US, indicating that some human capital may fail to transfer between foreign and US jobs.

We next look at the determinants of an immigrant's jobs in the US. The first column of Table 2 shows the results from a regression of an immigrant's job quality at US entry onto their pre-immigration characteristics. Demographic factors move in the expected direction: education and English skills are associated with higher-paying occupations. Immigrants with employer-sponsored visas place in higher-quality occupations, as do those who enter the country on valid visas. We control for a worker's home country job quality as a measure of that worker's skill level, allowing the effects of home job quality to vary based on whether an immigrant moved when less than 18 years old, since jobs pre-age 18 are likely less informative about skill levels. For immigrants who move at either age, having a higher-quality job at home is associated with a higher-quality job in the US.

Overall, the regression results show significant variation in the predicted initial job level by demographics. The regression predicts that a (hypothetical) "low-skill" immigrant who was in the lowest-quality job in their home country, has low English skills, and no education post-high school[9] would begin in the occupation located in the 25th percentile in the distribution of job qualities for immigrants observed in the NIS data. Repeating the exercise for a "high-skill" immigrant, with high English skills, a high level of education, and the highest possible home job, we find a predicted job quality in the 98th percentile. Much of this difference is driven by the home country job. If we take this "high-skill" immigrant and put them in the lowest-quality job at home, holding everything else constant, that immigrant's US initial occupation would

[9]We assume the rest of the demographic variables are at the mean for continuous variables and the mode for discrete variables.

now be at the 56th percentile of the distribution instead of the 98th.

Regressing the quality of the current job onto demographics and the initial US job (shown in the second column of Table 2) finds similar trends. Job quality increases with US work experience. We see a weaker relationship between home country occupation and job quality than for the initial job, which is unsurprising given that we also control for initial job quality. This regression suggests that the job growth rates of higher-skilled immigrants are faster than those of lower-skilled immigrants, since even when conditioning on initial job and time in the US, English skills still has a significant effect on the quality of the current job.

The descriptive statistics are informative for the overall degree of occupational upgrading in the sample. As usual, there are a variety of endogeneity concerns when trying to understand any causal relationships in the foregoing regressions. Above and beyond the usual concerns that immigrants with higher quality jobs have higher unobserved abilities and luckier job search outcomes, the sampling scheme introduces correlations between job mobility and missing data, as respondents only report information on their first and current US job. Workers who move jobs more often will have more missing years of observations, while we know the entire career path of a worker who never changes jobs. To deal with issues of selection into jobs and to estimate the role of eliminating remaining occupational frictions in the job choices of immigrants, we estimate a model of the labor market that can take into account both worker choices and the endogenous missing data.

3 Model

3.1 Setup and Initial Conditions

We develop and estimate a partial-equilibrium model of immigrant job choices and wages. In the model, immigrant i is in the US for periods $t = 0, 1, 2, ... T_i$, where time 0 is treated as the entry year and the terminal period T_i is exogenously given.[10] Each period, they can be working at a job, unemployed, or out of the labor force (OOLF). Agents are endowed with a set of observable characteristics X_{it} and a time-invariant discrete type τ, which is unobserved to the econometrician and drawn from discrete PMF $\Upsilon(\cdot)$ independently from the rest of the model. Each worker faces a set of potential jobs, with each job completely characterized by a unidimensional quality measure $\pi > 0$.

[10]In our model, period t is the number of years a person has been in the US. The terminal period is the year of the survey, so T_i is the number of years a person has been in the US at the time of the survey.

3.2 Model Timing and Job Choices

The model timing works as follows. At the start of period t, an immigrant's job status carried over from the previous period can be (1) employed in a job of quality $\pi_{i,t-1} > 0$, (2) unemployed, which we denote by $\pi_{i,t-1} = 0$, or (3) OOLF, which we denote by $\pi_{i,t-1} = -1$. For exposition, we first consider workers who ended the last period in the labor force. At the start of the period, the worker receives a shock with probability $O(\pi_{i,t-1}, X_{it}, \tau)$ that has them choose to exit the labor market. If they stay in the labor market, they next receive a shock that sends them into the unemployment pool with probability $q(X_{it})$. Next, with probability $p(X_{it})$,[11] the worker receives a new job offer, with the quality of the new offer π' drawn from the continuous offer distribution function $\Pi(\cdot|X_{it})$. Finally, if a worker was not fired and receives a job offer, they choose between their previous job and the new job offered. The chosen job becomes their recorded period t job and they are paid wages according to the wage function described below. Unemployed workers and workers not in the labor force workers are paid zero. Figure 2 shows the job transition timing for workers who were in the labor market at the start of the period.

The previous description only considered workers who were in the labor market at the start of the period. People can also potentially be OOLF in the previous period, and these workers choose whether or not to re-enter the labor market this period. If a worker is OOLF in the prior period, we assume they receive a shock with probability $RE(X_{it})$ that sends them back into the labor market. If the worker re-enters the labor market, they receive a job offer with probability $p(X_{it})$ as above. Because not all workers receive job offers each period, it is possible that a person can re-enter the labor market, but not get a job offer. In this case they would transition from OOLF to unemployed. In the initial period, $t = 0$, immigrants begin the period unemployed before receiving their OOLF and job shocks.[12]

3.3 Wages

The wages of employed workers are a function of their characteristics X_{it}, job quality π_{it}, and unobserved type τ. Log wages for person i in year t since entry are given by

$$\log[W_{it}(X_{it}, \pi_{it}, \tau)] = k_1(X_{it}, \pi_{it}, \tau) + k_2(t) + \varepsilon_{it}, \tag{1}$$

[11] The probability of a job offer $p(X_{it})$ is required to be the same for employed and unemployed workers for reasons discussed in Section 3.5.

[12] The OOLF probabilities depend on one's previous job. Since there is no previous job in the first period, we use their home country job quality.

where $k_1(\cdot)$ is a parameterized function of worker observables and unobservables, $k_2(t)$ is a parameterized time trend, and with the restrictions that a) ε_{it} is white noise statistically independent of the rest of the model, b) workers do not act on the ε_{it}, e.g., they cannot forecast ε_{it} when choosing their job between periods, and c) the deterministic component of wages is weakly increasing in π_{it} almost everywhere. Assumptions a) and b) ensure there is no selection on the idiosyncratic unobservable ε_{it}, ensuring that a given worker's wages are always increasing in job quality π across time, while assumption c) makes all workers have the same wage ranking of occupations. All three of these assumptions will be key for the worker's optimal policy, discussed below. To control for differential labor market dropout patterns across immigrants causing selection biases, we include our unobserved heterogeneity term τ in both the wage and labor market drop out rate processes.

3.4 Decision Problem

Workers make choices to maximize their lifetime expected discounted wages, with discount factor $\beta = 0.95$, leading to a value function

$$V_0(X_{i0}, \tau) = \max_{A^\star} E_0 \left[\sum_{t=0}^{T} \beta^t W_{it}\left(X_{it}, \pi^\star_{it}, \tau\right) \right], \tag{2}$$

where $A^\star$ is the accept/reject policy function for the worker given demographics X_{it}, type τ, and each possible job and shock history, and $\pi^\star_{it}$ are the (stochastic) outcomes induced by that policy and shock history.

Under our assumptions above, choosing the wage-maximizing job – which is also the job-quality-maximizing offer – in each period will suffice to maximize expected lifetime discounted wages. The effects of accepting an occupational offer are limited to a higher wage in the current period. There are no dynamic effects in the wage function of accepting a job, and accepting the new job would not change the process of future shocks and job offers, so all option values stay the same. Because the policy function is "Accept all offers above the current job", we can write the transition process of the observables in a closed form.[13]

First, we need to define the functions that are a part of the job transition function. Let the job quality of an outside offer be π^{Offer}_{it}. We define the job offer distribution function as

$$\Pi(z|X_{it}) = \Pr\left(\pi^{\text{Offer}}_{it} \leq z|X_{it}\right) \tag{3}$$

Let Offer_{it}, Fired_{it}, Dropout_{it}, and Re-enter_{it} be the realizations of the outside offer, firing

[13] Note this is the same policy function as it would be assuming $\beta = 0$, complete myopia.

shocks, labor force dropout shocks, and labor force re-entry shocks, respectively, all with support $\{0,1\}$. We denote

$$\begin{aligned} p(X_{it}) &= \Pr(\text{Offer}_{it} = 1|X_{it}) && (4) \\ q(X_{it}) &= \Pr(\text{Fired}_{it} = 1|X_{it}) && (5) \\ O(\pi_{i,t-1}, X_{it}, \tau) &= \Pr\left(\text{Dropout}_{it} = 1|\pi_{i,t-1}, X_{it}, \tau\right) && (6) \\ RE(X_{it}) &= \Pr\left(\text{Re-enter}_{it}|\pi_{i,t-1} = -1, X_{it}\right) && (7) \end{aligned}$$

Equation (4) gives the probability of getting a job offer, and equation (5) gives the probability of being fired. Equation (6) gives the probability of exiting the labor market, and equation (7) gives the probability that someone who is out of the labor market chooses to re-enter the labor market.

With our model setup, the observed job transition function $f^{\tau}\left(\pi_{it}|\pi_{i,t-1}, X_{it}\right)$ can be written as a Markov process conditional on the previous job, observables, and the unobserved type. There are 10 possible sets of previous state/current state transitions, e.g. unemployment to working, OOLF to OOLF, etc. For example, if the worker is employed at a job of quality $\pi_{i,t-1}$, receives a firing shock (happens with probability $q(X_{it})$) but does not receive a job offer (happens with probability $1-p(X_{it})$), the worker will be unemployed this period, $\pi_{it} = 0$.

A full table of the transitions and shocks that lead to those outcomes is shown in Table 3. The first row of the table shows potential transitions when the person was OOLF in the previous period. If they do not get a labor market re-entry shock, they remain OOLF. If they receive a re-entry shock, they re-enter the labor market and will be unemployed if they do not receive a job offer, and will accept the offer if they do. The second row shows the case when a person was unemployed in the prior period. If they receive an OOLF shock, they exit the labor market. If they do not receive a OOLF shock, then they remain unemployed if they do not receive a job offer, and accept the offer if they receive one. The third row shows people who were employed at a job with quality $\pi_{i,t-1}$ in the previous period. If they receive a labor market dropout shock, they exit the labor market. The remainder of the cases in this row are when they do not receive an OOLF shock. First, if they are fired and do not receive a new job offer, they move to unemployment. To stay at their current job, they first must not be fired. Then there are two possible scenarios where they stay at the job. The first is if they do not receive a job offer, and the second is if they receive an offer but it has lower quality than their current job. The fourth column is when this person moves to a new job. This happens if (1) they are not fired and receive a job offer that is higher quality than their current job, or (2) they are fired and receive a new job offer. This last case demonstrates how our model accounts for downwards job moves. This occurs when a worker is fired and then immediately gets an offer, and while there is an intervening unemployment "spell" it does not last long enough to be observed.

The econometrician observes job qualities without error and wages with independent additive measurement error. We do not see data on the firing shocks, offer shocks, or offered jobs, but only see accepted offers. An observable sample path generated by the model (for a worker who stays in the labor market) is shown in Figure 3(a). The model can generate workers spending multiple periods in the same job, upwards and downwards job transitions, movements into and out of unemployment, and labor force dropouts and re-entries. However, our particular data will look not look like this since we do not observe full labor market histories; the sampling scheme only records the occupation, wage, and duration of the first job and final job in the labor market, with all information on jobs between those missing. Given the data's sampling scheme, we would not observe this example immigrant's full career: The observed data we would see given this underlying occupational path are shown in Figure 3(b).

3.5 Discussion

The job offer distribution is one of our primary model objects of interest. In our counterfactuals, we predict what workers would do in response to a policy that affected their first job in the US, so we need to know how workers make choices in that new situation. Without knowing the offer distribution, we will not be able to make these predictions. Since we do not observe actual offers or firings, we want to be sure our data can distinguish different offer distributions, given the rest of our model. We can in fact show that with enough data, we could pin down the shape of the offer distribution along with job offer and firing rates. As we do not estimate the model non-parametrically, we reserve our proof of non-parametric identification for Appendix C, and discuss our parametric assumptions in Section 4.2 below.

This non-parametric identification result could be lost by augmenting the model with additional realistic features of the immigrant labor market. Flinn and Heckman (1982) show that in general it is not possible to identify the job offer distribution in search models using only accepted job offers. Here, we can do so because our model implies that the reservation job quality is simply the current job. If individuals had dynamic incentives where the relative value of jobs of different qualities was not stable, this result could no longer be true and the reservation job would be unknown. Three assumptions that would be difficult to relax without affecting the optimality of static wage maximization are 1) the offer probability and offer distributions are the same no matter the worker's employment status or job quality, 2) there are no switching costs, and 3) there are no wage returns to tenure that would be lost by switching between jobs of different quality. Adding any of these characteristics to the model would allow for the possibility that workers would reject some offers from jobs of higher quality than their current job.

The benefit of relaxing these assumptions would be some gain in realism, but at the cost of losing provable non-parametric identification, and thereby also losing some ability to understand what patterns of observed job switches drive our estimates of the job offer process.

4 Estimation

We estimate the model parameters using simulated maximum likelihood (SML). Our model delivers a Markov specification of the likelihood of job choices in each period given the previous job state. However, in the data, we do not observe job choices in every period, and instead only see each person's first and current US job. With this missing data on jobs, computing the likelihood of the observed data requires evaluation of high-dimensional integrals. SML estimation uses simulations to approximate these integrals, and delivers consistency and asymptotic normality results that allow us to do standard inference. In the remainder of this section, we discuss the model parameterization and explain the construction of our likelihood function. The full likelihood derivation and formal identification analysis are in Appendices B and C, respectively.

4.1 English skills

When making job choices, workers are assumed to know their characteristics X_{it}, which include their English skills. However, in the data, we only see a person's English skills at the time of the survey, which may be different than their English skills in prior periods. We use the self-reported data to classify each person's English skills as high or low at the time of the survey. To allow for the evolution of English skills over time, we estimate the English skill acquisition process. We assume that each person enters the US with low English skills. Each period, we assume they can acquire English skills with some probability that depends on their characteristics, and we estimate this transition process jointly with the model parameters. The English skill transition parameters are identified using the variation in English skills at the time of the survey combined with information on how long a person has been in the US.

4.2 Parameterization

We show in Appendix C that the offer rates, firing rates, and offer distribution are non-parametrically identified for any given set of exogenous characteristics X_{it}. However, to estimate the model with reasonable power given our sample size, we specify the parametric relationship between observables and the distribution of both job shock rates and the offer distribution. We allow for two unobserved types of workers, $\tau \in \{0,1\}$, drawn with probability υ_0 and $1-\upsilon_0$, respectively. We parameterize the offer probability $p(X_{it})$, the firing probability $q(X_{it})$, the OOLF shock $O\left(\pi_{i,t-1}, X_{it}, z_i, \tau\right)$, and the labor market re-entry shock $RE(X_{it})$ with the single-index functional forms as

$$\begin{aligned}
p(X_{it}) &\equiv \Phi\left(\alpha_0 + X_{it}'\alpha\right) && (8)\\
q(X_{it}) &\equiv q && (9)\\
O\left(\pi_{i,t-1}, X_{it}, z_i, \tau\right) &\equiv \Phi\left(\gamma_0 + X_{it}'\gamma + \gamma_\pi \pi_{i,t-1} + \gamma_z z_i + \gamma_\tau \cdot \mathbf{1}\{\tau = 1\}\right) && (10)\\
RE(X_{it}) &= r\ , && (11)
\end{aligned}$$

with $\Phi(x) \equiv \frac{1}{2} + \frac{1}{2}\tanh(x)$, a function bounded between 0 and 1 for all x. Equation (8) gives the probability a person gets a job offer each period, which we assume depends on their characteristics. In equation (9), the probability of being fired is constant across workers. We have estimated the model allowing for this to depend on characteristics, but we determined there is not enough statistical power to show a relationship between demographics and firing rates. Instead, for simplicity, we restricted the firing rate to be constant in our final specification. The OOLF process in equation (10) includes demographic characteristics and a person's prior job quality, as well as a term that allows the probability of exiting the labor market to shift based on unobserved type. We also include an additional covariate z_i in equation (10), a dummy variable for whether the immigrant has any children under 18. This variable serves as an exclusion restriction since it is not included in the wage equation. As usual, an exclusion restriction is not strictly necessary for identification given the non-linear model, but would be required for fully non-parametric identification, as discussed in Appendix C.3. In equation (11), the rate of re-entering the labor market is estimated as a constant, as given our data structure[14] we do not have much power to distinguish what characteristics prevent people from leaving the labor market versus helping them re-enter.

We parameterize the job quality offer distribution $\Pi(\cdot|X_i)$ as

$$\begin{aligned}
\pi' &\sim \text{Truncated LN}\left(\mu(X_{it}), \sigma^2\right) && (12)\\
\text{Support}\left(\pi'\right) &= [4, 60] && (13)\\
\mu(X_{it}) &= \psi_0 + X_{it}'\psi. && (14)
\end{aligned}$$

New job offers are drawn from a truncated lognormal distribution, where the mean job offer

[14] As discussed in Appendix C, in the NIS data we only observe OOLF status at the time of the survey. We must indirectly infer previous OOLF status using our model.

depends on a person's characteristics. Initial estimates suggested demographics do not play a major role in estimates of the standard deviation of job offers σ, so the parameter was assumed to be homogeneous.

We parameterize the wage equation as

$$\log(W_{it}) = \beta_0 + X_{it}'\beta + \gamma \cdot \pi_{it} + \beta_\tau \cdot \mathbf{1}\{\tau = 1\} + \delta_1 t + \delta_2 t^2 + \varepsilon_{it}. \tag{15}$$

Wages depend on exogenous characteristics X_{it} as well as unobserved type τ, which shifts log wages by a constant. As discussed in Section 2.1, our empirical measure of job quality π_{it} is the average US wage in that occupation at the time the worker starts the job. We also include a quadratic time trend, where t denotes the number of years a person has been in the US.

As explained in Section 4.1, we assume that all workers enter the US with low English skills and transition to high English skills with some probability each period. Denote $\tilde{X}_{it}$ as the set of all characteristics excluding English skills, and let E_{it} be a dummy variable for having high English skills. Then the full set of characteristics in each period can be written as $X_{it} = \{\tilde{X}_{it}, E_{it}\}$. The probability of moving from low to high English skills is defined as:

$$\Pr(E_{it} = 1 | \tilde{X}_{it}, E_{i,t-1} = 0) = \text{Logistic}(\gamma_0 + \gamma_1 \cdot \mathbf{1}\{t = 0\} + \gamma' \tilde{X}_{it}) \tag{16}$$

We allow for a shifter in the first period, so that the "base rate" of high English skills at US entry can match the data rather than just depend on the per-period chance of moving to high English each period while in the US. We assume that having high English skills is an absorbing state, so $\Pr(E_{it} = 1 | \tilde{X}_{it}, E_{i,t-1} = 1) = 1$.

The functional forms in this section all depend on X_{it}, an individual's characteristics, which can include the year that a person is in the labor market. In our model, we use t to denote the number of years a person has lived in the US, which can affect wages or the job offer distribution. We also include cohort effects in both wages and the job offer distribution. We do not, however, control for calendar year effects for any job outcome, meaning we do not control for fluctuations over the business cycle. We realize this is a limitation of our analysis, but in Section 5.5 we will show that our model still does a good job of fitting the data.

4.3 Likelihood

The key component in the likelihood function is the probability of transitioning between jobs each period. We let the function $f^\tau(\pi_{it} | \pi_{i,t-1}, X_{it})$ denote the likelihood of transitioning from job $\pi_{i,t-1}$ to job π_{it} with demographic characteristics X_{it} and unobserved type τ. This function incorporates the decision on whether or not to leave the labor market, unemployment shocks, the job offer distribution, and the decision on whether to accept or reject a job offer. These one-period likelihoods only depend on the previous job, and not the entire sequence of job offers. Appendix B explains how the function f^{τ} is calculated. To write the likelihood function, we also need the pdf of wage outcomes, which comes from the normal distribution and is denoted as $g^{\tau}(\cdot)$.

For each worker, we see their occupation and wage in their first and final job in the US, and their English skills are only asked at the time of the survey, T_i. We also know the years they worked at each of those two jobs, as well as the year they entered the US, which could be different from the year that the first job started. This leads to three sources of missing data. First, when a person enters the US, before they begin their first job, we do not know whether they were unemployed or OOLF in the years before that job, but simply know they were not working. The second set of missing data comes between the initial and final US job for each person. Since we know when the first job ends and the final job begins, we know if and when we are missing job data for each respondent. Third, we do not know English skills in any but the final period. To compute the likelihood, we will have to account for these missing data.

Denote the quality of the first and final job as π_i^A and π_i^B, respectively, the wages at both of these jobs as W_i^A and W_i^B, and English skills at the time of the survey as X_{i,T_i}.[15] Since we observe job durations, we also know the year the first job started (ys_i^A) and finished (yf_i^A), and the year the final job started (ys_i^B). Denote individual i's observed data as

$$\Omega_{i,\text{obs}} \quad = \quad \left\{\pi_i^A, \pi_i^B, ys_i^A, yf_i^A, ys_i^B, W_i^A, W_i^B, X_{i,T_i}\right\} . \tag{17}$$

The set of missing data consists of employment status each period before the first job, jobs between the first and second job, and English skills prior to the survey year:

$$\Omega_{i,\text{miss}} = \left\{\pi_{i,0}, \pi_{i,1} ..., \pi_{i,ys_i^A-1}, \pi_{yf_i^A+1}, \pi_{yf_i^A+2}, ... \pi_{ys_i^B-1}, X_{i,0}, ..., X_{i,T_i-1}\right\}. \tag{18}$$

If we knew both the observed data and the missing data, their joint likelihood can be written using the model-predicted transition kernels f^{τ}, the wage pdf g^{τ}, and the English transition probabilities, which we write as $\Pr(X_{it})$:

$$L_i\left(\theta|\Omega_{i,\text{obs}}, \Omega_{i,\text{miss}}, \tau_i\right) \quad = f^{\tau}\left(\pi_{i0}|X_{i0}\right) \prod_{t=1}^{ys_i^A-1} f^{\tau}\left(\pi_{it}|\pi_{i,t-1}, X_{it}\right)$$

[15] For ease of notation, we assume English is the only element of X_{it}, but any number of elements of X_{it} with missing data could use the same estimation procedure.

$$
\begin{aligned}
&\times && f^{\tau}\left(\pi_i^A|\pi_{i,ys_i^A-1},X_{i,ys_i^A}\right)\prod_{t=ys_i^A+1}^{yf_i^A} f^{\tau}\left(\pi_i^A|\pi_i^A,X_{it}\right)\\
&\times && \prod_{t=yf_i^A+1}^{ys_i^B-1} f^{\tau}\left(\pi_{it}|\pi_{i,t-1},X_{it}\right)\\
&\times && f^{\tau}\left(\pi_i^B|\pi_{i,ys_i^B-1},X_{i,ys_i^B}\right)\prod_{t=ys_i^B+1}^{T_i} f^{\tau}\left(\pi_i^B|\pi_i^B,X_{it}\right)\\
&\times && g^{\tau}\left(W_i^A|\pi_i^A,X_{i,ys_i^A}\right)\cdot g^{\tau}\left(W_i^B|\pi_i^B,X_{i,ys_i^B}\right)\\
&\times && \Pr\left(X_{i,0}\right)\cdot\prod_{t=1}^{T_i}\Pr\left(X_{i,t}|X_{i,t-1}\right) \qquad (19)
\end{aligned}
$$

Equation (19) is written in a way that splits up the observed and missing data. The first line in equation (19) gives the likelihood of the first set of missing data, which is the person's employment status from entry until the first observed job. The second line is the likelihood of observing a worker in a job of quality π_i^A for the observed duration. The third line gives the likelihood of the missing job path between the first and final job. The fourth line is the likelihood of taking the current job and then for each period they remain at that job. The fifth line is the likelihood of the observed wage outcomes. Finally, the sixth line gives the likelihood of the path of English skills.

The likelihood in equation (19) is infeasible to calculate without observing the missing data. In particular, consider the first term in the second line of that equation. We cannot calculate the probability a person moves to a job with quality π_i^A, because we do not know their prior employment status. The same is true for when they start their current job, π_i^B, because in this case we sometimes do not know their job quality in the prior period. Similarly, we cannot calculate the probability of the English skill transitions without knowing the previous period's English skills. However, we can integrate out the missing data to calculate the likelihood of the observed data:

$$
L_i\left(\theta|\Omega_{i,\text{obs}},\tau_i\right)=\int L_i\left(\theta|\Omega_{i,\text{obs}},\Omega_{i,\text{miss}},\tau_i\right)d\Omega_{i,\text{miss}}\,, \qquad (20)
$$

where $d\Omega_{i,\text{miss}}$ indicates integrating over each element of $\Omega_{i,\text{miss}}$.

Finally, Equation (20) is conditional on unobserved type. To calculate the unconditional likelihood, we assume there are two types, which occur with probability ν and $(1-\nu)$, and integrate (sum) over types:

$$
\begin{aligned}
L_i\left(\theta|\Omega_{i,\text{obs}}\right) &= \nu L_i\left(\theta|\Omega_{i,\text{obs}},\tau_i=0\right)\\
&+ (1-\nu)L_i\left(\theta|\Omega_{i,\text{obs}},\tau_i=1\right)\,. \qquad (21)
\end{aligned}
$$

The full-sample likelihood is then given by the usual product across individuals:

$$L\left(\theta|\left\{\Omega_{i,\text{obs}}\right\}_{i=1}^{n}\right)=\prod_{i=1}^{N}L_i\left(\theta|\Omega_{i,\text{obs}}\right). \tag{22}$$

4.3.1 Calculation and estimation

Calculation of equation (22) requires evaluation of potentially high-dimensional integrals, with each worker with missing data requiring evaluating an integral with one dimension per missing period of data. We evaluate these integrals using the Monte Carlo integration technique of independent importance sampling (see Evans and Swartz, 2000). For each individual, we simulate S independent draws of their missing job paths and English skills, denoted as $\Omega^s_{i,\text{miss}}$, $s=1...S$. These are drawn from the Markov transition kernels $f^\tau(\pi_{it}|\pi_{i,t-1},X_{it})$ for jobs and the English process $\Pr\left(X_{i,t}|X_{i,t-1}\right)$ for English skills. For each simulated path, we fill in the missing job and English data using the simulated data and calculate the likelihood of the that full career path, $L_i\left(\theta|\Omega_{i,\text{obs}},\Omega^s_{i,\text{miss},}\tau_i\right)$, which is simple to calculate using the Markov kernels. Then, to get the approximate integrated likelihood, we average across the likelihoods of the simulated paths, weighting each career's likelihood by the reciprocal of the likelihood of the simulated data, denoted $L^m_i\left(\theta|\Omega^s_{i,\text{miss},}\tau_i\right)$.[16] The simulated individual likelihood function is then the integral across unobserved types:

$$L^S_i\left(\theta|\Omega_{i,\text{obs}}\right)=\nu\cdot\frac{1}{S}\sum_{s=1}^{S}\frac{L_i\left(\theta|\Omega_{i,\text{obs}},\Omega^s_{i,\text{miss},}\tau_i=0\right)}{L^m_i\left(\theta|\Omega^s_{i,\text{miss},}\tau_i=0\right)}+(1-\nu)\cdot\frac{1}{S}\sum_{s=1}^{S}\frac{L_i\left(\theta|\Omega_{i,\text{obs}},\Omega^s_{i,\text{miss}},\tau_i=1\right)}{L^m_i\left(\theta|\Omega^s_{i,\text{miss},}\tau_i=1\right)}, \tag{23}$$

and the criterion we maximize is

$$\hat{\theta}_{SML}=\arg\max_{\theta}\prod_{i=1}^{n}L^S_i\left(\theta|\Omega_{i,\text{obs}}\right). \tag{24}$$

Our estimator $\hat{\theta}_{SML}$ has the SML subscript since Monte Carlo integration of the likelihood is a particular case of the estimation method of Simulated Maximum Likelihood, allowing us to use the associated statistical machinery. As $n\to\infty$ and $S\to\infty$, $\hat{\theta}_{SML}\to_p\theta$, and if $\frac{\sqrt{n}}{S}\to 0$, the distribution of $\hat{\theta}_{SML}$ is asymptotically equivalent to the distribution of the actual maximum

[16]In the language of integration theory, we use the likelihood of the missing data as the "importance sampler," which requires dividing the integrand by this likelihood; again, see Evans and Swartz 2000 for a discussion of the importance sampler.

likelihood estimator $\hat{\theta}_{MLE}$ (Gourieroux and Monfort, 1996).[17]

5 Results

We estimate the model to find the wage and occupational transition parameters. The following subsections explain our parameter estimates.

5.1 Out-of-the-labor-force Process

A worker leaves the labor market with some probability each period, and workers who are out of the labor force re-enter the labor market with some probability. Table 4 shows the parameter estimates for this process. People who work in higher-quality jobs, as well as men, are less likely to leave the labor market. We surprisingly do not find a statistically significant effect of children on the labor market exit process. There are two unobserved worker types in the model, and we find that 21% of workers are what we call the “non-working” type, meaning that they are more likely to exit the labor market. The non-working type leaves the labor market with a probability of around 14% each period, depending on demographics.[18] The “working” types have low probabilities of leaving the labor market, about 1% per period, meaning that most of them do not exit in any period. We also allow for some probability of labor market re-entry each period, finding that someone who is OOLF has a 34% chance of re-entering the labor market in a period. This number seems unusually high, but we would expect our sample to give high values for this parameter since we drop people who are always out of the labor force (as we do not have information on their home country job), meaning the workers remaining in our sample will be closer to the extensive margin.

[17] When estimating the model, we increased the value of S until both changing the initial seed of the random number generator and increasing the number of simulations per worker gave nearly identical results, with $S = 400$ throughout our results.

[18] Calculated for a woman, who has kids, is the average age in our sample, and whose previous job paid $15 per hour.

5.2 English skills

In the data, we observe each person's English skills at the time of the survey, which we classify as high or low, and we do not know their English skills when they entered the US. Since it is likely that these changed over time, we estimate a stochastic process for low vs. high English skills. In particular, we assume that all people enter the US with low English skills, and can transition to having high English skills each period, which we assume is an absorbing state. Table 5 shows the parameters of the English skills transition process. We allow the probability of obtaining English skills to vary with education, home country GDP, and also allow for a shift in the constant term for a person's first period in the US. Immigrants without a college education have a 10% chance of entering the US with high English skills, and each period there is a 3% chance of moving from low to high English skills. For college-educated immigrants, there is a 25% chance of entering with high English skills, and a 10% chance of transitioning from low to high English skills each period after. Conditional on education, we do not find an effect of home country GDP on English skill probabilities.

5.3 Wage equation

The parameters of the wage function, equation (15), are shown in Table 6. We see that wages increase with the quality of a job, and people who worked in higher-quality jobs at home earn higher wages in the US. Wages also increase with experience in the US labor market. Most of the demographic effects work in the expected direction. We see a 37% difference in the wages of working versus non-working types, where unsurprisingly the group that is less likely to exit the labor market earns higher wages.

5.4 Occupational Transition Parameters

We estimate three sets of parameters relating immigrant demographics to the occupational transition process. These results are shown in Table 7. The first column reports the parameters governing job offer rates. Column (2) of Table 7 shows that we estimate a job loss rate of 8.4% each year. This estimate seems fairly high, but since the model allows for job-unemployment-job transitions that we do not observe, our estimate implies that around 4% of worker-years are observed as unemployed, which is in line with standard estimates. The third column of Table 7 shows estimates of the mean and variance of the truncated lognormal job offer distribution. People with higher-quality home occupations get better job offers, as do people with a college education and English skills. Employer-sponsored visas lead to better job offers, perhaps due to higher-quality workers receiving these visas and being more likely to get better job offers over time. Because we expect the home country-to-US transition to differ significantly from within-US job transitions, we include a dummy variable for the first period in the US in the offer distribution mean, and find that initial offers in the US are slightly higher (although not statistically significantly so) than later offers. We estimate the standard deviation of the job offer distribution as a constant.[19]

The magnitudes of our point estimates are demonstrated graphically in Figure 4. The graphs contain the median simulated occupational paths, varying one observable while holding all others constant. Plot (a) looks at the effect of education on occupations. The middle line is the median simulated occupational path for everyone in the sample. Then, we simulate occupational outcomes for each person, once assuming that everyone has a college education, and then again assuming that no one has a college education. Plots (b)-(d) show the results when repeating this exercise for English skills, whether or not an immigrant enters the US on a valid visa, and home country occupation (moving all immigrants to the 25th and then the 75th percentile home occupations in the sample). All of these factors have a large effect on occupations at US entry. English skills and high-quality home country occupation also lead to faster occupational upgrading over time in the US. There are potential measurement concerns about initial visa status, given that undocumented entry may be under-reported in the data. This could reduce our estimated effect of visa status at entry, perhaps explaining why visa status has the smallest effect of the four demographics in Figure 4.

5.5 Model Fit

A comparison of the predicted occupations between the model and the data is shown in Figure 5(a). For each immigrant we have up to two occupation observations over potentially many years. To get predicted outcomes from the model, we simulate each worker's whole career path 400 times given their pre-immigration characteristics. We then sample these career paths using the sampling scheme from the data; that is, we drop any information on occupations before the first occupation and between the first occupation and the occupation in 2003. The model fit of average trends is good, particularly because the model can replicate the non-monotonicity of average occupational paths over time seen in the data. The fact that the average job quality rises with experience for 6 years and then begin to fall is difficult to fit without cohort-specific hetero-

[19]We have estimated versions of the model where both the job loss rate and standard deviation depend on characteristics, but this did not lead to substantially different results, so we chose to estimate them as constants.

geneity, since our model predicts (on average) monotonic career paths. The hump shape arises because of the variation in the demographic composition of workers who have been in the US for 1 year versus 6, as older immigrants tend to have lower-skilled demographics, lower-quality home occupations, and lower earnings. These demographic differences offset the positive job quality growth that accrues with US work experience. We also show the standard deviation of job quality in the model and data, where the model again fits the shape of the data. Panel (b) shows the model fit for wages, again showing that we are able to match the non-monotonic trend of the data.

6 Quantifying the Role of Occupational Upgrading

In this section, we quantify the importance of occupational upgrading for the wage growth of immigrants using two counterfactuals. In the first, which we call the *home job* counterfactual, we consider the ceteris paribus effect of starting immigrants' careers in their home country jobs. In particular, we consider a situation where the worker's characteristics, outside job offers, and labor market shocks were held equal, but their first job in the US had the job quality of immigrant's final job in their home country. The worker then re-optimizes their career choices given this new first job. For example, in this counterfactual, doctors in their home countries would have their first job in the US be a doctor, construction workers in their home country would begin their US careers as construction workers, and so on, and their later occupations over their careers would reflect their new starting jobs. One could consider this counterfactual as eliminating between-country labor market frictions or as increasing the transferability of human capital across countries.

In the second counterfactual, which we call the *long-run job* counterfactual, we place each worker in their (counterfactual) average job in the US after 15 years. To do this, we simulate the model, and find the average job that each person would be in after 15 years in the US labor market. For example, if a person starts as a tailor but their job offer distribution and job offer rates were such that they eventually become a store manager, we simulate their wage path assuming they begin their US career as a store manager.

While neither of these counterfactuals are realistic policy options, we consider our results informative about the gains from policies that reduce occupational search frictions. If actual policies can only partially mitigate occupational frictions, our results provide an upper bound for the impacts of policies that help immigrants match to occupations that match their skill levels. There are likely other reasons, such as skill acquisition in a new economy, that prevent immigrants from immediately being placed in their long-run occupation or in their home country occupation.

These two counterfactuals are similar, given that they both move immigrants to higher quality jobs at entry, but provide different insights about the reductions in occupational search frictions. The long-run jobs counterfactual addresses the concern that the last home occupation may not properly reflect immigrants' true skill levels. For example, we might expect that similar job titles across countries do not reflect equivalent skill levels required, and moving workers to their home country occupation is not a good proxy for moving them to a job that is a good match for their skills. The long-run jobs counterfactual gives a way of determining the potential of workers in the US rather than their potential in their home country. In fact, our results below will show that the quality of many immigrants' long-run US job is substantially higher than their home job quality.

While these counterfactuals are useful to quantify the effects of policies that remove occupational frictions, it is difficult to interpret the absolute magnitudes of our results. To assist with this, we do two reference counterfactual exercises. In the first, which we call the *English skills* counterfactual, we assign each person high English skills at entry into the US, as compared to the baseline where people can transition to high English skills with some probability each period. This reference counterfactual helps give some idea of what we estimate a more easily implementable policy (language training) could accomplish, at least in the best case. In the second reference counterfactual, which we call the *no-upgrading* counterfactual, we simulate outcomes in the case where we do not allow for any occupational upgrading. This allows us to calculate the importance of occupational upgrading for immigrant wage growth, since we can see how much worse wage outcomes would have been if immigrants could not change their initial jobs.

There are a few important caveats regarding our counterfactual results. First, we use a partial equilibrium model in this paper and do not allow for general equilibrium effects of changes in job choices. When we perform our counterfactuals, we move immigrants to higher quality jobs at entry into the US labor market. There are potential equilibrium effects on wages as more workers in these high skilled jobs may put downward pressure on wages. We would not see this effect due to our partial equilibrium setup.[20] In addition, we only look at wage gains from moving immigrants to different jobs. There are also potentially non-pecuniary benefits to helping immigrants get into their preferred job that are not captured in our framework.

[20] Llull (2018) and Ma (2018) consider the role that immigration plays in the determination of equilibrium wages across occupations.

6.1 Counterfactuals: Average Occupation and Wage Effects

To implement the counterfactuals, we simulate 400 different sample paths of job offers and labor market shocks for each immigrant. For each simulated sample path, workers choose their best available job in each period. For the home job counterfactual, each immigrant's first job in the US is set to be the same as their home country job. For the long-run jobs counterfactual, we first simulate the model to find each worker's average job after 15 years, and then replace their first jobs in the US with those long run jobs.

To demonstrate the effects of these policies, in Figure 6(a) we show the average simulated job quality paths over time in the US in the baseline and in each counterfactual. Figure 6(b) shows the differences in job qualities between each counterfactual and the baseline. The baseline initial job has an average job quality of \$14.20 an hour, compared to quality of \$17.67 an hour in the home job counterfactual and \$19.02 an hour in the long run job. This shows that immigrants downgrade their job quality at US entry, but in the long run end up in jobs with higher quality than their home job. For the reference counterfactuals, the English skills counterfactual raises initial job quality to \$17.26, as the workers' first offers will be higher quality if they have high English skills. The no upgrading reference counterfactual does not change initial assignment by construction.

We next compare the job paths across careers in the different counterfactuals. By construction, the impact of the higher initial assignment in the home jobs counterfactual will fade over time as workers get better offers that dominate even their home job, and the long-run effect on jobs is small. In the long-run job counterfactual, since workers start their careers in their long-run job, there will be almost no positive occupational growth on average. In fact, we see negative job growth on average. This is because the initial job in this counterfactual is the person's long run job, which they potentially got after acquiring English skills. However, since in the early years a lot of people do not have strong English skills, if they lose their job they will be drawing a new job offer from a lower skill distribution.

The reference counterfactuals are helpful to understand the estimated size of our occupational upgrading counterfactuals. First, the English skills counterfactual has significantly larger effects on job quality than the home job and long run job counterfactuals, as even after 15 years this counterfactual still leads to higher quality jobs than the baseline. Even though individuals learn English as their careers progress, the dynamic effects on job offers from having good English early in the career makes the effects of this counterfactual persistent. Second, we also show the results from the no upgrading counterfactual; in this case, the average job quality over time is close to flat.[21] In the case of occupations, this line is unimportant, but we will calculate

[21] This line is not exactly flat due to labor market exit decisions and unemployment. Even though each person's job (when employed) never changes, there is still variation in who is working each period.

the wages associated with this no-upgrading path below.

The effects of the counterfactuals on workers' wages over time in the US are shown in Figure 7 and Table 8. To create this table and figure, we calculated the average wages across simulations and workers in each period. The difference between the baseline and the home job counterfactual is approximately $2.05 at entry. Over time, the effects of the counterfactual fall as workers' baseline and counterfactual jobs converge, and after 10 years there is only a $0.62 difference in wages between the baseline and the home jobs counterfactual. For the long-run job, we see a larger increase in counterfactual wages at entry, corresponding to an approximately $2.88 increase in wages. However, similar to the home country job counterfactual, the effects decrease over time as workers move up the occupational ladder, and we see a $0.70 difference in wages in the baseline and counterfactual at year 10. The English skills counterfactual has the largest effects on wages: $4.11 at entry, and $3.92 after 10 years. When looking at the simulation that shuts down occupational upgrading, we see that wage growth is 25% lower than in the baseline. This suggests that occupational mobility accounts for 25% of worker's baseline wage growth over their first 15 years in the US. Panel (b) of Figure 7 shows the difference in wages between each counterfactual and the baseline, demonstrating the reduction in the effects of the counterfactuals over time.

Overall, looking at the results of the long run and home jobs counterfactuals, we see small effects of these policies on the average occupational path of immigrants. The home jobs counterfactual raises wages by 17.9% at entry, but after 10 years the gain from the counterfactual is down to 2.5%. In the long-run jobs counterfactual, which is more of a best-case scenario for immigrants, wages increase by 25.1% at entry, but this gain is temporary and declines to 2.8% after 10 years in the US.

The job quality paths for each immigrant depend on their observed characteristics through the job offer distribution and the probabilities of different shocks, which suggests that the effects of reducing occupational upgrading frictions need not be identical across different subgroups of workers. Repeating our above counterfactual exercises for particular groups of workers bears this out. For example, if we look only at immigrants who were in the top 10% of the home job qualities, we find that the home job counterfactual increases wages by 38% at entry. We explore these differing returns to the counterfactuals in the next section.

6.2 Heterogeneous Returns

The small average effects of the counterfactuals do not mean that the effects are small for every immigrant. The first natural hypothesis is that higher-ability immigrants would gain more from our two main counterfactuals, since it seems likely that higher-skilled jobs may have more barriers to entry across national borders. In Figures 8 and 9 we explore how the returns to each counterfactual vary with skill level. To do this, we create a pre-immigration skill measure by using the estimated wage equation to calculate each immigrant's predicted average wage at entry in the US.[22] This predicted wage at entry serves as our proxy for each immigrant's skill level.

Figure 8 shows the conditional relationship between our constructed skill measure (x-axis) and average entry wages in the US under the baseline and counterfactual scenarios. Panel (a) of Figure 8 gives the absolute levels of wages in each situation. The fact that the baseline is 45 degrees in panel (a) is mechanical due to how we constructed the skill measure. The counterfactual gains compared to the baseline are shown in panel (b), and this displays the results more clearly. We see the largest gains for the highest skilled, particularly for the home job counterfactual. Figure 9 contains the same information as Figure 8 but considers wages after 10 years in the US labor market: The effects of the home job counterfactual fade over time; however, they are still increasing with skill. The gains from the long run job counterfactual are more compressed with skill level. The returns to our reference English skills counterfactual are increasing in skills at US entry, but decreasing in skills after 10 years in the US. Both at US entry and after 10 years in the US labor market, the English counterfactual would help low-skilled workers significantly more than our counterfactuals reducing occupational frictions, and the English counterfactual has significantly better distributional consequences since it is less strongly biased towards helping high-skilled workers.

In the previous exercise, we saw that the returns to reductions in occupational search frictions are increasing in skill level, which we measured using predicted wages. This aggregates across a number of different demographics, and we are interested in determining the role of each of these characteristics in this outcome. To do this, we define an individual-specific average treatment effect (ATE) of each policy, which we will then compare to different characteristics to understand which components are most important. To calculate the ATE, we take a given set of job offers and labor market shocks and see how much starting a worker in their home job or their long-run job changes their occupations and wages. We then calculate the individual's ATE as the average wage gap between their baseline and counterfactual over the different paths of the unobserved offers and shocks. Our individual-specific ATE will reflect how the home jobs counterfactual or the long-run jobs counterfactual affects that individual's lifetime outcomes, averaging out over the luck that drives many transitions in our model. Aggregating these individual-specific ATEs across different groups of workers can be used to construct a variety of treatment effects along the lines of those in Heckman and Vytlacil (2007).

[22] To calculate each person's predicted entry wage, we need to know their initial job and English skills, which we simulate using our model estimates.

Our goal here is to determine the conditional relationship between observable demographics and the individual-specific ATEs, which could be used, for example, to target the policy to workers with the highest expected returns. One approximate but informative way to decompose the individual-specific ATEs is a linear regression approach. First, we use the model to calculate the individual-specific ATE for the counterfactual in question for each of our 2,453 workers. Then we simply run OLS, regressing the individual-specific ATEs onto pre-immigration observable demographics, and interpret the estimated equation as the best linear predictor of the individual-specific ATEs.

The results from this decomposition procedure for the home job counterfactual are shown in Table 9. In column (1), the dependent variable is the ATEs for the home job counterfactual for the first job in the US, and in column (3) the dependent variable is the ATEs after 10 years in the US. The uninteresting result is that workers with higher home country occupation qualities have larger wage gains from the counterfactual; since we move them to their home occupation in the US, this gain is primarily mechanical. What is more interesting is the result that the other demographics reflecting high-skilled workers (e.g. education, employer-sponsored visa) are all associated with smaller returns to the home jobs counterfactual. This makes sense in the logic of the model: fixing home occupation, workers with higher observed ability likely have bad unobserved traits (whether persistent ability or transitory luck) given they are in the same quality of home occupation as the other workers with worse observables.

A less mechanical result relating the components of pre-immigration skills to the effects of the treatment is given by running the same regression as in columns (1) and (3) of Table 9 but not conditioning on home occupation, since home occupation directly affects counterfactual outcomes. The coefficients in this regression reflect what types of workers are likely to face large barriers in transferring their jobs to the US. These results are shown in columns (2) and (4) of Table 9. We find that college educated workers have large returns to the home jobs counterfactual (an additional $0.28 per hour at entry), while workers with employer-sponsored visas still see negative returns. The finding that immigrants with employer-sponsored visas see small returns to the home jobs counterfactual is consistent with our intuition about what type of immigrants face occupational frictions. If you already have a job with a firm in the US, it makes sense that it would be relatively easy to stay in a relatively similar occupation after moving to the US. On the other hand, highly educated individuals who do not have employer-sponsored visas are the types of immigrants people anecdotally discuss when they think about barriers to transferring human capital: for example, PhDs driving taxis. We indeed find this group has large returns to the home job counterfactual. Similarly, immigrants from high income countries have lower returns to the counterfactual, as do younger immigrants; both of these results seem consistent with immigrants from richer countries and young workers facing fewer frictions.

We repeat the decomposition exercise for the long-run jobs counterfactual and regress the individual-specific ATEs onto demographics, with the results shown in Table 10. In columns (1) and (2), the dependent variables are the ATEs for the first job and after 10 years in the US labor market, respectively. The ATEs in this case are positively related to almost all demographics: that is, we expect that higher-skilled individuals will have higher returns to removing all the occupational barriers they face. In the long run jobs counterfactual, for instance, having some college education is now strongly associated with a high ATE at entry. Interpreting this result in terms of our model estimates, recall that having high education shifts the mean of the offer distribution upwards significantly. This results in faster job quality growth and a higher quality job in year 15, and so the long-run jobs counterfactual has a larger effect for workers with high education. The importance of the higher quality long run job dominates the fact that these high skill individuals see less downgrading from their home occupations at US entry, leading to an overall positive relationship between skills and the counterfactual impacts. Concretely, compare immigrant A, who has a college education, to immigrant B, who does not. Assume both immigrant A and immigrant B worked in apparel sales in their home country. One situation consistent with our model estimates would be that immigrant A's first job in the US is as a cashier, which is a relatively small drop in job quality, and they end up as an upper-level manager in the long run, which is a job with higher quality than their home job. On the other hand, immigrant B initially works as a manual laborer in the US, and eventually moves back into apparel sales again in the long run. The effect of moving these workers to their home job at entry could be larger for immigrant B, while moving them to their long-run job would help immigrant A more.

Overall, both counterfactuals show large variation in the returns to helping immigrants overcome the occupational barriers they face. The results on the importance of different demographics are not identical, however, with the relative importance of individual traits depending on the counterfactual of interest. Our results suggest that policies to reduce occupational entry barriers would have the largest impact on higher skill immigrants.

7 Conclusion

In this paper, we quantified the role of occupational upgrading in the wage growth of immigrants to the US. To do this, we used panel data on the migration histories, labor market histories, and demographics of US immigrants from the New Immigrant Survey. We created a model of labor market search and estimated it on the NIS sample, and we considered counterfactuals which we interpret as an upper bound on the benefits of policies which aim to increase the rate of immigrants' occupational upgrading. In the first counterfactual, we considered the occupational and wage paths of immigrants if they begin their US careers in their home country job. In the second counterfactual, we analyzed the effects of moving immigrants to their model-predicted long-run job directly at US entry.

The overall returns to these policies are modest on average, with the gains focused at the high end of the skill distribution. The home jobs counterfactual raises wages by 18% at entry, and after 10 years the gain from the counterfactual is around 2.5%. In the long-run jobs counterfactual, wages increase by 25% at entry, but only by 2.8% after 10 years in the US. The effects of occupational upgrading depend on pre-immigration characteristics, with higher-skilled immigrants seeing the largest gains from faster occupational mobility. Considering only immigrants who come from the top 10% of highest-paying occupations in their home countries, the home job counterfactual raises wages by 38% at entry, a much larger effect than the overall average.

Our results have implications for both US immigration policy and future research into immigrant assimilation. Policies aiming to help immigrants return to the jobs they held in their home countries would have a significant impact only for high-skilled immigrants who already have the best time in the US labor market. Rather than policies which look to help immigrants find the right jobs, policies specifically focused on increasing the skills of low-skilled immigrants may have better distributional consequences. For future research, our results emphasize the relationship between occupational upgrading and skills: The higher-skilled the immigrant, the higher the estimated role of occupational upgrading in wage growth. Given many data sets used in the immigration literature (e.g., Weiss et al. (2003) and this paper) have samples where high-skilled immigrants are over-represented compared to the whole immigrant population, the effects of potential policies from these papers cannot be uncritically applied to immigrants of different skill levels.

References

Akresh, I. R. (2008). Occupational trajectories of legal US immigrants: Downgrading and recovery. *Population and Development Review*, 34 (3):435–456.

Autor, D. and Dorn, D. (2009). This job is "getting old": Measuring changes in job opportunities using occupational age structure. *American Economic Review: Papers and Proceedings*, 99 (2):45–51.

Berman, E., Lang, K., and Siniver, E. (2003). Language-skill complementarity: returns to immigrant language acquisition. *Labour Economics*, 10:265–290.

Borjas, G. J. (1985). Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics*, 3 (4):463–489.

Borjas, G. J. and Friedberg, R. M. (2009). Recent trends in the earnings of new immigrants to the United States. *Working Paper*.

Chiswick, B. (1978). The effect of Americanization on the earnings of foreign-born men. *Journal of Political Economy*, 86:897–921.

Cohen-Goldner, S. and Eckstein, Z. (2008). Labor mobility of immigrants: Training, experience, language, and opportunities. *International Economic Review*, 49 (3):837–872.

Cox, D., Little, J., and O'Shea, D. (2007). *Ideals, varieties, and algorithms*, volume 3. Springer.

de Matos, A. D. (2012). The careers of immigrants. *Working Paper*.

Duleep, H. O. and Dowhan, D. J. (2002). Insights from longitudinal data on the earnings growth of U.S. foreign-born men. *Demography*, 39(3):485–506.

Eckstein, Z. and Weiss, Y. (2004). On the wage growth of immigrants: Israel, 1990-2000. *Journal of the European Economic Association*, 2(4):665–695.

Evans, M. and Swartz, T. (2000). *Approximating integrals via Monte Carlo and deterministic methods*, volume 20. OUP Oxford.

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

Friedman, M. and Kuznets, S. (1954). Income from independent professional practice. *National Bureau of Economic Research Books*.

Gourieroux, C. and Monfort, A. (1996). *Simulation-based econometric methods*. Oxford University Press.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52(2):271–320.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Heckman, J. J. and Vytlacil, E. J. (2007). Econometric evaluation of social programs, part i: Causal models, structural models and econometric policy evaluation. *Handbook of econometrics*, 6:4779–4874.

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

Imai, S., Stacey, D., and Warman, C. (2018). From engineer to taxi driver? Language proficiency and the occupational skills of immigrants. Working Paper.

Jaso, G., Massey, D. S., Rosenzweig, M. R., and Smith, J. P. (2006). The New Immigrant Survey 2003 Round 1 (NIS-2003-1) Public Release Data.

King, M., Ruggles, S., Alexander, J. T., Flood, S., Genadek, K., Schroeder, M. B., and Vick, B. T. R. (2010). Integrated Public Use Microdata Series - Current Population Survey: Version 3.0.

Kleiner, M. (2000). Occupational licensing. *Journal of Economic Perspectives*, 14 (4):189–202.

LaLonde, R. J. and Topel, R. H. (1992). Assimilation of immigrants in the U.S. labor market. In Borjas, G. J. and Freeman, R. B., editors, *Immigration and the Work Force*. The University of Chicago Press.

Llull, J. (2018). Immigration, wages, and education: A labour market equilibrium structural model. *The Review of Economic Studies*, 85 (3):1852–1896.

Lubotsky, D. (2007). Chutes or ladders? A longitudinal analysis of immigrant earnings. *Journal of Political Economy*, 115(5):820–867.

Ma, J. (2018). High skilled immigration and the market for skilled labor: The role of occupational choice. *Working paper*.

Weiss, Y., Sauer, R. M., and Gotlibovski, M. (2003). Immigration, search, and the loss of skill. *Journal of Labor Economics*, 21(3):557–591.

Tables and Figures

Table 1: Summary Statistics

Variable	Mean
Age	38.16
Percent male	56.75
Years living in the US at end of first job	2.89
Years living in the US at time of survey	8.66
Percent with an employer sponsor	36.40
Percent with more than high school education	65.14
Percent with high English skills	41.01
Percent who entered US on a valid visa	77.01
Sample Size	2,453

Table 2: Determinants of Job Quality in the US

	(1) Initial job	(2) Current job
College	2.219***	0.719*
	(0.444)	(0.427)
English skills	2.053***	1.430***
	(0.368)	(0.361)
Home country GDP	-0.651	0.161
	(0.423)	(0.492)
Quality of home country job, moved when <18	0.220*	0.171
	(0.121)	(0.115)
Quality of home country job, moved when ≥ 18	0.227***	0.0799**
	(0.0352)	(0.0353)
Employer sponsored visa	4.467***	0.449
	(0.366)	(0.386)
Entered US on Valid Visa	1.082**	0.713
	(0.458)	(0.470)
US work experience (years)	-0.00632	0.488***
	(0.110)	(0.141)
US work experience squared	-0.00756	-0.0157**
	(0.00708)	(0.00695)
US work experience x home country GDP	-0.0288	0.00118
	(0.0740)	(0.0392)
Quality of first US job		0.769***
		(0.0293)
Constant	5.711***	-0.740
	(0.826)	(1.050)
Observations	1064	1081
Adjusted R^2	0.531	0.670

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the average wage (calculated using CPS data) in the occupation a person is working in, and this is also used as a measure of quality of the home country job and first US job. "College" equals 1 if a person has more than 12 years of education, and 0 otherwise. "English skills" equals 1 if a person reports high English skills at the time of the survey, and 0 otherwise. Controls for schooling in the US, gender, home experience, and home experience squared are included but not reported. We also include an interaction between home country GDP and home job quality.

Table 3: Model Job Transitions

Prior state	New state: OOLF $\pi_{it}=-1$	Unemployed $\pi_{it}=0$	Stay at job $\pi_{it}=\pi_{it-1}$	New job $\pi_{it}=\pi_{it}^{\text{Offer}}$
OOLF $\pi_{i,t-1}=-1$	$\text{Re-enter}_{it}=0$	$\text{Re-enter}_{it}=1$ and Offer $=0$		$\text{Re-enter}_{it}=1$ and Offer $=1$
Unemployed $\pi_{i,t-1}=0$	$\text{Dropout}_{it}=1$	$\text{Dropout}_{it}=0$ and Offer $=0$		$\text{Dropout}_{it}=0$ and Offer $=1$
Employed $\pi_{i,t-1}>0$	$\text{Dropout}_{it}=1$	$\text{Dropout}_{it}=0$ and Fired $=1$ and Offer $=0$	$\text{Dropout}_{it}=0$ and Fired $=0$ and $\left\{\text{Offer}_{it}=0 \text{ or } \left(\text{Offer}_{it}=1 \text{ and } \pi_{it-1}>\pi_{it}^{\text{Offer}}\right)\right\}$	$\text{Dropout}_{it}=0$ and $\left\{\left(\text{Fired}_{it}=0 \text{ and } \text{Offer}_{it}=1 \text{ and } \pi_{it-1}<\pi_{it}^{\text{Offer}}\right) \text{ or } \left(\text{Fired}_{it}=1 \text{ and } \text{Offer}_{it}=1\right)\right\}$

Table 4: Out-of-the-Labor-Force Parameter Estimates

	(1) Leave labor market	(2) Re-enter labor market
Last job quality	-0.68	
	(0.064)	
Male	-0.29	
	(0.14)	
Year born	0.13	
	(0.0062)	
Kids	-0.077	
	(0.16)	
Non-working type	2.53	
	(0.29)	
Constant	-1.49	0.34
	(0.28)	(0.021)
Probability (non-working type)	0.21	
	(0.016)	

Notes: Standard errors in parentheses. Column (1) gives the relationship between different factors and the probability a person exits the labor market each period. We use the function $\frac{1}{2} + \frac{1}{2}\tanh(x)$ to ensure the probabilities are between 0 and 1. Last job quality is the logged average wage (using CPS data) in the person's previous period occupation. For the first period, we use the home country occupation as the last occupation. "Kids" is a dummy variable that equals 1 if a person has kids and 0 otherwise. Column (2) gives the probability a person who is OOLF re-enters the labor market.

Table 5: English skills transition parameter estimates

College	1.24
	(0.60)
Home country GDP	-0.065
	(0.22)
Period 1 shifter	1.28
	(0.88)
Constant term	-3.57
	(0.55)

Notes: Standard errors in parentheses. We calculate the probability a person transitions from low to high English skills each period. We use the logistic function to ensure the probabilities are between 0 and 1. College equals 1 if a person has more than 12 years of education, and is 0 otherwise. We have the period 1 shifter to allow for a different constant term in the initial period.

Table 6: Log Wage Function Parameter Estimates

Current job quality	0.76
	(0.026)
Home job quality	0.081
	(0.010)
Male	0.041
	(0.020)
Year born	-0.000087
	(0.011)
College	0.082
	(0.027)
English skills	0.21
	(0.023)
Valid visa	0.086
	(0.034)
Home country GDP	0.013
	(0.0080)
Entry year	-0.040
	(0.014)
Years in US	0.086
	(0.0089)
Years in US squared	-0.26
	(0.072)
Non-working type	-0.37
	(0.035)
Constant	0.30
	(0.068)
Standard deviation	0.45
	(0.0074)

Notes: Standard errors in parentheses. This table gives the parameters of the log wage distribution. Current and home job quality is the average wage (calculated using CPS data) in a given occupation. Current job quality is logged, and all other continuous variables are scaled to have mean 0 and variance 1. “College” equals 1 if a person has more than 12 years of education, and 0 otherwise. “English skills” equals 1 if a person has strong English skills, and 0 otherwise.

Table 7: Job Offer rates, Job Loss Rates, and the Job Offer Distribution

	(1) Job offer rates	(2) Job loss rates	(3) Job offer distribution
Home job quality	-0.028		0.12
	(0.022)		(0.0059)
Male	0.032		0.063
	(0.044)		(0.014)
Year born	-0.0077		0.043
	(0.023)		(0.0071)
College	-0.28		0.27
	(0.059)		(0.019)
English skills	0.0032		0.22
	(0.054)		(0.020)
Employer-sponsored visa	0.14		0.13
	(0.048)		(0.015)
Valid visa	-0.42		0.23
	(0.077)		(0.022)
Home country GDP	-0.026		0.017
	(0.020)		(0.0071)
Entry year	0.35		-0.0054
	(0.025)		(0.0073)
Initial period shift			0.18
			(0.19)
Constant term	0.52	0.084	1.98
	(0.083)	(0.0045)	(0.022)
Standard deviation			0.38
			(0.0055)

Notes: Standard errors in parentheses. In column (1), the coefficients give the effect of each characteristic on the probability that a person gets a job offer each period. We use the function $\frac{1}{2} + \frac{1}{2}\tanh(x)$ to ensure the probabilities are between 0 and 1. Column (2) gives the estimated probability that a person loses their job each period. Column (3) shows the mean and standard deviation of the job offer distribution, which we assume is lognormal. All continuous variables are scaled to have mean 0 and variance 1. “College” equals 1 if a person has more than 12 years of education, and 0 otherwise. “English skills” equals 1 if a person has strong English skills, and 0 otherwise. The term “initial period shift” is the change in the constant term for the mean of the distribution.

Table 8: Counterfactual Effects on Wages

Years after US entry	(1) Baseline	(2) Home country job	(3) Long-run job	(4) No occupational mobility	(5) High English skills
0	11.47	13.52	14.35	11.47	15.58
5	18.05	19.67	20.16	16.80	22.69
10	25.30	25.92	26.00	22.18	29.22
15	29.22	29.46	29.41	25.08	32.52

Notes: In each counterfactual, we calculate the average wage for each person over 400 simulations. Column (1) is the baseline. Column (2) places each immigrant in their home country occupation in the first period in the US, and column (3) puts them in their long-run job at entry. Column (4) shows average wages when each person stays in their initial occupation each period. Column (5) shows outcomes if each person has high English skills

Table 9: Determinants of ATEs for the Home Job Counterfactual

	Entry		After 10 years	
	(1)	(2)	(3)	(4)
Home job quality	3.18***	–	1.08***	–
	(0.015)	–	(0.018)	–
Male	-0.42***	-0.07	-0.23***	-0.11**
	(0.026)	(0.115)	(0.031)	(0.049)
Year born	-0.43***	-0.67***	-0.26***	-0.34***
	(0.013)	(0.059)	(0.016)	(0.025)
College	-1.89***	0.28**	-0.65***	0.08
	(0.033)	(0.138)	(0.040)	(0.059)
Employer-sponsored visa	-1.25***	-0.31**	-0.75***	-0.42***
	(0.029)	(0.127)	(0.035)	(0.055)
Valid visa	-1.30***	-0.80***	-0.29***	-0.11
	(0.039)	(0.172)	(0.047)	(0.074)
Home country GDP	-0.12***	-0.19***	-0.01	-0.03
	(0.013)	(0.056)	(0.015)	(0.024)
Entry year	0.02	0.67***	-0.35***	-0.12***
	(0.015)	(0.067)	(0.019)	(0.029)
Constant	4.91***	2.57***	1.69***	0.90***
	(0.035)	(0.147)	(0.042)	(0.063)
R-squared	0.95	0.07	0.65	0.14
N	2453	2453	2453	2453

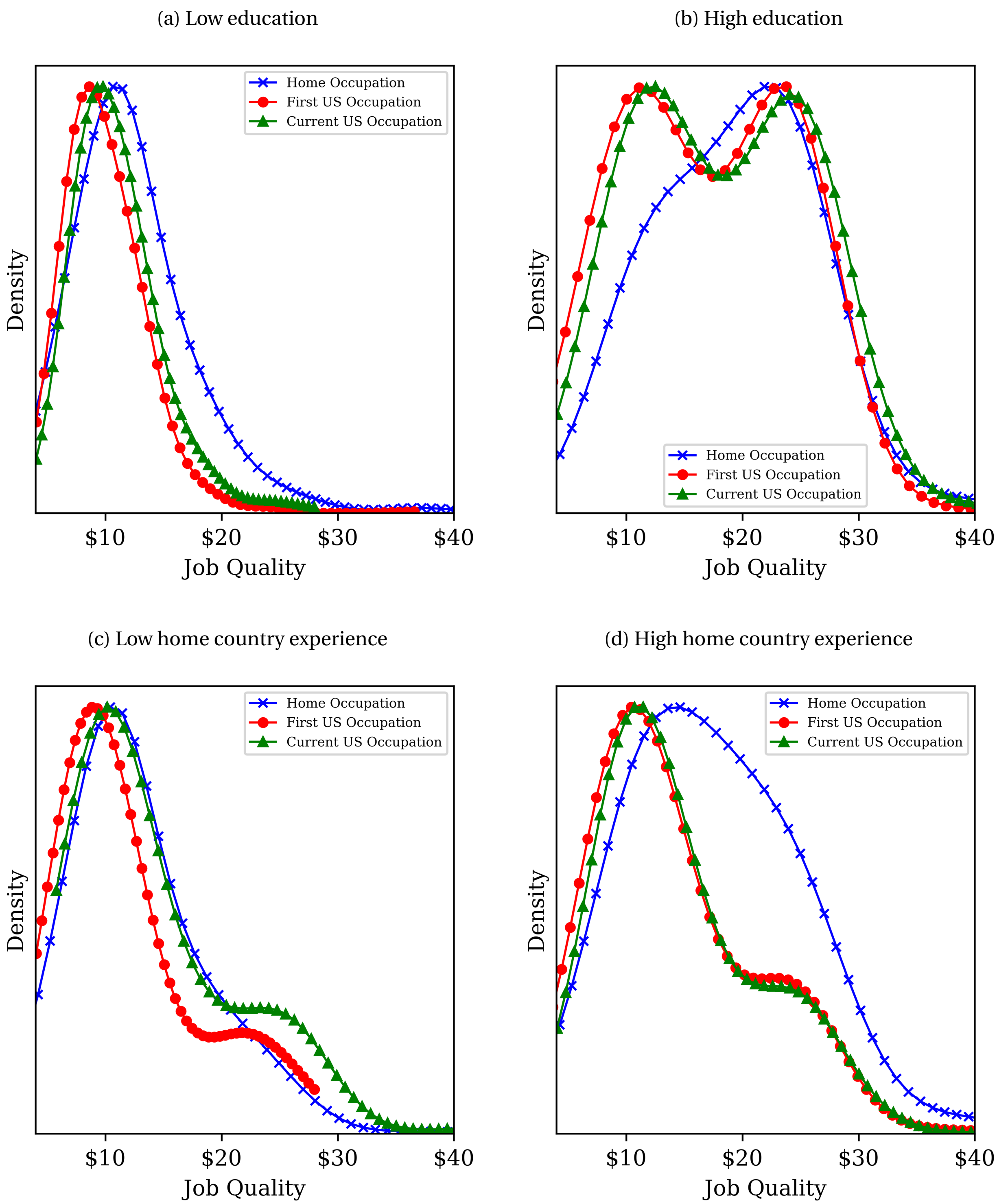
Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the ATE from the home job counterfactual for each person, which is calculated by taking the difference between the average wage in the counterfactual and the baseline for each person. Home job quality is calculated by taking the average wage, in the CPS, of people working in that occupation. "College" equals 1 if a person has more than 12 years of education, and 0 otherwise.

Table 10: Determinants of ATEs for the Long-Run Job Counterfactual

	Entry	After 10 years
	(1)	(2)
Home job quality	0.48***	0.02***
	(0.013)	(0.005)
Male	0.30***	0.02**
	(0.023)	(0.008)
Year born	0.17***	0.02***
	(0.012)	(0.004)
College	0.80***	0.47***
	(0.029)	(0.011)
Employer-sponsored visa	0.69***	-0.05***
	(0.026)	(0.009)
Valid visa	0.07**	0.48***
	(0.035)	(0.013)
Home country GDP	-0.002	0.03***
	(0.011)	(0.004)
Entry year	0.46***	-0.233***
	(0.014)	(0.005)
Constant	1.82***	0.05***
	(0.031)	(0.011)
R-squared	0.83	0.74
N	2453	2453

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the ATE from the long-run job counterfactual for each person, which is calculated by taking the difference between the average wage in the counterfactual and the baseline for each person. Home job quality is calculated by taking the average wage, in the CPS, of people working in that occupation. "College" equals 1 if a person has more than 12 years of education, and 0 otherwise.

Figure 1: Distributions of Home Country and US Occupations



Notes: Each plot shows the distribution of job qualities for the home, initial, and 2003 US occupation. Plot (a) shows for people with no college education, and plot (b) shows it for people with more than 12 years of education. Panel (c) shows people with 5 or fewer years of home experience, and panel (d) shows people with more than 15 years of home experience.

Figure 2: Timing of model

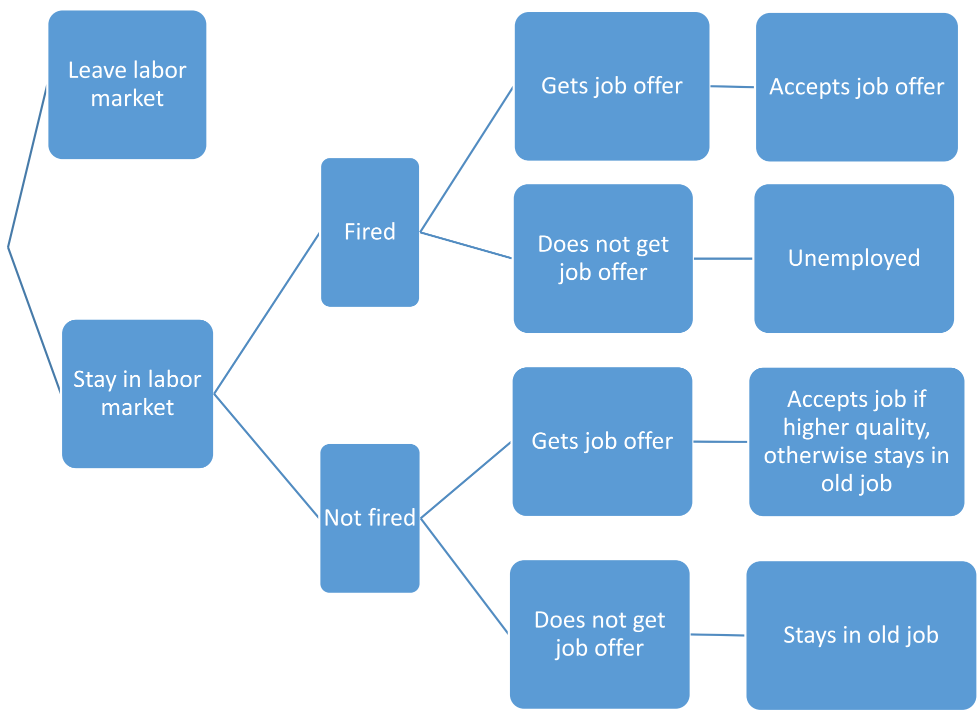
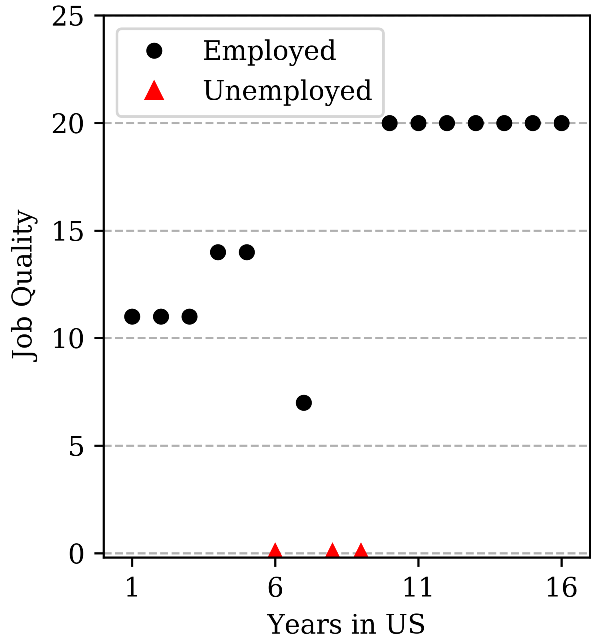
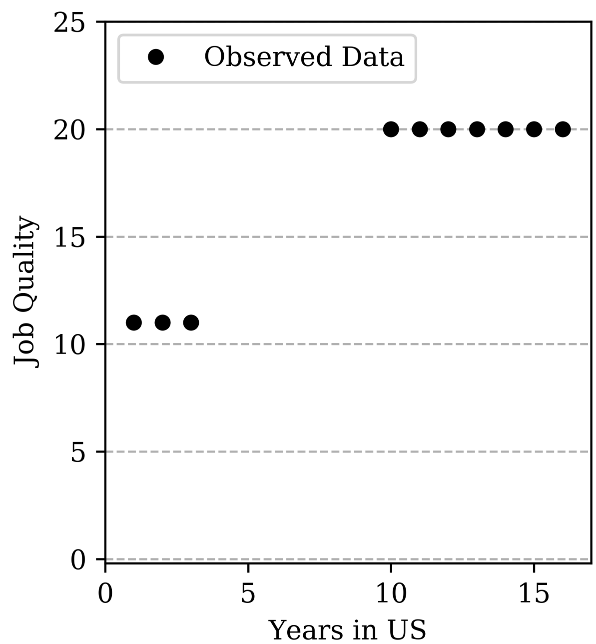


Figure 3: Example Career Paths in Model

(a) Example Career Path

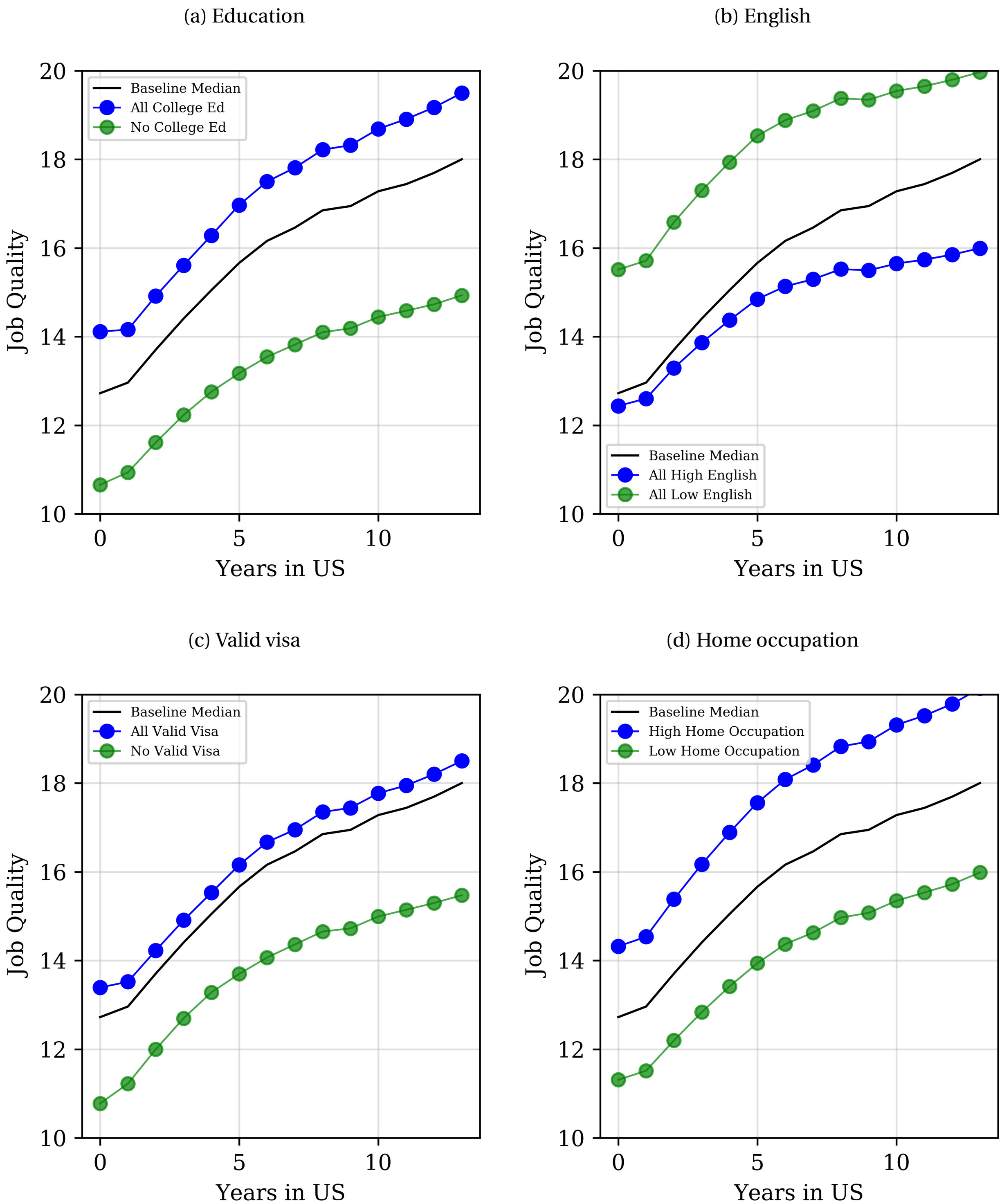


(b) Observed Data from Example Career



Notes: Plot (a) shows a potential career path for a worker. Plot (b) shows what the data would look like in this case given the NIS sampling scheme where we do not observe intermediate jobs.

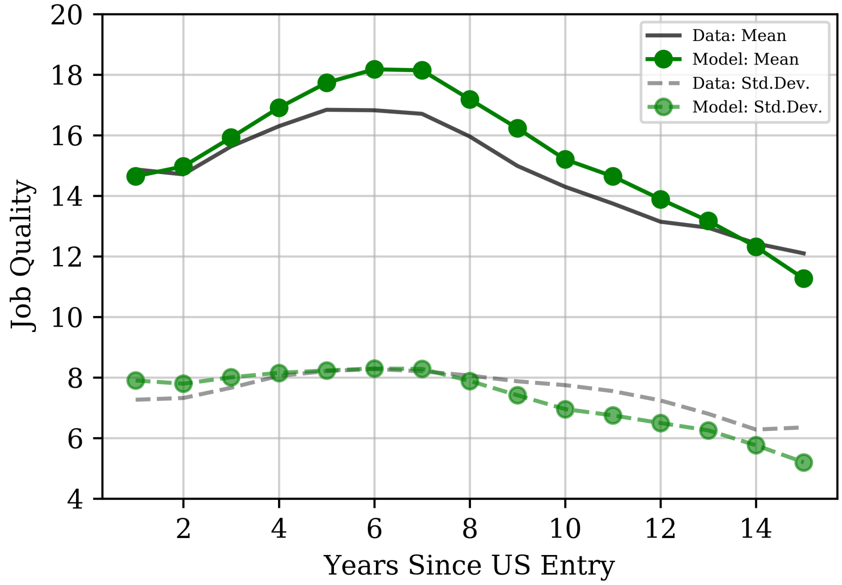
Figure 4: Effects of Demographic Characteristics on Occupational Outcomes

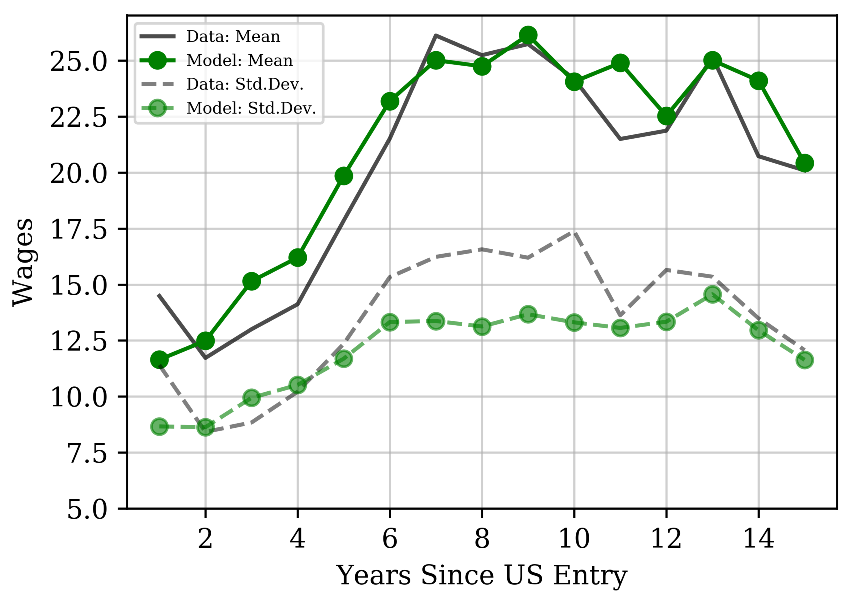


Notes: In each plot, the black line shows the median wage quality at a given number of years of experience in the US. In each of the plots, there are 2 other lines, which are created by changing one characteristic for everyone in the sample. Plot (a) shows everyone with and without some college education, plot (b) shows everyone with and without strong English skills, plot (c) shows everyone with and without a valid visa at entry, and plot (d) shows everyone a 75th percentile and then a 25th percentile quality home job.

Figure 5: Model Fit

(a) Occupations

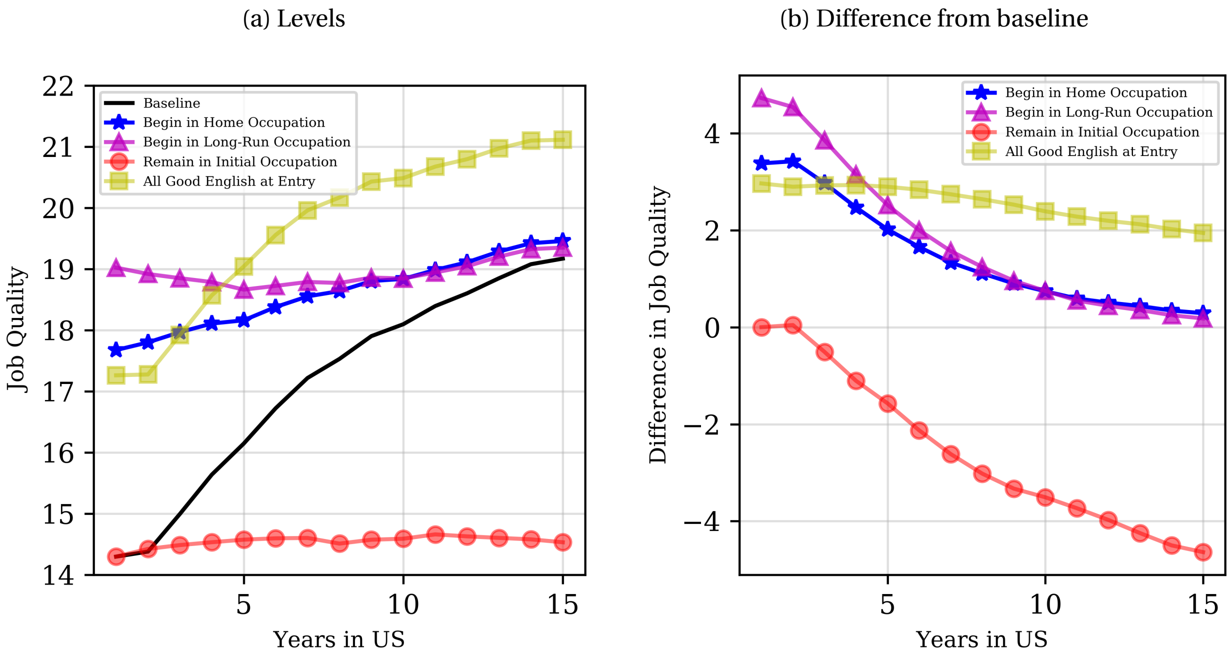




(b) Wages

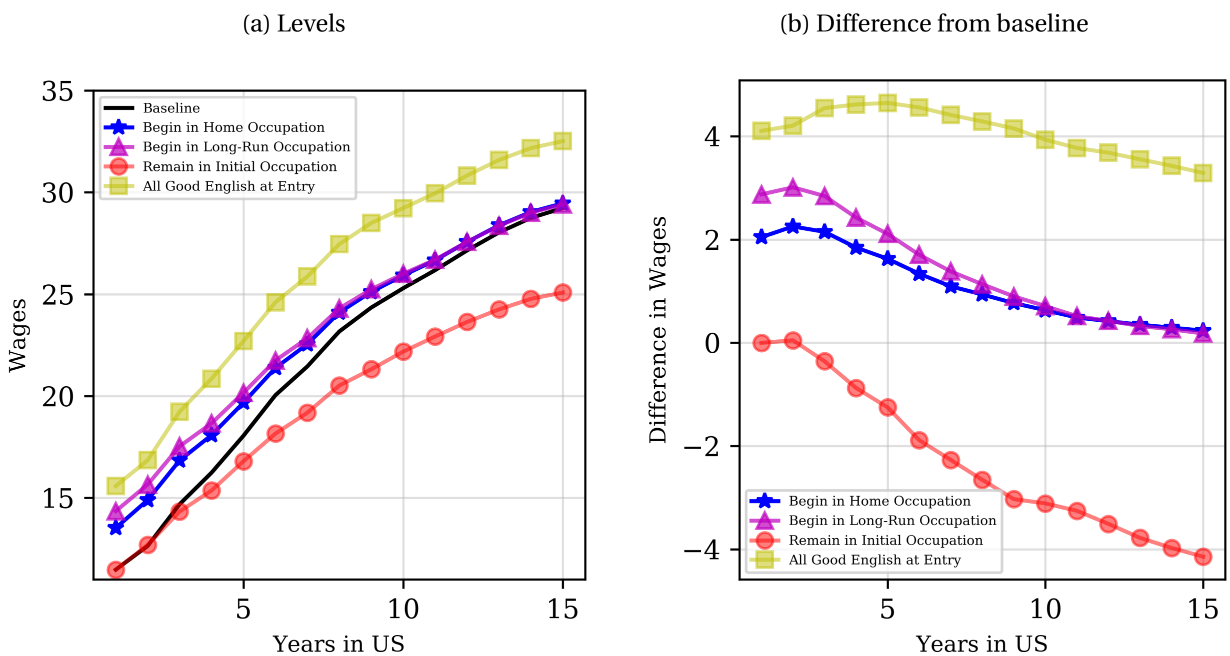
Notes: Plot (a) compares job qualities in the data and the model, taking the average for each year of experience in the US. Plot (b) compares wage outcomes in the data and model.

Figure 6: Counterfactual Effects on Job Quality



Notes: Sub-figure (a) shows the median job quality in each scenario. Sub-figure (b) shows the difference in the median job quality between the baseline and each counterfactual.

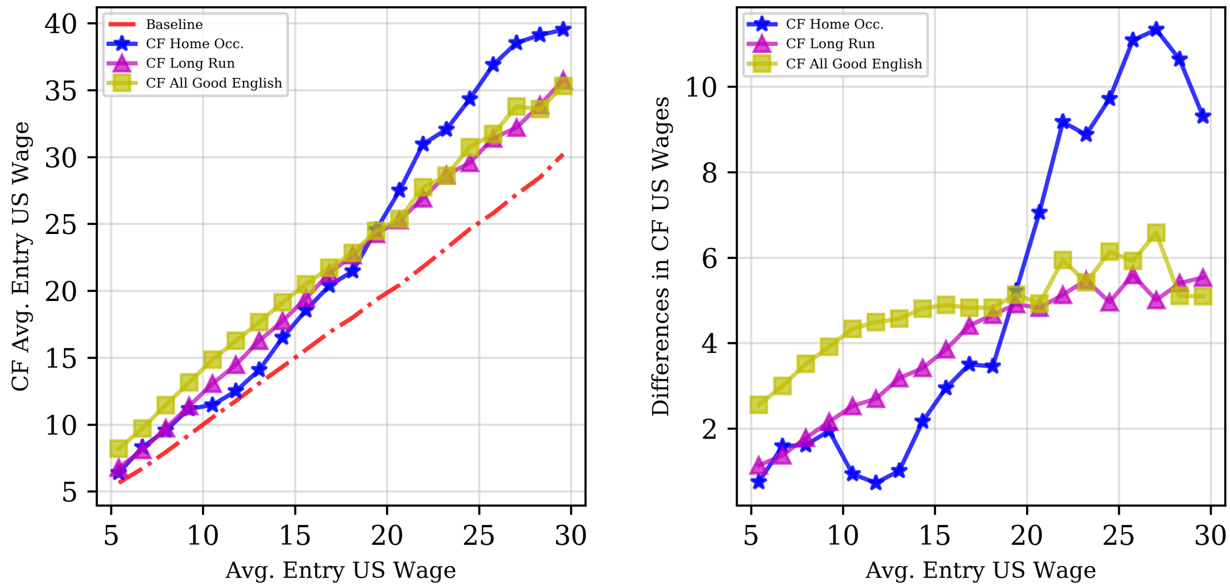
Figure 7: Counterfactual Effects on Wages



Notes: Wages in 2000 dollars. Sub-figure (a) shows median wages in each scenario. Sub-figure (b) shows the difference in median wages between the baseline and each counterfactual.

Figure 8: Returns to Counterfactuals by Predicted Entry Wage at US Entry

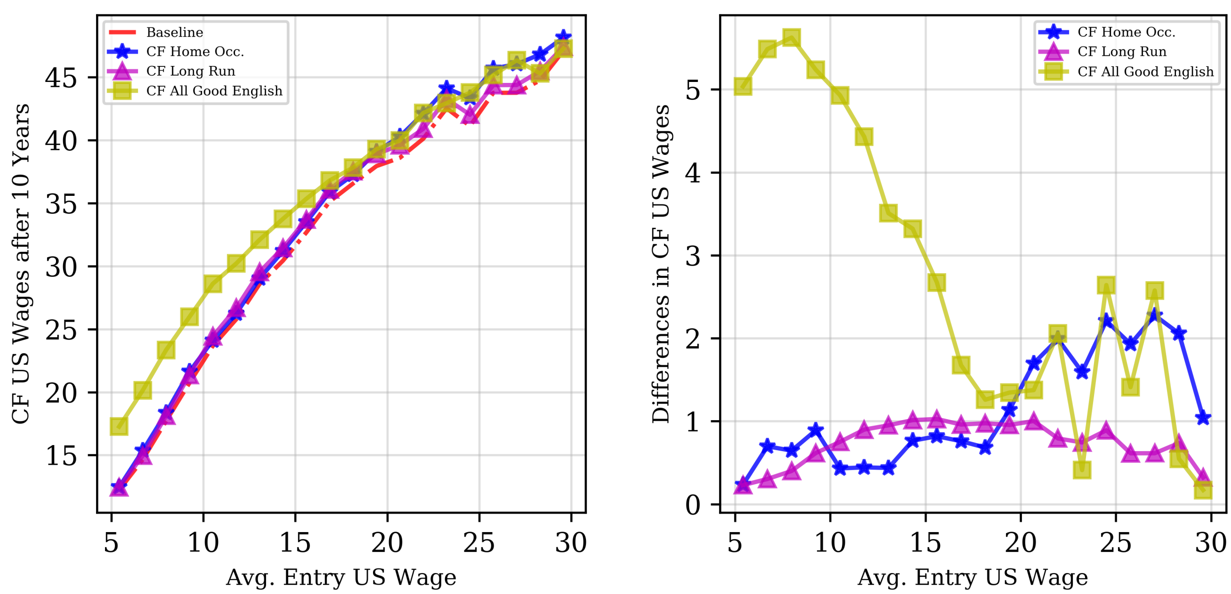
(a) Levels (b) Difference from baseline



Notes: Sub-figure (a) shows the median wages in each scenario. Sub-figure (b) shows the difference between the median wages in the baseline and the counterfactuals. Wages in 2000 US dollars. We calculate predicted US entry wages, as a measure of skill, using our model estimates.

Figure 9: Returns to Counterfactuals by Predicted Entry Wage after 10 Years in US

(a) Levels (b) Difference from baseline



Notes: Sub-figure (a) shows the median wages in each scenario. Sub-figure (b) shows the difference between the median wages in the baseline and the counterfactuals. Wages in 2000 US dollars. We calculate predicted US entry wages, as a measure of skill, using our model estimates.

Appendices for Online Publication:

A Estimation Sample

In this Appendix, we explain how we created our estimation sample. We start with a sample of 13,488 household heads or spouses, since not all relevant questions are asked to the remaining members of the household. We had to drop individuals from the sample for two main reasons: missing data or insufficient information on job transitions. In this section, we explain the specific reasons for each dropped observation.

We first dropped people who moved before 1984, which resulted in the loss of 1,193 observations. We then drop the people who were interviewed outside of the US. This results in a loss of 70 observations. We dropped another 125 people from the sample because they reported a primary job outside of the US, so they were not full participants in the US labor market at the time of the survey. We drop 191 observations where the birth year was missing, as we cannot calculate these people's age. Some people claim their last trip was to someplace other than the US. We drop these people since the NIS sample is supposed to consist of people currently in the US; this results in a loss of 1,076 observations. We drop 4,309 individuals who report having LPR status prior to immigration, and 61 people who report inconsistent years of US entry. There are 926 workers with some job information, but for whom much of the information is missing. We drop 63 observations because of missing information on schooling levels. In total, 504 people immigrated before age 18, meaning that their home country occupation is not informative about their skill levels. We drop these observations. In addition, we drop 4 observations because we do not know these people's home country, and an additional 2,011 observations because we do not know their home country occupation. We drop 31 observations where we do not have information on whether or not the immigrants entered the US on a valid visa. We drop 8 observations where we do not know their home country, and 464 observations where we do now know their English skills.

B Job Transition Kernels

We suppress the individual index i in this section. Recall from Section 3 that a worker's job in period t is characterized by quality π_t. We denote unemployment as its own job with some arbitrary value of π below the lower bound of the job offer distribution; denote unemployed workers as "working" at $\pi_t = 0$. This is consistent with the model, as unemployed workers will always accept an offer, as all actual employment offers have strictly positive π. Similarly, denote the status of out of the labor force with $\pi_t = -1$.

In Section 4.3, we calculated the likelihood function. To do this, we defined the function $f^{\tau}(\pi_t|\pi_{t-1}, X_t)$, which (using language from probability theory) is the transition kernel associated with observing a given pair of job outcomes in periods $t-1$ and t, conditional on worker characteristics X_t and unobserved type τ. These kernels incorporate labor market exit decisions, unemployment shocks, job offer rates, and the job offer distribution. In this section, we explain the construction of f^{τ}, and the main text explains how these are used in forming the full likelihood.

The cumulative distribution function of the job offer distribution is denoted as $\Pi(\cdot)$; for the associated probability density function, for simplicity, we use the notation $\partial\Pi(\pi) \equiv \frac{\partial\Pi(x)}{\partial x}|_{x=\pi}$ to denote the likelihood of a particular occupational draw π.

The conditional likelihoods can be calculated from the occupational transition equation, given in Table 3 in the paper. The model breaks down the likelihood of transitioning between occupations π_{t-1} and π_t, $f^{\tau}(\pi_t|\pi_{t-1}, X_t)$, into nine different cases depending on if the worker moves to a higher productivity firm, lower productivity firm, unemployment, etc. Recall that the job offer distribution is given by $\partial\Pi(\cdot)$, the probability of receiving a job offer is $p(\cdot)$, the probability of being fired is $q(\cdot)$, the probability of exiting the labor market is $O(\cdot)$, and the probability of re-entering the labor market is $RE(\cdot)$.[23] The cases are as follows:

1. A worker is out of the labor force in both periods $t-1$ and t. In this case, the worker did not receive a re-entry shock:

$$f^{\tau}(\pi_t = -1|\pi_{t-1} = -1, X_t) = 1 - RE(X_t) \tag{25}$$

2. A worker re-enters the labor force, but is unemployed. In this case, we know they got a re-entry shock, but did not get a job offer:

$$f^{\tau}(\pi_t = 0|\pi_{t-1} = -1, X_t) = RE(X_t) \times \left(1 - p(X_t)\right) \tag{26}$$

3. A worker re-enters the labor force, and moves to a job with quality π_t:

$$f^{\tau}(\pi_t|\pi_{t-1} = -1, X_t) = RE(X_t) \times p(X_t) \times \partial\Pi(\pi_t|X_t) \tag{27}$$

4. A worker moves from in the labor force to out of the labor force. In this case, we know they received an OOLF shock, and all other shocks are irrelevant. When $\pi_{t-1} \neq -1$,

$$f^{\tau}(\pi_t = -1|\pi_{t-1}, X_t) = O(\pi_{t-1}, X_t, \tau) \tag{28}$$

[23] We defined the out-of-the-labor-force probability to depend on one's previous job. In the first period, we do not have a previous-period job, so we use the home country job, which we denote as π^H.

5. A worker is unemployed both at the end of last period and at the end of the current one. This worker must have not received an offer in period t, yet still chose to remain in the labor market, so the likelihood is

$$f^{\tau}\left(\pi_t = 0|\pi_{t-1} = 0, X_t\right) = \left(1 - O\left(\pi_{t-1}, X_t, \tau\right)\right) \times \left(1 - p(X_t)\right). \tag{29}$$

6. A worker is employed at time $t-1$ but unemployed at t. In this case, the worker chose to remain in the labor market but must have been fired. We also know that they did not get a new job offer in this period, since all job offers are accepted when a person is unemployed. When $\pi_{t-1} > 0$, the likelihood can be written as

$$f^{\tau}\left(\pi_t = 0|\pi_{t-1}, X_t\right) = \left(1 - O\left(\pi_{t-1}, X_t, \tau\right)\right) \times q\left(X_t\right) \times \left(1 - p\left(X_t\right)\right). \tag{30}$$

7. A worker moves to a lower quality job but is not unemployed. This worker must have been fired, otherwise they would not have left their previous higher-productivity job. We also know that they received a job offer at productivity level π_t. When we see $0 < \pi_t < \pi_{t-1}$, the likelihood is

$$f^{\tau}\left(\pi_t|\pi_{t-1}, X_t\right) = \left(1 - O\left(\pi_{t-1}, X_t, \tau\right)\right) \times q(X_t) \times p(X_t) \times \partial\Pi\left(\pi_t|X_t\right). \tag{31}$$

8. A worker stays at the same job as in the previous period. In this case, we know that they did not leave the labor force, and that they were not fired, since the probability of getting a new offer at the same job quality is 0 with a continuous offer distribution. They either did not get a new offer, or they got an offer for a lower-quality job. When $\pi_{t-1} > 0$, the likelihood is

$$f^{\tau}\left(\pi_t = \pi_{t-1}|\pi_{t-1}, X_t\right) = \left(1 - O\left(\pi_{t-1}, X_t, \tau\right)\right) \times \left(1 - q\left(X_t\right)\right) \times \left(\left[1 - p\left(X_t\right)\right] + p\left(X_t\right) \cdot \Pi\left(\pi_{t-1}|X_t\right)\right). \tag{32}$$

9. A worker moves to a higher quality job. This case includes the scenario when a worker moves to a firm from unemployment. If we see a worker move to a higher quality job, we know that they got a job offer, and we know exactly what the offer was. In this case, it does not matter whether or not the worker was fired, since all that is relevant is that they received a higher quality job offer, which they would accept regardless of whether or not they had been fired. In this case, the likelihood of observing job $\pi_t > \pi_{t-1}$ is

$$f^{\tau}\left(\pi_t|\pi_{t-1}, X_t\right) = \left(1 - O\left(\pi_{t-1}, X_t, \tau\right)\right) \times p(X_t) \times \partial\Pi(\pi_t|X_t). \tag{33}$$

C Identification

C.1 Preliminaries

Some of the proofs throughout this section show that identification of model parameters is equivalent to the global uniqueness of a solution to a system of polynomial equations. Once this equivalence has been established, there exist computational tools to (potentially) solve for all solutions symbolically and thus show that the solution set is a singleton. Most of the time, showing by hand that these solutions are unique using basic solving/substituting is infeasible. As the equations are non-linear, a full-rank Jacobian is only a necessary but not sufficient condition for global uniqueness, in contrast with the linear case. The "Solve" function in Mathematica uses Gröbner bases to solve systems of polynomial equations, which (in relatively simple cases) can solve for all global solutions systems of polynomial equations (see Cox et al. (2007)). The Mathematica code deriving the solutions to any systems of equations below is available on request, and throughout we assume that uniqueness claimed by Mathematica is sufficient for true uniqueness.

C.2 English Process

All workers start out with low English skills. We denote $E_t = 0$ as low English skills and $E_t = 1$ as high English skills. Prior to the first period, they transition to high English skills with probability:

$$\Pr(E_0 = 1) = \Lambda\left(\gamma_0 + \gamma_1 + \gamma_2 ed + \gamma_3 gdp\right). \tag{34}$$

In the above equation, $\Lambda(\cdot)$ is the logistic function. In later periods, $t > 0$, the transitions between levels of English skills takes the form of a Markov process:

$$\Pr(E_t | E_{t-1}) \rightarrow \quad \begin{array}{cc|cc} & & \multicolumn{2}{c}{E_t} \\ & & 1 & 0 \\ \hline E_{t-1} & 1 & 1 & 0 \\ & 0 & \Lambda\left(\gamma_0 + \gamma_2 ed + \gamma_3 gdp\right) & 1 - \Lambda\left(\gamma_0 + \gamma_2 ed + \gamma_3 gdp\right) \end{array}$$

Note that the functional forms for the first and all future periods are similar and have the same parameters, with the only difference that we allow for a different constant in the first period. We are assuming that having strong English skills is an absorbing state.

Identification of the γ parameters relating observables to English transitions is complicated by the fact that we have only one observation on English skills per worker, and (by data construction) we do not see English skills in the first period for anyone. We can write the probability of not having high English skills by period t:

$$\Pr(E_t = 0) = \left(1 - \Lambda\left(\gamma_0 + \gamma_1 + \gamma_2 ed + \gamma_3 gdp\right)\right) \cdot \left(1 - \Lambda\left(\gamma_0 + \gamma_2 ed + \gamma_3 gdp\right)\right)^{t-1} \tag{35}$$

Considering only workers with the same set of ed and gdp, we can use constants $e_0 \equiv \gamma_0 + \gamma_1 + \gamma_2 ed + \gamma_3 gdp$ and $e_1 \equiv \gamma_0 + \gamma_2 ed + \gamma_3 gdp$ and rewrite equation (35) as

$$\Pr(E_t = 0) = (1 - e_0)(1 - e_1)^{t-1}. \tag{36}$$

Since the length of the worker's time in the US is exogenous, we can recover the left hand side of equation (36) from sample averages. With at least three periods of data, we can use periods $t = 2, 3$ and create two polynomial equations in two unknowns,

$$\begin{aligned} \Pr(E_2 = 0) &= (1 - e_0)(1 - e_1) \\ \Pr(E_3 = 0) &= (1 - e_0)(1 - e_1)^2 \end{aligned} \tag{37}$$

and Mathematica finds these have a unique solution (see section C.1 above). Since we can identify

$$e_0 \equiv \gamma_0 + \gamma_1 + \gamma_2 ed + \gamma_3 gdp \tag{38}$$

and

$$e_1 \equiv \gamma_0 + \gamma_2 ed + \gamma_3 gdp\,, \tag{39}$$

it is simple to identify $\gamma_1 = e_1 - e_0$ for any ed and gdp pair, and $\gamma_0, \gamma_2, \gamma_3$ can be recovered under linear independence of a constant vector, ed, and gdp as in the standard linear regression framework.

C.3 Out-of-the-Labor-Force Process

In this section, we show identification of the parameters that govern workers' transitions between unemployment, out-of-the-labor-force, and employment. We condition the argument below on a particular set of observable demographics X, and as the latent types are independent of those observables we can repeat the argument for each set of X. The data structure we use to recover these parameters has the worker's year of entry into the US, year of first employment, and then later data about employment and OOLF status. Of particular concern, we simply know the worker is "not working" from period 1 through the period of the first job, but we do not see whether they are unemployed or OOLF in those periods. Denote the "not working" state (for either reason) as -2, OOLF as -1, unemployment as 0, and employment as 1.[24]

The job history data for a particular individual might look like

$$(-2,-2,-2,1,0,0,1,-1...) \ ,$$

and by construction runs of -2's can only come at the beginning of the career. Recall from the model that in each period, workers are first hit with a labor market shock that sends them OOLF (if they are in the labor force) with probability $p_{\text{out},\tau}$ that depends on their unobserved type, $\tau \in \{0,1\}$ with probability ν and $1-\nu$ respectively. Because of this latent type, simply calculating $\Pr(\pi_t = -1|\pi_{t-1} = 1)$ will incorrectly estimate the probability that a given worker will drop out between periods. We only observe the actual values of π for later periods, and the distribution of worker types remaining in the market will not necessarily be given by ν and $1-\nu$. To show that we can recover the $p_{\text{out},\tau}$ for each type as well as the proportion of high types, ν, we use a two part argument. First, we show that we can recover the job offer probabilities consistently from data starting after the (endogenous) first job, as we can combine probabilities of different events that depend on the latent type in a way such that the type cancels out. Second, using the fact that we know the offer rate from unemployment and the fact that all offers from unemployment are accepted, we can use the proportion of workers who begin their career with various strings of -2's to recover the proportion of types and their respective probabilities of dropping out. Our proof intuitively depends on the fact that the hazard rate of leaving the -2 state and entering the labor market may not be constant across time, and the evolution of that hazard rate gives an indication about the proportions and OOLF probabilities of different types.

First, we derive the probability of a job offer p_{offer} from the observed data. For workers who we observe $\pi_{t-1} = 0$, we know that they take a job in the next period if a) they do not drop out, and b) they receive a job offer. The conditional probability of moving into a job is then

$$\Pr(\pi_t = 1|\pi_{t-1} = 0) = p_{\text{offer}}\left(1 - p_{\text{out},\tau}\right). \tag{40}$$

We do not know individual worker types, and we do not know the distribution of types in the labor market remaining at time t . However, we can calculate the conditional probability of a worker instead remaining in unemployment, which requires both the worker to not drop out of the labor force and the worker to not get a job offer:

$$\Pr(\pi_t = 0|\pi_{t-1} = 0) = \left(1 - p_{\text{offer}}\right)\left(1 - p_{\text{out},\tau}\right). \tag{41}$$

[24]For the purposes of this section, the identity of the particular job does not matter.

And now forming

$$\frac{\Pr(\pi_t = 1|\pi_{t-1} = 0)}{\Pr(\pi_t = 0|\pi_{t-1} = 0)} = \frac{p_{\text{offer}}\left(1 - p_{\text{out},\tau}\right)}{\left(1 - p_{\text{offer}}\right)\left(1 - p_{\text{out},\tau}\right)} = \frac{p_{\text{offer}}}{1 - p_{\text{offer}}}, \tag{42}$$

the unobserved type terms cancel out and we can simply solve this equation for p_{offer} given the observed left hand side.

Now consider the proportion of workers who are either OOLF or unemployed in period 1, so they have a career path through the first period of (-2). If their type is τ, the probability of being unemployed is $\left(1 - p_{\text{out},\tau}\right)\left(1 - p_{\text{offer}}\right)$, while the probability of being out of the labor force is simply $p_{\text{out},\tau}$. Writing this with more notation, we have in the first period

$$\Pr(\pi_1 = 0|\tau) = \left(1 - p_{\text{out},\tau}\right)\left(1 - p_{\text{offer}}\right) \tag{43}$$

$$\Pr(\pi_1 = -1|\tau) = p_{\text{out},\tau} \tag{44}$$

In later periods, we have transitions between states from the following pseudo-transition matrix (given type):

$$\Pr(\pi_t|\pi_{t-1},\tau) \rightarrow \quad \begin{array}{cc|cc|} & & \multicolumn{2}{c}{t} \\ & & -1 & 0 \\ \hline t-1 & -1 & 1 - p_{\text{in}} & p_{\text{in}}\left(1 - p_{\text{offer}}\right) \\ & 0 & p_{\text{out},\tau} & \left(1 - p_{\text{out},\tau}\right)\left(1 - p_{\text{offer}}\right) \\ \hline \end{array}$$

This is not a standard Markov matrix as it does not add up to 1, but with all remaining probability you end up employed, which is observed. What we are then looking for is the probability of any path that keeps you consecutively in the -2 "state". The probability of having at k consecutive -2s in the first k periods is given by summing over the probabilities of all paths that move only between -1 and 0:

$$\Pr(-2, k \text{ times}|\tau) = \sum_{\pi \text{ paths} \in \{-1,0\}^k} \Pr(\pi_1|\tau) \prod_{t=2}^{k} \Pr(\pi_t|\pi_{t-1},\tau).$$

Integrating out over types, which we know in the overall population are split ν and $1-\nu$ between 0 and 1, we get

$$\Pr(-2, k \text{ times}) = \nu \left[\sum_{\pi \text{ paths} \in \{-1,0\}^k} \Pr(\pi_1|\tau = 0) \prod_{t=2}^{k} \Pr(\pi_t|\pi_{t-1},\tau = 0) \right]$$

$$+ \quad (1-\nu)\left[\sum_{\pi \text{ paths}\in\{-1,0\}^k} \Pr(\pi_1|\tau=1)\prod_{t=2}^{k}\Pr(\pi_t|\pi_{t-1},\tau=1)\right] \tag{45}$$

where the "path" sum is over all possible paths that do not include working over these k periods, $(0,-1,0,...)$, $(-1,0,-1,-1,...)$ etc.

Since each probability in the sum is a polynomial in p_{in} and $p_{\text{out},\tau}$, the overall probabilities can be written as a polynomial in p_{in} and the two $p_{\text{out},\tau}$ values as well as the population high type probability ν. The set of unknowns are ν, p_{in}, and both $p_{\text{out},\tau}$, so we need at least 4 equations (e.g. $k=1,2,3,4$) to uniquely identify the four unknowns. Solving these equations in Mathematica (see C.1 above) indeed finds that there are only two solutions to this polynomial system, one the true set of values and the other simply rearranging the 0 and 1 type labels, which are obviously economically equivalent.

The specifics of the identification proof above are not particularly intuitive. The intuition behind this identification result is a discrete version of the analysis in Heckman and Singer (1984), where non-constant hazard rates can be used in duration models to identify time-invariant unobserved heterogeneity. One way to see the intuition for how the probabilities of different strings of -2's provide information on OOLF rates is the following example: assume for argument's sake that the true value of $p_{\text{out},\tau}$ were 0 for both types. Since everyone starts in state 0, the proportion if workers in -2 in the first period is simply those who do not get offers, $\left(1-p_{\text{offer}}\right)$. The proportion with 2 consecutive -2's is $\left(1-p_{\text{offer}}\right)^2$, etc. So if in the (infinite) data set we see that $\frac{\Pr(-2\,k\,\text{times})}{\Pr(-2\,k-1\,\text{times})}$ is changing over k, we know that our assumption that $p_{\text{out}}=0$ for both types must be wrong, and the fact that the data can reject some set of p_{out} are suggestive of the result (proved above) that these paths uniquely identify these probabilities.

C.4 Occupational Offer Distribution

Our econometric model of occupational transitions relates individual demographic characteristics X_t to job quality outcomes π_t through Table 3, where we assume that workers are fired with probability $q(X_t)$, receive a new job offer with probability $p(X_t)$, and offers are drawn from the distribution $\Pi(\pi|X_t)$. In this section, we show that we can non-parametrically identify the job transition function. Our sample is too small to use non-parametric estimators, but in estimation we used flexible functional forms and in principle we could use increasingly flexible functional forms as the amount of data increased.

For identification purposes, we assume the offer distribution $\Pi(\cdot)$ has bounded support with known lower bound $\underline{\pi}$ and known upper bound $\bar{\pi}$. If we did not have an upper bound on the job offer distribution, some of the arguments would have to be modified into formal identification-at-infinity arguments. As it is, our argument below uses workers who are at the worst possible job and best possible job.

First we show the job offer and firing rates are identified. For this section, we suppress the observable demographics X_t; we can repeat the argument for any given X_t. We write the job firing rate as q and the job offer rate as p. We also condition on the worker not dropping out of the labor force between two periods, which is observed from the data; denote this event $d = 0$.

Consider a worker who is at job π_0 in the initial period. We will observe them at the same job next period only if they do not lose their job, and if they received an offer, it was lower than π_0. The probability of this event is

$$\Pr\left(\pi_1 = \pi_0 | d = 0\right) = \left(1 - q\right)\left(\left(1 - p\right) + p\Pi\left(\pi_0 | d = 0\right)\right). \tag{46}$$

Now consider workers who have $\pi_0 = \bar{\pi}$, that is, the workers with the best jobs. The probability of them getting an offer lower than $\bar{\pi}$ is 1, so $\Pi\left(\bar{\pi}\right) = 1$ and this reduces to

$$\Pr\left(\pi_1 = \pi_0 | \pi_0 = \bar{\pi}, d = 0\right) = 1 - q. \tag{47}$$

This directly identifies q, the probability of job loss. Intuitively, we have data about how long it takes a worker to switch jobs, as well as a ranking of jobs. If we look at individuals only in the highest type of jobs, the only model mechanism for leaving this job for a worse job is firing, since they will never get a better offer to make a job-to-job move.

Once we have identified the probability of job loss q, we can use a similar argument to recover the probability of a job offer p. Consider workers at $\pi_0 = \underline{\pi}$, that is, the workers with the worst jobs. Since we know the probability of an offer below $\underline{\pi}$ is 0, $\Pi\left(\underline{\pi}\right) = 0$, and the probability of staying at their job is

$$\Pr\left(\pi_1 = \pi_0 | \pi_0 = \underline{\pi}, d = 0\right) = \left(1 - q\right)\left(1 - p\right). \tag{48}$$

Since we already know q, this probability gives us p. As above, if we look at individuals only in the worst type of jobs who did not lose their jobs, the only model mechanism for moving up is receiving an outside offer. We know that all upwards moves come with an offer, and that every time individuals stay in their job there was not an offer. We are able to identify the relative frequency of job offers versus the offer distribution, unlike in many versions of search models, because we assume we have data on rankings of jobs, so we can ex ante identify workers who are either unlikely or likely to receive better offers.

Lastly, once we know p and q, solving for $\Pi(\pi_0)$ in equation (46) gives

$$\Pi(\pi_0|d=0) = \frac{\Pr(\pi_1 = \pi_0|d=0)}{p(1-q)} - \frac{1-p}{p}. \tag{49}$$

The right hand side is simply data ($\Pr(\pi_1 = \pi_0|d=0)$) and known parameters. Given an original job, we can now determine the correct proportion of workers who had either been fired or not received an offer. Then we can use the proportion of remaining workers who did not move to identify the probability of getting an offer below that job. As long as workers in an infinite sample could be observed at every possible job in some period (which will be true given the model setup), the full distribution of Π can be traced by varying π_0 in equation (49).

For this identification argument, we only required a limited part of the data: the quality of the first job and one observation on whether the individual remained in that job or not. The duration of the first job, the quality of the final job, and the duration of the final job are all not strictly required for identification but increase the power of our estimators. Since the actual cross-section of workers is relatively small, the additional power of knowing the first and final job durations helps significantly for getting a reasonable amount of precision.

C.5 Wage Equation

Identification of the wage function is a particular application of the general strategy for selection corrections, e.g., Heckman (1979). A standard wage regression would confound wage changes across observables with different labor force dropout rates across those observable groups, and so a correction term could be generated to control for this differential dropout rate.

First, assume that we know the parameters of the occupational transition processes from Appendices C.3 and C.4 above, so we can calculate the probability of any set of occupational transitions from the model, both conditional and unconditional on latent type. Recall that $\pi = -1$ indicates being out of the labor market, and $O(\pi_{i,t-1}, X_{it}, z_i, \tau)$ is the probability of exiting the labor market in a given period. In this function, X_{it} are our included exogenous variables, z_i is our excluded (from the wage equation) exogenous variable (in particular, a dummy for having children), and τ is the latent type, equal to 0 or 1 with probabilities ν and $1-\nu$, respectively.

Denote the entire occupational history through period $t-1$ as $\pi_{(H)}$. From the data, we can calculate the probability that a worker with given observables and occupational history is still in the labor force at time t. We denote this as $\Psi\left(X_{it}, z_i, \pi_{(H)}\right)$. The equivalent object that conditions on the worker's latent type, on the other hand, is not known directly from the data but can be calculated from the model; denote it as $\Gamma_t\left(X_{it}, z_i, \pi_{(H)}, \tau\right)$.

Consider the wage equation:

$$\log(W_{it}) \equiv w_{it} = \beta_0 + X'_{it}\beta + \gamma \cdot \pi_{it} + \beta_\tau \cdot \mathbf{1}\{\tau = 1\} + \delta_1 t + \delta_2 t^2 + e_{it}. \tag{50}$$

Even with the assumption that e_{it} is independent of everything else in the model, the fact that we don't observe τ can still lead to a selection problem. For workers who are currently employed, the conditional mean of observed wages for workers of characteristics X_{it} and z_i in current occupation π_{it} with occupational history $\pi_{(H)}$ is

$$E_t\left[w_{it}|X_{it}, z_i, \pi_{it}, \pi_{(H)}\right] = \beta_0 + X'_{it}\beta + \gamma \cdot \pi_{it} + \beta_\tau \cdot E\left[\mathbf{1}\{\tau = 1\}|X_{it}, z_i, \pi_{it}, \pi_{(H)}\right] + \delta_1 t + \delta_2 t^2. \tag{51}$$

Given X_{it} and $\pi_{(H)}$, today's job quality π_{it} is independent of τ since it is drawn after the last dropout shock. But we may have $E_t\left[\mathbf{1}\{\tau = 1\}|X_{it}, z_i, \pi_{(H)}, \pi_{it}\right] = \Pr_t\left(\tau = 1|X_{it}, z_i, \pi_{(H)}, \pi_{it}\right)$ as a non-constant function of both t (since the proportion of types will change over time if they have differential dropout rates) or X_{it}, which would bias OLS-style estimation of the wage parameters. Using Bayes' rule and denoting the likelihood of observing π_{it} by g_t:

$$\begin{aligned}\Pr_t\left(\tau = 1|X_{it}, z_i, \pi_{(H)}, \pi_{it}\right) &= \frac{g_t\left(\pi_{it}|\tau = 1, X_{it}, z_i, \pi_{(H)}\right)\Pr(\tau = 1)}{g_t\left(\pi_{it}|X_{it}, z_i, \pi_{(H)}\right)} \\ &\equiv H_t\left(X_i, \pi_{it}, \pi_{(H)}, z_i\right). \end{aligned} \tag{52}$$

Since $g_t\left(\pi_{it}|\tau = 1, X_{it}, z_i, \pi_{(H)}\right)$ is known from the model and occupational transition parameters, $g_t\left(\pi_{it}|X_{it}, z_i, \pi_{(H)}\right)$ is simply data, and $\Pr(\tau = 1) = \nu$ is also known from Appendix C.3 above, we could calculate H and directly include it as a regressor in an augmented wage equation:

$$w^\star\left(X_{it}, \pi_{it}, z_i, \pi_{(H)}\right) \equiv \beta_0 + X'_{it}\beta + \gamma \cdot \pi_{it} + \beta_\tau \cdot H_t\left(X_{it}, \pi_{it}, \pi_{(H)}, z_i\right) + \delta_1 t + \delta_2 t^2 + e_{it}. \tag{53}$$

Now, the only unobservable is e_{it}, and the β are all identified. Given the non-linearity of H, the model is identified even without the exclusion restriction we put on z_i of not directly affecting wages. With z_i, the model would be identified even if H were linear in X_{it} and $\left(t, t^2\right)$ over the whole support of X_{it}. We could identify β_3 directly by taking the two different values of z_i when calculating H conditioning on some X_{it} and $\pi_{(H)}$, and then comparing the difference in the two induced values of H to the induced average change in $w^\star$.

Actually implementing this identification argument using the analogue principle in a two-step Heckman (1979)-type estimator is impractical given the size of the data, since H has the full occupational history as an argument. While it is possible to integrate out over the history, as a practical matter, this creates additional complications, and we simply estimate the wage function within the MLE estimator of the full model.