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

Rebecca Lessem* and Carl Sanders†

August 28, 2018

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, the Society for Economics Dynamics, the Society of Labor Economics, and the North American Summer Meeting of the Econometric Society 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 often take jobs below their true skill level 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 level 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 relatively simple training to navigate the occupational barriers in the US labor market, the shortage of existing programs to help them may be a missed opportunity. However, expanding these programs has costs, and while the 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 level. 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 existing 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 also 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 large effects for the most highly skilled immigrants, but almost no returns for the average immigrant, suggesting that policies aiming to reduce barriers to occupational entry have significant distributional effects.

The structural model we use to quantify the gains from a reduction in occupational barriers considers immigrants making occupational choices over their careers in the US. The wages they face are a function of observable skills, labor market experience, and their current occupation. Every period, workers either remain at their previous occupation, receive a shock into the unemployment pool, or get an outside offer from another occupation drawn from a job offer

[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.

distribution that depends on their skills. Given the jobs available to them, workers choose their career path of jobs to maximize expected wages. We parameterize 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 then 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. This counterfactual aims to evaluate the effect of eliminating occupational downgrading at entry. Moving immigrants to their home country job raises average wages by 5.3% at entry, and after 9 years the gain from the counterfactual is less than 1%, 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 steady-state job. In this scenario, wages increase by a substantial 25.7% on average at entry, but the wage gains decrease quickly over time, to only 4.3% after 9 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. Overall, our results suggest that policies aiming to increase immigrants' occupational upgrading would lead to small average earnings benefits for the immigrants in our sample.

The effects of the counterfactual are small for the average immigrant, but this small average effect masks significant heterogeneity in the returns across different 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 33% at entry and by 6% after 9 years, 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 these 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 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 to this strand of literature is quantifying the importance of occupational upgrading for the wage growth of immigrants in the US labor market. There is a small group of papers that study the role of occupational upgrading for immigrants, but they typically focus on documenting the existence of occupational upgrading rather than evaluating how useful policies to speed up 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 removing 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, occupational barriers 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.

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

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 November 2003, becoming (colloquially) "green card" holders.[4] The green card recipients, along with their spouses, were primarily interviewed at some point during the May-November 6-month period at the location where the LPR documentation was sent. The immigrants responded to questions about demographics, labor market outcomes, and their migration history.

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. All of the information is self-reported. The labor market questions include wage, occupation, and firm tenure, and were asked about a person's current (year 2003) job, but also retrospectively about the final job they worked in their country of origin and their first job in the US. We use the data to construct a panel of wages and occupations for immigrants in the US. For both the immigrant's first and current job in the US, we have the wage at the job, the main occupation of the job, and job tenure in years. This data structure always misses 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 job. In the estimation of the model, we will deal with this missing data issue. We treat other forms of missing data (e.g., no wage reported for some jobs) as exogenous.

One caveat on our data is that the demographic questions are only asked about the immigrant's household in 2003. Language acquisition after immigration likely plays an important role in wage growth, but we cannot see the level of an immigrant's English skills at US entry. We treat English skills as constant, which may bias the role of English skills in our analysis if people's English skills change over time. The role of language acquisition in immigration has been examined in the contexts of other data sets by other studies, such as Cohen-Goldner and Eckstein (2008) and Berman et al. (2003).

To create our sample, we restrict it to 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 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 occupation, which we call the *job quality* in the remainder of the paper. In our interpretation, a workers who moves to a higher-quality occupation moves up the occupational ladder.

We also characterize jobs according to the license requirements of each job. Licensing requirements can act as a barrier to entry into occupations even for domestic workers, as in Friedman and Kuznets (1954). Since licenses (e.g., medical licensing) effectively never transfer across countries, immigrants in jobs that are license-heavy may face additional frictions. Moreover, 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 on the number of licenses required for each occupation in the immigrant's state of residence and use these counts as control variables.[6]

2.2 Descriptive Statistics

Table 1 shows sample summary statistics from the NIS. The average immigrant is around 37 years old, and the sample is about 55% male. Since we have job information for a person's first and current job in the US, we show US experience at each of these points. The average immigrant had been in the US for 2.07 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 years. About 65% of the sample has had some schooling beyond high school, and around 40% of the sample reports high English skills. The NIS data also report whether or not an immigrant entered the US for the first time on a valid visa. About 78% of the sample entered using a valid visa, meaning that about 22% of our sample entered as an undocumented immigrant.[7]

[5] An alternative but similar method is to calculate the percentile of the 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] These data were collected by scraping the O*NET Online website and are available on request.

[7] 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.

Because the NIS sample only includes green card recipients, we know that this group eventually received legal visas before receiving LPR status, but we do not know when they received legal visa status. We also know the type of visa that each immigrant received, with 34% of the sample moving to the US on a visa sponsored by an employer. Visa type is a control rarely available in standard Census-based data sets, and immigrants who were able to obtain 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 occupational 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. Given our results from the model that occupational upgrading plays a minor role in the wage growth of low-skilled immigrants, the upwards bias in our estimates does not present a problem in interpreting our estimates as an upper bound.

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.

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 US 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 how long a person has been in the US. We see that recent immigrants tend to be in low-quality jobs at the time of the survey. Immigrants who have been in the US for longer also have a more similar distribution of job qualities to their home jobs than those who just entered, consistent with movement up the occupational ladder.

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 are those who enter the country on valid visas. We control for a worker's home country occupational 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 while 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. Taking a (non-existent) "low-skill" immigrant who was in the lowest-quality occupation in their home country with no English skills, no education post-high school, and the rest of the demographics at the averages (for continuous variables) or modes (for discrete variables), the regression predicts that immigrant would begin in the occupation located in the 14th 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, and the average/modal other demographics, the regression predicts that immigrants would end up in the 99th percentile job qualities of immigrants in the NIS. The point estimate for home job shows a strong relationship between home country jobs and US jobs even conditional on the other demographics. Taking this “high-skill” immigrant and putting them in the lowest-quality job at home, that immigrant’s US initial occupation would now be at the 45th percentile of the distribution instead of the 99th.

The regression of the quality of the current job onto demographics and the initial US job (shown in the second column of Table 2) follows similar trends. Job quality increases with 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 also suggests that the job growth rates of higher-skilled immigrants are faster than those of lower-skilled immigrants, since even conditioning on initial job and time in the US, many of the demographics still have significant effects.

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 of the endogeneity of job choices, the sampling scheme introduces additional problems. Since the data only report two jobs, workers who move jobs more often have more missing years of observations, while we know the entire career path of a worker who never changes jobs. Interpreting the regressions becomes more difficult with this endogenous missing data. To deal with issues of selection into jobs and the labor market 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 works in the US labor market for periods $t = 0, 1, 2, \dots T_i$, where the terminal period T_i is exogenously given. Agents are endowed with exogenous time-invariant observable characteristics X_i and a 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 $\pi > 0$.

3.2 Model Timing and Job Choices

The model timing works as follows. At the end of period t, an immigrant can be (1) employed in a job of quality $\pi_{it} > 0$, (2) unemployed, which we denote by $\pi_{it} = 0$, or (3) out of the labor force (OOLF), which we denote by $\pi_{it} = -1$. For exposition, we first consider workers who are in the labor force. Job transitions occur as follows between periods t and $t+1$: First, with probability $q(X_i)$, the worker receives a shock that sends them into the unemployment pool. Next, with probability $p(X_i)$,[8] the worker receives a new occupational offer, with the quality of the new offer π' drawn from the continuous offer distribution function $\Pi(\cdot|X_i)$. Finally, the worker chooses either their previous job (if they were not fired), unemployment, or the new occupational offer. The worker begins period $t+1$ with that choice as the state, and earns the associated wages during that period. Note that the job offer distribution does not depend on previous occupation. However, a person with a higher-quality initial job will be less likely to accept a new job offer, since it it less likely to be better than their current job. The endogenous distribution of accepted offers will then be a function of the previous job. Figure 2 shows the job transition timing for workers who were in the labor market at the start of the period. In the first period, each worker takes the offered job π_{i0}, drawn from an offer distribution $\Pi_0(\cdot|X_i)$, potentially different than the time $t > 0$ offer distribution Π above.

Before each period (including the first), each worker draws a shock that sends them out of the labor force with probability $O(\pi_{i,t-1}, X_i, \tau)$. This probability depends on a worker's previous job, characteristics, and unobserved type. We assume that being out of the labor force is an absorbing state, so the worker remains there until T_i.[9]

3.3 Wages

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

$$\log(W_{it}) = w(X_i, \pi_{it}, \tau) + h(t) + \varepsilon_{it}, \tag{1}$$

where w is a parameterized function of worker observables and unobservables, $h(t)$ is a parameterized time trend, and with the restrictions that a) ε_{it} is white noise statistically independent

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

[9]We make this assumption due to the structure of the data sampling. We only observe whether a person is out of the labor force at the time of the survey, as well as information on their first and current job in the US. We do not observe transitions in and out of the labor force, and cannot identify the probability of returning to the labor market.

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}, while assumption c) makes all workers have the same ranking of occupations.

Our setup controls for differential labor market dropout patterns across immigrants causing selection biases. Selection bias in the estimation of wage parameters could occur because workers of different unobserved types τ earn different wages, but also face different probabilities of dropping out of the labor force. For example, consider comparing the average wages of workers with identical observables at periods 1 versus 10. If the one type of workers has a higher per-period probability of dropping out of the labor market, along with lower wage levels, the average observed wages for employed immigrants in period 10 would tend to be higher than those of observably equivalent workers in period 1. This apparent time trend would be because more lower-wage workers dropped out of the labor market than higher-wage workers. Allowing for unobserved heterogeneity that affects both wages and labor market exit decisions prevents estimation of the wage equation without considering the other choices in the labor market, so we estimate the wage equation jointly with the rest of the model to ensure identification of the wage parameters.

3.4 Decision Problem

Workers make choices to maximize their lifetime expected discounted (with discount factor β) wages, leading to a value function

$$V_0(X_i, \tau) = \max_{A^\star} E_0 \left[\sum_{t=1}^{T} \beta^t W_t \left(X_i, \pi^\star_{it}, \tau \right) \right], \tag{2}$$

where $A^\star$ is the accept/reject policy function for the worker given X_i, 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. The usefulness of this assumption is that the policy function "Accept all offers above the current job" allows us to write the transition process of the observables in a closed form.

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_{it}^{Offer}, and let Offer_{it}, Fired_{it}, and Dropout_{it} be the realizations of the outside offer, firing shocks, and labor force dropout shocks, respectively. We denote

$$
\begin{aligned}
\Pi_0(z|X_i) &= \Pr\left(\pi_{i0} \leq z|X_i\right) && (3)\\
\Pi(z|X_i) &= \Pr\left(\pi_{it}^{\text{Offer}} \leq z|X_i\right) && (4)\\
p(X_i) &= \Pr\left(\text{Offer}_{it} = 1|X_i\right) && (5)\\
q(X_i) &= \Pr\left(\text{Fired}_{it} = 1|X_i\right) && (6)\\
O(\pi_{i,t-1}, X_i, \tau) &= \Pr\left(\text{Dropout}_{it} = 1|\pi_{i,t-1}, X_i, \tau\right) && (7)
\end{aligned}
$$

Letting $\Delta_{it} \equiv \left\{\text{Dropout}_{it}, \text{Fired}_{it}, \text{Offer}_{it}, \pi_{it}^{\text{Offer}}\right\}$, the observed job transition function can be written as a Markov process:

$$
\pi_{i,t+1}\left(X_i, \pi_{it}, \tau|\Delta_{it}\right) = \begin{cases}
-1 & \text{if Dropout}_{it} = 1 \text{ or } \pi_{it} = -1\\
0 & \text{if } \pi_{it} = 0,\ \text{Offer}_{it} = 0\\
0 & \text{if } \pi_{it} > 0,\ \text{Fired}_{it} = 1,\ \text{Offer}_{it} = 0\\
\pi_{it}^{\text{Offer}} & \text{if } \pi_{it} = 0,\ \text{Offer}_{it} = 1\\
\pi_{it}^{\text{Offer}} & \text{if } \pi_{it} > 0,\ \text{Fired}_{it} = 1,\ \text{Offer}_{it} = 1\\
\pi_{it} & \text{if } \pi_{it} > 0,\ \text{Fired}_{it} = 0,\ \text{Offer}_{it} = 0\\
\max\left\{\pi_{it}^{\text{Offer}}, \pi_{it}\right\} & \text{if } \pi_{it} > 0,\ \text{Fired}_{it} = 0,\ \text{Offer}_{it} = 1
\end{cases} \quad (8)
$$

In the cases where the worker was fired or previously unemployed, the worker accepts any offer, but will only accept better offers for job-to-job moves. Downwards job moves are possible, but only occur when a worker is fired and then immediately gets an offer; to us this will look like a downwards job-to-job move.

The econometrician is assumed to observe realized occupational qualities without error and wages with independent additive measurement error. We do not assume data on the firing shock, offer shocks, or offered jobs, but only see accepted offers. In our data we also do not observe full labor market histories, but 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.

A sample path generated by the model (for a worker who stays in the labor market) is shown in Figure 3 (a). Model-generated paths can match observed sample paths in terms of workers spending multiple periods in the same job, as well as both upwards and downwards occupational transitions, movements into and out of unemployment, and labor force dropouts. 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).

While each individual will have "jumpy" occupational paths, the average changes in occupational quality over time in the US for any given skill level and initial occupation are monotonic. Figure 4 shows smoothed versions of sample paths averaged over many workers and many simulations of each worker. Different lines correspond to different initial draws of job in the US (with a distribution given by the pdf on the left side of the figure). In the example, the "long run" occupation – the occupation where the expected next-period occupation is the same as the current one – for the immigrant's skills is a job with an average hourly wage of $20. Immigrants who receive low initial job offers start off significantly lower on the occupational ladder than those who receive high offers, but over time in the US they tend towards the same steady-state job distribution no matter their initial offer.

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 this prediction. 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 no matter the actual shape of the offer distribution. 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 4.1 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. For example, 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 our data, we only observe each person's current and first US job, so we do not observe job choices in every period. With 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 Appendix B and C, respectively.

4.1 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_i. 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 worker, $\tau \in \{0,1\}$, drawn with probability υ_0 and $1-\upsilon_0$, respectively. We parameterize the offer probability $p(X_i)$, the firing probability $q(X_i)$, and the OOLF shock $O(\pi_{it}, X_i, z_i, \tau)$ with the single-index functional forms as

$$p(X_i) \equiv \Phi\left(\alpha_0 + X_i'\alpha\right) \tag{9}$$

$$q(X_i) \equiv q \tag{10}$$

$$O(\pi_{it}, X_i, z_i, \tau) \equiv \Phi\left(\gamma_0 + X_i'\gamma + \gamma_\pi \pi_{it} + \gamma_z z_i + \gamma_\tau \cdot \mathbf{1}\{\tau = 1\}\right), \tag{11}$$

with $\Phi(x) \equiv \frac{1}{2} + \frac{1}{2}\tanh(x)$, a function bounded between 0 and 1 for all x. Note that in equation (10) 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 (11) includes 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 (11), 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.

We parameterize the wage equation as

$$\log(W_{it}) = \beta_0 + X_i'\beta + \gamma \cdot \pi_{it} + \beta_\tau \cdot \mathbf{1}\{\tau = 1\} + \delta_1 t + \delta_2 t^2 + \varepsilon_{it}. \tag{12}$$

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

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

$$\begin{aligned}
\pi' &\sim \text{Truncated LN}\left(\mu(X_i), \sigma^2\right) &(13)\\
\text{Support}\left(\pi'\right) &= [4, 60] &(14)\\
\mu(X_i) &= \psi_0 + X_{it}'\psi + \kappa_0 \cdot \mathbf{1}\{t = 1\}. &(15)
\end{aligned}$$

New job offers are drawn from a truncated lognormal distribution, where the mean job offer 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 allow for a first-period-specific constant κ_0 in the mean of the job offer distribution. This allows for the possibility that skills which help an immigrant's initial placement in the US may not be determinative of that person's success after arrival in the US; this additional parameter helps in matching the shape of the first couple years of a worker's career.

4.2 Likelihood

Because we do not observe the full sequence of jobs in the data, the likelihood of each worker's data $L_i(\text{data}; \theta)$ is infeasible to calculate. We use a SML approach, which uses a function $\tilde{L}_i$ (data, $u;\theta$) that is an unbiased simulator of L_i :

$$E_u\left[\tilde{L}_i\left(\text{data, } u;\theta\right)|\text{data}\right] = L_i\left(\text{data};\theta\right), \tag{16}$$

where u is a random variable with a known distribution conditional on the data. The SML estimator approximates the conditional expectation by taking the average of $\tilde{L}_i$ over a set of S simulated values of u, transforming the infeasible standard maximum likelihood (MLE) estimator

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \sum_{i=1}^{n} \log L_i\left(\text{data},\theta\right) \tag{17}$$

into

$$\hat{\theta}_{SML} = \arg\max_{\theta} \sum_{i=1}^{n} \log\left[\frac{1}{S}\sum_{s=1}^{S}\tilde{L}_i\left(\text{data, } u^s;\theta\right)\right]. \tag{18}$$

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 $\hat{\theta}_{MLE}$ (Gourieroux and Monfort, 1996).[10] The SML method is particularly useful when u is a latent variable in the true model, and conditioning on that latent variable would make the likelihood standard. In our case, if we had access to the data on missing jobs, the likelihood would be a straightforward Markov process.

We now explain the derivation of the simulator of our likelihood function. Let the function $f_0(\cdot)$ give the likelihood of being in a given job in the initial period. Recall that we assume that everyone who stays in the labor market gets a job offer in period 0, so f_0 represents the time 0 job offer distribution, while also incorporating the chance that a person chooses to leave the labor market. We use $f(\cdot)$ to denote the likelihood of the job quality choice in all future periods. 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. This likelihood only depends on the previous job, and not the entire sequence of job offers, due to the Markov structure of the model. Appendix B explains how the functions f and f_0 are calculated. We write the pdf of the wage function as $g(\cdot)$. All of these functions depend on unobserved type; we denote each likelihood conditioning on type with a τ superscript.

Our observed data include the job quality of the first and final job, denoted as π_{i0} and π_{iT}, respectively. We also observe wages at both of these jobs, denoted as W_{i0} and W_{iT}. We also know the year the first job finished (yf_0) and the year the final job started (ys_T). Denote the

[10] 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 $S = 300$ throughout our results.

data as

$$\Omega_i \quad = \quad \left\{\pi_{i0}, \pi_{iT}, yf_{i0}, ys_{iT}, W_{i0}, W_{iT}\right\} \; . \tag{19}$$

There are missing data in that we do not observe all jobs between the initial and the final job. Denote the set of missing jobs as $\Omega_i^M = \left\{\pi_{i,yf_{i0}+1}, \pi_{i,yf_{i0}+2}, ..., \pi_{i,ys_T-1}\right\}$. If we observed the missing occupational quality data, the full likelihood could be calculated as follows:

$$\begin{aligned} L_i^\tau(\Omega_i, \Omega_i^M | X_i; \theta) \quad = \quad & \underbrace{f_0^\tau(\pi_{i0}|X_i)}_{\text{Initial job offer}} \times \underbrace{f^\tau(\pi_{i0}|\pi_{i0}, X_i)^{yf_{i0}-1}}_{\text{Stays in initial job each period}} \\ \times \quad & \underbrace{\prod_{t=yf_{i0+1}}^{ys_i} f^\tau\left(\pi_{it}|\pi_{i,t-1}, X_i\right)}_{\text{Likelihood of unobserved jobs}} \\ \times \quad & \underbrace{f^\tau\left(\pi_{iT}|\pi_{t-1} = \pi_{iT}, X_i, \tau\right)^{T-ys_T-1}}_{\text{Stays in final job in each period}} \\ \times \quad & \underbrace{g^\tau(W_{i0}|\pi_{i0}, X_i) \cdot g^\tau(W_{iT}|\pi_{iT}, X_i)}_{\text{Wage pdf}}. \end{aligned} \tag{20}$$

The first line has two components: the probability of the first job offer, which is given by the function $f_0(\cdot)$, and the probability that a person stays in that job for the given number of periods. If we observed all job outcomes, the second line in equation (20) would give the likelihood contribution for the missing periods. Once a person lands in their final job, we know the probability they stay there each period, and this happens for $T - ys_{iT} - 1$ periods. The last line of equation (20) gives the wage contribution to the likelihood.

Equation (20) is infeasible to calculate because of the missing jobs. Let $h(\cdot)$ denote the joint density of job outcomes (given X_i) in all missing periods, and then write the likelihood of the observed data by integrating out the missing data. The likelihood of the observed data can be written as follows:

$$\begin{aligned} L_i^\tau(\Omega_i|X_i;\theta) \quad = \quad & f_0^\tau(\pi_{i0}|X_i) \times f^\tau(\pi_{i0}|\pi_{i0}, X_i)^{yf_{i0}-1} \\ \times \quad & \underbrace{\int \cdots \int f^\tau\left(\pi_{iT}|\pi_{ys_{iT}-1}, X_i\right) h^\tau\left(\pi_{ys_{iT}}, \pi_{ys_{iT}-1}, \cdots, \pi_{yf_{i0}+1}\right) d\pi_{ys_{iT}} d\pi_{ys_{iT}-1} \cdots d\pi_{yf_{i0}}}_{\text{Expected value of } f} \\ \times \quad & f^\tau\left(\pi_{iT}|\pi_{i,t-1} = \pi_{iT}, X_i\right)^{T-ys_{iT}-1} \\ \times \quad & g^\tau(W_{i0}|\pi_{i0}, X_i) \cdot g^\tau(W_{iT}|\pi_{iT}, X_i). \end{aligned} \tag{21}$$

Calculation of equation (21) directly is computationally challenging because of the multidimensional integral.

To overcome this integration problem, we simply replace L_i^τ by a function $\tilde{L}_i^\tau$ such that $E\left[\tilde{L}_i^\tau\right] = L_i^\tau$. To do this, note that the second line in equation (21) is the expectation of the function f over the missing jobs when the missing jobs are sampled from $h^\tau(\cdot)$. We then replace the second line in equation (21) by drawing S simulations of the missing jobs from the true model given the guessed parameters and taking the average value, noting that as $S \to \infty$,

$$\frac{1}{S}\sum_{s=1}^{S} f^\tau\left(\pi_{it}|\pi_{i,t-1}^s, X_i\right) \quad \to_p \quad E\left[f^\tau\left(\pi_{iT}|\pi_{ys_{iT}-1}, X_i\right)\right] . \tag{22}$$

Equation (21) is conditional on unobserved type. To calculate the unconditional likelihood, we assume there are two types, which occur with probability ν_0 and $(1-\nu_0)$, and integrate (sum) over types:

$$\begin{aligned} \tilde{L}_i\left(\Omega_i|X_i;\theta\right) &= \nu_0 \tilde{L}_i^{\tau=0}\left(\Omega_i|X_i;\theta\right) \\ &+ (1-\nu_0) L_i^{\tau=1}\left(\Omega_i|X_i;\theta\right) . \end{aligned} \tag{23}$$

Maximizing the full SML criterion function (equation 18) gives parameter estimates, and the likelihood can be used to calculate asymptotic standard errors and perform standard inference.

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 probability $O(\pi_{it}, X_i, z_i, \tau)$ each period. Table 3 shows the parameter estimates for this process. People who work in higher-quality occupations, as well as men, are less likely to leave the labor market, while people with children are more likely to exit than those without. There are two unobserved worker types in the model, and we find that 23% of workers are what we call the "non-working" type, meaning that they are more likely to exit the labor market. We find the non-working type has a per-period probability of over 95% of leaving the labor market, meaning that most of them get labor market exit shocks in the initial period. The "working" types have low probabilities of leaving the labor market, meaning that most of them do not exit in any period.

5.2 Wage Equation

The parameters of the wage function, equation (12), are shown in Table 4. 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 large but noisy difference in the wages of working versus non-working types, with types with a low probability of exiting the labor market estimated to earn 29% more than the exiting type, but with a standard error of 25%. The parameters of the OOLF shock distribution imply that the non-working types have a high probability of exiting the labor market in the initial period, meaning that we do not observe wages for most of them. Few wage observations in the dropout group makes it difficult to precisely estimate a wage difference across the two groups.

5.3 Occupational Transition Parameters

We estimate three sets of parameters relating immigrant demographics to the occupational transition process, given by equations (9), (10), and (15). These results are shown in Table 5. The first column reports the parameters governing job offer rates. People in higher-quality home occupations are more likely to get job offers each period, as are people with higher levels of education. The job loss rate, shown in column (2) of Table 5, is estimated as a constant. The model estimates imply about a 4% job loss rate each year. The third column of Table 5 shows the estimates of the mean and variance of the lognormal job offer distribution. People with higher-quality home occupations get better job offers, as do people with higher 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 the job offers in the first period could be different than other offers, we estimate a separate term for the constant in the mean of the job offer distribution for the first period. We estimate the standard deviation of the job offer distribution as a constant.[11]

The magnitudes of our point estimates are demonstrated graphically in Figure 5. 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 of repeating this exercise for English skills, whether or not an immigrant enters the US on a

[11]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.

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 occupation at US entry. English skills and high-quality home country occupation also lead to faster occupational upgrading over time in the US. Occupational upgrading paths differ dramatically between high- and low-skilled immigrants across all four measures, which is an indication that the counterfactuals will have heterogeneous effects across workers.

5.4 Model Fit

A comparison of the predicted occupations between the model and the data is shown in Figure 6(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 200 times given their pre-immigration characteristics. We then sample these career paths using the sampling scheme from the data; we drop any information on occupations 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 occupational quality rises with experience for 6 years and then begin to fall is difficult to fit without cohort-specific heterogeneity, since our model predicts (on average) monotonic career paths. We can match the hump shape largely because the changing demographics 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 occupations, and lower earnings, which offsets the positive occupational growth effects of increasing experience. Panel (b) shows the model fit for wages, again showing that we are able to match the non-monotonic shape 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 placing immigrants into their home country job when they enter the US labor market. In particular, we consider a situation where the worker's characteristics, outside job offers, and labor market shocks were held equal, but the initial job after entry in the US had the occupational quality of immigrant's last job in their home country. The worker then re-optimizes their career choices given this new initial condition. For example, in this counterfactual, doctors in their home countries would be doctors in the US immediately, construction workers in their home country would be construction workers in the US immediately, 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 long-run job in the US. To do this, we simulate the model, and find the average job that each person would be in the long-run steady state – that is, the job such that given their skills, on average in the next period they would be back at that job – and then place them in that job in their first period in the US. For example, if a person starts as a tailor but whose 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, regardless of their actual home occupation.

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.

The 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. On one hand, an immigrant worker may be significantly more productive in the US than in their home country, depending on the situation in the home country, but on the other, we might expect that similar job titles across countries do not reflect equivalent skill levels required. 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.

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 will be moving 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.[12] In addition, we look at only wage gains from moving immigrants to different jobs. There are also potentially non-pecuniary benefits to

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

helping immigrants get into their preferred job that are not captured in our framework.

Given our model estimates, the worker's unobserved type will not play a large role in our counterfactuals. Our estimates of the distribution of unobserved type effectively divides workers into two groups, those who leave the labor market immediately and those with a small chance of ever leaving. Because of this, in this section we restrict our analysis to the 2,471 immigrants who are observed in the labor market initially and of whom only a small percentage (<5%) will ever leave the labor market. We assume those workers have the strongly attached type. Interpretation of the counterfactuals will be implicitly conditional on being the type of worker with the low probability of leaving the labor force.

6.1 Counterfactuals: Average Occupation and Wage Effects

To implement the counterfactuals, we simulate 500 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. We average across the simulated counterfactual jobs in each period to generate occupational paths in both the baseline and in each counterfactual. For the home job counterfactual, each immigrant's first job in the US is the same as their home country job. For the long-run jobs counterfactual, we first simulate the model to find each worker's long-run job, and then place them in that job in the first period.[13]

To demonstrate the effects of these policies, in Figure 7(a) we show the average simulated occupational paths over time in the US in the baseline, the home job counterfactual, and the long-run job counterfactual. We also show occupational outcomes when we eliminate occupational upgrading, meaning that workers remain in their initial US job in all periods. We can start by comparing the initial US job in each scenario. The baseline initial job has an average wage of about $12 an hour, compared to $14 an hour in the home job counterfactual and $18 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 at home in the long run.

We can next compare the job paths across 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 there is no long-run effect on jobs. In the long-run job counterfactual, since workers start their careers in their long-run job, there will be almost no occupational growth on average, since workers will lose their jobs at a rate exactly offsetting their better job offers. Comparing the long-run jobs counterfactual to the baseline, we see that even after 10 years, workers are still not in their long-run job. The red occupational path (the no-upgrading counterfactual) shows the results if workers never received outside job offers; by construction the average occupation over time will be flat. In the

[13]To find the long-run job, we simulate workers out 20 years and take the average job across simulations as their mean long-run job quality.

case of occupations, this line is unimportant, but we will calculate the wages associated with this no-upgrading path below. Figure 7(b) shows the differences in occupation qualities between each counterfactual and the baseline. We see that the effect is largest for the long-run occupation counterfactual, and the effects of both counterfactuals shrink over time.

The effects of the counterfactuals on workers' wages over time in the US are shown in Figure 8 and Table 6. To create this table and figure, we calculated averages across the wage path of each simulation, feeding in the simulated occupation paths while holding everything else constant. The difference between the baseline and the home job counterfactual is only approximately $0.71 at entry. Over time, the effects of the counterfactual fall as workers' baseline and counterfactual jobs converge, and after 9 years there is only a $0.24 difference between the baseline and the counterfactual. For the long-run job, we see a larger increase in counterfactual wages at entry, corresponding to an approximately $3.40 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 $1.26 difference in wages in the baseline and counterfactual at year 9. When looking at the simulation that shuts down occupational upgrading, we see that wage growth is 30% lower than in the baseline. This suggests that occupational mobility accounts for 30% of worker's baseline wage growth over their first nine years in the US. Panel (b) 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 counterfactuals, we see small effects of these policies on the average occupational path of immigrants. The home jobs counterfactual raises wages by 5.3% at entry, and after 9 years the gain from the counterfactual is less than 1%. In the long-run jobs counterfactual, which is more of a best-case approach, wages increase by 25.7% at entry, but this gain is temporary and declines to only 4.3% after 9 years in the US.

The occupational 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 occupation qualities and calculate the average counterfactual effects, the home job counterfactual now increases wages by 33% at entry and by 6% after 9 years, and of course the long-run jobs counterfactual effect is even larger. 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 guess about what types of workers have large returns to these counterfactuals would be that higher-ability immigrants would gain more, since it seems likely that higher-skilled jobs may have more barriers to entry across national borders. In Figures 9 and 10 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, assuming they worked in their home country job. This serves as our proxy for each immigrant's skill level.

We then compare predicted entry wages to wages in the baseline and in the counterfactual. Figure 9 shows estimates of 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 9 gives the absolute levels of wages in each situation. The fact that the blue line (the home job counterfactual) is 45 degrees in panel (a) is mechanical due to how we constructed the skill measure. However, average wages in the home job counterfactual are above the baseline wages for skills above \$12, and below baseline wages for skills less than \$12. The counterfactual gains compared to the baseline are shown in panel (b), and this displays the results more clearly. The lowest skilled workers would actually lose from the home job counterfactual, and for most workers, who lie in the \$10-\$20 range, the gains are small, with larger gains for the highest skilled. The effects of the long run job are positive for everyone, and increasing in skill level. Figure 10 contains the same information as Figure 9 but considers wages after 10 years in the US labor market: The effects of the home job counterfactual fade over time and are compressed across the skill distribution so that even for the highest-skilled workers, the counterfactual effects have largely disappeared by the 10th year. The gains from the long run job counterfactual are still positive for everyone after 10 years, albeit small, and are increasing in skill level.

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 first 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 set of job offers and labor market shocks as given and see how much starting a particular worker in their home job or their long-run job changes their particular occupations and wages. We then take the average of the difference in wages in the baseline and the counterfactual for each person over the distribution 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).

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 treatment in question for each of our 2,417 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 7. In column (1) we consider the ATEs for the home job counterfactual at US entry as the dependent variable, and show the regression coefficients for the demographics. In column (3) we repeat the exercise with home job individual-level ATEs for year 10 as the dependent variable. 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 (education, high English skills, 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 observed lower ability workers.

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 7 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 7. We find that college educated workers have large returns to the home jobs counterfactual (an additional \$1.42 per hour), while workers with good English skills and employer-sponsored visas see smaller returns.

The finding that immigrants with good English skills and 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 can speak English well and 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 lack good English skills and 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, men have slightly lower returns to the counterfactual, as do older-age immigrants; both of these results seem consistent with men and young workers facing fewer frictions. Immigrants who entered in more recent cohorts (even conditional on age) also have slightly higher returns than earlier cohorts, perhaps indicating an increase in the difficulty of assimilation.

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 8. In columns (1) and (2), the dependent variable is the ATEs at entry and after 10 years in the US labor market, respectively. In comparison to the home jobs counterfactual, the size of the long-run jobs ATE is positively related to almost all demographics. In the long run jobs counterfactual, for instance, high English skills are now strongly associated with a high ATE. Interpreting this result in terms of our model estimates, recall that English skills shift the mean of the offer distribution upwards significantly. This results in faster job quality growth and a higher quality steady-state job, and so the long-run jobs counterfactual has a larger effect for higher English skill workers. The importance of the higher steady-state job in the counterfactual dominates the fact that these high English skill individuals see less downgrading from their home occupations at US entry, which is why the home jobs counterfactual had a smaller effect for high English skill workers. Concretely, compare immigrant A, who has high English skills, to immigrant B, who has low English skills. Assume both immigrant A and immigrant B worked as an accountant 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 secretary, which is a relatively small drop in occupational 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 becomes an accountant 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. The fact that the counterfactual effects are not uniform across immigrants suggests that targeting higher-skill immigrants with policies to reduce their occupational entry barriers would have the largest benefits.

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 steady-state 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 5.3% at entry, and after 9 years the gain from the counterfactual is less than 1%. In the long-run jobs counterfactual, wages increase by 25.7% at entry, but only by 4.3% after 9 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 33% at entry and 6% after 9 years, 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 with high-skilled immigrants 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.

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	37.2
Percent male	54.8
Years living in the US at end of first job	2.07
Years living in the US at time of survey	8.00
Percent with an employer sponsor	34.1
Percent with more than high school education	64.9
Percent with high English skills	39.7
Percent who entered US on a valid visa	77.9
Percent who leave the labor market	16.4
Percent with children	72.4
Sample Size	2,933

Table 2: Determinants of Job Quality in the US

	(1) Initial job	(2) Current job
College	2.397***	0.909***
	(0.372)	(0.344)
English skills	1.782***	1.726***
	(0.305)	(0.280)
Home country GDP	-0.245	0.115
	(0.226)	(0.156)
Quality of home country job, moved when <18	0.241**	0.116
	(0.0963)	(0.0906)
Quality of home country job, moved when ≥ 18	0.276***	0.0890***
	(0.0247)	(0.0210)
License requirements of home country job	3.418***	-3.034***
	(1.146)	(0.959)
Employer-sponsored visa	3.440***	0.508*
	(0.300)	(0.288)
Entered US on valid visa	1.835***	0.687*
	(0.408)	(0.405)
US work experience (years)	-0.177	0.582***
	(0.187)	(0.105)
US work experience squared	0.00605	-0.0187***
	(0.0172)	(0.00547)
US work experience x home country GDP	0.0232	-0.00875
	(0.0375)	(0.00985)
Quality of first US job		0.711***
		(0.0215)
Constant	1.562	2.488**
	(1.340)	(1.182)
Observations	1,843	1,721
Adjusted R^2	0.423	0.641

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, and 0 otherwise. License requirements is a dummy variable indicating whether or not there are licenses available in the occupation a person is working in. Controls for schooling in the US, gender, home experience, home experience squared, and home country GDP (as well as interactions between home country GDP and home job quality) included but not reported.

Table 3: Out-of-the-Labor-Force Parameter Estimates

Last occupation quality	-0.48
	(0.15)
Male	-0.29
	(0.16)
Year born	0.31
	(0.090)
Kids	0.29
	(0.17)
Non-working type	5.77
	(0.56)
Constant	-1.43
	(0.37)
Probability (non-working type)	0.23
	(0.0078)

Notes: Standard errors in parentheses. These terms give 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 occupation quality is the 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.

Table 4: Log Wage Function Parameter Estimates

Current occupation quality	0.60
	(0.022)
Home occupation quality	0.10
	(0.0090)
Male	0.020
	(0.017)
Year born	0.012
	(0.012)
College	0.11
	(0.024)
English skills	0.20
	(0.019)
Valid visa	0.096
	(0.029)
Home country GDP	-0.00049
	(0.022)
Entry year	0.0039
	(0.013)
Years in US	0.066
	(0.0063)
Years in US squared	-0.049
	(0.041)
Non-working type	-0.29
	(0.25)
Constant	0.86
	(0.32)
Standard deviation	0.46
	(0.0055)

Notes: Standard errors in parentheses. This table gives the parameters of the log wage distribution. Current and home occupation is the average wage (calculated using CPS data) in a given occupation. Current occupation 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 reports strong English skills, and 0 otherwise.

Table 5: Job Offer rates, Job Loss Rates, and the Job Offer Distribution

	(1) Job offer rates	(2) Job loss rates	(3) Job offer distribution
Home occupation quality	0.13		0.21
	(0.057)		(0.027)
Male	-0.14		0.15
	(0.074)		(0.038)
Year born	-0.011		0.14
	(0.052)		(0.024)
College	0.23		0.090
	(0.032)		(0.014)
English skills	-0.0058		0.19
	(0.027)		(0.011)
Employer-sponsored visa	0.020		0.071
	(0.035)		(0.016)
Valid visa	0.012		0.011
	(0.050)		(0.029)
Home country GDP	-0.063		0.28
	(0.057)		(0.028)
Entry year	0.097		0.25
	(0.063)		(0.032)
Initial period shift			0.12
			(0.030)
Constant term	-0.23	0.046	1.37
	(0.073)	(0.0027)	(0.040)
Standard deviation			0.60
			(0.0082)

Notes: Standard errors in parentheses. In columns (1) and (2), the coefficients give the effect of each characteristic on the probability that a person gets a job offer or loses their job each period. All continuous variables are scaled to have mean 0 and variance 1. We use the function $\frac{1}{2} + \frac{1}{2}\tanh(x)$ to ensure the probabilities are between 0 and 1. Column (3) shows the parameters for the mean and standard deviation of the job offer distribution, which we assume to be lognormal. “College” equals 1 if a person has more than 12 years of education, and 0 otherwise. “English skills” equals 1 if a person reports 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 6: Counterfactual Effects on Wages

Years after US entry	Baseline	Home country job	Long-run job	No occupational mobility
0	13.29	14.00	16.69	13.27
3	17.87	18.43	20.34	16.18
6	22.70	23.08	24.47	19.52
9	27.90	28.14	29.16	23.31

Notes: In each counterfactual, we calculate the average wage for each person over 500 simulations. The first column is the baseline. The second column places each immigrant in their home country occupation in the first period in the US, and the third column puts them in their long-run job at entry. The last column shows average wages when each person stays in their initial occupation each period.

Table 7: Determinants of ATEs for the Home Job Counterfactual

	Entry		After 10 years	
	(1)	(2)	(3)	(4)
Home occupation quality	4.01***		1.69***	
	(0.019)		(0.035)	
Male	-0.75***	-0.32**	-0.54***	-0.35***
	(0.033)	(0.15)	(0.061)	(0.085)
Year born	-0.43***	-0.71***	-0.26***	-0.38***
	(0.017)	(0.076)	(0.032)	(0.044)
College	-1.11***	1.42***	-0.60***	0.47***
	(0.043)	(0.18)	(0.079)	(0.11)
English skills	-1.30***	-0.73***	-0.91***	-0.67***
	(0.038)	(0.17)	(0.069)	(0.097)
Employer-sponsored visa	-1.97***	-0.87***	-1.22***	-0.76***
	(0.038)	(0.16)	(0.069)	(0.095)
Valid visa	-0.29***	0.38*	-0.14	0.15
	(0.051)	(0.22)	(0.093)	(0.13)
Home country GDP	-0.088***	-0.15**	-0.061**	-0.088**
	(0.016)	(0.071)	(0.029)	(0.041)
Entry year	-0.41***	0.26***	-0.27***	0.012
	(0.020)	(0.088)	(0.037)	(0.051)
Constant	3.43***	0.39**	1.71***	0.43***
	(0.045)	(0.19)	(0.083)	(0.11)
R-squared	0.95	0.087	0.55	0.11
N	2,471	2,471	2,471	2,471

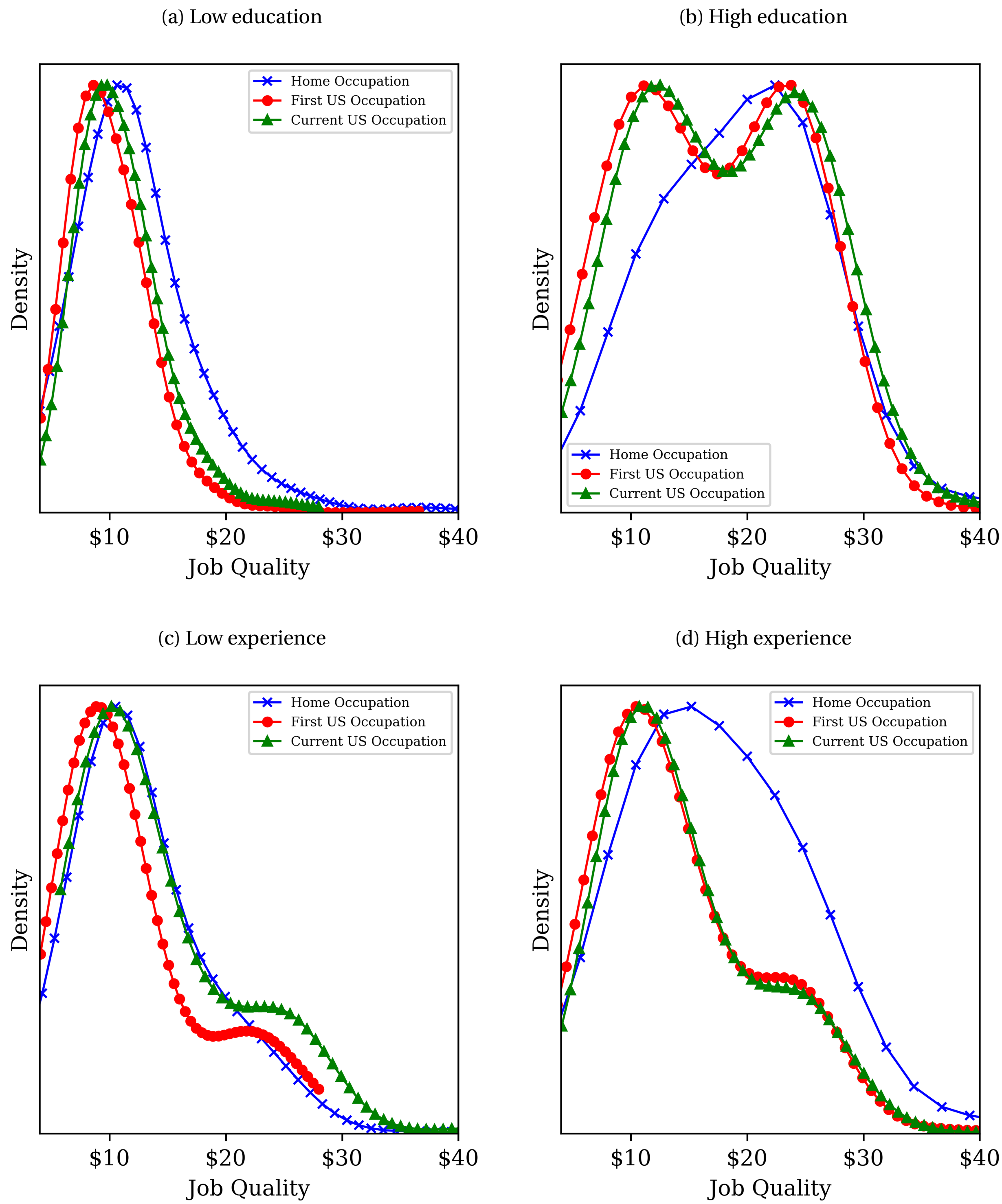
Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the ATE from the home job counterfactual for each person, which is calculated by taking the difference between the average wage in the counterfactual and the baseline for each person. Home occupation 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. "English skills" equals 1 if a person reports strong English skills, and 0 otherwise.

Table 8: Determinants of ATEs for the Long-Run Job Counterfactual

	Entry	After 10 years
	(1)	(2)
Home occupation quality	0.59***	0.40***
	(0.019)	(0.026)
Male	0.17***	0.18***
	(0.033)	(0.046)
Year born	0.24***	0.11***
	(0.017)	(0.024)
College	0.65***	0.42***
	(0.043)	(0.060)
English skills	1.39***	0.74***
	(0.038)	(0.053)
Employer-sponsored visa	0.15***	0.25***
	(0.037)	(0.052)
Valid visa	-0.45***	0.13*
	(0.050)	(0.071)
Home country GDP	0.048***	-0.0010
	(0.016)	(0.022)
Entry year	0.89***	0.18***
	(0.020)	(0.028)
Constant	1.24***	1.00***
	(0.045)	(0.063)
R-squared	0.95	0.42
N	2,471	2,471

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 occupation 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. “English skills” equals 1 if a person reports strong English skills, 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 a high school education, and plot (b) shows it for people with more than 12 years of education.

Figure 2: Timing of model

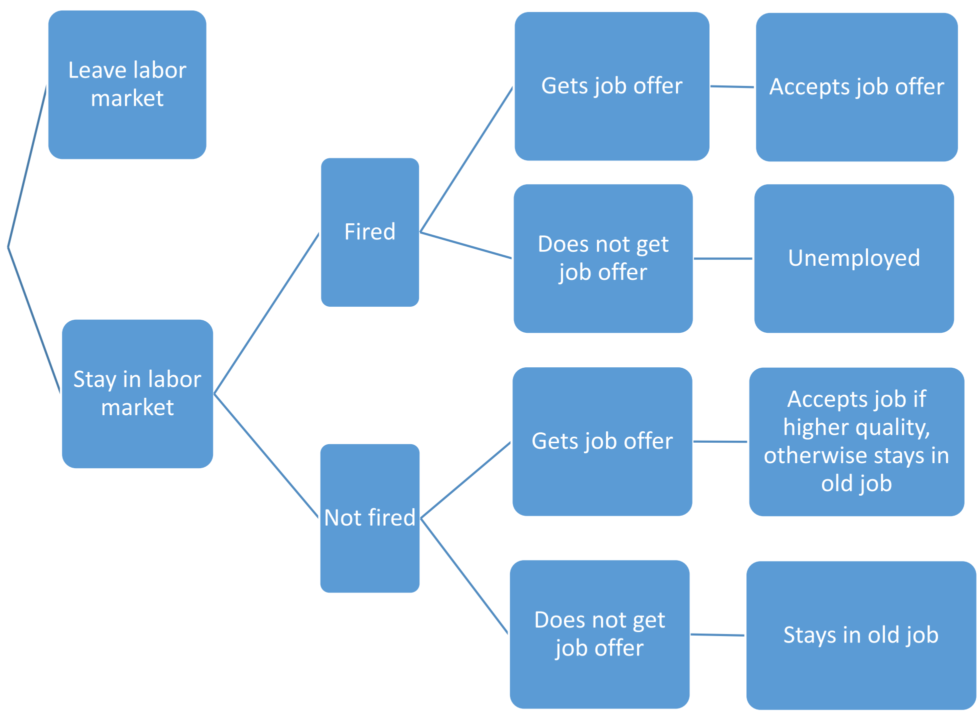
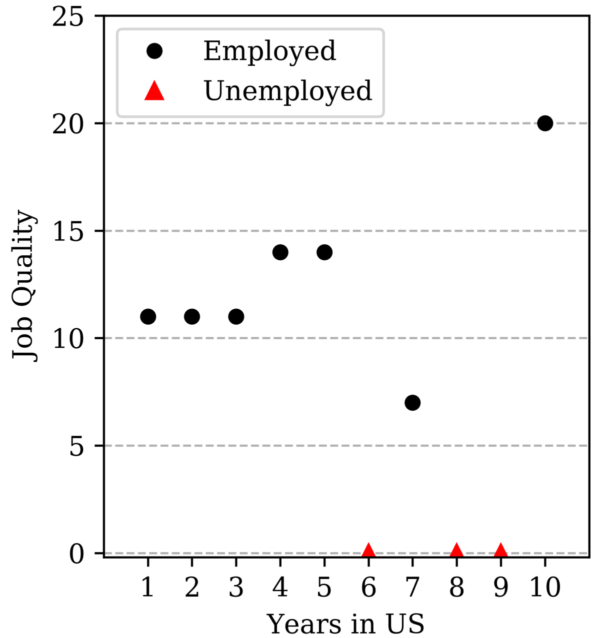
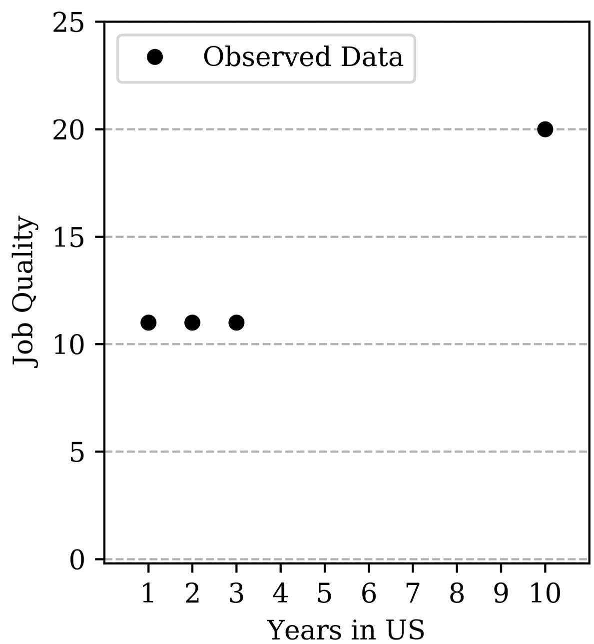


Figure 3: Example Career Paths in Model

(a) Example Career Path

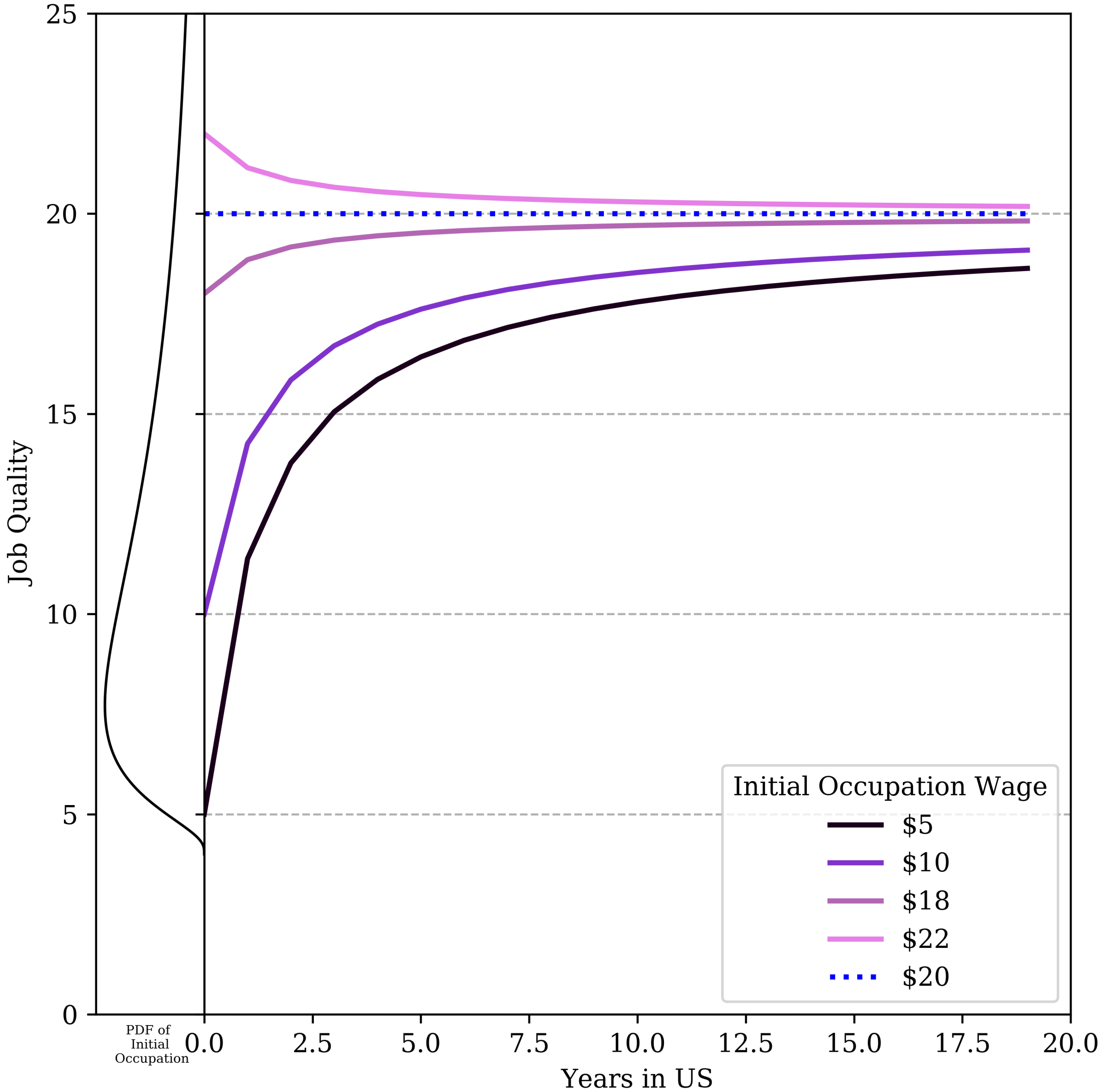


(b) Observed Data from Example Career



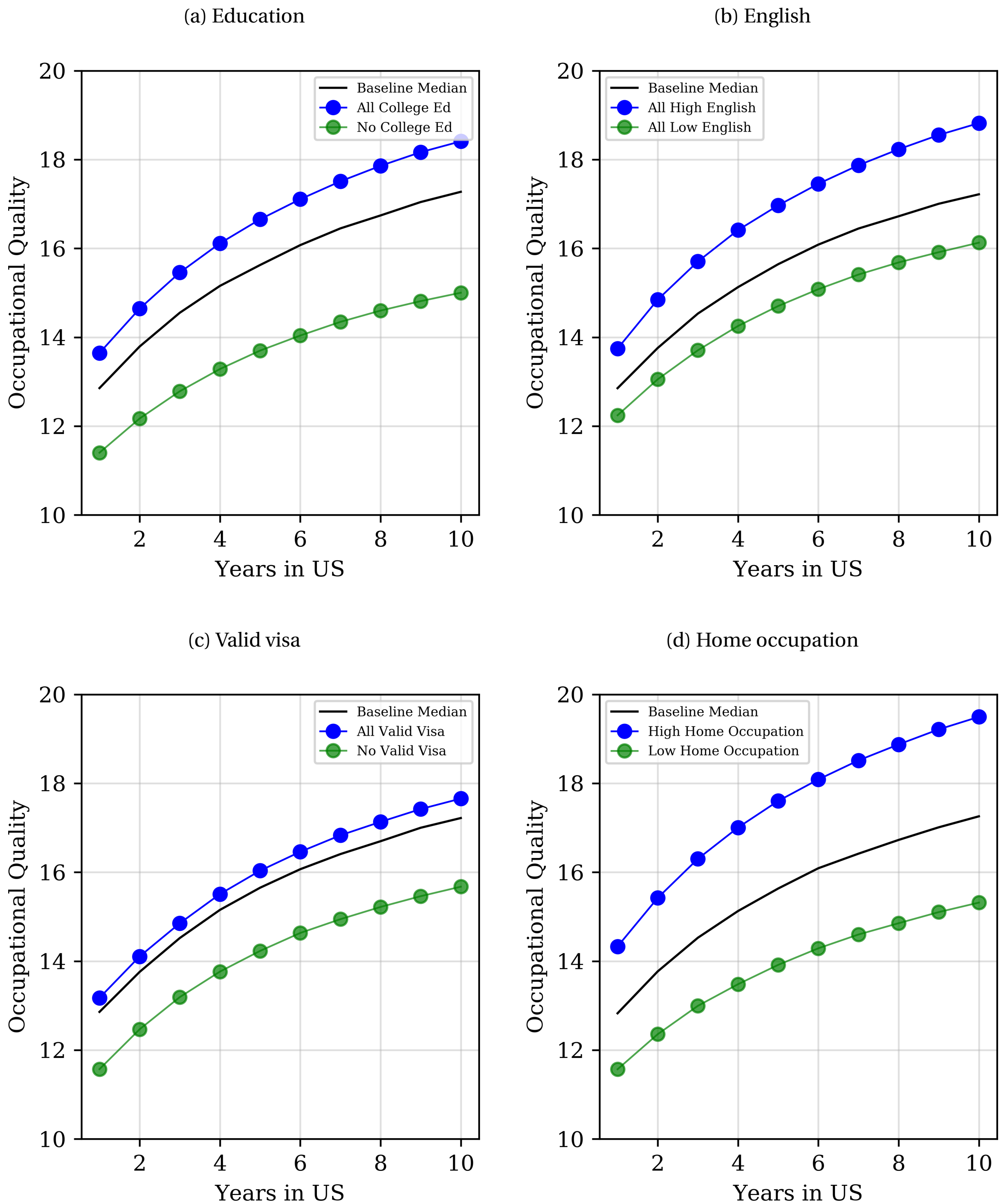
Notes: Plot (a) shows a potential career path for a worker. Plot (b) shows what the data would look like in this case given the NIS sampling scheme where we do not observe intermediate jobs.

Figure 4: Effect of Initial Job on Career Path



Notes: This figure shows smoothed career paths in the model for different initial job qualities in the US.

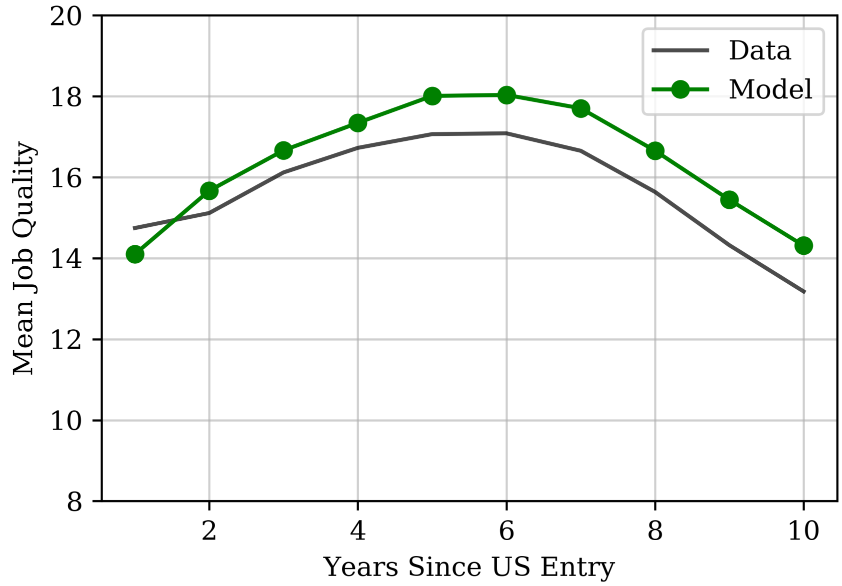
Figure 5: Effects of Demographic Characteristics on Occupational Outcomes

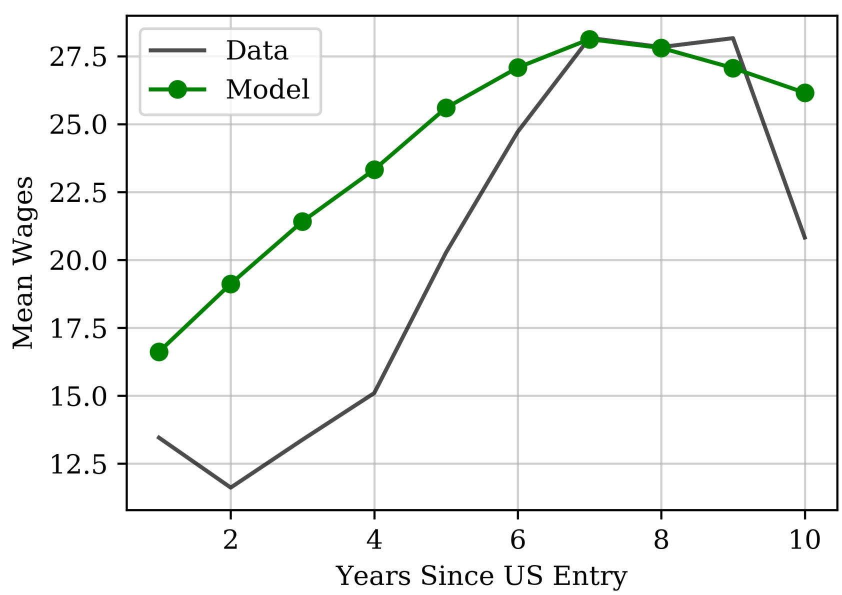


Notes: In each plot, the black line shows the median wage quality at a given number of years of experience in the US. In each of the plots, there are 2 other lines, which are created by changing one characteristic for everyone in the sample. Plot (a) shows everyone with and without some college education, 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 occupation.

Figure 6: Model Fit

(a) Occupations



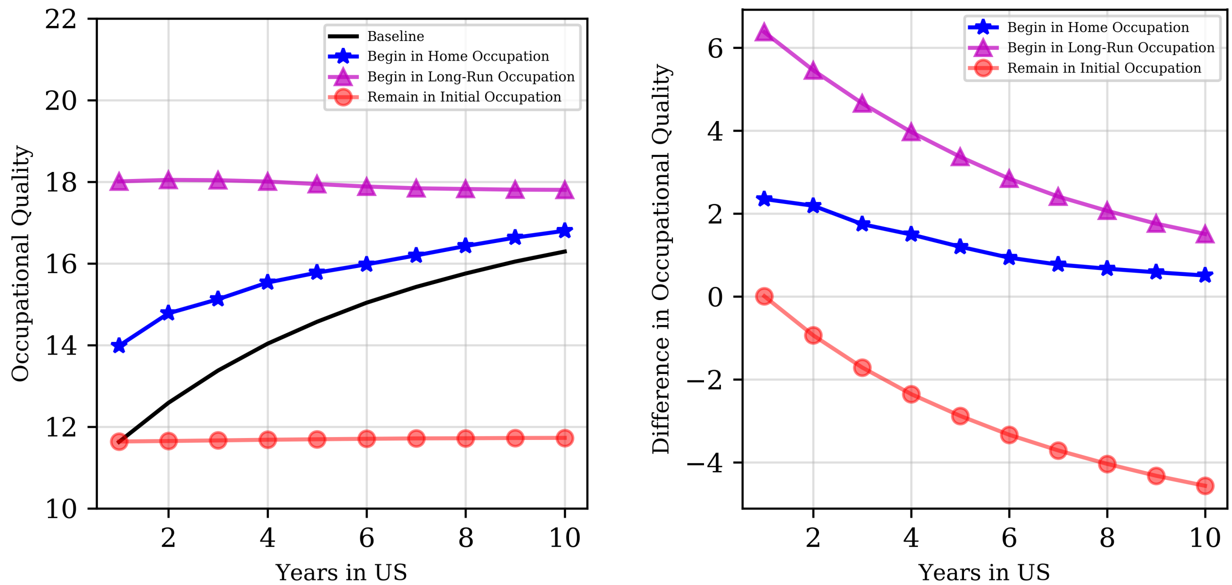


(b) Wages

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 7: Counterfactual Effects on Occupational Quality

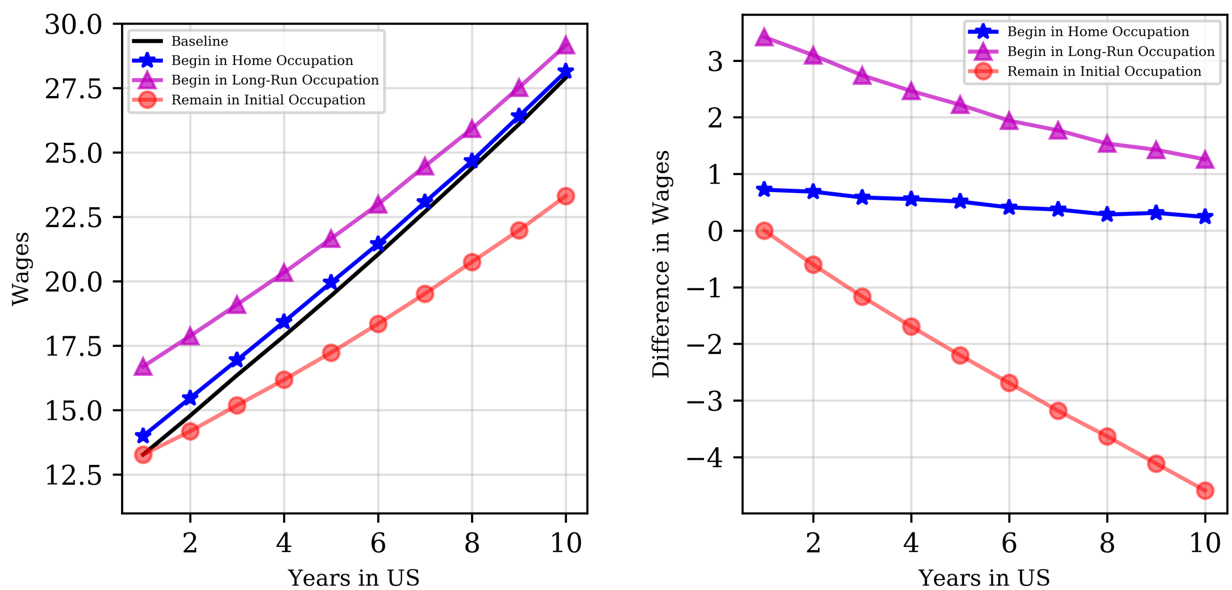
(a) Levels (b) Difference from baseline



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

Figure 8: Counterfactual Effects on Wages

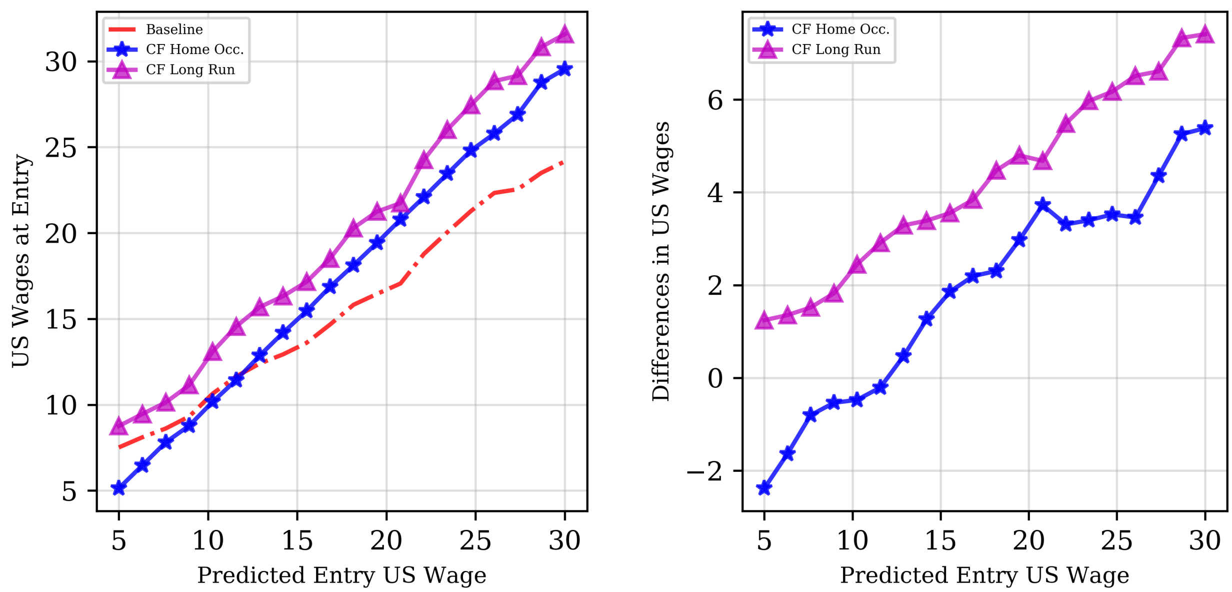
(a) Levels (b) Difference from baseline



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

Figure 9: Returns to Counterfactuals by Predicted Entry Wage at US Entry

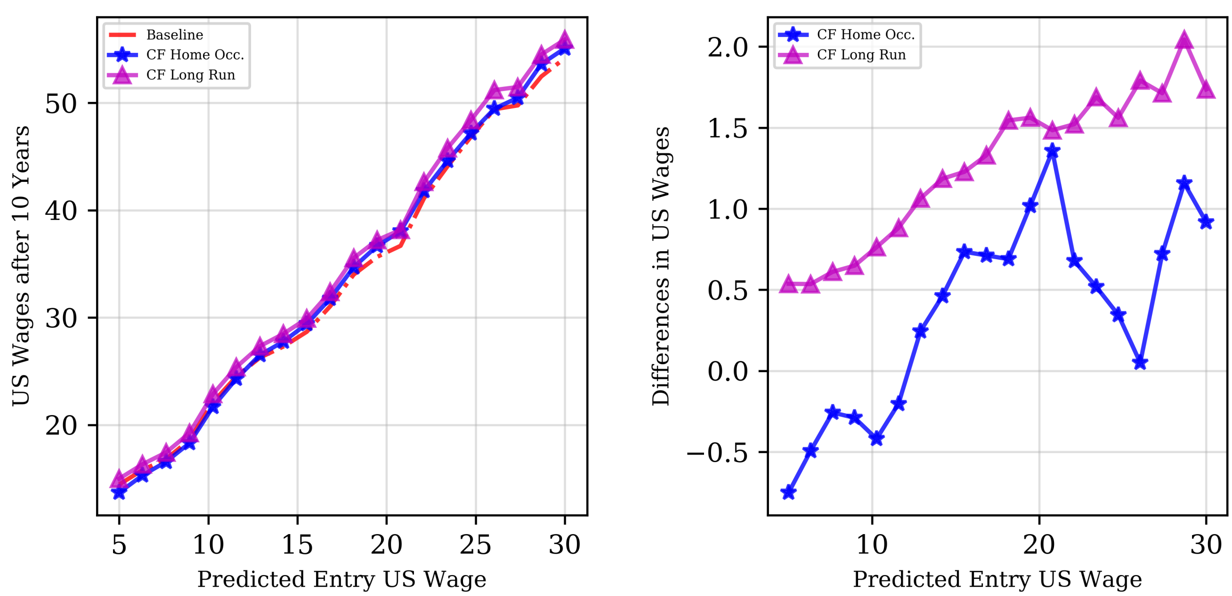
(a) Levels | (b) Difference from baseline



Notes: Sub-figure (a) shows the median wages. 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 10: 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. 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.

For Online Publication:

A Estimation Sample

In this Appendix, we explain how we created our estimation sample. The NIS starts with a sample of 26,078 individuals. We only keep household heads or spouses, since not all relevant questions are asked to the remaining members of the household. This causes us to drop 14,214 observations. 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 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 1,106 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. Finally, we drop 8 observations where the initial year in the US was prior to 1984. We drop 29 observations where we do not have information on whether or not the immigrants entered the US on a valid visa.

B Job Transition Kernels

Recall from Section 3 that jobs are characterized by quality π_{it}. We denote unemployment as its own job with some arbitrary value of π below the lower bound of the job offer distribution; call unemployed workers "employed" at $\pi_t = 0$. This is consistent with the model, as unemployed workers will always accept an offer, as all actual employment offers have superior π. Similarly, denote the status of out of the labor force with $\pi_t = -1$.

In Section 4.2, we calculated the likelihood function. To do this, we defined the functions $f_0^\tau(\pi_0|X)$ and $f^\tau(\pi_t|\pi_{t-1}, X)$, which (using language from probability theory) are the transition kernels associated with observing a given pair of job outcomes in periods $t-1$ and t. 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 f_0^{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 $\Pi(\cdot)$; for the associated probability density function, for simplicity, we use the notation $\partial\Pi \equiv \frac{\partial\Pi(x)}{\partial x}$ to denote the likelihood of a particular occupational draw x.

The first component of the likelihood is the initial job offer. We assume each worker gets a job offer in the first period. We need to consider the probability that the worker drops out of the labor market, which we denote using the function $O(\cdot)$, and, if they stay in the labor market, the likelihood of their given draw, given by $\partial\Pi_0(\cdot)$. These functions were defined in equations (3) and (7).

$$f_0^{\tau}(\pi_0|X) = \begin{cases} \partial\Pi_0(\pi_0|X)\left(1-O\left(\pi^H, X, \tau\right)\right) & \text{if } \pi_0 > -1 \\ O\left(\pi^H, X, \tau\right) & \text{if } \pi_0 = -1 \end{cases} \tag{24}$$

Note that 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.

The later conditional likelihoods can be calculated from the occupational transition equation, given in equation (8). The model breaks down the likelihood of transitioning between occupations π_{t-1} and π_t, $f^{\tau}(\pi_t|\pi_{t-1})$, into six different cases depending 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 a job offer is $p(\cdot)$, and the probability of being fired is $q(\cdot)$, as given by equations (4)-(6). The cases are as follows:

1. 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

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

2. 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

$$f^{\tau}\left(\pi_t = 0|\pi_{t-1} = 0, X\right) = \left(1 - O\left(\pi_{t-1}, X, \tau\right)\right) \times \left(1 - p(X)\right). \tag{25}$$

3. 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. The likelihood in this case is the product of the probability of staying in the labor market, receiving a job offer, and receiving the specific offer that is observed. In this case, the likelihood of observing job $\pi_t > \pi_{t-1}$ is

$$f^{\tau}\left(\pi_t|\pi_{t-1}, X\right) = \left(1 - O\left(\pi_{t-1}, X, \tau\right)\right) \times p(X) \times \partial\Pi(\pi_t|X). \tag{26}$$

4. 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

$$f^{\tau}\left(\pi_t|\pi_{t-1}, X\right) = \left(1 - O\left(\pi_{t-1}, X, \tau\right)\right) \times q(X) \times p(X) \times \partial\Pi(\pi_t|X). \tag{27}$$

5. 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 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

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

6. A worker leaves the labor force. In the case they were out of the labor force the previous period, the likelihood is 1. In the case they were not, we know they received an OOLF shock, and all other shocks are irrelevant. When $\pi_{t-1} \neq -1$,

$$f^{\tau}\left(\pi_t = -1|\pi_{t-1}, X\right) = O\left(\pi_{t-1}, X, \tau\right)$$

C Identification

C.1 Out-of-the-Labor-Force Process

In this section, we assume access to panel data of workers with dummy variables indicating whether or not they have ever moved from in the labor force to the absorbing state of out of the labor force, as well as the occupations in each period in the labor force and a vector of time-invariant characteristics X_i. With large N and fixed T, this gives us access to statistics such as $\Pr\left(\text{worker with characteristics } X_i \text{ drops out by time } t\right)$, etc.

First, consider a simpler problem where there are two types of workers, type H having measure ν and type L having measure $1-\nu$, and each type has a constant per-period probability of permanently dropping out of the labor force. Denote the converse probabilities, the probability of not receiving the dropout shock within a period, as p_H and p_L, respectively. Our goal is to show that ν, p_H, and p_L are identified.

We denote the overall probability of a random worker remaining in the labor market for at least the first period as r_1 , and it is calculated as

$$r_1 \equiv \nu \cdot p_H + (1-\nu) \cdot p_L \equiv r_1\ , \tag{28}$$

since with probability ν the worker is of the high type H and then will remain with probability p_H, and similarly for the low type. By the same logic, the probability of remaining in the labor force for at least two periods is

$$r_2 \equiv \nu \cdot p_H^2 + (1-\nu) \cdot p_L^2 \tag{29}$$

since each type of worker has to avoid the dropout shock twice in a row. The three-period version is of course

$$r_3 \equiv \nu \cdot p_H^3 + (1-\nu) \cdot p_L^3. \tag{30}$$

We assume that we have access to r_1, r_2, and r_3 in the data. With an independent sample of workers observed for at least three periods, the sample probabilities of remaining in the labor force at least 1, 2, and 3 periods will converge to r_1, r_2, and r_3 as the number of workers goes to infinity.

If we can show that knowledge of r_1, r_2, and r_3 uniquely defines ν, p_H, and p_L (up to switching the labels H and L), then the OOLF type probabilities and type-specific hazard rates are identified in this two-type model. This requires establishing global uniqueness of the solution to a system of three polynomials in three unknowns. The fact that there are three equations in three unknowns is necessary but not sufficient for uniqueness of solutions, and given the nonlinearity of the model, showing the Jacobian is full rank is only sufficient for local identification.

Buchenburger's Algorithm (see Cox et al., 2007) is a finite-step algorithm that can find all roots of this system of equations, but in practice there is no guarantee the number of steps is small enough to be computationally feasible. In a relatively small system such as this, Mathematica has no trouble showing that known values of r_1, r_2, and r, with the additional normalization that $p_H > p_L$, are sufficient to establish globally unique values of ν, p_H, and p_L.

The final step is to extend identification to including observable characteristics in p_H and p_L in a manner consistent with our model. The simplest argument is to consider a set of workers with identical X_i and the same π_{it} for at least the first three periods. The OOLF process

$$O(\pi_{it}, X_i, \tau_i) = \Phi\left(\gamma_0 + X_i'\gamma + \gamma_\pi \pi_{it} + \gamma_\tau \cdot \mathbf{1}\{\tau_i = 1\}\right), \tag{31}$$

can take on two different unobserved values for each given set of observables (X_i and π_{it}), but this unobserved value is constant within individuals. Calling these two probabilities p_H and p_L, we can identify the unobserved type probabilities and the dropout probabilities per type. Given identification of ν, p_H, and p_L for these values of X_i and π_{it}, this means the left hand side of equation (11) is known and we can write

$$\Phi^{-1}\left(O(\pi_{it}, X_i, \tau_i)\right) = \gamma_0 + X_i'\gamma + \gamma_\pi \pi_{it} + \gamma_\tau \cdot \mathbf{1}\{\tau_i = 1\}.$$

By repeating the argument above for a set of workers with observable vectors $[1, X_i, \pi_{it}]$ that form a full rank matrix, the γ are all uniquely determined and the OOLF parameters are identified.

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. Here the time-varying hazard rates for labor force dropout can best be seen in an extreme example: If $p_H = 1$ and $p_L = 0$, the initial unconditional (on type) probability of a worker dropping out is $1-\nu$ in the first period, and in all later periods no additional workers drop out. Contrast this with $p_H = 1-\nu$ and $p_L = 1-\nu$, so the probability of dropout in period 1 is still $\nu \cdot (1-\nu) + (1-\nu) \cdot (1-\nu) = 1-\nu$, but workers would continue to drop out with a constant hazard each period. This example shows the essential point that without panel data on at least 2 periods we would never be able to tell the two models apart, and indeed without three periods we would not have enough equations to identify all of ν, p_H, and p_L. Our extension to observed characteristics in the dropout probabilities depends on enough workers having similar observables and careers so we can look at the within-group pattern of dropout rates to determine the type probabilities and dropout probabilities, then comparing across workers with different observables to see the effects of observables on within-group dropout rates.

C.2 Occupational Offer Distribution

Our econometric model of occupational transitions relates individual demographic characteristics X_i to occupational productivity outcomes π_{it} through equation (8), where we assume that workers are fired with probability $q(X_i)$, receive a new job offer with probability $p(X_i)$, and offers are drawn from the distribution $\Pi(\pi|X_i)$. 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 as flexible functional forms as possible 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_i; we can repeat the argument for any given X_i. 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

$$\Pr(\pi_1 = \pi_0|d = 0) = \left(1 - q\right)\left(\left(1 - p\right) + p\Pi(\pi_0|d = 0)\right). \tag{32}$$

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 $\Pi(\bar{\pi}) = 1$ and this reduces to

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

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 not moving jobs is

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

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 ((32)) gives

$$\Pi(\pi_0 | d = 0) = \frac{\Pr(\pi_1 = \pi_0 | d = 0)}{p\left(1-q\right)} - \frac{1-p}{p}. \tag{33}$$

The right hand side is simply data ($\Pr(\pi_1 = \pi_0 | d = 0)$) and known parameters. Given an original job, we can now remove 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 are 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 (33).

For this identification argument, we only required a limited part of the data: the type of the first job and one observation on whether the individual remained in that job or not. The duration of the first job, the type 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.3 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 can be generated to control for this differential dropout.

First, assume that the OOLF process is identified from Appendix C.1 above, so we have a known dropout probability conditional on both observables and latent type. Recall that $\pi = -1$ indicates exiting the labor market, and $O(\pi_{i,t-1}, X_i, z_i, \tau)$ is the probability of exiting the labor market in a given period. In this function, X_i 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 $\Psi\left(X_i, z_i, \pi_{(H)}\right)$. 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. We denote it as $\Gamma_t\left(X_i, z_i, \pi_{(H)}, \tau\right)$, and it is calculated as follows:

$$\Gamma_t\left(X_i, z_i, \pi_{(H)}, \tau\right) = \prod_{s=1}^{t-1}\left(1 - O\left(\pi_{is}, X_i, z_i, \tau\right)\right) . \tag{34}$$

Consider the wage equation:

$$\log\left(W_{it}\right) \equiv w_{it} = \beta_0 + X_i'\beta + \gamma \cdot \pi_{it} + \beta_\tau \cdot \mathbf{1}\{\tau = 1\} + \delta_1 t + \delta_2 t^2 + e_{it}. \tag{35}$$

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_i and z_i in current occupation π_{it} with occupational history $\pi_{(H)}$ is

$$E_t\left[w_{it}|X_i, z_i, \pi_{it}, \pi_{(H)}\right] = \beta_0 + X_i'\beta + \gamma \cdot \pi_{it} + \beta_\tau \cdot E\left[\mathbf{1}\{\tau = 1\}|X_i, z_i, \pi_{it}, \pi_{(H)}\right] + \delta_1 t + \delta_2 t^2. \tag{36}$$

Given X_i and $\pi_{(H)}$, today's job quality π_{it} is independent of τ since it is drawn after the last dropout shock. Also, τ is independent of unemployment given X. Instead of conditioning on the job quality, we can just condition on the fact that the worker is employed, which we can write as $\pi_{it} > 0$:

$$E_t\left[w|X_i, z_i, \pi_{it}, \pi_{(H)}\right] = \beta_0 + X_i'\beta + \gamma \cdot \pi_{it} + \beta_\tau E\left[\mathbf{1}\{\tau = 1\}|X_i, z_i, \pi_{(H)}, \pi_{it} > 0\right] + \delta_1 t + \delta_2 t^2. \tag{37}$$

But we may have $E_t\left[\mathbf{1}\{\tau = 1\}|X_i, z_i, \pi_{(H)}, \pi_{it}\right] = \Pr_t\left(\tau = 1|X_i, z_i, \pi_{(H)}, \pi_{it}\right)$ 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_i, which would bias estimation of the wage parameters. Using Bayes' rule:

$$\begin{aligned}
\Pr_t\left(\tau = 1|X_i, z_i, \pi_{(H)}, \pi_{it}\right) &= \frac{\Pr_t\left(\pi_{it} > 0|\tau = 1, X_i, z_i, \pi_{(H)}\right)\Pr(\tau = 1)}{\Pr_t\left(\pi_{it} > 0|X_i, z_i, \pi_{(H)}\right)} \\
&= \Gamma_t\left(X_i, z_i, \pi_{(H)}, \tau = 1\right) \cdot \nu \\
&\equiv H_t\left(X_i, \pi_{(H)}, z_i\right). \qquad (38)
\end{aligned}$$

Since Γ and ν are known from Section C.1 above and Ψ is directly observable, we could calculate

H and directly include it as a regressor in an augmented wage equation:

$$w^{\star}\left(X_i, \pi_{it}, z_i, \pi_{(H)}\right) \equiv \beta_0 + X_i'\beta + \gamma \cdot \pi_{it} + \beta_\tau \cdot H_t\left(X_i, \pi_{(H)}, z_i\right) + \delta_1 t + \delta_2 t^2 + e_{it}. \quad (39)$$

Now, the only unobservable is e_i, 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_i and $\left(t, t^2\right)$ over the whole support of x_i. We could identify β_3 directly by taking the two different values of z_i when calculating H conditioning on some x_i 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.